ระบบคอมพิวเตอร์ช่วยวินิจฉัยเนื้องอกสมองบนพื้นฐานการประมวลผลภาพ

A COMPUTER-AIDED DIAGNOSIS SYSTEM OF BRAIN TUMOR BASED ON IMAGE PROCESSING

พงษ์ศักดิ์ พิลาอุ่น

PONGSAK PILAOON

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2567

KMITL-2025-EN-D-018-044

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A COMPUTER-AIDED DIAGNOSIS SYSTEM OF BRAIN TUMOR BASED ON IMAGE PROCESSING

PONGSAK PILAOON

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

DOCTOR OF ENGINEERING IN ELECTRICAL ENGINEERING

SCHOOL OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2024

KMITL-2025-EN-D-018-044

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2024

SCHOOL OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	ระบบคอมพิวเตอร์ช่วยวินิจฉัยเนื้องอกสมองบนพื้นฐานการประมวลผลภาพ
นักศึกษา	นายพงษ์ศักดิ์ พิลาอุ่น
รหัสประจำตัว	61601029
ปริญญา	วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า
พ.ศ.	2567
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผู้ช่วยศาสตราจารย์ ดร.นพดล มณีรัตน์

บทคัดย่อ

เนื้องอกสมองกลิโอมา (Glioma) ชนิดร้ายแรงสุดเรียกว่ากลัยโอบลาสโตมามัลติฟอร์ม (Glioblastoma Multiforme) มีการแพร่กระจายอย่างรวดเร็ว การตรวจพบอย่างรวดเร็วและรักษาอย่างทันท่วงทีจะเพิ่มโอกาสรอดชีวิตได้ การสร้างภาพด้วยเรโซแนนซ์แม่เหล็ก (Magnetic Resonance Imaging) เป็นเทคนิคมาตรฐานที่ไม่รุกรานและใช้อย่างกว้างขวางในคลินิกเพื่อวินิจฉัยเนื้องอกสมอง เจ้าหน้าที่ทางการแพทย์วินิจฉัยเนื้องอกสมองโดยใช้ภาพ MRI เพื่อแปลความตำแหน่งที่ตั้ง ขนาดเนื้องอก และระดับเกรดเนื้องอกเพื่อการรักษาที่แม่นยำและได้ผลดีที่สุด เนื่องจากการวินิจฉัยโรคในปัจจุบันต้องใช้เวลาค่อนข้างมากและมีแนวโน้มที่จะเกิดความผิดพลาดจากการมองเห็นของมนุษย์ได้ และรวมทั้งต้องดำเนินการโดยผู้ที่มีความเชี่ยวชาญทางการแพทย์ซึ่งเป็นข้อจำกัดและขาดแคลนในพื้นที่ห่างไกล ดังนั้นงานวิจัยนี้มุ่งเน้นที่จะพัฒนาเครื่องมือที่แม่นยำและมีประสิทธิภาพเพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์ในการจำแนกเนื้องอกสมองกลิโอมาโดยเฉพาะ GBM ชนิดร้ายแรง เกรด 4 โดยใช้เทคนิคการมองเห็นของคอมพิวเตอร์ ซึ่งในงานวิจัยนี้ได้นำเสนอกระบวนวิธีการสองแบบ

วิธีการแรกใช้ภาพ MRI จากกลุ่มข้อมูล Repository of Molecular Brain Neoplasia Data (REMBRANDT) ประกอบด้วยภาพสมองปกติ 44 ภาพและเนื้องอกสมองอีก 111 ภาพ ซึ่งเป็นกลุ่มข้อมูลขนาดเล็กและไม่สมดุลกัน ใช้การเตรียมกระบวนการโดยปฏิบัติการทางสัณฐานวิทยา (Morphological operation) เพื่อแยกแยะส่วนกะโหลกออกจากพื้นที่สมอง จากนั้นภาพพื้นที่สมองที่แยกออกถูกใช้เพื่อการแบ่งออกเป็นส่วนซึ่งอยู่บนพื้นฐานของวิธีการใช้ค่าเทรชโฮลด์ค่าเดียวกันทั้งภาพ (Global Threshold) พื้นที่รูปภาพเนื้องอกที่ได้หลังการแบ่งแยกนำไปสกัดฟีเจอร์โดยใช้เมทริกซ์สัมประสิทธิ์ระดับสีเทา (Gray Level Co-occurrence Matrix, GLCM) ในขั้นต่อไปจะได้รับค่าฟีเจอร์และจัดเรียงรูปแบบตารางเพื่อใช้เป็นกลุ่มข้อมูลการอบรมและทดสอบ ตัวจำแนกความแตกต่างแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM) เชิงเส้นถูกนำมาใช้งานเพื่อการจัดหมวดหมู่ และได้รับความแม่นยำ (Accuracy) 96.875 เปอร์เซ็นต์ ความไว (Sensitivity) 95 เปอร์เซ็นต์ และความเฉพาะเจาะจง (Specificity) 100 เปอร์เซ็นต์ ตามลำดับ พร้อมกันนี้ได้นำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักการโครงข่ายประสาทเทียมคอนโวลูชันแบบ GoogLeNet และ AlexNet กับกลุ่มข้อมูลเดียวกัน เพื่อช่วยให้การสกัดฟีเจอร์อย่างอัตโนมัติ โดยได้รับความเที่ยงตรง 80.85 เปอร์เซ็นต์ และ 93.62 เปอร์เซ็นต์ ตามลำดับ

วิธีการที่สองนั้นได้รับภาพถ่ายทางการแพทย์เพิ่มเติมเพื่อปรับปรุงสมรรถนะการทำงานและใช้หลักการเรียนรู้เชิงลึกโดยโครงข่ายประสาทเทียมคอนโวลูชัน (Convolution Neural Network) เพื่อจัดหมวดหมู่ ภาพถ่าย MRI จากกลุ่มข้อมูล REMBRANDT และ BraTS2021 ซึ่งประกอบด้วย 1,856 ภาพถูกนำมาใช้เพื่อการปรับปรุงความแม่นยำในการทำนายที่ดีขึ้น โดยในขั้นตอนเตรียมกระบวนการปฏิบัติการทางสัณฐานวิทยาและการปรับสมดุลฮิสโตแกรมแบบปรับตัวแบบจำกัดคอนทราส (Contrast Limited Adaptive Histogram Equalization, CLAHE) ถูกนำมาใช้เพื่อแยกส่วนกะโหลกออกและเพิ่มคอนทราสของภาพก่อนดำเนินการขั้นถัดไป เพื่อเปรียบเทียบและปรับปรุงประสิทธิภาพการทำนาย งานวิจัยส่วนนี้นำเสนอการเตรียมกลุ่มข้อมูลสองแบบคือแบบไม่สมดุลและสมดุลกันโดยใช้การเพิ่มปริมาณภาพอบรม แต่ละวิธีถูกอบรมโดยใช้หลักการถ่ายโอนการเรียนรู้แบบ CNN 8 แบบด้วยกัน ผลลัพธ์การทำนายแสดงให้เห็นว่ากลุ่มข้อมูลที่สมดุลกันทำให้สรรถนะการทำนายได้รับการปรับปรุงอย่างมีนัยสำคัญ โดยอุปกรณ์จัดหมวดหมู่ที่ดีสุดคือ InceptionV3 ที่ใช้กลุ่มข้อมูลสมดุลกันทั้งหมด 2,056 รูปภาพ ซึ่งใช้เวลาอบรมกลุ่มข้อมูล 25 นาที 15 วินาที และทดสอบกลุ่มข้อมูล 372 รูปภาพด้วยเวลาเพียง 5.878 วินาทีเท่านั้น ซึ่งได้รับความแม่นยำถึง 99.73 เปอร์เซ็นต์ ความไว 99.61 เปอร์เซ็นต์ ความเฉพาะเจาะจง 100 เปอร์เซ็นต์ และพื้นที่ภายใต้เส้นโค้ง Receiver Operating Characteristic (AUC) เท่ากับ 1 ตามลำดับ นอกจากนั้นได้ศึกษาเพิ่มเติมการนำหลักการเรียนรู้แบบผสมผสาน (Ensemble Learning) ซึ่งอยู่บนพื้นฐานของการโหวตเสียงส่วนมากเพื่อปรับปรุงประสิทธิภาพการทำนายผล และได้นำเสนอ Gradient weighted Class Activation Mapping (Grad-CAM) เพื่อแสดงภาพการมองเห็นและจำแนกความแตกต่างพื้นที่รูปภาพ MRI ที่ได้ทำนายผลแล้วว่าส่วนไหนที่มีผลต่อการทำนายมากสุด วิธีที่นำเสนอจัดหมวดหมู่เนื้องอกสมองได้อย่างแม่นยำ และในอนาคตข้างหน้าเครื่องมือที่แม่นยำ มีประสิทธิภาพ และจำแนกเนื้องอกสมองได้ตั้งแต่เกรด 1 ถึง เกรด 4 จะถูกพัฒนาอย่างต่อเนื่อง รวมทั้งจะพิจารณาทดสอบผลกระทบต่อการปฏิบัติทางการแพทย์ (Clinical Implications) เพื่อดูว่าผลจากงานวิจัยมีความหมายอย่างไรต่อการดูแลรักษาผู้ป่วยในทางปฏิบัติ เพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์ในการวินิจฉัยเนื้องอกสมองกลิโอมาโดยใช้การประมวลผลภาพ (Image Processing)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis	A COMPUTER-AIDED DIAGNOSIS SYSTEM OF BRAIN TUMOR BASED ON IMAGE PROCESSING
Student	Mr. Pongsak Pilaoon
Student ID.	61601029
Degree	Doctor of Engineering
Program	Electrical Engineering
Year	2024
Thesis Advisor	Assist. Prof. Dr. Noppadol Maneerat

ABSTRACT

Glioma brain tumors arise from glial cells in the brain and spinal cord and have been classified into four grades. Gliomas grade IV also known as glioblastoma multiforme (GBM) are the most malignant primary brain tumor and have very rapid growth. The earlier detection of glioma brain tumor and immediate treatment will increase the survival opportunity. Magnetic resonance imaging (MRI) has been a standard technique that is non-invasive and widely used in clinics for brain tumor diagnosis. The medical staff diagnosed the brain tumor using the MRI image to interpret the location, blob size, and tumor grading. The current disease diagnosis is time-consuming and tends to errors happening by human visual sight and shall be performed by medical expertise which is a constraint in upcountry or remote areas. Therefore, this research aims to develop accurate and effective tools to assist the medical staff in the classification of glioma brain tumors especially GBM grade IV that is the most malignant using artificial intelligence (AI) and computer vision techniques.

In this research, we have proposed 2 methodologies for brain tumor classification. The first method MRI image from the REMBRANDT dataset contains 44 normal and 111 brain tumors that are small and imbalance datasets. The morphological operation preprocessing was performed to remove the skeleton from the brain region. Then the separated brain region image was used for segmentation based on the global threshold method. The obtained tumor blob extracted the feature using Gray Level Co-occurrence Matrix (GLCM) feature extraction in the next step. We received the extracted features in tabulated format for training and testing datasets. A linear SVM classifier was implemented for the classification task and finally, we obtained an accuracy of 96.875 percent, sensitivity of 95 percent, and specificity of 100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

percent respectively. Meanwhile, with the existing dataset, we have implemented the convolution neural network GoogLeNet and AlexNet to classify the glioma brain tumor. Features are automatically extracted instead of handcraft and obtained an accuracy of 80.85 percent and 93.62 percent respectively.

The second methodology, the medical image dataset, was obtained increasingly to improve performance. We implemented the deep learning Convolution Neural Networks for classification tasks. The MRI images from REMBRANDT and BraTS2021 datasets that contain 1,856 images were utilized to improve the accuracy. In the preprocessing stage, the morphological operation and contrast-limited Adaptive Histogram Equalization were performed for the removal skeleton and enhancement of the contrast of images. The two approaches were conducted comprising imbalanced and balanced datasets using image augmentation. Each approach was trained based on transfer learning CNN with various 8 types of transfer learning. The prediction result showed that the balanced dataset significantly improved performance. The best classifier is InceptionV3 with a balanced dataset that obtained an accuracy of 99.73 percent, a sensitivity of 99.61 percent, a specificity of 100 percent, and an area under the Receiver Operating Characteristic curve (AUC) of 1 respectively. In the meantime, the training duration for a balanced dataset comprising 2,056 images was 25 minutes and 15 seconds, while the testing inference time for a dataset of 372 images was 5.878 seconds. The ensemble learning using majority hard and soft voting was further implemented for investigating the performance of prediction. Additionally, we proposed a gradient-weighted class activation mapping (Grad CAM) to visualize the discriminative image region of the predicted MRI image. The proposed methodology has precisely classified glioma brain tumors and we envision the development of accurate and effective tools to assist medical staff in classifying glioma brain tumors using computer vision techniques. In the future, we plan to enhance the classifier to detect glioma tumors ranging from grade I to IV. The clinical implications of this development will be applied to assist medical staff and improve the practical diagnosis of glioma brain tumors through image processing.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จได้ด้วยความกรุณาจากอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. นพดล มณีรัตน์ ที่ให้ความช่วยเหลือ ให้คำชี้แนะ ช่วยแก้ปัญหาตลอดจนให้ความรู้และประสบการณ์ที่ดีแก่ข้าพเจ้า

ขอขอบคุณคณะกรรมการทุกท่านที่ให้คำชี้แนะแนวทาง เพื่อพัฒนาและแก้ไขงานวิจัยในวิทยานิพนธ์เล่มนี้

สำหรับคุณงามความดีอันใดที่เกิดจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบให้กับบิดามารดาซึ่งเป็นที่รักและเคารพยิ่ง ตลอดจนครูอาจารย์ที่เคารพทุกท่านที่ได้ถ่ายทอดความรู้และประสบการณ์ที่ดีให้แก่ข้าพเจ้า

พงษ์ศักดิ์ พิลาอุ่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ III

กิตติกรรมประกาศ V

สารบัญ VI

สารบัญตาราง X

สารบัญรูป XI

รายการคำย่อและสัญลักษณ์ XIII

บทที่ 1 บทนำ 1

1.1 ข้อมูลทางการแพทย์ 1

1.2 การใช้ระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค 3

1.3 ความมุ่งหมายและวัตถุประสงค์ของงานวิจัย 4

1.4 ขอบเขตงานวิจัย 5

1.5 ขั้นตอนการวิจัย 6

1.6 รายละเอียดเนื้อหาในแต่ละบท 6

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 7

2.1 ข้อมูลทางการแพทย์เกี่ยวกับเนื้องอกสมองกลิโอมา 7

2.1.1 กลุ่มข้อมูล REMBRANDT 7

2.1.2 กลุ่มข้อมูล BraTS2021 8

2.2 ทฤษฎีที่เกี่ยวข้องกับการจัดหมวดหมู่เนื้องอกสมองด้วยการเรียนรู้ของเครื่องจักร 8

2.2.1 การสกัดฟีเจอร์โดยใช้วิธีการเมทริกซ์การเกิดร่วมของระดับสีเทา 8

2.2.2 การจัดหมวดหมู่ด้วยการเรียนรู้เครื่องจักรซัพพอร์ตเวกเตอร์แมชชีน 11

2.2.2.1 อัลกอริทึม Optimal hyperplane 11

2.2.2.2 ไฮเปอร์เพลนแบบซอฟท์มาจิน Soft Margin 12

2.3 ทฤษฎีที่เกี่ยวข้องกับการจัดหมวดหมู่เนื้องอกสมองโดยใช้การเรียนรู้เชิงลึก 13

2.3.1 การประมวลผลภาพดิจิทัล (Digital Image Processing) 13

2.3.1.1 ปฏิบัติการทางสัณฐานวิทยา (Morphological Operation) 13

2.3.1.2 การลบหรือหักล้างภาพ (Image Subtraction) 13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.3.1.3 Contrast Limited Adaptive Histogram Equalization 14
2.3.1.4 การแปลงเรขาคณิตของภาพ 15
2.3.2 การเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน (CNNs) 16
2.3.2.1 การทำพูลลิง (Pooling) 19
2.3.2.2 Stochastic Gradient Descent with Momentum 20
2.3.3 การเรียนรู้แบบถ่ายโอน (Transfer Learning) 21
2.3.3.1 VGG16 22
2.3.3.2 VGG19 23
2.3.3.3 ResNet50 / ResNet101 24
2.3.3.4 Xception 26
2.3.3.5 InceptionV3 27
2.3.3.6 InceptionResNetV2 30
2.3.3.7 DenseNet201 31
2.3.4 Gradient weighted Class Activation Mapping (Grad-CAM) 32
2.4 งานวิจัยที่เกี่ยวข้อง 33
บทที่ 3 การวินิจฉัยเนื้องอกสมองด้วยการเรียนรู้เครื่องจักรและโครงข่ายประสาทเทียม CNN 35
3.1 การวินิจฉัยเนื้องอกสมองกลิโอมาด้วยการเรียนรู้ของเครื่องจักร 35
3.1.1 กระบวนวิธีการทดลอง 35
3.1.2 การออกแบบการทดลอง (Experimental Setup) 37
3.1.3 ขั้นตอนการเตรียมกระบวนการ 37
3.1.3.1 การปฏิบัติการทางสัณฐานวิทยา (Morphological) 38
3.1.3.2 การลบหรือหักล้างภาพ (Image Subtraction) 38
3.1.4 การแบ่งออกเป็นส่วน (Segmentation) 39
3.1.5 การสกัดฟีเจอร์ด้วยวิธีเมทริกซ์การเกิดร่วมของระดับสีเทา 40
3.1.6 การจัดหมวดหมู่ด้วยการเรียนรู้แบบซัพพอร์ตเวกเตอร์แมชชีน 41
3.2 การวินิจฉัยเนื้องอกสมองกลิโอมาด้วยโครงข่ายประสาทเทียม CNN 43
3.2.1 กระบวนวิธีการทดลอง 43
3.2.2 ขั้นตอนการเตรียมกระบวนการโดยสร้างพื้นที่จัดเก็บรูปภาพส่วนขยาย 46
3.2.3 การจัดหมวดหมู่เนื้องอกสมองโดยใช้ GoogLeNet 46
3.2.3.1 โครงข่ายที่ได้รับการอบรมไว้ก่อนแบบ GoogLeNet 47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

3.2.3.2 อบรมโครงข่ายและการทำนายเอาท์พุท 47
3.2.4 การจัดหมวดหมู่เนื้องอกสมองโดยใช้ AlexNet 47
3.2.4.1 เลเยอร์โครงข่ายที่ได้รับการอบรมแล้วแบบ AlexNet 49
3.2.4.2 การอบรมโครงข่ายและทำนายเอาท์พุท 49
บทที่ 4 การวินิจฉัยเนื้องอกสมองกลิโอมาที่มีกลุ่มข้อมูลขนาดใหญ่ด้วยโครงข่ายประสาทเทียมคอนโวลูชัน CNNS 51
4.1 กระบวนวิธีการทดลอง 51
4.2 การออกแบบตั้งค่าการทดลอง 52
4.2.1 กระบวนวิธีการที่ 1 การอบรมโดยใช้กลุ่มข้อมูลไม่สมดุลกัน 55
4.2.2 กระบวนวิธีการที่ 2 การอบรมโดยใช้กลุ่มข้อมูลที่สมดุลกัน 58
4.2.2.1 Morphological Operation Preprocessing 60
4.2.2.2 Contrast Limited Adaptive Histogram Equalization 62
4.2.2.3 การเสริมภาพเพิ่มจำนวนกลุ่มข้อมูลอบรม (Augmentation) 63
4.2.2.4 การอบรมโดยใช้โครงข่ายประสาทเทียมคอนโวลูชัน 64
4.2.3 กระบวนวิธีการที่ 3 การจัดหมวดหมู่โดยใช้ Majority Voting 66
4.2.4 การสร้างภาพการมองเห็นพื้นที่ซึ่งถูกแยกแยะโดยใช้ GRAD-CAM 66
4.3 การประเมินผลสมรรถนะ 68
4.3.1 การวิเคราะห์เชิงสถิติโดยใช้ p-value 70
บทที่ 5 การทดลองและการวิเคราะห์ผลการทดลอง 72
5.1 ผลการทดลองการวินิจฉัยเนื้องอกสมองที่มีกลุ่มข้อมูลขนาดเล็ก 72
5.1.1 ผลการทดลองการวินิจฉัยเนื้องอกสมองด้วยซัพพอร์ตเวกเตอร์แมชชีน 72
5.1.2 ผลการทดลองวินิจฉัยเนื้องอกสมองด้วยการเรียนรู้แบบถ่ายโอน 73
5.2 ผลการทดลองการวินิจฉัยเนื้องอกสมองที่กลุ่มข้อมูลมีจำนวนมากขึ้นโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน 76
5.2.1 ผลการทดลองวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน 76
5.2.2 ผลการทดลองวิธีการที่ 2 กลุ่มข้อมูลที่ซึ่งสมดุลกัน 81
5.2.3 ผลการทดลองวิธีการที่ 3 การเรียนรู้แบบ Ensemble Learning 90
5.2.4 ผลการวิเคราะห์เชิงสถิติการทำนายผลด้วย p-value 90
5.2.5 ผลการทดลองการมองเห็นพื้นที่รูปภาพซึ่งถูกแยกแยะ 91

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

5.2.6 ผลลัพธ์การทดลองเปรียบเทียบกับงานเกี่ยวข้องที่ผ่านมา 92

บทที่ 6 สรุปผลการทดลอง ข้อเสนอแนะ และแนวทางพัฒนา 96

6.1 สรุปผลการทดลอง 96

6.2 ข้อเสนอแนะ และแนวทางพัฒนาในอนาคต 97

เอกสารอ้างอิง 99

ภาคผนวก ก บทความที่ได้รับการตีพิมพ์และเผยแพร่ 104

ประวัติผู้เขียน 145

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ หน้า

ตารางที่ 2.1 อัลกอริทึม Stochastic Gradient Descent with Momentum (SGDM)..... 21
ตารางที่ 2.2 คุณลักษณะของโครงข่ายการเรียนรู้แบบถ่ายโอนที่ใช้ในงานวิจัยนี้ 21
ตารางที่ 2.3 โครงสร้างเลเยอร์ของโครงข่ายแบบ VGG แต่ละแบบ 23
ตารางที่ 2.4 สถาปัตยกรรมของ ImageNet โดยโครงสร้างของบล็อกถูกระบุในวงเล็บ..... 26
ตารางที่ 2.5 โครงสร้างสถาปัตยกรรมแต่ละเลเยอร์ของ InceptionV3 30
ตารางที่ 3.1 รูปภาพ MRI ที่ถูกแปลงเป็นรูปแบบ JPEG และใช้ในการทดลอง..... 37
ตารางที่ 3.2 ตัวอย่างของฟีเจอร์ที่ถูกสกัดแยกออกมาโดยใช้วิธี GLCM..... 41
ตารางที่ 4.1 รูปภาพ MRI ที่ถูกแปลงเป็นฟอร์แมท JPEG ที่ได้จากแต่ละฐานข้อมูล..... 53
ตารางที่ 4.2 เทคนิคการเสริมภาพสำหรับรูปภาพสมองปกติซึ่งเป็นคลาสส่วนน้อย 54
ตารางที่ 4.3 การผสมผสานรูปภาพสำหรับกลุ่มข้อมูลของแต่ละวิธีการ 54
ตารางที่ 4.4 ตัวอย่าง Ground Truth Label กลุ่มข้อมูลเพื่อใช้อบรมโครงข่าย 64
ตารางที่ 5.1 ผลลัพธ์การวินิจฉัยเนื้องอกด้วยซัพพอร์ตเวกเตอร์แมชชีน..... 73
ตารางที่ 5.2 ผลลัพธ์การทำนายเนื้องอกสมองด้วยการเรียนรู้แบบถ่ายโอน 74
ตารางที่ 5.3 ผลลัพธ์การเปรียบเทียบกับงานวิจัยที่ผ่านมาสำหรับกลุ่มข้อมูลขนาดเล็ก..... 74
ตารางที่ 5.4 ผลลัพธ์การอบรมกลุ่มข้อมูลด้วยวิธีการที่ 1 กลุ่มข้อมูลที่ไม่สมดุลกัน 76
ตารางที่ 5.5 ผลลัพธ์สมรรถนะการจัดหมวดหมู่ของกระบวนวิธีการที่ 1 กลุ่มข้อมูลที่ไม่สมดุล..... 80
ตารางที่ 5.6 ผลลัพธ์การอบรมกลุ่มข้อมูลของกระบวนวิธีการที่ 2 กลุ่มข้อมูลที่สมดุลกัน..... 81
ตารางที่ 5.7 ผลลัพธ์สมรรถนะการจัดหมวดหมู่ด้วยวิธีการที่ 2 กลุ่มข้อมูลที่สมดุลกัน 87
ตารางที่ 5.8 ผลลัพธ์สมรรถนะการจัดหมวดหมู่ด้วยวิธีการที่ 3 การเรียนรู้แบบผสมผสาน..... 90
ตารางที่ 5.9 การวิเคราะห์เชิงสถิติสำหรับการทำนายโดยใช้ p-value 91
ตารางที่ 5.10 การเปรียบเทียบผลลัพธ์การทดลองกับงานในลักษณะเดียวกันที่ผ่านมา..... 93

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ หน้า

รูปที่ 1.1 ภาพเฉือนตามแนวแกนของภาพ MRI ของเนื้องอกกลิโอมาระดับรุนแรง 2
รูปที่ 1.2 ภาพการวิเคราะห์เนื้องอกสมองที่เกิดขึ้นในตำแหน่งต่างกัน 4
รูปที่ 2.1 วิธีการสร้างเมทริกซ์ Co-occurrence 9
รูปที่ 2.2 ภาพประกอบการอธิบายซัพพอร์ตเวกเตอร์แมชชีน 11
รูปที่ 2.3 ตัวอย่างเตรียมกระบวนการ 4 ขั้นตอนเพื่อให้ได้ภาพพื้นที่สมองเท่านั้น 14
รูปที่ 2.4 อัลกอริทึม CLAHE เพื่อปรับปรุงคอนทราสของภาพเนื้องอกสมอง 15
รูปที่ 2.5 วิธีการเสริมภาพสำหรับสมองปกติซึ่งเป็นกลุ่มข้อมูลส่วนน้อย 16
รูปที่ 2.6 ตัวอย่างของคอนโวลูชันสองมิติซึ่งไม่มีการสลับเคอร์เนล (Without Kernel-Flipping).... 18
รูปที่ 2.7 ตัวอย่างของคอนโวลูชันที่ใช้ค่าเคอร์เนล 3 x 3..... 18
รูปที่ 2.8 ตัวอย่างค่า Stride เท่ากับ 1..... 18
รูปที่ 2.9 ส่วนประกอบโดยทั่วไปของเลเยอร์โครงข่ายประสาทเทียมคอนโวลูชัน..... 19
รูปที่ 2.10 การทำงานของฟังก์ชัน Max Pooling 19
รูปที่ 2.11 หลักการ Stochastic Gradient Descent with Momentum (SGDM)..... 20
รูปที่ 2.12 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ VGG16..... 22
รูปที่ 2.13 ตัวอย่างสถาปัตยกรรมโครงข่ายสำหรับ ImageNet 25
รูปที่ 2.14 ฟังก์ชัน Residual เพื่อเพิ่มความลึก สำหรับ ImageNet 26
รูปที่ 2.15 โครงสร้างสถาปัตยกรรมของ Xception..... 27
รูปที่ 2.16 การแยกตัวประกอบ (Factorization) ของโมดูลใน InceptionV3 28
รูปที่ 2.17 การแยกตัวประกอบเชิงพื้นที่ (Spatial Factorization) ในโมดูล Inception 29
รูปที่ 2.18 การขยายฟิลเตอร์เพื่อใช้หลักการขนาดมิติที่สูงขึ้นในโมดูล Inception..... 29
รูปที่ 2.19 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ InceptionV3..... 29
รูปที่ 2.20 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ InceptionResNetV2.... 31
รูปที่ 2.21 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ Densenet201..... 31
รูปที่ 3.1 โฟลว์ชาร์ตการวินิจฉัยเนื้องอกด้วยการเรียนรู้แบบซัพพอร์ตเวกเตอร์แมชชีน..... 36
รูปที่ 3.2 ตัวอย่างของภาพ MRI ถูกแปลงเป็นฟอร์แมท JPEG..... 37
รูปที่ 3.3 ตัวอย่างของภาพพื้นที่สมองสำหรับขั้นเตรียมกระบวนการ 38
รูปที่ 3.4 แผนภูมิฮิสโตแกรมของภาพสมองในขั้นตอนการแบ่งออกเป็นส่วน..... 39
รูปที่ 3.5 ภาพที่ถูกแบ่งออกเป็นส่วนของทั้งเนื้องอกสมองและสมองปกติ..... 40
รูปที่ 3.6 ภาพประกอบการอธิบายซัพพอร์ตเวกเตอร์แมชชีน 41
รูปที่ 3.7 Scatter Plot และ Confusion Matrix ของกลุ่มข้อมูลที่ได้รับการอบรมแล้ว 42

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ หน้า
รูปที่ 3.8 โฟลว์ชาร์ตการวินิจฉัยเนื้องอกสมองด้วยการเรียนรู้แบบถ่ายโอน 45
รูปที่ 3.9 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชั่น (CNN) 46
รูปที่ 3.10 ผลลัพธ์การวิเคราะห์และชาร์ทการอบรมของโครงข่าย GoogLeNet 48
รูปที่ 3.11 การวิเคราะห์เลเยอร์และชาร์ทการอบรมโครงข่าย AlexNet 50
รูปที่ 4.1 บล็อกไดอะแกรมการจัดหมวดหมู่เนื้องอกสมองกลิโอมาด้วย CNNs 52
รูปที่ 4.2 กลุ่มข้อมูลที่รวบรวมสำหรับการจัดหมวดหมู่เนื้องอกสมองด้วย CNNs 53
รูปที่ 4.3 บล็อกไดอะแกรมสำหรับการอบรมกลุ่มข้อมูลซึ่งไม่สมดุลกัน 56
รูปที่ 4.4 โฟลว์ชาร์ตการอบรมโครงข่ายโดยใช้กลุ่มข้อมูลซึ่งไม่สมดุลกัน 57
รูปที่ 4.5 บล็อกไดอะแกรมการอบรมกลุ่มข้อมูลซึ่งสมดุลกันโดยการเสริมภาพ 58
รูปที่ 4.6 โฟลว์ชาร์ตการอบรมโครงข่ายโดยใช้กลุ่มข้อมูลซึ่งสมดุลกันด้วยการเสริมรูปภาพ 59
รูปที่ 4.7 ขั้นตอนการปฏิบัติการทางสัณฐานวิทยา 60
รูปที่ 4.8 ขั้นตอนของอัลกอริทึม CLAHE 63
รูปที่ 4.9 โฟลว์ชาร์ตขั้นตอนการอบรมโมเดลโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน 65
รูปที่ 4.10 Ensemble Learning โดยใช้การโหวต Majority Voting 66
รูปที่ 4.11 บล็อกไดอะแกรมเพื่อการมองเห็นรูปภาพการทำนายโดยใช้อัลกอริทึม Grad-CAM 67
รูปที่ 4.12 ผลการทำนายที่แสดงในรูปแบบ Confusion Chart 69
รูปที่ 4.13 วิธีคำนวณ p-value สำหรับ t-test การทดสอบทางเดียวด้านซ้าย 70
รูปที่ 4.14 ผลการคำนวณ p-value สำหรับ InceptionV3 71
รูปที่ 5.1 ผลการทำนายที่แสดงในรูปแบบ Confusion Chart 75
รูปที่ 5.2 ประสิทธิผลของการทำนายโดยใช้ GoogLeNet และ AlexNet 75
รูปที่ 5.3 เปรียบเทียบเวลาที่ใช้ในการอบรมซึ่งใช้กลุ่มข้อมูลไม่สมดุลกัน 77
รูปที่ 5.4 ตัวอย่างความก้าวหน้าการอบรมที่พล็อตไว้สำหรับวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน .. 78
รูปที่ 5.5 ตัวอย่างกราฟเส้นโค้ง ROC ที่ถูกพล็อตไว้สำหรับวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน..... 79
รูปที่ 5.6 แผนภูมิรูปภาพเปรียบเทียบผลลัพธ์การทำนายเมื่อใช้กับกลุ่มข้อมูลซึ่งไม่สมดุลกัน 80
รูปที่ 5.7 เปรียบเทียบเวลาที่ใช้ในการอบรมซึ่งใช้กลุ่มข้อมูลที่สมดุลกัน 82
รูปที่ 5.8 ความก้าวหน้าการอบรมสำหรับวิธีการที่ 2 กลุ่มข้อมูลซึ่งสมดุลกัน 86
รูปที่ 5.9 แผนภูมิรูปภาพเปรียบเทียบผลลัพธ์การทำนายเมื่อใช้กับกลุ่มข้อมูลซึ่งสมดุลกัน 87
รูปที่ 5.10 Confusion Chart ของการทำนายผลโดยใช้กลุ่มข้อมูลซึ่งสมดุลกัน 88
รูปที่ 5.11 ตัวอย่างพล็อตกราฟเส้นโค้ง ROC สำหรับวิธีการที่ 2 กลุ่มข้อมูลซึ่งสมดุลกัน 89
รูปที่ 5.12 อัลกอริทึม Grad-CAM และ LIME โดยใช้ InceptionV3 กับกลุ่มข้อมูลสมดุลกัน..... 92

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายการคำย่อและสัญลักษณ์

CAD	Computer-Aided Diagnosis
MRI	Magnetic Resonance Imaging
CNN	Convolution Neural Network
Grad-CAM	Gradient-weighted Class Activation Mapping
SVM	Support Vector Machine
CLAHE	Contrast Limited Adaptive Histogram Equalization
SGDM	Stochastic Gradient Descent with Momentum
GBM	Glioblastoma Multiforme
LIME	Local Interpretable Model-agnostic Explanation
DICOM	Digital Imaging and Communications in Medicine
NIfTI	Neuroimaging Informatics Technology Initiative
GLCM	Gray-Level Co-occurrence Matrix
REMBRANDT	Repository of Molecular Brain Neoplasia Data
BraTS	Brain Tumor Segmentation
ROC	Receiver Operating Characteristic
AUC	Area Under the ROC Curve

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

งานวิจัยนี้นำเสนอการวินิจฉัยเนื้องอกสมองบนพื้นฐานเทคนิคคอมพิวเตอร์วิชัน ซึ่งในบทนี้กล่าวถึงข้อมูลทางการแพทย์เกี่ยวกับโรคเนื้องอกสมอง การใช้ระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคจากภาพถ่ายทางการแพทย์ ความมุ่งหมายและวัตถุประสงค์ ขอบเขตและขั้นตอนดำเนินงาน และรายละเอียดเนื้อหาในแต่ละบทเพื่ออธิบายภาพรวมทั้งหมดของงานวิจัย

1.1 ข้อมูลทางการแพทย์

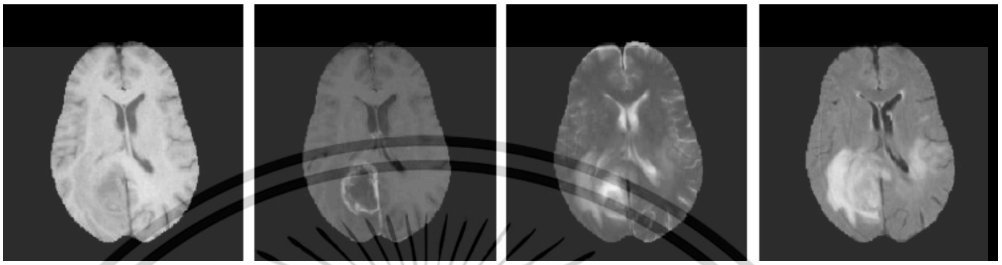
ด้วยความชุกชุมที่น้อยกว่า 1 เปอร์เซ็นต์ ในประชากรตะวันตก เนื้องอกสมองไม่ถือเป็นโรคที่เกิดทั่วไป อย่างไรก็ตามโรคนี้อยู่ในกลุ่มมะเร็งที่ทำให้ถึงแก่ชีวิตได้มากสุด การศึกษาเมื่อไม่นานมานี้ประมาณการในสหรัฐอเมริกาอัตราการเกิดขึ้นของเนื้องอกสมองที่เกิดขึ้นจากเซลล์ในสมองเอง (Primary Brain Tumor) อยู่ที่ 0.025 เปอร์เซ็นต์ โดยประมาณ หนึ่งในสามเป็นเนื้องอกชนิดร้ายแรง และที่เหลือเป็นเนื้องอกไม่อันตรายหรือไม่แน่ชัดว่าอาจจะรุนแรง เนื้องอกสมองแบ่งประเภทโดยอ้างอิงตามลักษณะการเกิดหรือระดับความรุนแรง เนื้องอกสมองที่เกิดจากเซลล์ในสมองเองเกิดขึ้นเองจากความผิดปกติของเซลล์ในสมอง ในขณะที่เนื้องอกสมองชนิดแพร่กระจาย (Metastatic Brain Tumor) เกิดจากมะเร็งอื่นในร่างกายที่แพร่มา เช่น ปอด ลำไส้ใหญ่ และทรวงอก เป็นต้น

องค์การอนามัยโลก (WHO) แบ่งเนื้องอกสมองเป็นสี่ระดับด้วยกันโดยอ้างอิงตามการรุกรานและความร้ายแรงของโรค กลิโอมา (Glioma) เป็นเนื้องอกสมองที่เกิดจากเซลล์ในสมองในวัยผู้ใหญ่และเชื่อว่าเป็นสัดส่วนถึง 70 เปอร์เซ็นต์ ของเนื้องอกสมองที่เกิดจากเซลล์ในสมองเองชนิดร้ายแรงที่เกิดจากเซลล์เกลีย (Glial Cells) ซึ่งมันถูกแบ่งออกเป็นสี่ระดับด้วยกันและเกรดที่สูงสุดคือเกรด 3 และเกรด 4 ซึ่งถือเป็นเนื้องอกชนิดร้ายแรงซึ่งนำไปสู่การสูญเสียชีวิตได้ กลิโอมาเกรดสี่เรียกอีกอย่างว่า กลิโอบลาสโทมามัลติฟอร์ม (GBM) ซึ่งเป็นเนื้องอกสมองที่เกิดจากเซลล์ในสมองเองระดับที่ร้ายแรงที่สุดและเติบโตลุกลามอย่างรวดเร็ว เนื้องอกนี้แทรกซึมและแพร่กระจายไปพร้อมกับทางเดินเส้นประสาทที่เป็นสีขาว (The White Matter Fiber Tracts) และแสดงรูปร่างท่อที่ผิดปกติและแกนที่เน่าตายเฉพาะส่วน และพบอาการบวมน้ำเกิดขึ้นรอบๆ ซึ่งนำไปสู่การเกิดก้อนในสมองซึ่งมีอิทธิพลต่อเนื้อเยื่อส่วนดีของสมอง (Healthy Tissues) [1], [2] ระยะเวลามีชีวิตรอดโดยเฉลี่ยสำหรับผู้ป่วย GBM คือหนึ่งปี ตัวเลือกการรักษาสำหรับกลิโอมารวมทั้งการผ่าตัด การบำบัดด้วยรังสี เคมีบำบัด หรือรวมกันหลายวิธี

การสร้างภาพด้วยเรโซแนนซ์แม่เหล็ก (MRI) เป็นเทคนิคมาตรฐานที่ไม่รุกรานและใช้อย่างกว้างขวางในคลินิกเพื่อวินิจฉัยเนื้องอกสมอง เพื่อสร้างให้เห็นความแตกต่างและตัดกันของภาพเนื้อเยื่อส่วนดีและเนื้องอก การปรับเปลี่ยนเวลากระตุ้น (Excitation) และเวลาทำซ้ำ (Repetition) ถูกนำมาใช้เพื่อทำเครื่องมืออเนกประสงค์สำหรับโครงสร้างรูปภาพที่สนใจแบบต่างๆ วิธีการทางคลินิก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในปัจจุบัน ลำดับรูปภาพ MRI ต่างกันถูกนำมาใช้สำหรับวินิจฉัยเนื้องอกสมอง ลำดับหรือรูปแบบดังกล่าวประกอบด้วย T_1-weighted MRI (T_1), T_1-weighted MRI with Contrast Enhancement (T_{1C}), T_2-weighted MRI (T_2), T_2-weighted MRI with Fluid Attenuated Inversion Recovery (T_{2FLAIR}) รูปภาพที่ 1.1 แสดงภาพเฉือนตามแนวแกนของลำดับหรือรูปแบบ MRI มาตรฐานของผู้ป่วยเนื้องอกกลิโอมา



ก. T_1 ข. T_{1C} ค. T_2 ง. T_{2FLAIR}

รูปที่ 1.1 ภาพเฉือนตามแนวแกนของภาพ MRI ของเนื้องอกกลิโอมาระดับรุนแรง

MRI แบบ T_1-weighted ใช้เพื่อวิเคราะห์โครงสร้างและอธิบายประกอบอย่างง่ายของเนื้อเยื่อที่ปกติ ด้วยภาพ MRI แบบ T_1-weighted พร้อมทั้งการเพิ่มการตัดกันของภาพ (Contrast Enhancement) ทำให้ขอบเขตเนื้องอกปรากฏชัดเจนสว่างกว่า เนื้อส่วนที่เน่าตายและพื้นที่เนื้องอกถูกแบ่งแยกออกอย่างชัดเจน ในขณะที่ภาพ MRI แบบ T_2-weighted พื้นที่บวมน้ำที่ซึ่งล้อมรอบเนื้องอกปรากฏสว่างขึ้น ส่วนภาพ MRI แบบ T_{2FLAIR} (FLAIR) เป็นขั้นตอนพิเศษเพื่อช่วยแบ่งแยกพื้นที่บวมน้ำจากน้ำหล่อเลี้ยงสมองและไขสันหลัง (Cerebrospinal Fluid, CSF) โดยการเพิ่ม Time to Echo (TE) และ Repetition Time (TR) ด้วยวิธีการนี้เนื้องอกสมองผิดปกติยังคงความสว่างแต่น้ำหล่อเลี้ยงสมองและไขสันหลังที่ปกติถูกทำให้ลดทอนลงและทำให้เกิดความมืดในภาพ ขั้นตอนนี้ละเอียดอ่อนมากเพื่อวิธีการทางพยาธิวิทยาและทำให้ความแตกต่างระหว่างเนื้องอกและ CSF แยกแยะได้ง่ายขึ้น

เนื้องอกสมองกลิโอมาวินิจฉัยโดยเจ้าหน้าที่ทางการแพทย์ซึ่งใช้เวลาค่อนข้างมากและควรจะทำโดยผู้ชำนาญการเท่านั้น ซึ่งค่อนข้างเป็นข้อจำกัดในชนบทหรือพื้นที่ห่างไกล ดังนั้นงานวิจัยนี้มีเป้าหมายเพื่อพัฒนาเครื่องมือที่แม่นยำและมีประสิทธิภาพสำหรับเป็นผู้ช่วยเจ้าหน้าที่ทางการแพทย์ในการแบ่งแยกจัดหมวดหมู่เนื้องอกสมองกลิโอมาโดยใช้ปัญญาประดิษฐ์ (Artificial Intelligence, AI) และเทคนิคการมองเห็นของคอมพิวเตอร์หรือคอมพิวเตอร์วิชัน (Computer Vision)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

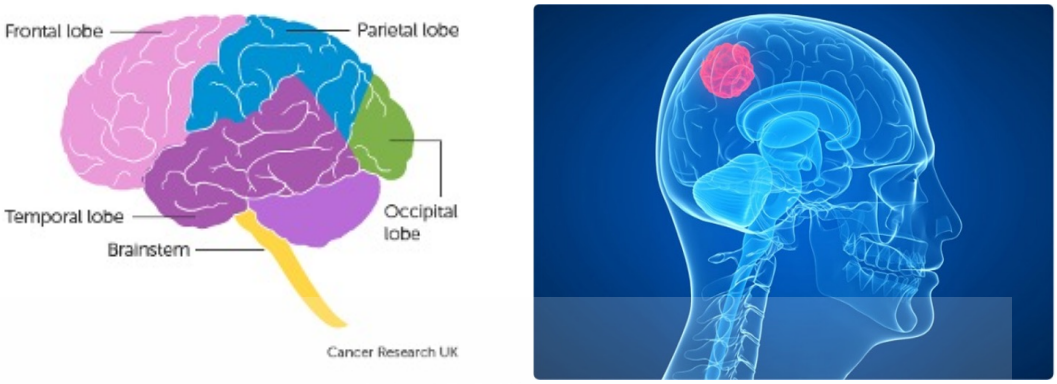
1.2 การใช้ระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค

วิธีการปัจจุบันสำหรับวินิจฉัยเนื้องอกสมองกลิโอมาซึ่งถูกดำเนินการโดยเจ้าหน้าที่ทางการแพทย์ตรวจสอบรูปภาพ MRI ด้วยตาเปล่าซึ่งมีแนวโน้มที่จะเกิดเหตุการณ์ความผิดพลาดได้โดยสายตามนุษย์ ตัวอย่างเบื้องต้นอ้างอิงตามวิธีการปัจจุบันในสถาบันการแพทย์หรือคลินิก เจ้าหน้าที่ทางการแพทย์หรือนักรังสีวิทยา (Radiologist) จำนวนมากกว่าหนึ่งท่านจะวิเคราะห์เนื้องอกสมองจากภาพสแกน MRI ซึ่งรวมทั้งภาพตัดเฉือนตามแนวแกนทั้งสามแกน (x, y, z) และท้ายที่สุดจะได้รับข้อมูลสรุปตำแหน่งที่ตั้งของเนื้องอก เช่น สมองกลีบขมับ (Temporal Lobe) หรือ สมองส่วนหน้า (Frontal Lobe) และขนาดแผลหรือก้อน อ้างอิงตามรูปที่ 1.2 และตัดสินใจว่าเป็นเนื้องอกระดับหรือเกรดใด กระบวนวิธีการนี้ต้องดำเนินการโดยผู้ชำนาญการและนักรังสีวิทยาที่มีประสบการณ์ซึ่งเป็นข้อจำกัดในชนบทหรือพื้นที่ห่างไกล ดังนั้นระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคถูกนำเสนอเพื่อการคัดกรองเนื้องอกสมองกลิโอมาโดยใช้วิทยาการคอมพิวเตอร์ ซึ่งมีแนวโน้มประสิทธิภาพเพิ่มขึ้นทั้งทางด้านเทคโนโลยีฮาร์ดแวร์และซอฟท์แวร์ ระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคถูกนำมาใช้สำหรับการคัดกรองเบื้องต้นและเป็นผู้ช่วยเจ้าหน้าที่ทางการแพทย์ในการวินิจฉัยเนื้องอกสมองกลิโอมา เพื่อรักษาเพิ่มเติมซึ่งจะช่วยเพิ่มโอกาสรอดชีวิตของผู้ป่วย

วิเคราะห์ตัวอย่างการวินิจฉัยเนื้องอกสมองจากกลุ่มข้อมูล Repository of Molecular Brain Neoplasia Data (REMBRANDT) เพื่อความเข้าใจยิ่งขึ้นของวิธีการวินิจฉัยเนื้องอกสมองในปัจจุบัน กลุ่มข้อมูลประกอบด้วยภาพ MRI ที่ได้รับจาก Thomas Jefferson University Institution (TJU) ข้อมูลทางคลินิกถูกวิเคราะห์และรวบรวมจากนักรังสีวิทยา 3 ท่าน ผู้ซึ่งตรวจสอบภาพถ่ายทางการแพทย์ MRI เพื่อวินิจฉัยชนิดและระดับหรือเกรดของเนื้องอกสมอง เช่น Astrocytoma เกรด 3 หรือ GBM เกรด 4 เป็นต้น นักรังสีวิทยาใช้ภาพ MRI ที่เป็นภาพตัดเฉือนตามแนวแกนเพื่อวินิจฉัยโรค ท้ายที่สุดข้อมูลการวินิจฉัยโรคถูกสรุปข้อมูลในรูปแบบตารางของผู้ป่วยแต่ละราย

ในปัจจุบันเมื่อไม่นานมานี้การศึกษาวิจัยหลายงานได้นำมาใช้งานจริงสำหรับระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค (Computer-Aided Diagnosis, CAD) เพื่อช่วยเจ้าหน้าที่ทางการแพทย์ในการตัดสินใจว่าเป็นโรคชนิดใด ความผิดพลาดที่เกิดจากมนุษย์มีแนวโน้มจะเกิดขึ้นเมื่อการวินิจฉัยโดยการมองด้วยตาเปล่า ดังนั้นระบบ CAD ช่วยแพทย์ได้ในการตัดสินใจขั้นสุดท้ายและเน้นย้ำผลการวินิจฉัยอีกครั้งก่อนดำเนินการรักษาต่อไป ด้วยการวินิจฉัยที่แม่นยำ ผู้ป่วยจะได้รับการรักษาอย่างทันท่วงทีและเพิ่มอัตราการรอดชีวิตของผู้ป่วยได้อีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. ภาพส่วนต่างๆของสมอง ข. ตัวอย่างก้อนเนื้องอกสมองที่เกิดขึ้น

รูปที่ 1.2 ภาพการวิเคราะห์เนื้องอกสมองที่เกิดขึ้นในตำแหน่งต่างกัน

1.3 ความมุ่งหมายและวัตถุประสงค์ของงานวิจัย

การวินิจฉัยเนื้องอกสมองโดยทั่วไปดำเนินการโดยนักรังสีวิทยาหรือแพทย์ผ่านการตีความหมายภาพถ่ายทางการแพทย์ MRI อย่างไรก็ตาม ความเที่ยงตรงแม่นยำของผลลัพธ์วินิจฉัยขึ้นอยู่กับความเชี่ยวชาญและประสบการณ์ของเจ้าหน้าที่ทางการแพทย์ ซึ่งเป็นสิ่งที่ท้าทายและขาดแคลนในพื้นที่ชนบท เพื่อแก้ปัญหาเกี่ยวกับเรื่องนี้ระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคได้ถูกพัฒนาขึ้นเพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์ ในการทำให้แน่ใจในผลการวินิจฉัยที่แม่นยำและเป็นการตรวจสอบอีกครั้งก่อนการรักษาโรคได้ถูกต้อง โดยระบบ CAD จะกำจัดความเป็นไปได้ที่จะเกิดความผิดพลาดจากการมองเห็นของมนุษย์ และทำให้ผลการวินิจฉัยเนื้องอกสมองเที่ยงตรงขึ้น สิ่งนี้จะช่วยเพิ่มอัตรารอดชีวิตของผู้ป่วย เนื่องจากผู้ป่วยเหล่านั้นจะได้รับการรักษาอย่างทันท่วงทีและถูกวิธีตรงตามโรคที่เกิดขึ้น

ในงานวิจัยนี้มีเป้าหมายเพื่อพัฒนาระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคที่แม่นยำและใช้งานได้จริงสำหรับการช่วยเหลือการวินิจฉัยเนื้องอกสมอง โดยนำเสนอกระบวนวิธีการ 2 แบบสำหรับการคัดกรองเนื้องอกสมองจากภาพถ่ายทางการแพทย์ MRI ดังต่อไปนี้

1. การจัดหมวดหมู่เนื้องอกสมองกลิโอมาที่มีกลุ่มข้อมูลน้อยและไม่สมดุล (Small and Imbalance Dataset) โดยใช้การจัดหมวดหมู่ด้วยการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM) และเทียบกับการใช้การเรียนรู้แบบถ่ายโอน (Transfer Learning)
2. เนื้องอกสมองกลิโอมาที่เพิ่มจำนวนกลุ่มข้อมูลขึ้นและจัดหมวดหมู่โดยใช้การเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน (CNNs) หลากหลายแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งแต่ละวิธีการเหมาะสมกับข้อจำกัดที่ต่างกัน และมีข้อเด่นข้อด้อยที่ต่างกันในการจัดหมวดหมู่เนื้องอกสมองอีกด้วย ซึ่งขึ้นอยู่กับทรัพยากรคอมพิวเตอร์และขนาดกลุ่มข้อมูล

1.4 ขอบเขตงานวิจัย

งานวิจัยนี้มุ่งเน้นการค้นคว้าเพื่อนำระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค ช่วยเหลือเจ้าหน้าที่ทางการแพทย์ในการวินิจฉัยเนื้องอกสมองกลิโอมาอย่างถูกต้องแม่นยำ โดยได้รับข้อมูลภาพถ่ายทางการแพทย์ของเนื้องอกสมอง GBM เกรด 4 จากฐานข้อมูล 2 แหล่งด้วยกันประกอบด้วย กลุ่มข้อมูลที่ 1 คือ Repository of Molecular Brain Neoplasia Data ซึ่งประกอบด้วยภาพ MRI จากผู้ป่วย 130 ท่านรวมทั้งข้อมูลการวิเคราะห์ทางเทคนิค โดยภาพดังกล่าวอยู่ในรูปแบบ DICOM และถูกแปลงเป็น JPEG และเก็บไว้ในฐานข้อมูลเพื่อดำเนินการต่อไป ส่วนกลุ่มข้อมูลที่ 2 คือ Brain Tumor Segmentation 2021 (BraTS) ซึ่งประกอบด้วยภาพเนื้องอกสมองกลิโอมาที่ได้รับการยืนยันผลวินิจฉัย โดยภาพทั้งหมดอยู่ในรูปแบบ Neuroimaging Informatics Technology Initiative (NIfTI) และถูกแปลงเป็น JPEG เพื่อใช้งานต่อไป งานวิจัยนี้มุ่งเน้นการใช้ระบบคอมพิวเตอร์เพื่อช่วยจัดหมวดหมู่เนื้องอกสมองทดแทนการวินิจฉัยด้วยตาเปล่าของเจ้าหน้าที่ทางการแพทย์ โดยเฉพาะอย่างยิ่งเนื้องอกสมองชนิด GBM เกรด 4 ซึ่งร้ายแรงที่สุด เพื่อลดความผิดพลาดและเพิ่มความแม่นยำในการวินิจฉัยโรคเนื้องอกสมอง โดยนำเสนอสองวิธีการคือการจัดหมวดหมู่ (Classification) โดยใช้วิธีการเรียนรู้ของเครื่องจักรแบบดั้งเดิม (Traditional Machine Learning) ด้วยซัพพอร์ตเวกเตอร์แมชชีน ซึ่งต้องใช้การแบ่งออกเป็นส่วน (Segmentation) และการสกัดฟีเจอร์ออกมา (Feature Extraction) ด้วยวิธีการทำด้วยมือ (Handcrafted) ก่อนที่จะป้อนข้อมูลในรูปแบบตารางเพื่อการจัดหมวดหมู่โดยใช้การเรียนรู้ของเครื่องจักรต่อไป วิธีการนี้ใช้เวลาค่อนข้างมากในการดำเนินการแต่ละขั้นตอนแต่มีจุดเด่นคือไม่จำเป็นต้องใช้ทรัพยากรคอมพิวเตอร์ที่สูงในการประมวลผล ส่วนอีกวิธีคือการจัดหมวดหมู่ด้วยการเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน (Convolution Neural Network) ด้วยวิธีการนี้ฟีเจอร์จะถูกสกัดออกมาอย่างอัตโนมัติ เพื่อใช้ตัดสินใจว่าส่วนไหนเป็นเนื้องอกหรือสมองปกติด้วยการใช้อัลกอริทึมพิจารณาแต่ละพิกเซลอย่างละเอียดทำให้ได้รับผลลัพธ์ที่รวดเร็วและแม่นยำ แต่ต้องอาศัยทรัพยากรคอมพิวเตอร์ที่สูงในการประมวลผลเช่นกัน รวมทั้งเพื่อให้เกิดความน่าเชื่อถือ ได้นำเสนอ Gradient-weighted Class Activation Mapping (Grad-CAM) เพื่อแสดงแผนที่ความร้อน (Heatmap) ที่มีผลต่อการตัดสินใจมากที่สุด เพื่อเปรียบเทียบว่าจุดไหนที่อัลกอริทึมใช้ตัดสินใจว่าเป็นภาพเนื้องอกหรือสมองปกติ

ผลลัพธ์ความเที่ยงตรงของทั้งสองวิธีการได้รับการเปรียบเทียบและวิเคราะห์เพื่อหาวิธีปรับปรุงให้ดียิ่งขึ้นเพื่อความเหมาะสมในการนำไปใช้งานจริงในอนาคตต่อไป หรือเพื่อการพัฒนาการจัดหมวดหมู่เนื้องอกสมองกลิโอมาที่มีความแม่นยำยิ่งขึ้น และเพื่อป้องกันปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งเกินไป (Overfitting) รวมทั้งพัฒนาให้ตรวจจับเนื้องอกสมองเกรด 1 ถึง 4 ได้ และจะศึกษาทดสอบผลกระทบต่อการปฏิบัติทางการแพทย์ (Clinical Implications) ในอนาคตข้างหน้าต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ขั้นตอนการวิจัย

1. ศึกษาเกี่ยวกับเนื้องอกสมองกลิโอมา ระดับความรุนแรง และการวินิจฉัยโรค เพื่อพัฒนาการหาวิธีนำระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค เพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์

2. ศึกษาค้นคว้างานวิจัยที่เกี่ยวกับการจัดหมวดหมู่เนื้องอกสมองที่มีความน่าเชื่อถือ เพื่ออ้างอิงและศึกษาต่อยอดเพิ่มเติมให้ได้รับผลลัพธ์ที่มีความแม่นยำ น่าเชื่อถือและเป็นที่ยอมรับได้

3. ค้นคว้ารวบรวมกลุ่มข้อมูล (Dataset) รูปภาพทางการแพทย์ MRI ของผู้ป่วยเนื้องอกสมองรวมทั้งสมองปกติเพื่อดำเนินการศึกษาการจัดหมวดหมู่แบบไบนารีต่อไป

4. ศึกษาวิจัยและนำมาใช้งานจริงสำหรับการจัดหมวดหมู่ (Classification) โดยใช้วิธีการเรียนรู้ของเครื่องจักรแบบดั้งเดิมด้วยซัพพอร์ตเวกเตอร์แมชชีน

5. ศึกษาวิจัยและนำมาใช้งานจริงสำหรับการจัดหมวดหมู่ด้วยการเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน

6. สรุปผลเปรียบเทียบผลลัพธ์ของทั้งสองวิธีการและงานวิจัยที่ผ่านมา รวมทั้งการพัฒนาปรับปรุงระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคให้ดียิ่งขึ้นในอนาคต

1.6 รายละเอียดเนื้อหาในแต่ละบท

วิทยานิพนธ์นี้ประกอบด้วยเนื้อหาที่เกี่ยวข้องกับการจัดหมวดหมู่เนื้องอกสมองกลิโอมาเพื่อพัฒนาระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค โดยรายละเอียดเนื้อหาแต่ละบทประกอบด้วย

บทที่ 1 กล่าวถึงข้อมูลทางการแพทย์ที่เกี่ยวข้องกับโรคเนื้องอกสมองกลิโอมา ระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค ความมุ่งหมายและวัตถุประสงค์งานวิจัย ขอบเขต และขั้นตอนของงานวิจัย

บทที่ 2 กล่าวถึงทฤษฎีที่เกี่ยวข้อง และทบทวนงานวิจัยที่น่าเชื่อถือและเกี่ยวข้องกับการจัดหมวดหมู่เนื้องอกสมองเพื่อใช้พัฒนาระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค โดยอยู่บนพื้นฐานของการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน และโครงข่ายประสาทเทียมแบบคอนโวลูชัน

บทที่ 3 กล่าวถึงหลักการและขั้นตอนการจัดหมวดหมู่เนื้องอกสมองกลิโอมาที่มีกลุ่มข้อมูลน้อยและไม่สมดุล (Small and Imbalance Dataset) โดยใช้การจัดหมวดหมู่ด้วยการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน และเทียบกับการใช้การเรียนรู้แบบถ่ายโอน

บทที่ 4 กล่าวถึงหลักการและกระบวนวิธีการจัดหมวดหมู่เนื้องอกสมองกลิโอมาที่มีกลุ่มข้อมูลขนาดใหญ่ขึ้น โดยใช้การเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน

บทที่ 5 กล่าวถึงผลการทดลองและการวิเคราะห์ผลการทดลองที่ได้รับจากกระบวนวิธีการแต่ละแบบ

บทที่ 6 สรุปผลการทดลองที่ได้จากการทดลองวิจัย ข้อเสนอแนะ ตลอดจนแนวทางการพัฒนาเกี่ยวกับงานวิจัยนี้ในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้กล่าวถึงเนื้อหาเกี่ยวกับทฤษฎีที่เกี่ยวข้องกับระบบคอมพิวเตอร์ช่วยวินิจฉัยเนื้องอกสมองบนพื้นฐานเทคนิคคอมพิวเตอร์วิชัน กลุ่มข้อมูลภาพถ่ายทางการแพทย์ที่ใช้สำหรับการจัดหมวดหมู่ ทฤษฎีเกี่ยวกับการเตรียมกระบวนการรูปภาพ การจัดหมวดหมู่ด้วยการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน เนื้อหาของการเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน และการเรียนรู้แบบถ่ายโอนที่ทันสมัยและเป็นที่นิยมใช้งานจัดหมวดหมู่แบบต่างๆ หลักการมองเห็นพื้นที่รูปภาพซึ่งถูกแยกแยะเพื่อเน้นย้ำผลลัพธ์การทำนายผล และทบทวนวรรณกรรมงานวิจัยที่เกี่ยวข้อง

2.1 ข้อมูลทางการแพทย์เกี่ยวกับเนื้องอกสมองกลิโอมา

เนื้องอกสมองกลิโอมาที่เกิดขึ้นจากเซลล์ในสมองเอง เกิดจากเซลล์เกลียที่มีรูปร่างคล้ายดาวที่อยู่ในสมองและไขสันหลัง ซึ่งปกติทำหน้าที่หลายอย่างทั้งจัดหาสารอาหารให้แก่เนื้อเยื่อประสาท สร้างการไหลเวียนของเลือดในสมอง และซ่อมแซมกระบวนการสมองและไขสันหลังเมื่อได้รับบาดเจ็บหรือติดเชื้อ ซึ่งมันถูกแบ่งออกเป็นสี่ระดับ กลิโอมาเกรดสี่เรียกอีกอย่างว่า กลิโอบลาสโทมา มัลติฟอร์ม (GBM) ซึ่งเป็นเนื้องอกสมองที่เกิดจากเซลล์ในสมองเองระดับที่ร้ายแรงที่สุดและเติบโตลุกลามอย่างรวดเร็ว การตรวจพบอย่างรวดเร็ว และได้รับการรักษาอย่างทันท่วงทีจะส่งผลต่ออัตราการรอดชีวิต เทคนิคมาตรฐานการวินิจฉัยเนื้องอกสมองคือ ภาพถ่ายเรโซแนนซ์แม่เหล็ก (Magnetic Resonance Imaging, MRI) ซึ่งเป็นวิธีที่ไม่รุกรานและถูกจัดเตรียมให้เห็นความแตกต่างและตัดกันของภาพเนื้อเยื่อส่วนดีและเนื้องอก และเทคนิคนี้ใช้อย่างกว้างขวางในคลินิกเพื่อวินิจฉัยเนื้องอกสมอง การวินิจฉัยเนื้องอกสมองกลิโอมาต้องทำโดยแพทย์หรือนักรังสีวิทยาที่เชี่ยวชาญดังที่กล่าวมาแล้ว โดยใช้การอ่านวิเคราะห์ภาพถ่ายทางการแพทย์ MRI เพื่อวิเคราะห์ตำแหน่ง ขนาดเส้นผ่านศูนย์กลาง และระดับความรุนแรงของเนื้องอกเพื่อการรักษาอย่างทันท่วงที ทั้งนี้เพื่อให้ทราบผลวินิจฉัยที่แน่นอนต้องอาศัยการตรวจชิ้นเนื้อทางพยาธิวิทยาเพื่อยืนยันอีกครั้ง [1], [2]

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค เพื่อสนับสนุนช่วยเหลือเจ้าหน้าที่ทางการแพทย์วินิจฉัยโรคได้แม่นยำและรวดเร็วขึ้น โดยได้นำเสนอ 2 กระบวนวิธีการจัดหมวดหมู่เนื้องอกสมอง ซึ่งวิธีที่หนึ่งคือใช้การเรียนรู้ของเครื่องจักร และอีกวิธีหนึ่งคือใช้การเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมแบบคอนโวลูชัน ซึ่งจะกล่าวถึงรายละเอียดและทฤษฎีของทั้งสองกระบวนวิธีการต่อไป

2.1.1 กลุ่มข้อมูล REMBRANDT

The Repository of Molecular Brain Neoplasia Data เป็นฐานข้อมูลแบบปิดซึ่งมีเป้าหมายเพื่ออำนวยความสะดวกและค้นพบโดยการเชื่อมโยงระหว่างข้อมูลทางด้านคลินิกและข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะจีโนม (Genome) ซึ่งประกอบด้วยลักษณะปรากฏทางคลินิก (Clinical Phenotype) จำนวน 13,472 ลักษณะ และแสดงในแบบรูปภาพ MRI หลายขั้นตอน (Multi Sequence) ของผู้ป่วยจำนวน 130 ท่าน [3] รูปภาพ MRI แสดงในรูปแบบฟอร์แมท DICOM และได้รับจากสถาบันการแพทย์สองแห่งคือ Thomas Jefferson University และ Henry Ford Hospital

2.1.2 กลุ่มข้อมูล BraTS2021

Brain Tumor Segmentation เป็นฐานข้อมูลรูปภาพ MRI ของเนื้องอกสมองกลิโอมา ที่รวบรวมจากหลายสถาบันการแพทย์ พร้อมทั้งได้รับการวินิจฉัยยืนยันทางด้านรังสีวิทยา ถูกนำไปใช้เพื่ออบรม (Training) ทำให้สมบูรณ์ (Validation) และทดสอบ (Testing) สำหรับการแข่งขันการแบ่งแยกออกเป็นส่วนของเนื้องอกสมอง ปี 2021 โดยรูปภาพ MRI จัดเตรียมในรูปแบบไฟล์ NIfTI และแบ่งย่อยอธิบายเป็น Native (T_1), Post-Contrast T_1-weighted (T_{1Gd}), T_2-weighted (T_2) และ T_2 Fluid Attenuated Inversion Recovery (T_{2FLAIR}) รวมทั้งถูกรวบรวมจากมาตรการทางคลินิกที่หลากหลายและอุปกรณ์แสกนภาพจากหลายสถาบันการแพทย์ [4], [5], [6]

2.2 ทฤษฎีที่เกี่ยวข้องกับการจัดหมวดหมู่เนื้องอกสมองด้วยการเรียนรู้ของเครื่องจักร

โดยงานวิจัยส่วนนี้นำหลักการเรียนรู้ของเครื่องจักรเพื่อจัดหมวดหมู่เนื้องอกสมองที่มีกลุ่มข้อมูลจำนวนน้อย เนื่องจากแต่ละขั้นตอนทำด้วยมือ (Handcrafted) จึงมีความยุ่งยากในขั้นตอนทำงาน แต่ใช้ทรัพยากรคอมพิวเตอร์ในการประมวลผลน้อยเมื่อเทียบกับการเรียนรู้เชิงลึก เนื้อหาทฤษฎีที่สำคัญที่เกี่ยวข้องเกี่ยวกับการจัดหมวดหมู่เนื้องอกสมองที่มีกลุ่มข้อมูลจำนวนไม่มากนัก ด้วยการเรียนรู้ของเครื่องจักรจะกล่าวดังรายละเอียดต่อไปนี้

2.2.1 การสกัดฟีเจอร์โดยใช้วิธีการเมทริกซ์การเกิดร่วมของระดับสีเทา

การสกัดฟีเจอร์หลังจากได้ภาพที่ผ่านการเตรียมกระบวนการและแบ่งออกเป็นส่วนแล้ว ก่อนที่จะนำข้อมูลจัดเรียงรูปแบบตารางเพื่อเป็นตัวแทนของแต่ละภาพ และป้อนข้อมูลตารางเพื่อจัดหมวดหมู่ด้วยการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีนต่อไป โดยในงานนี้ใช้วิธีการเมทริกซ์การเกิดร่วมของระดับสีเทา (Gray-Level Co-occurrence Matrix, GLCM) เพื่อสกัดฟีเจอร์ เป็นวิธีการทางสถิติของพื้นผิวเพื่อพิจารณามิติสัมพันธ์ (Spatial Relations) ของพิกเซล อธิบายเพิ่มเติมดังรูปที่ 2.1 พิจารณารูปภาพ f และเมทริกซ์ G ซึ่งมีอิลิเมนท์ g_{ij} เป็นจำนวนครั้งที่พิกเซลจับคู่กันด้วยความเข้ม (Intensities) z_i และ z_j เกิดขึ้นในรูปภาพ f ในตำแหน่งที่ระบุ ซึ่งเมทริกซ์ที่สร้างขึ้นในลักษณะนี้เรียกว่าเมทริกซ์การเกิดร่วมของระดับสีเทา จะเห็นได้ว่าอิลิเมนท์ (1, 1) ของ G เท่ากับ 1 เนื่องจากมีการเกิดขึ้นเพียง 1 ครั้งในรูปภาพ f ที่ซึ่งพิกเซลมีค่าเท่ากับ 1 และมีพิกเซลข้างเคียงในด้านฝั่งขวามือมีค่าพิกเซลเท่ากับ 1 ในลักษณะเดียวกันอิลิเมนท์ (1, 2) ของ G เท่ากับ 2 เนื่องจากในรูปภาพ f นั้นค่าพิกเซลเท่ากับ 1 และมีพิกเซลข้างเคียงฝั่งขวามือเท่ากับ 2 นั้นมี 2 ครั้งที่เกิดขึ้น อิลิเมนท์อื่นของเมทริกซ์ G คำนวณได้ในลักษณะเดียวกัน จำนวนของระดับความเข้มที่เป็นไปได้ใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปภาพจะเป็นตัวกำหนดขนาดของเมทริกซ์ G สำหรับภาพ 8 บิท (ระดับที่เป็นไปได้คือ 256) ขนาดของเมทริกซ์ G จะเท่ากับ 256 x 256 บางครั้งเพื่อลดภาระในการคำนวณ วิธีการความถี่ย่านความเข้มเพื่อให้เมทริกซ์ G จัดการได้ง่าย เช่นความเข้ม 256 กำหนดให้ระดับความเข้ม 32 ระดับเริ่มต้นเท่ากับ 1 และความเข้ม 32 ระดับถัดไปเท่ากับ 2 และทำแบบเดียวกันไป จะได้รับเมทริกซ์การเกิดร่วมขนาด 8 x 8

กำหนดให้จำนวนรวมของพิกเซลที่จับคู่กัน (n) เท่ากับผลรวมของอิลิเมนท์ G (n = 16 สำหรับตัวอย่างรูปที่ 2.1) และจำนวนของ p_{ij} หาได้โดย

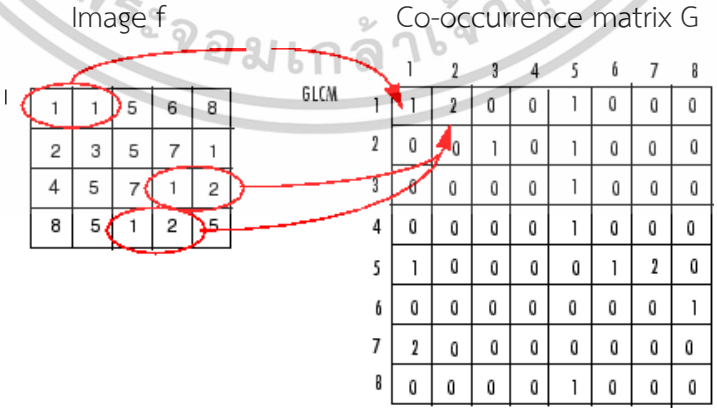
$$p_{ij} = g_{ij}/n \tag{2.1}$$

ซึ่งค่านี้เป็นประมาณการความน่าจะเป็นที่การจับคู่จะมีค่าเท่ากับ (z_i, z_j) ความน่าจะเป็นนี้จะอยู่ในช่วง [0, 1] และผลรวมเท่ากับ 1

$$\sum_{i=1}^{K}\sum_{j=1}^{K} p_{ij} = 1 \tag{2.2}$$

โดยที่ค่า K คือขนาดแถวหรือคอลัมน์ของเมทริกซ์ G

รูปแบบความเข้มของพื้นผิวตรวจจับได้โดยการเลือกตำแหน่งที่เหมาะสม และวิเคราะห์อิลิเมนท์ G โดยกลุ่มของการจำกัดความใช้สำหรับกำหนดคุณสมบัติของ G คือ ค่าสถิติเรียกว่า ความต่างกัน (Contrast), ความสัมพันธ์กัน (Correlation), พลังงาน (Energy) และ ความเป็นเนื้อเดียวกัน (Homogeneity) [7] ได้รับจากการคำนวณตามสมการที่ (2.7) – (2.10) และพารามิเตอร์ใช้ในการหาความสัมพันธ์กัน ระบุดังสมการที่ (2.3) – (2.6)



รูปที่ 2.1 วิธีการสร้างเมทริกซ์ Co-occurrence

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$m_r = \sum_{i=1}^{K} i \sum_{j=1}^{K} p_{ij} \tag{2.3}$$

$$m_c = \sum_{j=1}^{K} j \sum_{i=1}^{K} p_{ij} \tag{2.4}$$

$$\sigma_r^2 = \sum_{i=1}^{K} (i - m_r)^2 \sum_{j=1}^{K} p_{ij} \tag{2.5}$$

$$\sigma_c^2 = \sum_{j=1}^{K} (j - m_c)^2 \sum_{i=1}^{K} p_{ij} \tag{2.6}$$

ความสัมพันธ์กันใช้วัดความสัมพันธ์ของพิกเซลนั้นกับข้างเคียง ซึ่งค่านี้อยู่ในช่วง 1 ถึง -1 ซึ่งค่านี้จะไม่ถูกระบุหรือหาค่าไม่ได้ถ้าค่าเบี่ยงเบนมาตรฐานเท่ากับ 0

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} \frac{(i - m_r)(j - m_c) p_{ij}}{\sigma_r \sigma_c} \tag{2.7}$$

$$\sigma_r \neq 0; \sigma_c \neq 0$$

ความต่างกันของพื้นผิว ใช้วัดความเข้มของความต่างระหว่างพิกเซลและพื้นผิวข้างเคียงในรูปภาพ ซึ่งค่านี้อยู่ในช่วง 0 (เมื่อ G เป็นค่าคงที่) จนถึง $(K - 1)^2$

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} (i - j)^2 p_{ij} \tag{2.8}$$

ความสม่ำเสมอ (Uniformity) หรือเรียกอีกอย่างว่าพลังงาน ใช้เพื่อวัดความเป็นเนื้อเดียวกัน และอยู่ในช่วง [0, 1] ซึ่งค่านี้เป็น 1 สำหรับรูปภาพคงที่

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} {p_{ij}}^2 \tag{2.9}$$

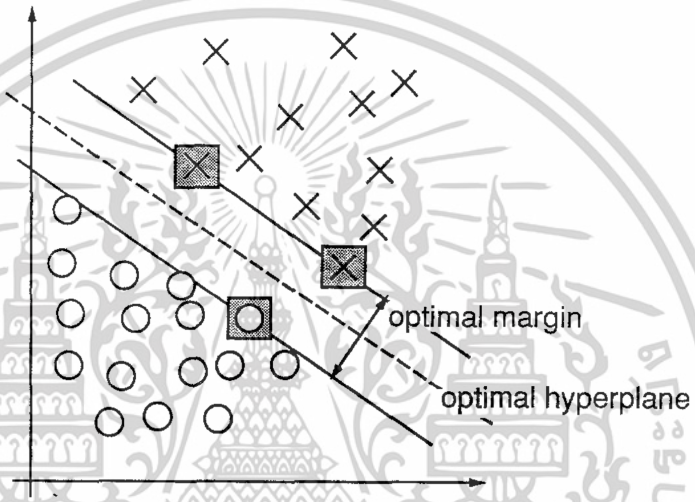
ความเป็นเนื้อเดียวกัน ใช้วัดระยะที่ใกล้สุดของการกระจายอิลิเมนท์ G ในรูปแบบทแยงมุม (Diagonal) และอยู่ในช่วง [0, 1] โดยค่าสูงสุดจะได้รับเมื่อ G เป็นเมทริกซ์แบบทแยงมุม (Diagonal Matrix)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} \frac{p_{ij}}{1 + |i - j|} \quad (2.10)$$

2.2.2 การจัดหมวดหมู่ด้วยการเรียนรู้เครื่องจักรซัพพอร์ตเวกเตอร์แมชชีน

เมื่อได้ข้อมูลฟีเจอร์ที่สกัดออกมาด้วยวิธีการเมทริกซ์การเกิดร่วมของระดับสีเทา และจัดเรียงในรูปแบบตาราง ข้อมูลเหล่านี้ซึ่งเป็นตัวแทนของแต่ละภาพจะถูกจัดหมวดหมู่ด้วยการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน (SVM) [8]



รูปที่ 2.2 ภาพประกอบการอธิบายซัพพอร์ตเวกเตอร์แมชชีน

2.2.2.1 อัลกอริทึม Optimal hyperplane

กลุ่มฉลากข้อมูลการอบรม (Labeled Training) อยู่ในรูปแบบ

$$(y_1, x_1), \ldots, (y_l, x_l), \quad y_i = \in \{-1, 1\} \quad (2.11)$$

เส้นแบ่งแยกสามารถเขียนในรูปแบบสมการเชิงเส้นในรูปของเวกเตอร์ w และค่าคงที่ b ดังนี้

$$w \cdot x_i + b \geq 1 \quad \text{if } y_i = 1,$$
$$w \cdot x_i + b \leq -1 \quad \text{if } y_i = -1, \quad (2.12)$$

สมการที่ (2.12) เขียนใหม่ในรูปแบบทำให้เป็นมาตรฐาน (normalized) ได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$y_i(w \cdot x_i + b) \geq 1 \quad \text{โดยที่ } i = 1, \ldots, l \tag{2.13}$$

จะได้ไฮเปอร์เพลนที่เหมาะที่สุด (Optimal Hyperplane)

$$w_0 \cdot x + b_0 = 0 \tag{2.14}$$

2.2.2.2 ไฮเปอร์เพลนแบบซอฟท์มาจิน Soft Margin

พิจารณากรณีที่กลุ่มข้อมูลอบรมไม่สามารถถูกแบ่งแยกได้โดยสมบูรณ์และไม่มีข้อผิดพลาด ดังนั้นต้องแบ่งแยกกลุ่มข้อมูลอบรมเพื่อให้ข้อผิดพลาดน้อยเท่าที่จะเป็นไปได้ เพื่อที่จะทำให้ถูกต้อง ค่าตัวแปรที่เป็นจำนวนไม่ติดลบได้นำมาใช้ $\zeta_i \geq 0$, i = 1, ... , l และทำให้ฟังก์ชันมีค่าน้อยสุด (Minimize the Functional) ดังสมการ (2.15)

$$\Phi(\xi) = \sum_{i=1}^{l} \xi_i^{\sigma} \tag{2.15}$$

สำหรับการทำให้มีค่าน้อยของ $\sigma > 0$ และข้อจำกัด (Constraint) ต่อไปนี้

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \quad \text{โดยที่ } i = 1, \ldots, l \tag{2.16}$$

$$\xi_i \geq 0 \quad \text{โดยที่ } i = 1, \ldots, l \tag{2.17}$$

เพื่อแยกแยะส่วนที่เหลือของกลุ่มข้อมูลทดสอบ สามารถสร้างไฮเปอร์เพลนการแยกแยะที่เหมาะสมที่สุด (Optimal Separating Hyperplane) โดยทำให้เป็นรูปแบบมาตรฐาน (minimized function)

$$\frac{1}{2}w^2 + CF\left(\sum_{i=1}^{l} \xi_i^{\sigma}\right) \tag{2.18}$$

โดยกำหนดให้ข้อจำกัดดังสมการที่ (2.16), (2.17) ที่ซึ่ง F(u) คือ Monotonic Convex Function และ C คือค่าคงที่

ด้วยค่าคงที่ที่มีค่ามากและค่า σ ที่มีค่าน้อยอย่างเพียงพอ เวกเตอร์ w_0 และค่าคงที่ b_0 ที่เป็นรูปแบบมาตรฐาน (2.18) ภายใต้ข้อจำกัดในสมการที่ (2.16), (2.17) เป็นข้อบ่งชี้ว่าไฮเปอร์เพลนนั้นได้สร้างจำนวนข้อผิดพลาดของกลุ่มข้อมูลทดสอบน้อยที่สุด และแยกแยะอิลิเมนท์ที่เหลือด้วยมาร์จินสูงสุด ทำให้สามารถพิจารณาไฮเปอร์เพลนที่อยู่ในรูปแบบมาตรฐานได้ดังนี้

$$\frac{1}{2}\left(w \cdot w + \sum_{i=1}^{l} \xi_i^2\right) \tag{2.19}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 ทฤษฎีที่เกี่ยวข้องกับการจัดหมวดหมู่เนื้องอกสมองโดยใช้การเรียนรู้เชิงลึก

2.3.1 การประมวลผลภาพดิจิทัล (Digital Image Processing)

การประมวลผลภาพดิจิทัลหมายถึงการประมวลผลภาพดิจิทัลสองมิติด้วยคอมพิวเตอร์ ซึ่งภาพดิจิทัลหมายถึงอาร์เรย์ (Array) สองมิติของจำนวนจริงหรือจำนวนเชิงซ้อนที่ถูกแทนด้วยจำนวนบิตที่จำกัด เช่นภาพที่แทนด้วยแปดบิต เป็นภาพที่มีระดับความเข้มเท่ากับ $2^8 = 256$ ระดับ โดยที่ 0 แทนระดับความเข้มต่ำสุดซึ่งเทียบเท่ากับสีดำในภาพ และ 255 แทนระดับความเข้มสูงสุดซึ่งเทียบเท่ากับสีขาวในภาพ การประมวลผลภาพดิจิทัลถูกนำมาประยุกต์ใช้อย่างกว้างขวาง เช่นการวิเคราะห์ภาพถ่ายทางอากาศ การประมวลผลภาพทางการแพทย์ งานด้านหุ่นยนต์ และการตรวจสอบอัตโนมัติของชิ้นส่วนทางอุตสาหกรรม เป็นต้น [9]

2.3.1.1 ปฏิบัติการทางสัณฐานวิทยา (Morphological Operation)

ในงานวิจัยนี้เพื่อเตรียมภาพก่อนการจัดหมวดหมู่ได้นำวิธีปฏิบัติการทางสัณฐานวิทยาใช้เพื่อแยกส่วนกะโหลกศีรษะและสิ่งปลอมปนออกจากภาพส่วนสมอง การปฏิบัติการทางสัณฐานวิทยาแบบเปิด (Opening) คือ การกัดกร่อนรูปภาพ (Erodes) จากนั้นขยายรูปภาพที่ถูกกัดกร่อนออก (Dilates) โดยใช้โครงสร้างอิลิเมนท์เดียวกันทั้งสองปฏิบัติการ การปฏิบัติการทางสัณฐานวิทยาแบบเปิดใช้ประโยชน์เพื่อนำวัตถุเล็กๆ และเส้นบางๆ ออกและเหลือไว้เฉพาะรูปร่างหรือขนาดวัตถุที่ใหญ่ในรูปภาพนั้น ซึ่งในขั้นตอนเตรียมกระบวนการนี้โครงสร้างแบบดิสค์ (Disk Structuring) ถูกนำมาใช้โดยปฏิบัติการสัณฐานวิทยาแบบเปิดของเซต A ด้วยโครงสร้างอิลิเมนท์ B แสดงดังสมการที่ (2.20)

$$A \circ B = (A \ominus B) \oplus B \tag{2.20}$$

โดยกำหนดให้ $\ominus$ คือ ปฏิบัติการ Erosion

และ $\oplus$ คือ ปฏิบัติการ Dilation

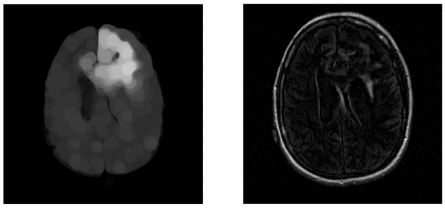
ปฏิบัติการสัณฐานวิทยาแบบเปิดของเซต A โดย B หมายถึง Erosion ของ A โดย B และจากนั้นตามด้วยการ Dilation ของผลลัพธ์โดย B [7], [10] ด้วยกระบวนการดังกล่าวส่วนกะโหลกถูกแยกออกมาโดยการสัณฐานวิทยาแบบเปิด และเหลือพื้นที่สมองเท่านั้น อย่างไรก็ตามบางพิกเซลของสมองถูกแยกออกไปด้วย ดังนั้นฟังก์ชันการลบพิกเซลถูกใช้เพื่อแยกส่วนพื้นที่สมองออกจากส่วนกะโหลก

2.3.1.2 การลบหรือหักล้างภาพ (Image Subtraction)

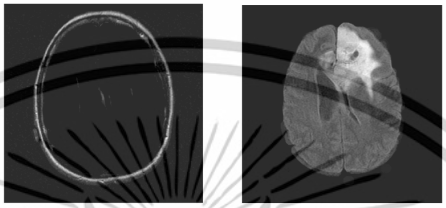
การลบพิกเซลถูกใช้เพื่อแยกส่วนกะโหลกออกจากพื้นที่สมองเพื่อเตรียมความพร้อม ขั้นแรกภาพที่ได้ผ่านกระบวนการสัณฐานวิทยาแบบเปิดแล้วถูกหักล้างออกจากภาพต้นแบบ เพื่อให้ได้ภาพกะโหลกและส่วนแปลกปลอมพื้นที่สมองเท่านั้น จากนั้นภาพที่ผ่านกระบวนการสัณฐานวิทยาแบบเปิดแล้วถูกลบออกจากภาพกะโหลกและสิ่งแปลกปลอมที่ได้จากขั้นตอนก่อนหน้า เพื่อให้ได้รับภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กะโหลกอย่างเดียว ท้ายที่สุดภาพกะโหลกศีรษะที่ถูกแยกออกมาจะถูกลบจากภาพต้นแบบอีกครั้ง และได้รับพื้นที่สมองเท่านั้น โดยภาพการแยกพื้นที่สมองออกจากกะโหลกแสดงดังรูปที่ 2.3



ก. ภาพที่ผ่านการ Opening ข. ภาพต้นแบบลบด้วยภาพ Opening



ค. รูปภาพส่วนกะโหลก ง. ภาพพื้นที่สมองที่แยกออกมา

รูปที่ 2.3 ตัวอย่างเตรียมกระบวนการ 4 ขั้นตอนเพื่อให้ได้ภาพพื้นที่สมองเท่านั้น

2.3.1.3 Contrast Limited Adaptive Histogram Equalization

หลังจากได้ภาพที่ผ่านกระบวนการสัณฐานวิทยา จนกระทั่งได้ภาพส่วนสมองอย่างเดียวแล้ว จะนำไปสู่กระบวนการปรับปรุงภาพ (Image Enhancement) คือความแตกต่างระหว่างบริเวณขาวกับดำหรือคอนทราส (Contrast) เพื่อทำให้การแสดงกราฟิกมีความชัดเจนมากขึ้นสำหรับการแสดงและการวิเคราะห์

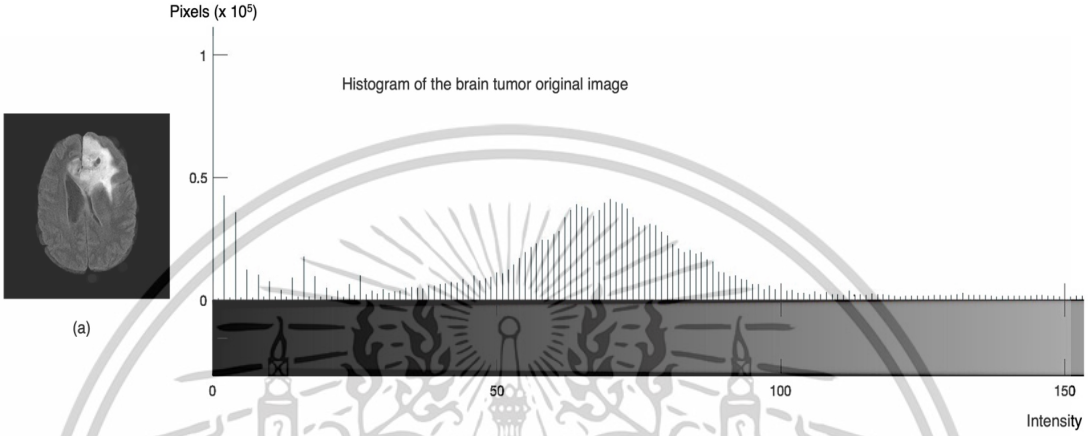
งานวิจัยนี้นำเสนอการปรับปรุงภาพโดยใช้อัลกอริทึม CLAHE ซึ่งปรับปรุงและพัฒนาจาก Adaptive Histogram Equalization (AHE) เพื่อปรับปรุงคอนทราสของภาพ [11], [12] กลุ่มข้อมูลที่ได้รับมาจากแหล่งที่ต่างกันทำให้คอนทราสหรือความเข้ม (Intensity) หลากหลายไปด้วย ดังนั้นการปรับปรุงคอนทราสโดยใช้อัลกอริทึม CLAHE ช่วยปรับปรุงและทำภาพให้เป็นรูปแบบเดียวกัน (Normalize) ก่อนเข้าสู่กระบวนการจัดหมวดหมู่ ซึ่งการใช้อัลกอริทึม CLAHE โดยใช้ MATLAB ประกอบด้วยสามส่วนหลักคือ Tile Generation, Histogram Equalization และ Bilinear Interpolation และในส่วน Histogram Equalization ประกอบด้วยห้าขั้นตอนคือ Histogram Computation, Excess Calculation, Excess Distribution, Excess Redistribution และ Scaling และ Mapping โดยใช้ A Cumulative Distribution Function (CDF) ที่แสดงในสมการที่ (2.21)

$$S_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \qquad k = 0, 1, 2, \ldots, L-1 \tag{2.21}$$

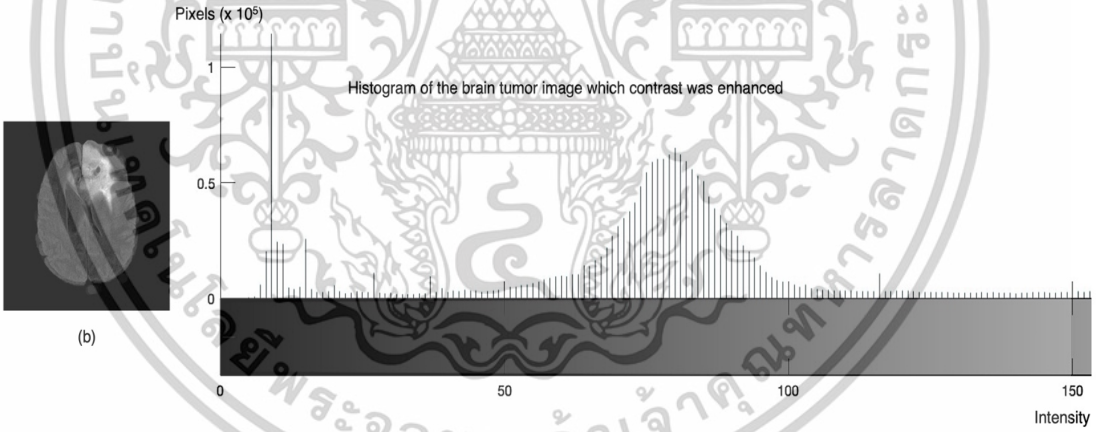
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ซึ่ง T (r_k) คือการแปลงฮิสโตแกรมให้เท่ากัน (Histogram Equalization Transformation), MN คือจำนวนรวมของพิกเซลในรูปภาพ, n_j คือ จำนวนของพิกเซลที่มีความเข้มเท่ากับ r_j และ L คือจำนวนของระดับความเข้มที่เป็นไปได้ในรูปภาพ

ภาพต้นแบบและภาพที่ปรับปรุงคอนทราสของภาพถ่ายเนื้องอกสมองพร้อมทั้งแผนภูมิฮิสโตแกรม (Histogram) สำหรับแต่ละภาพแสดงดังรูปที่ 2.4



ก. ภาพต้นแบบและฮิสโตแกรม



ข. ภาพเนื้องอกสมองที่ปรับปรุงคอนทราสและฮิสโตแกรม

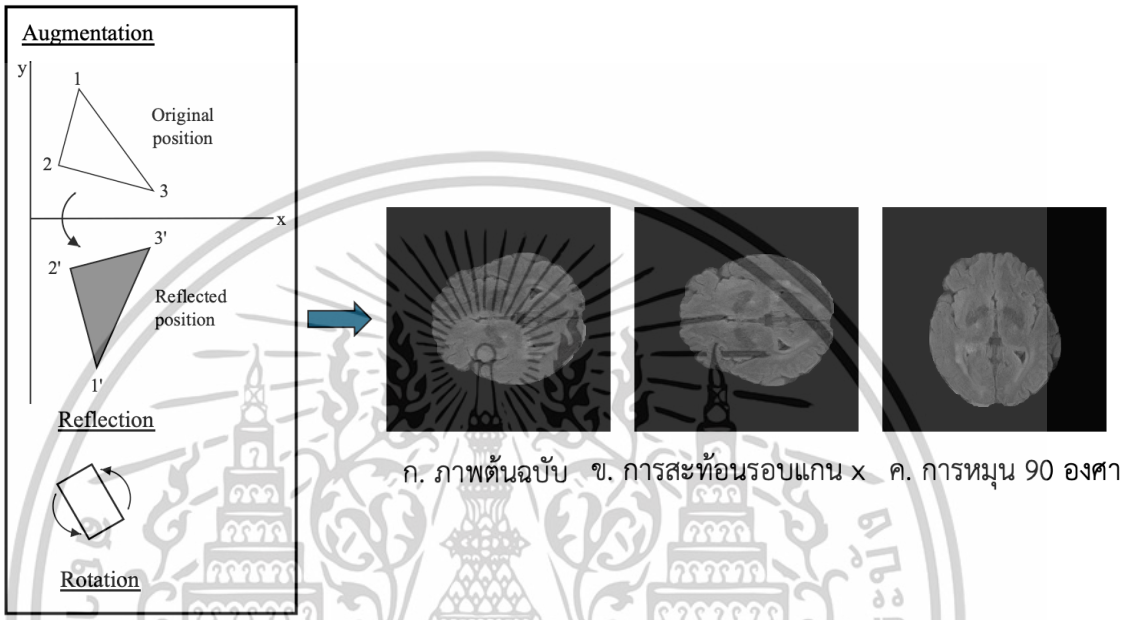
รูปที่ 2.4 อัลกอริทึม CLAHE เพื่อปรับปรุงคอนทราสของภาพเนื้องอกสมอง

2.3.1.4 การแปลงเรขาคณิตของภาพ

เพื่อเพิ่มจำนวนของภาพสมองปกติซึ่งเป็นกลุ่มที่มีจำนวนภาพน้อยกว่า (Minority Class) รวมทั้งสร้างกลุ่มข้อมูลที่สมดุล ดังนั้นวิธีเพิ่มจำนวนภาพ (Image Augmentation) ถูกนำมาใช้สำหรับกลุ่มข้อมูลอบรม (Training Dataset) โดยใช้วิธีการง่ายๆ คือ หมุนภาพและภาพสะท้อน สำหรับกลุ่มข้อมูลส่วนน้อยซึ่งสำหรับงานวิจัยนี้คือภาพสมองปกติ รูปถ่าย MRI สมองปกติจาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฐานข้อมูล REMBRANDT จะถูกเสริมภาพถ่ายโดยใช้การหมุนภาพ ในขณะที่ภาพสมองปกติของฐานข้อมูล BraTS2021 ซึ่งมีจำนวนน้อยที่สุดจะถูกเสริมภาพโดยใช้เทคนิคการหมุนและสะท้อนภาพเพื่อเพิ่มปริมาณภาพสำหรับกลุ่มข้อมูลอบรม รูปที่ 2.5 แสดงวิธีการเสริมรูปภาพสำหรับแต่ละฐานข้อมูล นอกจากนั้นภาพได้ถูกเสริมเพิ่มเติมด้วยวิธีสร้างสัญญานรบกวนและเบลอภาพ ซึ่งจะกล่าวในบทต่อไป



ก. ภาพต้นฉบับ ข. การสะท้อนรอบแกน x ค. การหมุน 90 องศา

รูปที่ 2.5 วิธีการเสริมภาพสำหรับสมองปกติซึ่งเป็นกลุ่มข้อมูลส่วนน้อย

2.3.2 การเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน (CNNs)

โครงข่ายประสาทเทียมคอนโวลูชันที่ได้รับการอบรมไว้ก่อนแล้ว (Pre-Trained Network) นำมาใช้ประโยชน์เพื่องานจัดหมวดหมู่ ซึ่งจุดเด่นของการเรียนรู้แบบถ่ายโอน (Transfer Learning) คือ โครงข่ายได้เรียนรู้ฟีเจอร์มากมายที่เป็นตัวแทนของรูปภาพที่หลากหลาย โครงข่ายรับภาพเป็นอินพุทและจับคู่ตรงกับฉลากที่เป็นจริง (Ground Truth Label) ในที่นี้คือสมองปกติหรือเนื้องอกกลิโอมา จากนั้นข้อมูลนี้จะถูกอบรมโดยโครงข่ายเพื่อสกัดฟีเจอร์ออกมาอย่างอัตโนมัติต่อไป โครงข่ายที่อบรมแล้วนำไปใช้เพื่อทำนายกลุ่มข้อมูลทดสอบและหาความแม่นยำและประเมินผล การปรับแต่งโครงข่ายที่ได้รับการอบรมไว้ก่อนทำงานได้ง่ายและเร็วกว่าการอบรมโครงข่ายจากไม่มีอะไรเลยหรือการเริ่มตั้งแต่ต้น (From Scratch) โดยโครงข่ายการเรียนรู้แบบถ่ายโอนนี้จะประกอบด้วยสามส่วนหลักคือ เลเยอร์ของโครงข่ายที่ได้รับการอบรมไว้ก่อนแล้ว กลุ่มข้อมูลที่ใช้เพื่ออบรม (Training Dataset) และสุดท้าย คืออัลกอริทึมออพชัน (Algorithm Options)

โดยคำว่าโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) หมายถึงโครงข่ายที่นำหลักการทางคณิตศาสตร์ของการปฏิบัติการที่เรียกว่าคอนโวลูชันมาใช้ ซึ่งเป็นการดำเนินการเชิงเส้นแบบพิเศษ โครงข่ายคอนโวลูชันเป็นโครงข่ายประสาทเทียมอย่างง่ายที่ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คอนโวลูชันในการคูณเมทริกซ์ทั่วไปอย่างน้อยตั้งแต่หนึ่งเลเยอร์ขึ้นไป การปฏิบัติการแบบคอนโวลูชันโดยทั่วไปแทนที่ด้วยสัญลักษณ์ดอกจัน (Asterisk) และจำกัดความด้วยสมการที่ (2.22) ที่ซึ่ง *x* อ้างอิงถึงอินพุทและค่าอาร์กิวเมนต์ (Argument) ที่สอง *w* คือเคอร์เนลหรือฟิลเตอร์ ค่าเอาท์พุทคือ *s(t)* บางครั้งกล่าวถึงฟีเจอร์แมพ (Feature Map) [13]

$$s(t) = (x * w)(t) \tag{2.22}$$

ถ้าสมมุติว่า *x* และ *w* ถูกระบุอยู่บนจำนวนเต็มของช่วงเวลา *t* ดังนั้นจะสามารถระบุค่าคอนโวลูชันแบบไม่ต่อเนื่อง (Discrete Convolution) ดังต่อไปนี้

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)\, w(t-a) \tag{2.23}$$

กรณีใช้กับรูปภาพสองมิติ กำหนดให้ *I* คืออินพุท และต้องการใช้ค่าเคอร์เนลหรือฟิลเตอร์สองมิติ *K* อีกด้วย

$$S(i,j) = (I * K)(i,j) = \sum_{m}\sum_{n} I(m,n)K(i-m, j-n) \tag{2.24}$$

เนื่องจากคอนโวลูชันเปลี่ยนที่ได้โดยผลลัพธ์ไม่เปลี่ยนแปลง ดังนั้นสามารถเขียนสมการได้อีกอย่างคือ

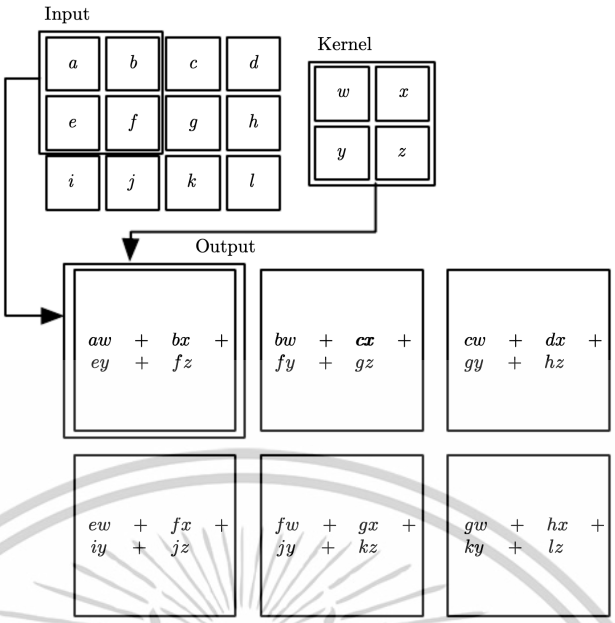
$$S(i,j) = (K * I)(i,j) = \sum_{m}\sum_{n} I(i-m, j-n)K(m,n) \tag{2.25}$$

โดยปกติแล้วไลบรารีโครงข่ายประสาทเทียมใช้งานฟังก์ชันที่เรียกว่า Cross-Correlation ที่ซึ่งเหมือนกับคอนโวลูชันแต่ว่าไม่มีการพลิกสลับเคอร์เนล (Kernel) ซึ่งสามารถเขียนได้ดังสมการที่ (2.26)

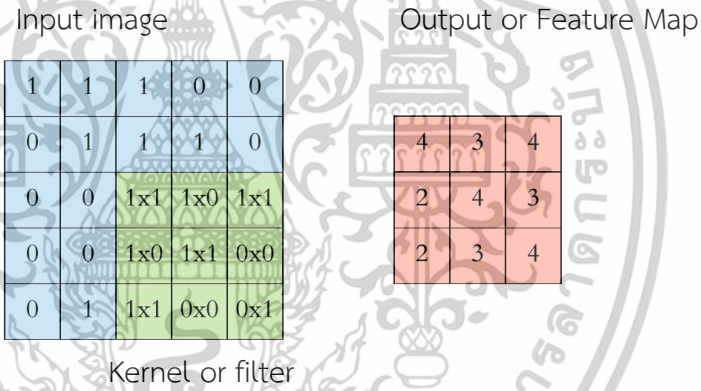
$$S(i,j) = (I * K)(i,j) = \sum_{m}\sum_{n} I(i+m, j+n)K(m,n) \tag{2.26}$$

เพื่ออธิบายเพิ่มเติมเกี่ยวกับคอนโวลูชัน (โดยไม่มีการพลิกกลับของเคอร์เนล) แสดงได้ดังรูปที่ 2.6 ซึ่งเป็นการประยุกต์ใช้กับเทนเซอร์แบบสองมิติ จากภาพเอาท์พุทเทนเซอร์ (Output Tensor) ที่อยู่ซ้ายมือบนสุดเป็นผลลัพธ์ที่ได้จากการนำเคอร์เนลประยุกต์ใช้ร่วมกับอินพุทเทนเซอร์ (Input Tensor) ซ้ายมือบนสุดที่ซึ่งสอดคล้องกัน ในส่วนรูปที่ 2.7 แสดงการนำเคอร์เนลแบบ 3 x 3 เพื่อเลื่อนผ่านภาพอินพุทด้วยค่า Stride เท่ากับ 1 ซึ่งแสดงได้ดังรูปที่ 2.8 ที่จะเลื่อนผ่านรูปทีละพิกเซล ผลลัพธ์เอาท์พุทฟีเจอร์แมพที่ได้จะเท่ากับผลรวมของ Input * Kernel

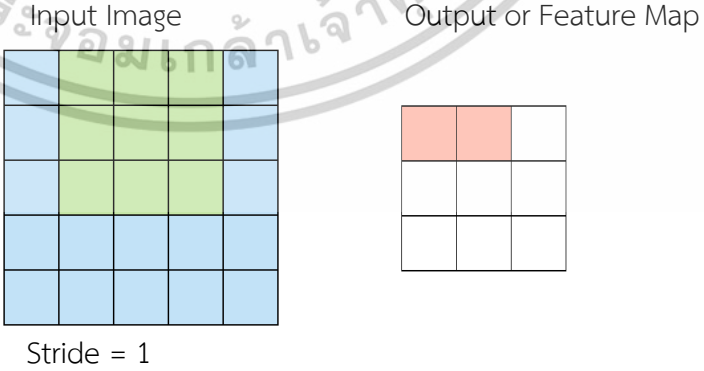
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 ตัวอย่างของคอนโวลูชันสองมิติซึ่งไม่มีการสลับเคอร์เนล (Without Kernel-Flipping)



รูปที่ 2.7 ตัวอย่างของคอนโวลูชันที่ใช้ค่าเคอร์เนล 3 x 3

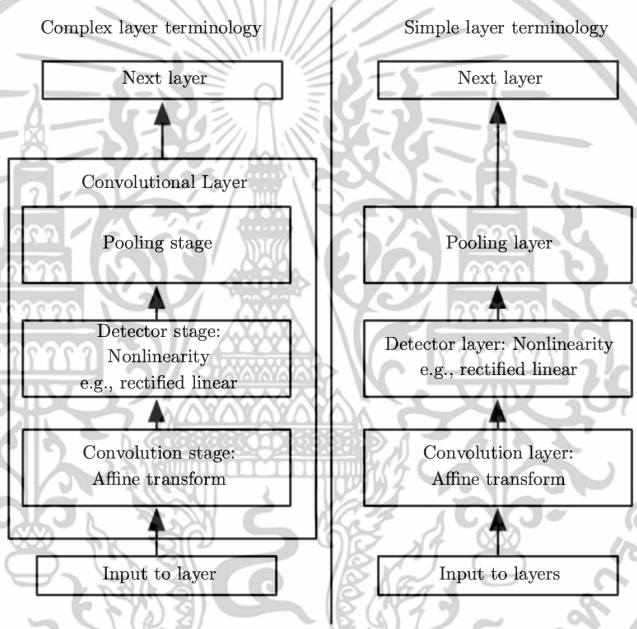


รูปที่ 2.8 ตัวอย่างค่า Stride เท่ากับ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

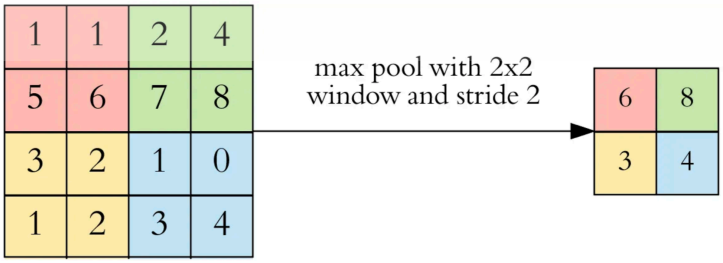
2.3.2.1 การทำพูลลิง (Pooling)

โดยทั่วไปแล้วเลเยอร์ของโครงข่ายคอนโวลูชันประกอบด้วย 3 ขั้นตอนดังรูปที่ 2.9 ขั้นตอนแรกเลเยอร์ดำเนินการหลายคอนโวลูชันขนานกันไปเพื่อสร้างกลุ่มของตัวกระตุ้นแบบเชิงเส้น ในส่วนขั้นตอนที่สองตัวกระตุ้นเชิงเส้นแต่ละตัววิ่งไหลผ่านฟังก์ชันการกระตุ้นที่ไม่เป็นเชิงเส้น เช่น ฟังก์ชัน Rectified Linear Activation (ReLU) เรียกขั้นตอนนี้อีกอย่างว่าขั้นตอนการตรวจจับ (Detector Stage) ขั้นตอนที่ 3 ใช้ฟังก์ชันพูลลิง (Pooling) เพื่อปรับปรุงเลเยอร์เอาท์พุทเพิ่มเติม โดยฟังก์ชันพูลลิงแทนที่เอาท์พุทที่ตำแหน่งแน่นอนโดยใช้ข้อมูลสรุปเปรียบเทียบกับเอาท์พุทข้างเคียง เช่น Max Pooling จะรายงานเอาท์พุทสูงสุดภายในพื้นที่ข้างเคียงสี่เหลี่ยมดังรูปที่ 2.10 ฟังก์ชันพูลลิงอื่นซึ่งนิยมใช้คือ Average, L^2 Norm หรือ Weighted Average



(ซ้าย) คอนโวลูชันถูกมองเป็นจำนวนย่อยของเลเยอร์ซับซ้อนที่สัมพันธ์กัน
(ขวา) โครงข่ายคอนโวลูชันถูกมองเป็นจำนวนที่ใหญ่กว่าของเลเยอร์อย่างง่าย

รูปที่ 2.9 ส่วนประกอบโดยทั่วไปของเลเยอร์โครงข่ายประสาทเทียมคอนโวลูชัน

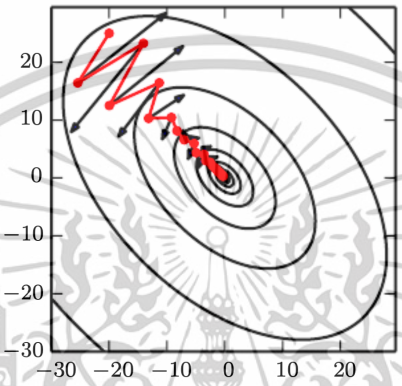


รูปที่ 2.10 การทำงานของฟังก์ชัน Max Pooling

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2.2 Stochastic Gradient Descent with Momentum

อัลกอริทึมออพชันแบบ Stochastic Gradient Descent with Momentum (SGDM) ถูกนำมาใช้ ด้วยอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.0001 โดยอัลกอริทึมออพชันแบบ SGDM สามารถแสดงภาพเพื่อความเข้าใจดังรูปที่ 2.11 และสามารถอธิบายในรูปสมการคณิตศาสตร์ (2.27) ถึง (2.29) ที่ซึ่ง ϵ คืออัตราการเรียนรู้ $x^{(i)}$ คือภาพอินพุท $y^{(i)}$ คือลาเบล (Label) ของรูปภาพ และ θ คือค่าน้ำหนัก (Weight) [13]



รูปที่ 2.11 หลักการ Stochastic Gradient Descent with Momentum (SGDM)

ขั้นตอนแรกต้องคำนวณค่าประมาณกราเดียน (Gradient Estimate)

$$g = \frac{1}{m} \nabla_{\theta} \sum_{i=1}^{m} L\left(f\left(x^{(i)}; \theta\right), y^{(i)}\right) \tag{2.27}$$

จากนั้นคำนวณค่าความเร็วอัพเดท

$$\nu \leftarrow \alpha\nu - \epsilon \nabla_{\theta} \left(\frac{1}{m} \sum_{i=1}^{m} L\left(f\left(x^{(i)}; \theta\right), y^{(i)}\right) \right) \tag{2.28}$$

ที่ซึ่ง α คือ ค่าพารามิเตอร์โมเมนตัม และค่าอัพเดทน้ำหนัก (Weight) จะถูกคำนวณ

$$\theta = \theta + \nu \tag{2.29}$$

จะเห็นได้ว่าค่าความเร็ว υ ได้จากการรวม Gradient Element g, $(\upsilon = \alpha\upsilon - \epsilon g)$
ซึ่งหลักการ SGDM นี้สามารถอธิบายเพิ่มเติมเพื่อความเข้าใจยิ่งขึ้นได้ดังตารางที่ 2.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.1 อัลกอริทึม Stochastic Gradient Descent with Momentum (SGDM)

Stochastic Gradient Descent with Momentum (SGDM)

Require: Learning Rate ϵ, Momentum Parameter α .

Require: Initial Parameter θ, Initial Velocity v .

while Stopping criterion not met **do**

Sample a minibatch of m examples from the training set $\{ x^{(1)}, \ldots, x^{(m)} \}$ with corresponding targets $y^{(i)}$.

Compute gradient estimate: $g = \frac{1}{m} \nabla_{\theta} \sum_{i=1}^{m} L(f (x^{(i)}; \theta) , y^{(i)})$

Compute velocity update: $v = \alpha v - \epsilon g$

Apply update: $\theta = \theta + v$

end while

2.3.3 การเรียนรู้แบบถ่ายโอน (Transfer Learning)

ในงานวิจัยนี้ได้นำเสนอเพื่อใช้งานจริงสำหรับเลเยอร์โครงข่ายที่อบรมไว้ก่อนแล้วซึ่งประกอบด้วย VGG16, VGG19 [14], ResNet50, 101 [15], Xception [16], InceptionV3 [17], InceptionResNetV2 [18], และ Densenet201 [49] คุณลักษณะของแต่ละโครงข่ายอธิบายดังตารางที่ 2.2 และภาพอินพุททั้งหมดจะถูกแปลงเป็นภาพชนิด RGB ก่อนที่จะใช้สำหรับการอบรมโดยอัลกอริทึมโครงข่าย

ตารางที่ 2.2 คุณลักษณะของโครงข่ายการเรียนรู้แบบถ่ายโอนที่ใช้ในงานวิจัยนี้

CNNs	Layers	Parameters (x 10^6)	Input Size
VGG16	16	138	224 x 224
VGG19	19	144	224 x 224
ResNet50	50	25.6	224 x 224
ResNet101	101	44.6	224 x 224
Xception	71	22.9	299 x 299
InceptionV3	48	23.9	299 x 299
InceptionResNetV2	164	55.9	299 x 299
Densenet201	201	8.1	224 x 224

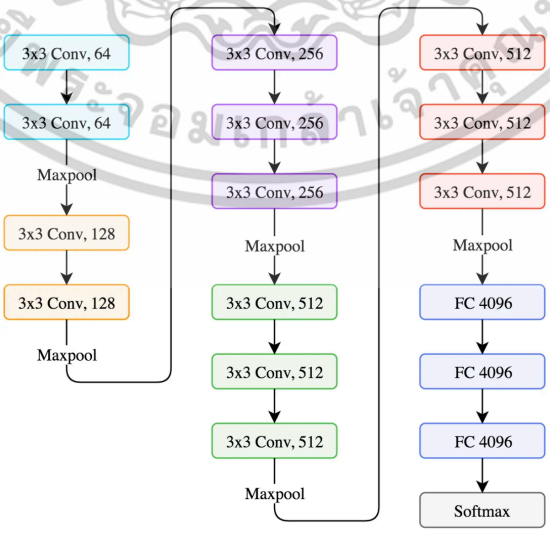
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.3.1 VGG16

VGG นั้นชื่อเต็มคือ Visual Geometry Group ซึ่งเป็นส่วนหนึ่งของคณะวิทยาศาสตร์และวิศวกรรมศาสตร์แห่งมหาลัยออกฟอร์ด โดย VGG นี้ได้นำไปประยุกต์ใช้งานจดจำใบหน้าและจัดหมวดหมู่รูปภาพ เริ่มตั้งแต่ VGG16 จนถึง VGG19 ความตั้งใจเริ่มแรกของนักวิจัยคือ เพื่อทำความเข้าใจว่าความลึกหรือจำนวนชั้นของโครงข่ายคอนโวลูชันมีผลต่อความแม่นยำการจัดหมวดหมู่และจดจำรูปภาพที่มีกลุ่มข้อมูลขนาดใหญ่อย่างไร เพื่อหลีกเลี่ยงพารามิเตอร์ที่มากไปดังนั้นคอนโวลูชันเคอร์เนลขนาดเล็กแบบ 3 x 3 ถูกนำมาใช้ในทุกเลเยอร์ [14]

โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network, CNN) ประกอบด้วยอินพุทเลเยอร์ เอาท์พุทเลเยอร์ และเลเยอร์ที่ซ่อนไว้ (Hidden Layers) อีกหลายชั้น โดย VGG16 เป็น CNN ชนิดหนึ่งที่ได้รับการยอมรับว่าทำงานได้ดีเยี่ยมในทางคอมพิวเตอร์วิชัน โดยคุณลักษณะถูกกำหนดโดยความลึกของมันซึ่งประกอบด้วย 16 เลเยอร์ โดยเป็นคอนโวลูชันเลเยอร์ 13 ชั้น Max Pooling เลเยอร์ 5 ชั้นและเป็นเลเยอร์แบบ Dense หรือ Fully Connected อีก 3 ชั้น ถึงแม้จะรวมทั้งหมดเท่ากับ 21 เลเยอร์ แต่ทว่ามีเลเยอร์แบบ Weight เพียง 16 เลเยอร์เท่านั้น กล่าวคือเป็นเลเยอร์พารามิเตอร์ที่เรียนรู้ได้

VGG16 ต้องใช้รูปภาพอินพุทขนาด 224 x 224 และ 3 RGB Channel โครงสร้างโดยรวมประกอบด้วยเลเยอร์คอนโวลูชัน 5 กลุ่มและตามด้วย Max Pooling ในขณะที่แต่ละเลเยอร์คอนโวลูชันของ AlexNet ประกอบด้วยคอนโวลูชันโอเปอเรชันหนึ่งเดียวเท่านั้น และขนาดของเคอร์เนลคือ 7 x 7 ส่วน VGGNet นั้นแต่ละเลเยอร์คอนโวลูชันจะประกอบด้วยคอนโวลูชันโอเปอเรชัน (Convolution Operations) 2 – 4 อย่างด้วยกัน การปรับปรุงที่เห็นได้อย่างชัดเจนของ VGGNet คือลดขนาดของคอนโวลูชันเคอร์เนล (3 x 3) และเพิ่มจำนวนของคอนโวลูชันเลเยอร์ โครงสร้างของ VGG16 แสดงได้ดังรูปที่ 2.12



รูปที่ 2.12 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ VGG16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.3.2 VGG19

VGG19 เป็นการปรับปรุงเพิ่มเติมของโมเดล VGG ซึ่งจะประกอบด้วย 19 เลเยอร์ แบ่งเป็น 16 เลเยอร์คอนโวลูชัน 3 เลเยอร์ Fully Connected 5 เลเยอร์ Max Pooling และ 1 เลเยอร์ Softmax ซึ่งโครงสร้างของ VGG19 เหมือนกับ VGG16 ที่กล่าวก่อนหน้านี้ ต้องใช้รูปภาพอินพุทขนาด 224 x 224 และ 3 RGB Channel โดยคอนโวลูชันเคอร์เนลขนาดเล็กแบบ 3 x 3 ถูกนำมาใช้ในทุกเลเยอร์ แต่สิ่งที่ต่างกันคือความลึกที่เพิ่มขึ้น เพื่อปรับปรุงความแม่นยำในการจัดหมวดหมู่ให้ดีขึ้น โดยตารางที่ 2.3 แสดงความลึกของ VGG แต่ละแบบที่ถูกเพิ่มขึ้นตามคอลัมน์ A ถึง E ซึ่งเลเยอร์ถูกเพิ่มขึ้นแสดงเป็นอักษรตัวหนา พารามิเตอร์เลเยอร์คอนโวลูชันถูกแสดงในรูปแบบ “conv(Receptive Field Size) - (Number of Channels)” ส่วนฟังก์ชัน ReLU Activation ไม่ได้แสดงเพื่อความกะทัดรัด คอลัมน์ E คือ VGG19 ซึ่งถูกปรับปรุงจาก VGG16 โดยเลเยอร์คอนโวลูชันกลุ่มที่ 3, 4 และ 5 จะเพิ่มคอนโวลูชันขึ้นกลุ่มละ 1 เลเยอร์เป็น 4 x Convolution [14]

ตารางที่ 2.3 โครงสร้างเลเยอร์ของโครงข่ายแบบ VGG แต่ละแบบ

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.3.3 ResNet50 / ResNet101

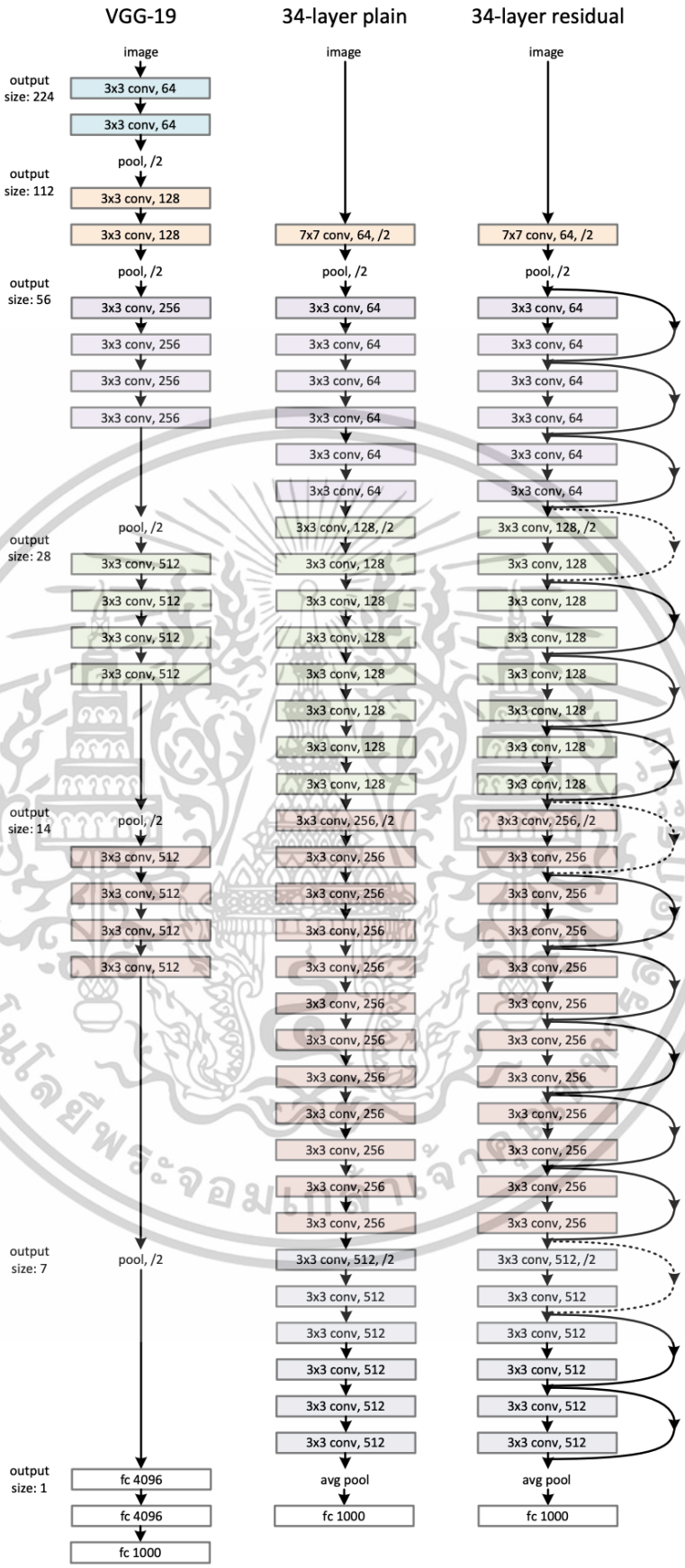
เริ่มแรกนักวิจัยเข้าใจว่ายิ่งโครงข่ายลึกจะเพิ่มความแม่นยำในการจัดหมวดหมู่ภาพ แท้ที่จริงแล้วเมื่อโครงข่ายเพิ่มขึ้นถึงระดับหนึ่งจะทำให้สมรรถนะในการจัดหมวดหมู่ลดลง ซึ่งสิ่งนี้เป็นปัญหาคอขวดอย่างหนึ่งของโครงข่ายแบบ VGG เนื่องจากไม่สามารถสร้างโครงข่ายที่จำนวนชั้นที่ต้องการได้เพราะจะเริ่มสูญเสียความสามารถบางอย่างไป

ปัญหาที่สำคัญอย่างหนึ่งที่ ResNet แก้ไขได้ดีคือปัญหาเรื่อง Vanishing Gradient ที่จะเกิดขึ้นในโครงข่ายคอนโวลูชันที่ลึกเกินไป โดยการเพิ่มทางลัดหรือ Shortcut เข้าไปในโครงข่ายซึ่งจะทำให้โครงข่ายเป็น Residual Network ตามรูปที่ 2.13 แสดงโครงข่ายแบบธรรมดา (Plain Network) แบบ VGG-19 และ เพิ่มเลเยอร์เป็น 34 ชั้นเพื่อทดสอบเปรียบเทียบกับโครงข่ายที่มี 34 เลเยอร์พารามิเตอร์เท่ากันแต่เป็นการต่อแบบ Residual Network โดยรายละเอียดของภาพประกอบด้วย (ซ้าย) โมเดล VGG-19 ซึ่งใช้เพื่ออ้างอิง (กลาง) โครงข่ายแบบธรรมดาประกอบด้วยเลเยอร์พารามิเตอร์ 34 เลเยอร์ (ขวา) โครงข่ายแบบ Residual ประกอบด้วยเลเยอร์พารามิเตอร์ 34 เลเยอร์ โดย Shortcut ที่เป็นเส้นประแสดงถึงขนาดมิติที่เพิ่มขึ้น ส่วนรายละเอียดเพิ่มเติมและ Variant อื่นที่เกี่ยวข้องแสดงดังตารางที่ 2.4

จะเห็นได้ว่าจุดต่อทางลัด (Shortcut Connection) ถูกเพิ่มในแต่ละคู่ของฟิลเตอร์ขนาด 3 x 3 ตามที่แสดงในรูปที่ 2.13 (ขวามือ) ผลลัพธ์ที่ได้จากการใช้โครงข่ายแบบ Residual 34 เลเยอร์ สังเกตการณ์พบว่าความผิดพลาดการอบรมลดลง สิ่งนี้บ่งชี้ว่าปัญหา Degradation ที่เกิดจากความลึกที่เพิ่มขึ้นถูกทำให้ดีขึ้นโดยใช้โครงข่ายแบบ Residual นอกจากนั้นเมื่อเปรียบเทียบกับโครงข่ายแบบธรรมดาที่มี 34 เลเยอร์เท่ากัน ผลการทดลองแสดงให้เห็นว่า Residual Network ลดค่าความผิดพลาดลงได้ 3.5% นั่นยืนยันถึงประสิทธิภาพของ Residual Network เมื่อประยุกต์ใช้งานกับระบบโครงข่ายที่มีความลึกมากได้เป็นอย่างดี [15]

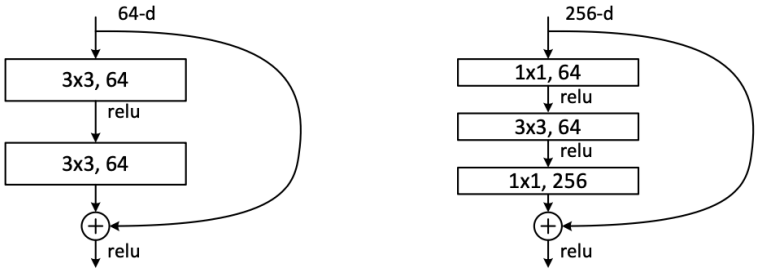
ทีมนักวิจัยปรับปรุงออกแบบโครงสร้างบล็อกแบบ Bottleneck โดยแต่ละ Residual Function ใช้ชั้นเลเยอร์แบบ 3 ชั้นแทนแบบ 2 ชั้นเลเยอร์ดังรูปที่ 2.14 ซึ่งเลเยอร์แบบ 3 ชั้นนี้ประกอบด้วยคอนโวลูชัน 1 x 1, 3 x 3 และ 1 x 1 ตามลำดับ โดย ResNets แบบ 101 เลเยอร์ ถูกสร้างขึ้นโดยใช้มากกว่า 3 บล็อกเลเยอร์ รายละเอียดดังตารางที่ 2.4 ผลลัพธ์ที่ได้แม่นยำกว่าแบบ 34 เลเยอร์รวมทั้งไม่พบปัญหา Degradation ถึงแม้จะเพิ่มความลึกของโครงข่ายขึ้นก็ตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 ตัวอย่างสถาปัตยกรรมโครงข่ายสำหรับ ImageNet

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ซ้าย) โครงสร้างบล็อกสำหรับ ResNet-34 (ขวา) โครงสร้างบล็อกแบบ Bottleneck สำหรับ ResNet-50/101/152

รูปที่ 2.14 ฟังก์ชัน Residual เพื่อเพิ่มความลึก สำหรับ ImageNet

ตารางที่ 2.4 สถาปัตยกรรมของ ImageNet โดยโครงสร้างของบล็อกถูกระบุในวงเล็บ

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

ResNet ที่มี 50 เลเยอร์ แทนที่แต่ละสองเลเยอร์บล็อกที่อยู่ในแบบ 34 เลเยอร์ด้วยเลเยอร์บล็อกแบบคอขวด (Bottleneck Block) สามเลเยอร์ ทำให้ได้ผลลัพธ์คือ ResNet50 โมเดลนี้มี 3.8 Billion Floating Point Operations per second (FLOPs)

ResNet ที่มี 101 เลเยอร์ สร้างเลเยอร์แบบนี้โดยใช้มากกว่าสามเลเยอร์บล็อกตามตารางที่ 2.4 สิ่งที่เห็นได้ชัด ถึงแม้ว่าความลึกเพิ่มขึ้นอย่างมีนัยสำคัญ ซึ่งมี 7.6 Billion FLOPs แต่ก็ยังมีความซับซ้อนน้อยกว่า VGG-16/19 ซึ่งประกอบด้วย 15.3/19.6 Billion FLOPs

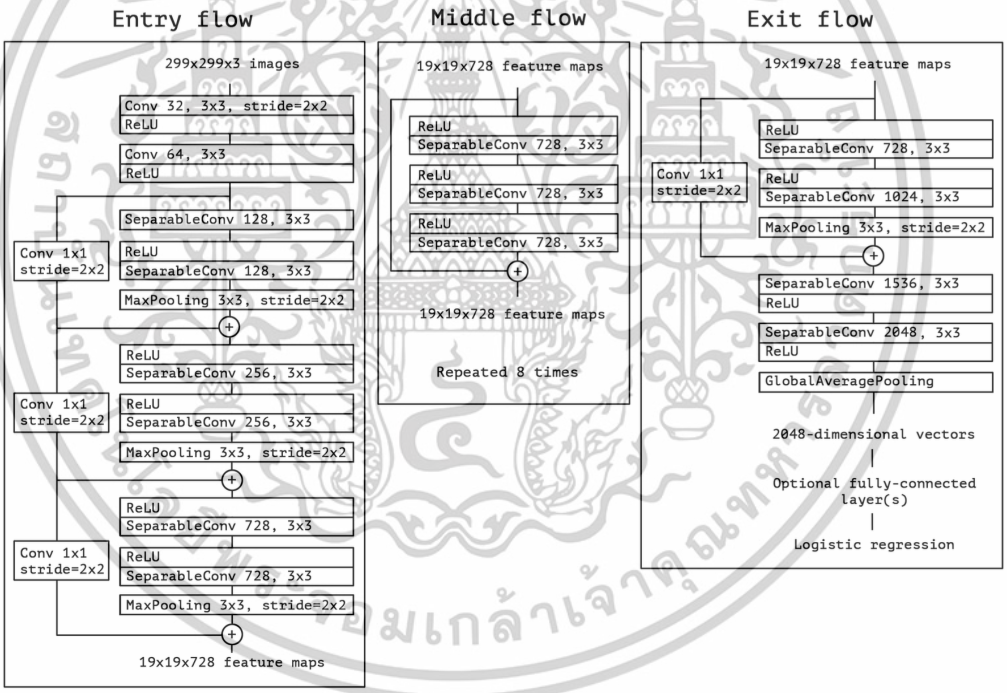
2.3.3.4 Xception

โครงข่ายประสาทเทียมแบบ Xception ได้รับการพัฒนาโดยนักวิจัยของ Google โดยดัดแปลงจาก InceptionV3 เพื่อให้ได้ผลลัพธ์การจำแนกหมวดหมู่ที่ดีขึ้น นักวิจัยสมมุติฐานว่าการ Mapping ของ Cross Channels Correlations และ Spatial Correlations ในฟีเจอร์แมพของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงข่ายประสาทเทียมคอนโวลูชันนั้นสามารถแยกออกจากกันได้อย่างสิ้นเชิง (Entirely Decoupled) เนื่องจากข้อสันนิษฐานนี้เป็นรูปแบบที่เข้มข้นกว่าสถาปัตยกรรมแบบ Inception ดังนั้น คำว่า Xception นั้นเป็นชื่อย่อมาจาก Extreme Inception ซึ่งอยู่บนพื้นฐานของ Depthwise Separable Convolution Layers (A Depthwise Convolution Followed by a Pointwise Convolution) [16] โครงสร้างสถาปัตยกรรมแสดงดังรูปที่ 2.15 ซึ่งประกอบด้วยเลเยอร์คอนโวลูชันจำนวน 36 ชั้นและแบ่งเป็น 14 โมดูลหลักๆ แต่ละโมดูลถูกเชื่อมต่อด้วย Residual Connection อยู่รอบตัวโมดูล ยกเว้นโมดูลแรกและโมดูลสุดท้าย จากรูปข้อมูลจะไหลผ่านจาก Entry Flow และผ่านไปยัง Middle Flow ซึ่งจะเหมือนเดิมทั้งหมดแปดครั้ง และสุดท้ายข้อมูลจะไหลออกไปยัง Exit Flow

กล่าวโดยสรุปสถาปัตยกรรมของ Xception คือ การต่อเป็นชั้นแบบเชิงเส้นของเลเยอร์ Depthwise Separable Convolution ร่วมกับการต่อแบบ Residual Network เพื่อแก้ปัญหาเรื่อง Vanishing Gradient ดังที่กล่าวมาแล้วใน ResNet ด้วยสถาปัตยกรรมแบบนี้ทำให้การทำความเข้าใจและแก้ไขปรับปรุงทำได้ง่ายขึ้น



รูปที่ 2.15 โครงสร้างสถาปัตยกรรมของ Xception

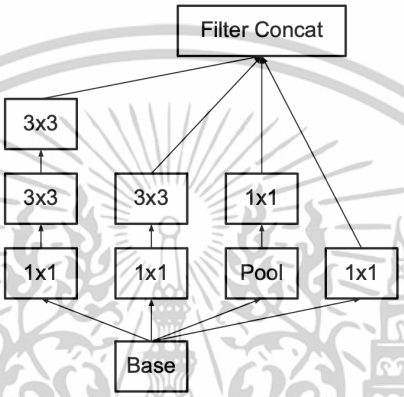
2.3.3.5 InceptionV3

InceptionV3 เป็นโครงข่ายประสาทเทียมคอนโวลูชันที่พัฒนาต่อจากโมเดลพื้นฐาน InceptionV1 หรือที่ถูกแนะนำในชื่อ GoogLeNet ซึ่งได้รับการพัฒนาขึ้นโดยทีมงานของ Google ในปี 2014 แทนที่จะเพิ่มความลึกของเลเยอร์ InceptionV1 เสนอแนวคิดการใช้หลายฟิลเตอร์ที่ขนาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่างกันในระดับเลเยอร์ต่างๆ ดังนั้นเลเยอร์จะถูกต่อขนานออกไปทำให้กว้างขึ้นแทน โดย InceptionV3 ได้รับการปรับปรุงเพื่อให้ทำงานมีประสิทธิภาพดีขึ้นกว่ารุ่นก่อนหน้า การเปลี่ยนแปลงโครงสร้างเลเยอร์ประกอบด้วย [17]

การแยกตัวประกอบเพื่อลดขนาดคอนโวลูชันลง (Factorization) ตัวกรองเชิงพื้นที่ (Spatial Filter) ขนาด 5 x 5 ซึ่งใช้งานกับ InceptionV1 ถูกทดแทนด้วย 3 x 3 จำนวนสองคอนโวลูชันเพื่อลดเวลาและต้นทุนการคำนวณลงรวมทั้งปรับปรุงสมรรถนะให้ดีขึ้น ดังแสดงในรูปที่ 2.16

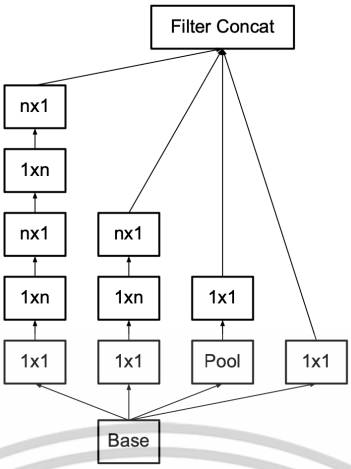


รูปที่ 2.16 การแยกตัวประกอบ (Factorization) ของโมดูลใน InceptionV3

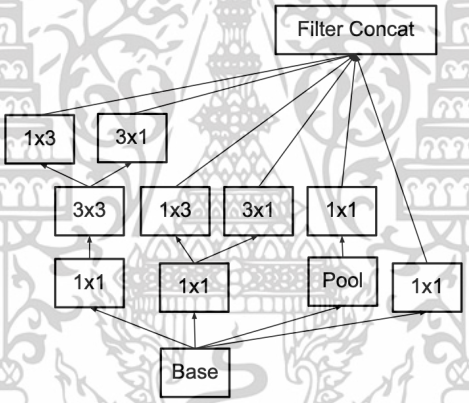
การแยกตัวประกอบเชิงพื้นที่ (Spatial Factorization) ในรูปแบบคอนโวลูชันที่ไม่สมมาตร (Asymmetric) เพื่อลดต้นทุนการคำนวณดังนั้นคอนโวลูชันขนาดสมมาตร n x n จะถูกแทนที่ด้วยคอนโวลูชัน 1 x n และตามด้วย n x 1 ดังแสดงในรูปที่ 2.17 และในโมดูล Inception มีการขยายฟิลเตอร์เพื่อใช้หลักการขนาดมิติที่สูงขึ้น (High Dimensional Representations) แสดงได้ดังรูปที่ 2.18 ซึ่งจะช่วยให้การอบรมโครงข่ายเร็วขึ้น

โครงสร้างเลเยอร์ของ InceptionV3 แสดงได้ดังรูปที่ 2.19 โดยมีขนาดภาพอินพุทเท่ากับ 299 x 299 x 3 และขนาดเอาท์พุทของแต่ละโมดูลจะใช้เป็นขนาดอินพุทของโมดูลถัดไป และรายละเอียดแต่ละคอนโวลูชันเลเยอร์ถูกแสดงในตารางที่ 2.5 ซึ่งจะเห็นได้ว่าโครงข่ายในส่วน Inception ประกอบด้วย 3 โมดูลตามที่แสดงในรูปที่ 2.16, 2.17 และ 2.18 ถึงแม้ว่าขนาดโครงข่ายมีความลึกถึง 42 เลเยอร์ แต่ทว่าต้นทุนการคำนวณมากกว่า GoogLeNet แค่ 2.5 เท่าเท่านั้น รวมทั้งได้รับผลประสิทธิภาพที่ดีกว่า VGGNet อีกด้วย

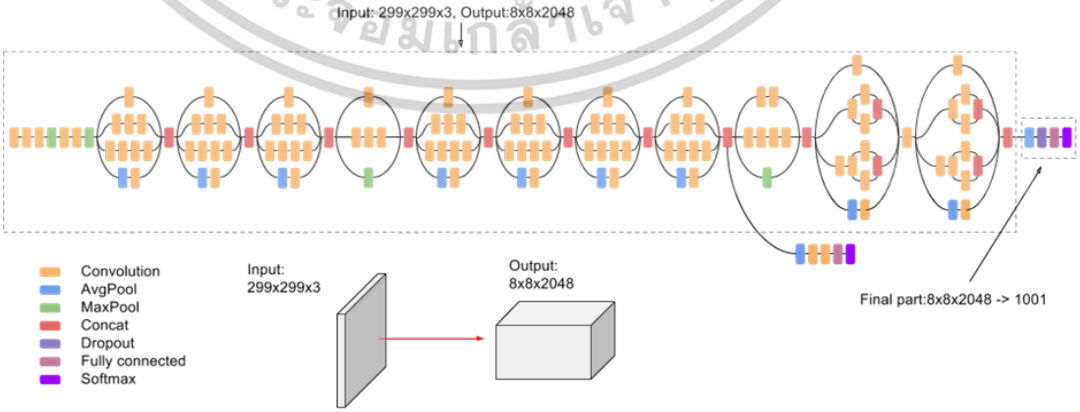
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.17 การแยกตัวประกอบเชิงพื้นที่ (Spatial Factorization) ในโมดูล Inception



รูปที่ 2.18 การขยายฟิลเตอร์เพื่อใช้หลักการขนาดมิติที่สูงขึ้นในโมดูล Inception



รูปที่ 2.19 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ InceptionV3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

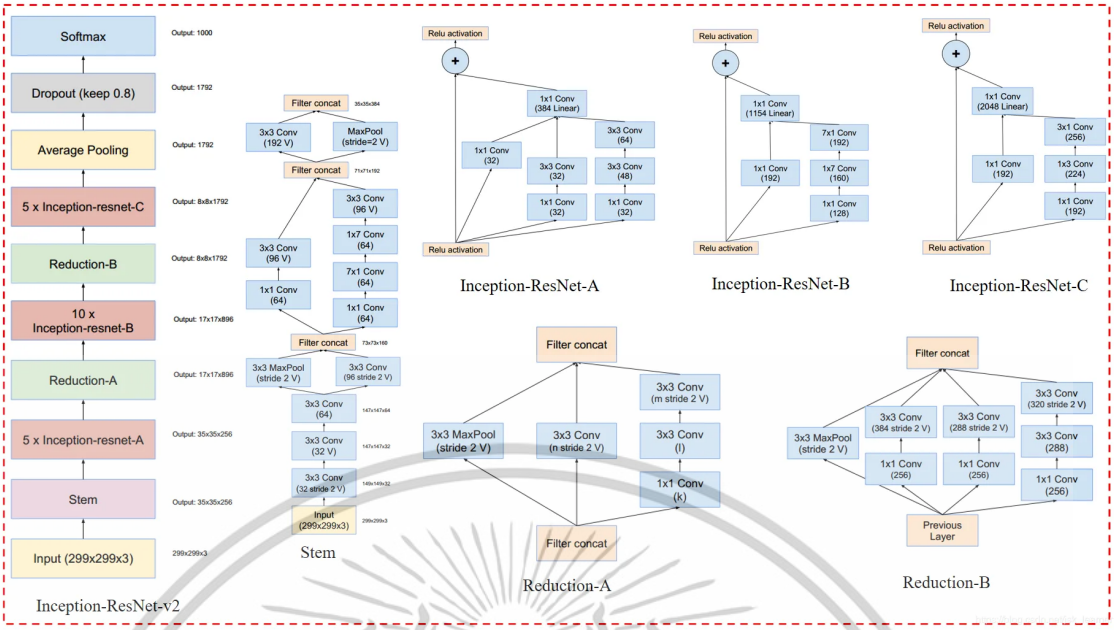
ตารางที่ 2.5 โครงสร้างสถาปัตยกรรมแต่ละเลเยอร์ของ InceptionV3

Type	Patch / Stride Size	Input Size
Conv	3×3/2	299×299×3
Conv	3×3/1	149×149×32
Conv padded	3×3/1	147×147×32
Pool	3×3/2	147×147×64
Conv	3×3/1	73×73×64
Conv	3×3/2	71×71×80
Conv	3×3/1	35×35×192
3 × Inception	Module 1 (ตามรูปที่ 2.16)	35×35×288
5 × Inception	Module 2 (ตามรูปที่ 2.17)	17×17×768
2 × Inception	Module 3 (ตามรูปที่ 2.18)	8×8×1280
Pool	8 × 8	8 × 8 × 2048
Linear	Logits	1 × 1 × 2048
Softmax	Classifier	1 × 1 × 1000

2.3.3.6 InceptionResNetV2

ผู้วิจัยได้ปรับปรุงโครงข่ายเพิ่มเติมจาก Inception ให้ดียิ่งขึ้น โดยเป็นการผสมผสานกับ Residual Network ซึ่งเป็นโครงข่ายที่จัดหมวดหมู่ได้ผลลัพธ์ที่แม่นยำดังที่กล่าวในหัวข้อ 2.3.3.3 ซึ่ง InceptionResNetV2 นี้ได้รับการปรับปรุงเพิ่มเติมโดยใช้โครงข่ายแบบ InceptionV4 ซึ่งเป็นเวอร์ชัน (Version) ล่าสุดและนำมาผสานกับ Residual Network (InceptionResNetV1 พัฒนาต่อยอดจาก InceptionV3) โดยนำข้อดีของแต่ละแบบมาใช้งาน [18] รูปที่ 2.20 แสดงโครงสร้าง InceptionResNetV2 ซึ่งประกอบด้วยหลายส่วนย่อย เช่น Stem, Inception-ResNet-A, B และ C รวมทั้ง Reduction A, B

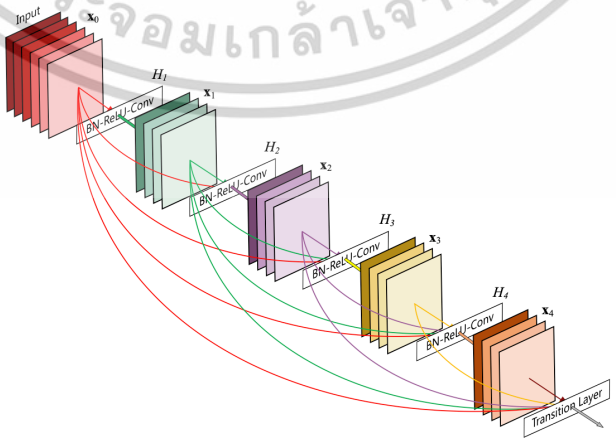
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.20 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ InceptionResNetV2

2.3.3.7 DenseNet201

โครงสร้างสถาปัตยกรรมแบบป้อนไปข้างหน้า (Feedforward) ของ CNN ที่เชื่อมต่อแต่ละเลเยอร์ไปยังทุกๆ เลเยอร์อื่นสำหรับแต่ละเลเยอร์ ฟีเจอร์แมพ (Feature Map) ของทุกเลเยอร์ก่อนหน้าถูกใช้เป็นอินพุท และฟีเจอร์แมพของตัวมันเองถูกใช้เป็นอินพุทสำหรับทุกเลเยอร์ย่อยต่อไป DenseNet มีข้อดีหลายอย่าง ทั้งช่วยลดปัญหาเกี่ยวกับ Vanishing Gradient จุดแข็งเกี่ยวกับการแพร่กระจายของฟีเจอร์ สนับสนุนการนำฟีเจอร์กลับมาใช้ใหม่ และจำนวนพารามิเตอร์ที่ใช้ลดลงอย่างมาก โดย DenseNet ได้รับการปรับปรุงอย่างมีนัยสำคัญเมื่อเทียบกับอุปกรณ์จัดหมวดหมู่ที่ทันสมัยในลักษณะเดียวกัน ในขณะที่ต้องการพารามิเตอร์และเวลาการคำนวณที่น้อยกว่าเพื่อได้รับสมรรถนะที่สูงกว่า [49] โครงสร้างสถาปัตยกรรมแสดงดังรูปที่ 2.21



รูปที่ 2.21 โครงสร้างเลเยอร์ของโครงข่ายประสาทเทียมคอนโวลูชันแบบ Densenet201

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.4 Gradient weighted Class Activation Mapping (Grad-CAM)

การนำมาใช้งานจริงสำหรับ Gradient weighted Class Activation Mapping (Grad-CAM) เพื่อการแสดงให้เห็นผลลัพธ์การทำนายของเครื่องจัดหมวดหมู่ในขั้นตอนสุดท้ายเพื่อการแปรความหมายให้เกิดความโปร่งใสยิ่งขึ้นด้วยการสร้างการอธิบายเชิงรูปภาพ อัลกอริทึมการมองเห็นสามารถแยกแยะระหว่างสองคลาสได้แม่นยำขึ้น เปิดเผยความน่าเชื่อถือ (Trustworthiness) ของเครื่องจัดหมวดหมู่ รวมทั้งระบุความเอนเอียงของกลุ่มข้อมูลได้ด้วย [19], [20] Grad-CAM ใช้ข้อมูลของกราเดียนท์ (Gradient) เพื่อไหลผ่านไปยังเลเยอร์สุดท้ายของคอนโวลูชันของ CNN ดังนั้นจึงสามารถใช้เพื่อมองเห็นทุกการกระตุ้น (Activation) และมีผลต่อการตัดสินใจในโครงข่ายเชิงลึกได้ เพื่อที่จะได้รับแผนที่ตำแหน่งที่แยกแยะแต่ละคลาส Grad-CAM นั้น ขั้นแรกระดับคะแนนของคลาส (Class Score) สำหรับเอาท์พุทคลาส c ในโมเดล CAM หรือในที่นี้ให้คำจำกัดความเป็น y^c ซึ่งจะหาได้ดังสมการที่ (2.30)

$$y^c = \sum_k \alpha_k^c \frac{1}{Z} \sum_i \sum_j A_{ij}^k \tag{2.30}$$

ที่ซึ่ง α_k^c คือค่าน้ำหนัก (Weight) ที่สอดคล้องกับกับคลาส c สำหรับฟีเจอร์แมพ k

A_{ij}^k คือค่าพิกเซลที่ตำแหน่งในฟีเจอร์แมพ k

Z เท่ากับจำนวนรวมของพิกเซลในฟีเจอร์แมพ

อัลกอริทึม Grad-CAM ถูกพัฒนาเพื่อประโยชน์หรือเอาชนะข้อจำกัดของ CAM ที่ซึ่งไม่ได้ขึ้นอยู่กับการเรียนรู้เกี่ยวกับน้ำหนัก (Weight) สำหรับการสร้างแผนที่แบบเชิงเส้น น้ำหนัก Grad-CAM คำนวณได้ตามสมการที่ (2.31)

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{2.31}$$

ฟังก์ชัน ReLU ได้ถูกนำมาประยุกต์ผสมผสานกับน้ำหนักของฟีเจอร์แมพเพื่อคงไว้ซึ่งความสัมพันธ์ที่เป็นค่าบวกในแผนที่ตัวกระตุ้นสุดท้ายเท่านั้น ท้ายที่สุดแผนที่ตำแหน่งที่ถูกแยกแยะของคลาส Grad-CAM ถูกคำนวณตามสมการที่ (2.32)

$$L_{Grad-CAM}^c = ReLU(\sum_k \alpha_k^c A^k) \tag{2.32}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 งานวิจัยที่เกี่ยวข้อง

เมื่อไม่นานมานี้มีงานวิจัยหลากหลายที่เกี่ยวข้องกับการแบ่งแยก (Segmentation) และจัดหมวดหมู่ (Classification) เนื้องอกสมองโดยใช้การเรียนรู้ของเครื่องจักร (Machine Learning) และการเรียนรู้เชิงลึก (Deep Learning) ได้มีการนำเสนอ Hein Tun Zaw และคณะนำเสนอการแบ่งออกเป็นส่วนของภาพถ่ายทางการแพทย์เนื้องอกสมองจากฐานข้อมูล REMBRANDT โดยใช้วิธี Maximum Entropy Threshold เพื่อแยกพื้นที่เนื้องอก ปฏิบัติการสัณฐานวิทยา (Morphology) เพื่อเตรียมกระบวนการ และฟีเจอร์ความเข้มถูกสกัดออกมาสำหรับการจัดหมวดหมู่ซึ่งอยู่บนพื้นฐานของการเรียนรู้เครื่องจักรแบบ Naïve Bayes [21] และสำหรับงานวิจัยที่ผ่านมาเกี่ยวกับการจัดหมวดหมู่เนื้องอกสมองโดยใช้การเรียนรู้เครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน ได้พัฒนาต่อจากงานของ Hein Tun Zaw เพื่อปรับปรุงเพิ่มเติมให้ได้รับความแม่นยำเพิ่มขึ้น โดยใช้ภาพจากฐานข้อมูลเดียวกันคือ REMBRANDT กลุ่มข้อมูลนำมาใช้เพื่อแบ่งออกเป็นส่วนโดยใช้วิธี Global Thresholding และฟีเจอร์ถูกสกัดออกโดยใช้วิธี Gray Level Co-occurrence Matrix ฟีเจอร์ที่ได้รับในรูปตารางข้อมูลถูกป้อนเป็นอินพุทเพื่อจัดหมวดหมู่โดยซัพพอร์ตเวกเตอร์แมชชีน และได้รับผลลัพธ์ความแม่นยำที่ดี [22] เพื่อนำเสนอการแบ่งออกเป็นส่วนและสกัดฟีเจอร์ได้อย่างอัตโนมัติของเนื้องอกสมอง อีกหนึ่งงานที่ผ่านมาได้นำเสนอการจัดหมวดหมู่เนื้องอกสมองโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน กลุ่มข้อมูลขนาดเล็กและไม่สมดุลกันถูกจัดหมวดหมู่อย่างอัตโนมัติโดยใช้การเรียนรู้แบบถ่ายโอน ที่มีชื่อว่า GoogLeNet และ AlexNet ซึ่งได้รับผลลัพธ์ความแม่นยำที่ดี [23] Hossain และคณะนำเสนอการแบ่งออกเป็นส่วนของเนื้องอกสมองโดยใช้วิธีอัลกอริทึมคลัสเตอร์ Fuzzy C-means และตามด้วยใช้เครื่องจัดหมวดหมู่แบบการเรียนรู้เครื่องจักรแบบดั้งเดิมเพื่อเปรียบเทียบกับ CNN [24] N.M. Dipu และคณะนำเสนอการจัดหมวดหมู่เนื้องอกสมองโดยใช้ YOLOv5 และการเรียนรู้เชิงลึกแบบ FastAI กับภาพถ่ายทางการแพทย์จากฐานข้อมูล BraTS2018 ซึ่งประกอบด้วย MRI 1,992 ภาพ [25] S.Pereira และคณะนำเสนอวิธีการแบ่งออกเป็นส่วนอย่างอัตโนมัติของเนื้องอกสมองซึ่งอยู่บนพื้นฐาน CNN 3x3 kernels ขนาดเล็ก และคิดค้นใช้งานการทำความเข้มให้เหมือนกัน (Normalization) ในขั้นตอนเตรียมกระบวนการ [26] Arora และ Sharma นำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบไบนารีโดยใช้การเรียนรู้เครื่องจักรแบบดั้งเดิมหลากหลายแบบ รวมทั้งเปรียบเทียบกับการเรียนรู้แบบถ่ายโอน กล่าวคือ VGG16, InceptionV3 และ ResNet [27] Dubey และคณะนำเสนอการตรวจจับเนื้องอกสมองโดยใช้ Stationary Wavelet Transform (SWT) และอัลกอริทึม Watershed สำหรับการแบ่งออกเป็นส่วน ฟีเจอร์ถูกสกัดออกมาโดยใช้วิธี GLCM ท้ายที่สุดใช้โครงข่ายนิวรอล (Neural Network) เพื่อจัดหมวดหมู่ [28] Gaikward และคณะนำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบไบนารี ขั้นตอนเตรียมกระบวนการและเสริมจำนวนรูปภาพได้นำมาใช้สำหรับกลุ่มข้อมูลรูปภาพก่อนนำไปจัดหมวดหมู่โดยใช้การเรียนรู้แบบถ่ายโอน VGG16 [29] Kumar Baranwal และคณะนำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบหลายคลาสโดยใช้การเรียนรู้เครื่องจักรแบบดั้งเดิมแบบซัพพอร์ตเวกเตอร์แมชชีน และเทียบกับ CNN [30] Hareem Kibriya และ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะนำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบหลายคลาสโดยใช้การเรียนรู้แบบถ่ายโอน กล่าวคือ ResNet18 และ GoogLeNet รวมทั้งการนำมาใช้ของซัพพอร์ตเวกเตอร์แมชชีน เพื่อปรับปรุงความแม่นยำด้วย [31] Aboussaleh และคณะนำเสนอการแบ่งออกเป็นส่วนและทำนายเนื้องอกสมองโดยใช้โครงข่าย CNN ที่สร้างขึ้นเองเพื่อหลีกเลี่ยงกระบวนการเตรียมฉลากของรูปภาพ โดยแบ่งขั้นตอนออกเป็นสองเฟส ลำดับแรกภาพ MRI ถูกจัดหมวดหมู่เป็นเนื้องอกหรือปกติโดยโมเดล CNN ฉลากที่แท้จริงถูกใช้ในขั้นตอนนี้ ขั้นที่สองการแบ่งออกเป็นส่วนถูกประยุกต์ใช้กับรูปภาพที่มีเนื้องอก โดยใช้ฟีเจอร์ที่สกัดได้จากเลเยอร์คอนโวลูชันสุดท้ายจากขั้นตอนแรก ท้ายที่สุดเพื่อปรับปรุงผลลัพธ์ขั้นตอนหลังกระบวนการถูกประยุกต์ใช้ [32] Wazif Reza และคณะนำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบหลายคลาสโดยใช้สถาปัตยกรรม VGG16 ที่ได้รับการปรับแต่ง [33] Ayesha Younis และคณะนำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบไบนารีโดยใช้การเรียนรู้เชิงลึกหลากหลายแบบ กล่าวคือ CNN, VGG16 และ Ensemble Learning [34] Hanming Hu และคณะนำเสนอการตรวจจับและวินิจฉัยเนื้องอกสมองโดยใช้การเรียนรู้เชิงลึก CNNs หลายแบบเพื่อทดสอบประสิทธิภาพและความแม่นยำเทียบกันซึ่งประกอบด้วย VGG16/19, AlexNet, GoogLeNet, ResNet และนำเสนอ YOLO เพื่อทำเครื่องหมายตำแหน่งที่ตั้งจริงของเนื้องอก [35] S. Asif และคณะนำเสนอการตรวจจับเนื้องอกสมองโดยใช้การเรียนรู้แบบถ่ายโอนที่ใช้อัลกอริทึมออพติไมเซอร์ (Optimizers) แบบต่างๆ กล่าวคือ ADAM, SGD และ RMSprop ผลลัพธ์การทดลองแสดงให้เห็นว่าสถาปัตยกรรม Xception ร่วมกับออพติไมเซอร์แบบ ADAM ให้ผลลัพธ์ที่ดีกว่าโมเดลแบบอื่นโดยได้รับความเที่ยงตรง (Accuracy) 99.67 เปอร์เซ็นต์ ความไว (Sensitivity) 99.68 เปอร์เซ็นต์ ความแม่นยำ (Precision) 99.68 เปอร์เซ็นต์ และความเฉพาะเจาะจง (Specificity) 99.66 เปอร์เซ็นต์ตามลำดับ [36] S. Mohsen และคณะนำเสนอการจัดหมวดหมู่เนื้องอกสมอง Glioma และ Pituitary โดยใช้ VGG19 และ ResNext101_32x8d โดยกลุ่มข้อมูลประกอบด้วยภาพ MRI 900 ภาพสำหรับแต่ละชนิดเนื้องอกสมอง นอกจากนี้เทคนิคภาพเดี่ยวแบบความละเอียดสูง (SISR) ได้นำมาใช้เพื่อปรับปรุงความละเอียดภาพก่อนการจัดหมวดหมู่ ผลลัพธ์การทดลองแสดงให้เห็นว่าความแม่นยำที่ได้รับจาก VGG19 คือ 99.89 เปอร์เซ็นต์ และ ResNext101_32x8d คือ 100 เปอร์เซ็นต์ ตามลำดับ [37]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การวินิจฉัยเนื้องอกสมองด้วยการเรียนรู้เครื่องจักรและโครงข่ายประสาทเทียม CNN

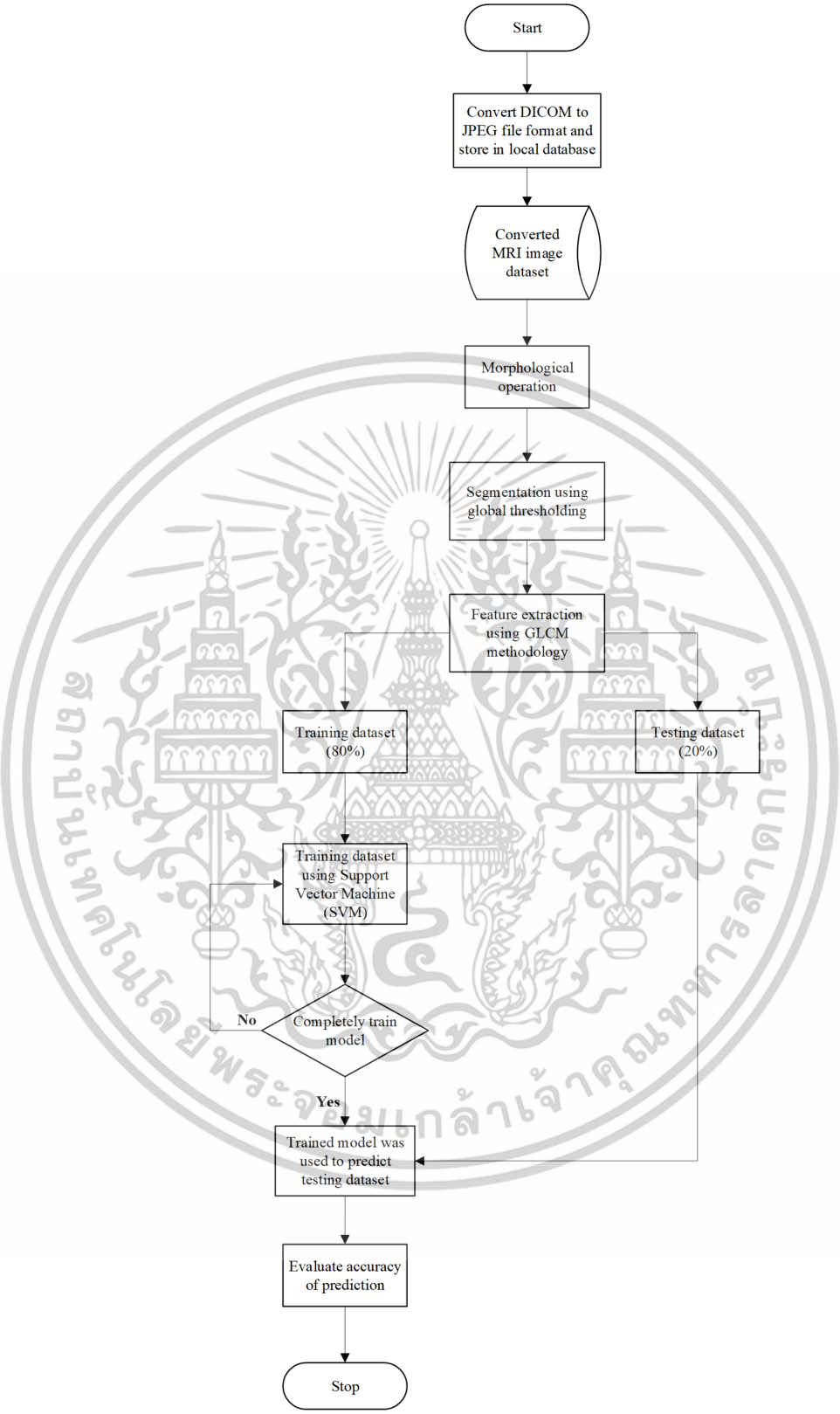
ในบทนี้กล่าวถึงการจัดหมวดหมู่เนื้องอกสมองที่มีกลุ่มข้อมูลจำนวนไม่มากนัก โดยใช้หลักการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน เพื่อเปรียบเทียบกับการเรียนรู้แบบถ่ายโอน หรือเรียกอีกอย่างว่าการเรียนรู้เชิงลึกแบบ CNN ซึ่งในส่วนนี้นำเสนอสองแบบด้วยกัน คือ GoogLeNet และ AlexNet เพื่อเปรียบเทียบผลลัพธ์ข้อดีและข้อเสียของอุปกรณ์จัดหมวดหมู่แต่ละประเภท สำหรับวิธีการจัดหมวดหมู่ด้วยซัพพอร์ตเวกเตอร์แมชชีน การเตรียมกระบวนการ แบ่งออกเป็นส่วน และสกัดฟีเจอร์ เป็นขั้นตอนที่ต้องทำด้วยมือ ในขณะที่การเรียนรู้แบบถ่ายโอนในส่วนนี้ภาพถ่ายทางการแพทย์ไม่ผ่านการเตรียมกระบวนการ และใช้กลุ่มข้อมูลจำนวนน้อยและไม่สมดุลกัน

3.1 การวินิจฉัยเนื้องอกสมองกลิโอมาด้วยการเรียนรู้ของเครื่องจักร

3.1.1 กระบวนวิธีการทดลอง

ในบทนี้กล่าวถึงการจำแนกเนื้องอกสมองกลิโอมาระดับรุนแรงที่เรียกว่า Glioblastoma (GBM) โดยใช้การเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน ซึ่งเป็นการศึกษาและปรับปรุงความแม่นยำจากงานวิจัยที่ผ่านมาซึ่งนำเสนอโดย Hein Tun Zaw และคณะ [21] ที่สามารถตรวจจับเนื้องอกส่วนกลีบขมับได้ดีเป็นพิเศษ แต่งานวิจัยนี้ปรับปรุงเพื่อการตรวจจับเนื้องอกสมองชนิด GBM ที่เกิดขึ้นในทุกส่วนของสมอง โดยงานวิจัยนี้ได้นำเทคนิคการประมวลผลภาพมาใช้รวมทั้งนำเสนอการเรียนรู้ของเครื่องจักรที่ได้รับการควบคุมดูแลเพื่อแบ่งแยกหมวดหมู่ภาพทางการแพทย์ ซึ่งขั้นตอนการดำเนินการแสดงดังโฟลว์ชาร์ต (Flowchart) รูปที่ 3.1 เริ่มแรกกลุ่มข้อมูลที่ได้รับจากฐานข้อมูล REMBRANDT ซึ่งจะแปลงรูปภาพจาก DICOM เป็น JPEG ฟอร์แมทและภาพที่ได้จะถูกเก็บไว้ในฐานข้อมูลส่วนบุคคลเพื่อทำงานต่อไป ปฏิบัติการสัณฐานวิทยานำมาใช้เพื่อแยกส่วนพื้นที่สมองออกมาเพื่อง่ายต่อการทำงานถัดไป จากนั้นนำภาพพื้นที่สมองที่ได้นำมาแยกออกเป็นส่วน (Segmentation) โดยใช้วิธีเทรชโฮลด์ค่าเดียวกันทั้งภาพ (Global Threshold) เพื่อให้ได้ภาพส่วนก้อนเนื้องอกอย่างเดียว ก่อนที่จะนำข้อมูลที่ได้ไปสกัดฟีเจอร์โดยใช้วิธี Gray Level Co-occurrence Matrix (GLCM) ค่าฟีเจอร์เหล่านี้จะถูกจัดเรียงในรูปแบบตารางและแบ่งออกเป็น 80 เปอร์เซ็นต์ สำหรับกลุ่มข้อมูลอบรม และ 20 เปอร์เซ็นต์ สำหรับกลุ่มข้อมูลทดสอบ การเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน ถูกนำมาใช้กับกลุ่มข้อมูลอบรมและได้รับโมเดลที่จะเป็นเครื่องมือในการจัดหมวดหมู่หรือทำนายผลสำหรับกลุ่มข้อมูลทดสอบ เมื่อได้รับผลลัพธ์การทำนายพารามิเตอร์ความแม่นยำจะถูกประเมินเพื่อวัดผลประสิทธิภาพของโมเดลการทำนายในขั้นตอนสุดท้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.1 โฟลว์ชาร์ตการวินิจฉัยเนื้องอกด้วยการเรียนรู้แบบซัพพอร์ตเวกเตอร์แมชชีน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

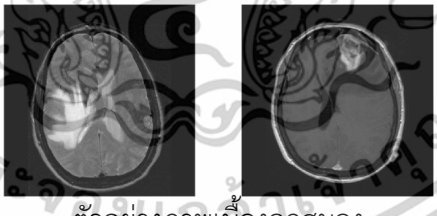
3.1.2 การออกแบบการทดลอง (Experimental Setup)

รูปภาพ MRI จากฐานข้อมูล REMBRANDT ถูกแปลงเป็น JPEG ฟอร์แมทและเก็บในฐานข้อมูลส่วนบุคคล (Local Database) รูปภาพรวมทั้งหมดที่ใช้ในการทดลองนี้จำนวน 155 รูปภาพ ขนาดของภาพ 630 x 630 พิกเซล ซึ่งเป็นภาพสมองปกติ 44 ภาพ เนื้องอกสมองกลิโอมาระดับรุนแรงเกรด 3 หรือ 4 ซึ่งจะเรียกอีกอย่างว่า Glioblastoma (GBM) จำนวน 111 ภาพ รูปภาพเหล่านี้นำไปผ่านขั้นตอนเตรียมกระบวนการ (Preprocessing) สกัดฟีเจอร์ออกมาและถูกจัดเรียงในรูปแบบตารางก่อนที่จะได้รับการจัดสรรเป็นกลุ่มข้อมูลอบรมและทดสอบในอัตราส่วน 80 เปอร์เซ็นต์ และ 20 เปอร์เซ็นต์ ตามลำดับ รายละเอียดจำนวนรูปภาพดังที่กล่าวมาแสดงดังตารางที่ 3.1 ตัวอย่างภาพเนื้องอกสมอง GBMs และสมองปกติถูกแสดงไว้ในรูปที่ 3.2

ตารางที่ 3.1 รูปภาพ MRI ที่ถูกแปลงเป็นรูปแบบ JPEG และใช้ในการทดลอง

Dataset	Normal	Glioma	Total	Training	Testing	Dimensions
REMBRANDT	44	111	155	123	32	630 x 630 pixel

ตัวอย่างภาพสมองปกติ



ตัวอย่างภาพเนื้องอกสมอง

รูปที่ 3.2 ตัวอย่างของภาพ MRI ถูกแปลงเป็นฟอร์แมท JPEG

3.1.3 ขั้นตอนการเตรียมกระบวนการ

ขั้นตอนนี้ภาพ MRI ถูกแปลงเป็นฟอร์แมท JPEG และค่าความละเอียดภาพถูกปรับแต่งเป็น 630 x 630 พิกเซล การปฏิบัติการทางสัณฐานวิทยาและการหักล้างพิกเซล (Pixel Subtraction) ถูกใช้เพื่อการแยกส่วนกะโหลกออกจากพื้นที่สมอง ขั้นตอนแรกรูปภาพถูกแปลงเป็นเกรย์สเกลและใช้วิธีการแบบปฏิบัติการทางสัณฐานวิทยาแบบเปิด (Opening Morphological) ด้วยโครงสร้างแบบดิสค์ถูกใช้เพื่อแยกวัตถุเล็กๆ และเส้นบางๆ ออกจากภาพในขณะที่ยังคงรักษารูปร่างและขนาดวัตถุที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

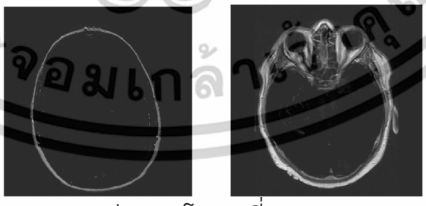
ใหญ่กว่าในภาพ ฟังก์ชันการลบภาพพื้นหลังถูกใช้ในขั้นตอนต่อมาเพื่อแยกกะโหลกออกจากพื้นที่สมอง ด้วยพื้นที่สมองที่หลงเหลืออยู่เท่านั้นจะเป็นการดีกว่าสำหรับขั้นตอนการแบ่งออกเป็นส่วน (Segmentation) โดยงานวิจัยส่วนนี้ใช้โปรแกรม MATLAB R2020b ในการแบ่งออกเป็นส่วนและการจัดหมวดหมู่ของเนื้องอกสมอง GBM

3.1.3.1 การปฏิบัติการทางสัณฐานวิทยา (Morphological)

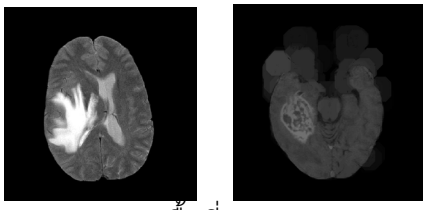
การปฏิบัติการทางสัณฐานวิทยาแบบเปิด (Opening Morphological) คือ การทำ Erosion และตามด้วย Dilation โดยใช้โครงสร้างอิลิเมนท์แบบเดียวกันสำหรับทั้งสองปฏิบัติการ งานวิจัยนี้ในขั้นตอนเตรียมกระบวนการใช้ปฏิบัติการทางสัณฐานวิทยาแบบเปิดและโครงสร้างแบบดิสค์ โดย Opening ของเซต A ด้วยโครงสร้างอิลิเมนท์ B ถูกระบุดังสมการที่ (2.20) ที่กล่าวมาแล้วก่อนหน้านี้ ด้วยกระบวนการดังกล่าวส่วนกะโหลกถูกแยกออกมาโดยการ Opening และเหลือพื้นที่สมองเท่านั้น อย่างไรก็ตามบางพิกเซลของสมองถูกแยกออกไปด้วย ดังนั้นฟังก์ชันการลบพิกเซลถูกใช้เพื่อแยกส่วนพื้นที่สมองออกจากส่วนกะโหลก

3.1.3.2 การลบหรือหักล้างภาพ (Image Subtraction)

การลบพิกเซลถูกใช้เพื่อแยกส่วนกะโหลกออกจากพื้นที่สมองเพื่อเตรียมความพร้อมในขั้นตอนการแบ่งออกเป็นส่วนต่อไป ขั้นแรกภาพที่ได้ผ่านกระบวนการ Opening แล้วถูกหักล้างออกจากภาพต้นแบบเพื่อให้ได้ภาพกะโหลกและส่วนแปลกปลอมพื้นที่สมองเท่านั้น จากนั้นภาพที่ Opening แล้วถูกลบออกจากภาพกะโหลกและสิ่งแปลกปลอมที่ได้จากขั้นตอนก่อนหน้าเพื่อให้ได้รับภาพกะโหลกอย่างเดียว ท้ายที่สุดภาพกะโหลกศีรษะที่ถูกแยกออกมาจะถูกลบจากภาพต้นแบบอีกครั้งและได้รับพื้นที่สมองเท่านั้นซึ่งพร้อมสำหรับการแบ่งออกเป็นส่วนในขั้นตอนถัดไป โดยภาพการแยกพื้นที่สมองออกจากกะโหลกแสดงดังรูปที่ 3.3



ภาพส่วนกะโหลกที่ถูกแยกออกมา



ภาพพื้นที่สมอง

รูปที่ 3.3 ตัวอย่างของภาพพื้นที่สมองสำหรับขั้นเตรียมกระบวนการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

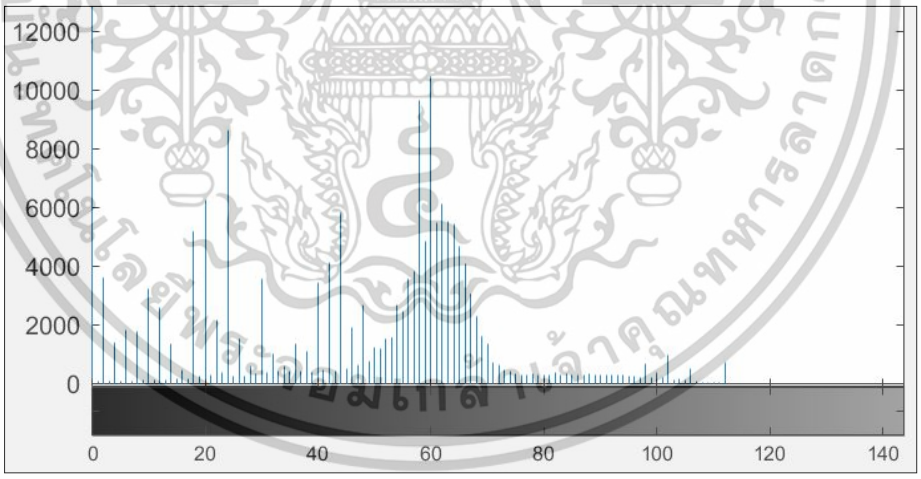
3.1.4 การแบ่งออกเป็นส่วน (Segmentation)

ในงานวิจัยนี้วิธีการเทรชโฮลด์ค่าเดียวกันทั้งภาพถูกนำมาใช้เพื่อการแบ่งออกเป็นส่วนและจำแนกความแตกต่างพื้นที่เนื้องอกในสมองและภาพพื้นหลังเพื่อดำเนินการต่อไป ค่าเทรชโฮลด์ที่เหมาะสมถูกเลือกโดยการตรวจสอบด้วยการสังเกตด้วยตาเปล่าของภาพฮิสโตแกรม วิธีการเพื่อสกัดแยกวัตถุออกจากพื้นหลังโดยการเลือกค่าเทรชโฮลด์ที่เหมาะสมที่สุด ทุกๆ ตำแหน่ง (x, y) ที่ซึ่งฟังก์ชัน f (x, y) มีค่ามากกว่า T ถูกเรียกว่าจุดตำแหน่งของวัตถุ และในทางตรงข้ามถ้ามีค่าน้อยกว่า T ถูกเรียกว่าจุดพื้นหลัง [7], [10] ภาพเทรชโฮลด์ g (x, y) ถูกระบุเป็นสมการได้ดังสมการที่ (3.1)

$$g(x, y) = 1 \text{ if } f(x, y) \geq T \qquad (3.1)$$

$$g(x, y) = 0 \text{ if } f(x, y) < T$$

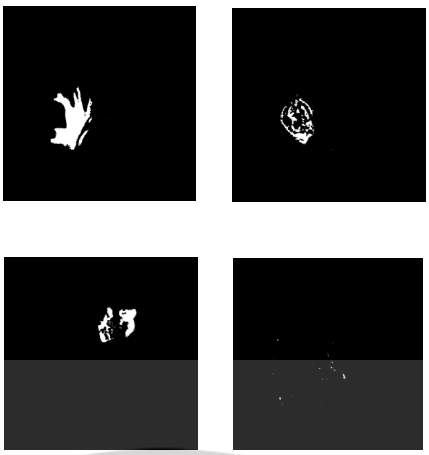
แผนภาพฮิสโตแกรมของการเตรียมกระบวนการภาพสมองถูกแสดงดังรูปที่ 3.4 โดยจากการสังเกตด้วยตาเปล่าพบว่าจุดตำแหน่งพื้นหลังคือค่าเทรชโฮลด์น้อยกว่า 85 และในทางตรงกันข้ามค่าที่มากกว่าคือวัตถุเนื้องอกสมอง โดยทั่วไปค่าเทรชโฮลด์ถูกทำเป็นนอมอลไลซ์โดยมีย่านอยู่ที่ [0, 1] ดังนั้นจะได้ค่าเทรชโฮลด์ที่เหมาะสมเท่ากับ 0.33 ซึ่งได้มาจาก 85/255 นั่นเอง



รูปที่ 3.4 แผนภูมิฮิสโตแกรมของภาพสมองในขั้นตอนการแบ่งออกเป็นส่วน

ค่าเทรชโฮลด์ถูกปรับแต่งละเอียดจนกระทั่งได้รับผลลัพธ์ที่ดีสุดโดยเนื้องอกจะเห็นได้ชัดเจนเพื่อจำแนกออกเป็นส่วนของภาพค่าเทรชโฮลด์ถูกใช้งานในฟังก์ชัน Imbinarize และผลลัพธ์การแบ่งออกเป็นส่วน (Segmentation) ถูกแสดงในรูปที่ 3.5 ภาพที่ถูกแบ่งแยกของทั้งเนื้องอกสมองและสมองปกติเตรียมพร้อมสำหรับการสกัดแยกฟีเจอร์ออกมาในขั้นตอนต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 ภาพที่ถูกแบ่งออกเป็นส่วนของทั้งเนื้องอกสมองและสมองปกติ

3.1.5 การสกัดฟีเจอร์ด้วยวิธีเมทริกซ์การเกิดร่วมของระดับสีเทา

การสกัดแยกออกมาของฟีเจอร์ คือ กระบวนการเพื่อเตรียมข้อมูลตัวแทนเพื่อใช้สำหรับการจัดหมวดหมู่แบบไบนารีในตัวจัดหมวดหมู่โดยใช้การเรียนรู้ของเครื่องจักร Gray Level Co-occurrence Matrix (GLCM) คือ วิธีการทางสถิติของการตรวจสอบพื้นผิวที่พิจารณาความสัมพันธ์เชิงพื้นที่ของพิกเซลและการสกัดแยกฟีเจอร์ออกมาด้วยวิธีเมทริกซ์การเกิดร่วมของระดับสีเทา ถูกนำมาใช้ในงานวิจัยนี้ ขั้นตอนแรกภาพเกรย์สเกลที่ถูกแบ่งออกเป็นส่วนถูกแปลงเป็นรูปเมทริกซ์โดยใช้ฟังก์ชันก์ graycomatrix จากนั้นได้รับค่าทางสถิติหลายค่าโดยใช้ฟังก์ชันก์ graycoprops ค่าทางสถิติเหล่านี้จะเตรียมข้อมูลเกี่ยวกับพื้นผิวของภาพ โดยค่าสถิติเรียกว่า Contrast, Correlation, Energy และ Homogeneity โดยจะถูกอธิบายเพิ่มเติมดังนี้

Contrast วัดค่าความเข้มที่ตัดกันระหว่างพิกเซลและส่วนข้างเคียงตลอดทั้งภาพ โดยค่าคอนทราสจะเป็น 0 สำหรับภาพแบบคงที่หรือพื้นผิวเดียวกันทั้งหมด

Correlation วัดค่าความสัมพันธ์กันระหว่างพิกเซลและพื้นผิวข้างเคียง และวัดตลอดทั้งภาพ ย่านวัดของค่านี้คือ [1, -1] ซึ่งสอดคล้องกับค่าความสัมพันธ์บวกและลบอย่างสมบูรณ์แบบ โดยค่าความสัมพันธ์กันจะไม่ใช่จำนวนตัวเลขในกรณีภาพแบบคงที่หรือพื้นผิวเดียวกันทั้งหมด

Energy วัดค่าของความสม่ำเสมอซึ่งอยู่ในย่าน [0, 1] โดยค่านี้จะมีค่าเท่ากับ 1 สำหรับภาพแบบคงที่หรือพื้นผิวเดียวกันทั้งหมด

Homogeneity วัดค่าการกระจายของอิลิเมนท์ใน GLCM ค่านี้เท่ากับ 1 สำหรับเส้นทแยงมุม GLCM

ฟีเจอร์เมทริกซ์การเกิดร่วมของระดับสีเทาทั้งหมดถูกสกัดแยกออกมาและสรุปในรูปแบบตาราง และแสดงตัวอย่างในตารางที่ 3.2 โดยกำหนดให้คลาสไบนารีของฟีเจอร์ถูกระบุเป็น 1 สำหรับภาพเนื้องอกสมองและ 0 สำหรับภาพสมองปกติและนำไปสู่กระบวนการเพิ่มเติมเพื่อจัดหมวดหมู่โดยใช้การเรียนรู้ของเครื่องจักร

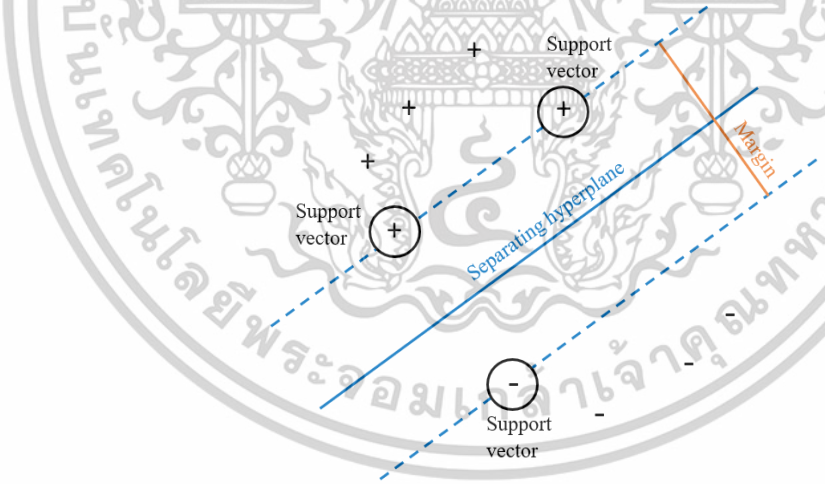
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 ตัวอย่างของฟีเจอร์ที่ถูกสกัดแยกออกมาโดยใช้วิธี GLCM

Contrast	Correlation	Energy	Homogeneity	Class
0.0725	0.9064	0.9827	0.9987	1
0.064	0.9463	0.9743	0.9988	1
0.0265	0.9528	0.988	0.9995	1
0.0027	0.5412	0.9998	0.9999	0

3.1.6 การจัดหมวดหมู่ด้วยการเรียนรู้แบบซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน คือ การเรียนรู้ของเครื่องจักรชนิดที่ต้องดูแลควบคุมซึ่งใช้กันอย่างกว้างขวางสำหรับการจัดหมวดหมู่แบบไบนารีในงานวิจัยเกี่ยวกับการวินิจฉัยโรคจากภาพทางการแพทย์ [38], [39], [40], [41] ซัพพอร์ตเวกเตอร์แมชชีนจัดหมวดหมู่ข้อมูลโดยการหาเส้นไฮเปอร์เพลนที่ดีสุดเพื่อใช้แบ่งแยกกลุ่มข้อมูลออกจากกันกับอีกกลุ่มอื่น ไฮเปอร์เพลนที่ดีสุดหมายถึงค่ามาร์จินที่มากสุดระหว่างสองคลาส และซัพพอร์ตเวกเตอร์ คือ จุดที่ใกล้กับเส้นไฮเปอร์เพลนที่สุด [8] โดยคำนิยามดังกล่าวสามารถดูรูปที่ 3.6 ประกอบการอธิบาย กำหนดให้ค่าบวกคือข้อมูลที่อยู่ในกลุ่ม 1 และค่าลบคือข้อมูลที่อยู่ในกลุ่ม -1



รูปที่ 3.6 ภาพประกอบการอธิบายซัพพอร์ตเวกเตอร์แมชชีน

เมื่อกำหนดให้กลุ่มข้อมูล $D = \{(x_i, y_i), i = 1 - n\}$ และ $y_i = \in \{-1, 1\}$,
ดังนั้นซัพพอร์ตเวกเตอร์แมชชีนถูกระบุเป็นสมการได้ดังนี้

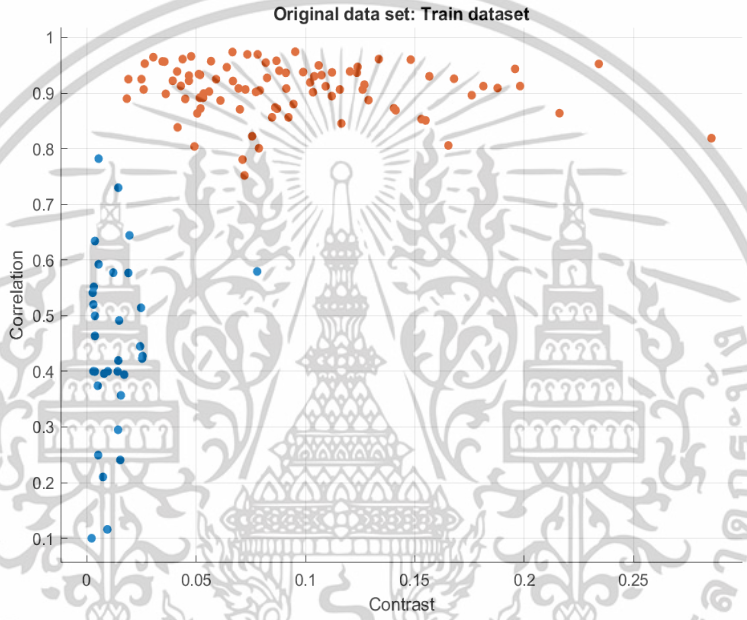
$$\min_{w,b,\zeta} \frac{1}{2}\|w\|^2 + C\sum_{i}^{n}\zeta_i \tag{3.2}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

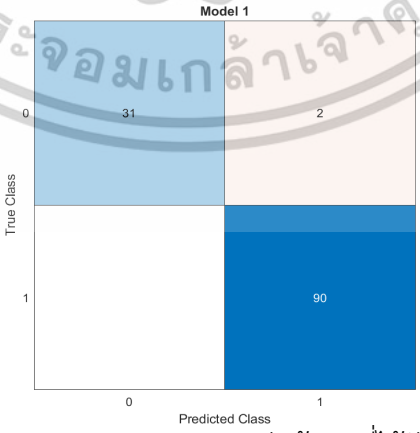
ขึ้นอยู่กับ $y_i(w \cdot x_i + b) \geq 1 - \zeta_i, \zeta_i \geq 0, i = 1 - n,$

ที่ซึ่ง ζ_i คือตัวแปร Slack Variables และ C คือพารามิเตอร์ Regularization Parameter

ฟีเจอร์ที่ถูกสกัดแยกออกมาในรูปแบบตารางถูกป้อนเป็นอินพุทเข้าไปยังตัวจัดหมวดหมู่แบบซัพพอร์ตเวกเตอร์แมชชีน โดยกลุ่มข้อมูลฟีเจอร์ประกอบไปด้วยภาพเนื้องอกสมอง 111 ฟีเจอร์และสมองปกติอีก 44 ฟีเจอร์ ลำดับแรกกลุ่มข้อมูลฟีเจอร์ถูกจัดเรียงอย่างสุ่มเพื่อการเตรียมความพร้อมและแบ่งกลุ่มเป็นข้อมูลเพื่ออบรมและทดสอบด้วยอัตราส่วน 80 : 20 ตามลำดับ ดังนั้นจะได้กลุ่มข้อมูลเพื่อการอบรม 123 ฟีเจอร์ซึ่งจะถูกป้อนไปยังตัวจัดหมวดหมู่ซัพพอร์ตเวกเตอร์แมชชีน และจากนั้นจะได้ผลลัพธ์ซึ่งแสดงในรูป Scatter Plot และ Confusion Matrix ตามที่แสดงในรูปที่ 3.7



ก. Scatter Plot ของกลุ่มข้อมูลที่ได้รับการอบรมแล้ว



ข. Confusion Matrix ของกลุ่มข้อมูลที่ได้รับการอบรมแล้ว

รูปที่ 3.7 Scatter Plot และ Confusion Matrix ของกลุ่มข้อมูลที่ได้รับการอบรมแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากนั้นโมเดลที่ได้รับการอบรมถูกนำไปใช้เพื่อทดสอบกลุ่มข้อมูลที่เหลืออีก 32 ฟีเจอร์เพื่อทำนายและจัดหมวดหมู่ว่ารูปภาพดังกล่าวเป็นสมองปกติหรือเนื้องอกสมองต่อไป และผลลัพธ์ที่ได้ถูกอธิบายเพิ่มเติมในบทที่ 5

3.2 การวินิจฉัยเนื้องอกสมองกลิโอมาด้วยโครงข่ายประสาทเทียม CNN

3.2.1 กระบวนวิธีการทดลอง

ในส่วนนี้การเรียนรู้เชิงลึกของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN) ได้นำมาใช้ประโยชน์เพื่อการจำแนกแบบไบนารีของเนื้องอกสมอง GBM กลุ่มข้อมูลจากฐานข้อมูล REMBRANDT ซึ่งประกอบด้วยรูปภาพ MRI ของเนื้องอกสมองชนิด GBM เกรด 4 และสมองปกติรวมทั้งหมด 155 ภาพถูกใช้เพื่อทดสอบในงานวิจัยนี้ การถ่ายโอนการเรียนรู้โดยโครงข่ายที่ได้รับการอบรมไว้แล้วที่มีชื่อว่า GoogLeNet และ AlexNet ได้รับการนำเสนอเพื่อจำแนกเนื้องอกสมอง GBM และภาพสมองปกติ ข้อดีและจุดเด่นของการจำแนกโดยการถ่ายโอนการเรียนรู้ คือ การแบ่งออกเป็นส่วนแบบแมนนวลและการสกัดฟีเจอร์จะเป็นกระบวนการแบบอัตโนมัติซึ่งจะลดความผิดพลาดจากการทำด้วยมือมนุษย์ ผลลัพธ์การทำนายโดยใช้โครงข่ายที่ได้รับการอบรม GoogLeNet ความแม่นยำเท่ากับ 80.85 เปอร์เซ็นต์ และ AlexNet ได้รับความแม่นยำ 93.62 เปอร์เซ็นต์ จะเห็นได้ว่าการจำแนกเนื้องอกสมอง GBM โดยการเรียนรู้เชิงลึกได้รับผลลัพธ์ความแม่นยำที่ดีและสามารถพัฒนาเพื่อนำมาใช้ได้จริงในอนาคต เพื่อช่วยเหลือบุคลากรทางการแพทย์ในการวินิจฉัยโรคพบอย่างรวดเร็ว เพื่อการรักษาเพิ่มเติมและเพิ่มโอกาสและอัตราการรอดชีวิตของผู้ป่วย การปรับแต่งแบบแมนนวลสำหรับการแบ่งออกเป็นส่วนและการสกัดฟีเจอร์ได้รับการปรับปรุงโดยเปลี่ยนเป็นการจำแนกแบบอัตโนมัติโดยใช้การเรียนรู้เชิงลึกโดยโครงข่ายที่อบรมไว้แล้ว สิ่งนี้คือข้อดีและจุดเด่นหลักของการทดลองนี้ ในบทต่อไปจะทดลองเพื่อใช้งานกับกลุ่มข้อมูลภาพที่มีขนาดใหญ่ขึ้นเพื่อปรับปรุงความแม่นยำ ทดสอบความทนทานและป้องกันปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งเกินไป

วัตถุประสงค์ของงานวิจัยนี้คือการศึกษาพัฒนาต่อเนื่องและมองหากระบวนวิธีการใหม่เพื่อจัดหมวดหมู่เนื้องอก GBM โดยอ้างอิงจากงานวิจัยที่ผ่านมา ซึ่งนำเสนอโดย Hein Tun Zaw และคณะ [21] ที่สามารถตรวจจับเนื้องอกส่วนกลีบขมับได้ดีเป็นพิเศษ ในขณะที่งานวิจัยนี้ปรับปรุงเพื่อการตรวจจับเนื้องอกสมองชนิด GBM ที่เกิดขึ้นในทุกส่วนของสมองและนอกเหนือจากนั้นการแบ่งออกเป็นส่วนและการสกัดฟีเจอร์ออกมาถูกทำอย่างอัตโนมัติโดยใช้การเรียนรู้เชิงลึก และการเรียนรู้เชิงลึกนี้ถูกนำมาใช้งานจริงในงานวิจัยที่หลากหลายเพื่อวินิจฉัยโรคจากรูปภาพทางการแพทย์ [24], [25], [26], [27], [43], [44] ในงานวิจัยนี้โครงข่ายประสาทเทียมแบบคอนโวลูชันที่ได้รับการอบรมไว้แล้วที่ชื่อว่า GoogLeNet และ AlexNet ถูกนำมาใช้งานจริงเพื่อจัดหมวดหมู่รูปภาพ MRI จากฐานข้อมูล REMBRANDT เพื่อจำแนกว่าเป็นเนื้องอกสมอง GBM หรือภาพสมองปกติ โดยภาพที่จัดเตรียมไว้ทั้งหมด 155 ภาพและจัดเรียงใน Datastores และฉลากของรูปภาพถูกจัดเตรียมและ

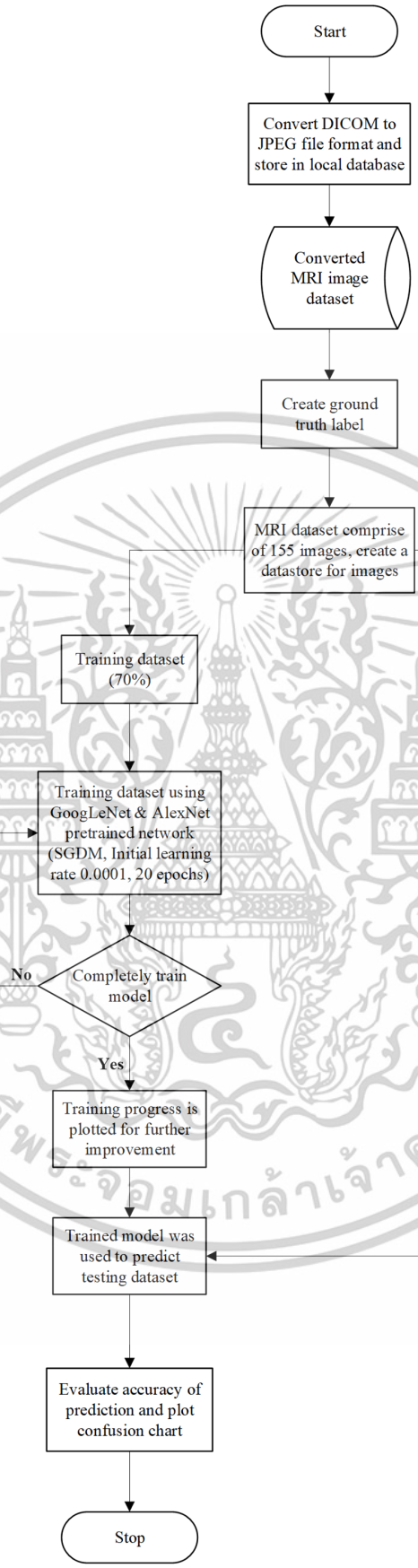
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบุว่าเป็นเนื้องอกสมองหรือสมองปกติ กลุ่มข้อมูลถูกแบ่งเป็นอบรมและทดสอบ (Training, Testing Dataset) ด้วยอัตราส่วน 70 และ 30 เปอร์เซ็นต์ ตามลำดับ และเพื่อความพร้อมในการอบรมของโครงข่ายแต่ละแบบ ขนาดภาพอินพุทต้องมีการเตรียมการก่อน โดยโครงข่าย GoogLeNet ต้องใช้ขนาดภาพอินพุท 224 x 224 x 3 ในขณะที่โครงข่าย AlexNet ต้องใช้ขนาดภาพอินพุท 227 x 227 x 3 ซึ่ง 3 คือค่า Color Channel RGB

กระบวนวิธีการสำหรับการจัดหมวดหมู่แบบไบนารีของเนื้องอกสมอง GBM โดยใช้การถ่ายโอนการเรียนรู้แสดงโฟลว์ชาร์ตได้ดังรูปที่ 3.8 ซึ่งขั้นตอนประกอบด้วยภาพอินพุท MRI ที่ถูกแปลงจากรูปภาพทางการแพทย์ DICOM เป็นไฟล์ JPEG เพื่อดำเนินการเพิ่มเติม ข้อมูลที่เก็บไว้ (Datastore) และฉลากรูปภาพถูกสร้างขึ้นให้สอดคล้องกันซึ่งประกอบด้วยกลุ่มข้อมูล 155 รูปภาพ ซึ่งจะรวมทั้งเนื้องอกสมอง GBM และสภาพสมองปกติ นอกจากนั้นรูปภาพถูกแบ่งออกอย่างสุ่มเป็นกลุ่มข้อมูลเพื่อการอบรมและทดสอบ จากนั้นขั้นตอนการเตรียมกระบวนการกลุ่มข้อมูลได้รับการปรับแต่งเพิ่มเติม (Augmentation) เพื่อปรับขนาดภาพอินพุทอย่างเหมาะสมสำหรับแต่ละโครงข่ายที่ถูกอบรมไว้ก่อนแล้ว การถ่ายโอนการเรียนรู้ ที่มีชื่อว่า GoogLeNet และ AlexNet ได้ถูกนำมาใช้งานจริงเพื่อจัดหมวดหมู่เนื้องอกสมอง โดยส่วนประกอบการเรียนรู้แบบถ่ายโอนคือเลเยอร์ของโครงข่าย (Network Layer) ข้อมูลที่ใช้อบรม (Training Data) และอัลกอริทึมเพิ่มเติม (Algorithm Option) ซึ่งจะถูกปรับแต่งอย่างเหมาะสมก่อนที่จะเข้าสู่ขั้นตอนการอบรมโครงข่าย ต่อจากนั้นข้อมูลจะถูกอบรมโดยโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ได้รับการอบรมไว้ก่อนแล้วรวมทั้งความก้าวหน้าการอบรมถูกพล็อตเพื่อจุดประสงค์ในการติดตามมอนิเตอร์ ในขั้นตอนสุดท้ายโครงข่ายที่ถูกอบรมถูกใช้ประโยชน์เพื่อการทำนายกลุ่มข้อมูลเพื่อการทดสอบที่ถูกแบ่งไว้ก่อนหน้านี้และเมทริกซ์คอนฟิวชัน (Confusion Matrix) จะถูกพล็อตและวิเคราะห์ความแม่นยำด้วยเช่นกัน

ภาพ MRI ทั้งหมดที่ใช้ในการทดลองนี้ได้นำมาจาก The Repository of Molecular Brain Neoplasia Data ซึ่งประกอบด้วยภาพ MRI สแกนจากผู้ป่วย 130 ท่านรวมทั้งข้อมูลการวิเคราะห์ทางคลินิก โดยภาพ MRI ดังกล่าวอยู่ในรูปแบบ DICOM ฟอร์แมท [3] กลุ่มข้อมูลภาพเหล่านี้ถูกแปลงเป็นไฟล์ภาพแบบ JPEG และเก็บไว้ในฐานข้อมูลเพื่อดำเนินการต่อไป โดยภาพที่ถูกแปลงรวมทั้งหมดคือ 155 ภาพ ซึ่งประกอบด้วยเนื้องอกสมอง GBMs ในตำแหน่งแตกต่างกันทั้งหมด 111 ภาพ และสมองปกติอีก 44 ภาพซึ่งจะถูกใช้ในการทดลองนี้ ภาพเนื้องอกสมอง GBMs และสมองปกติถูกแสดงไว้ในรูปที่ 3.2 พื้นที่จัดเก็บข้อมูล (Datastore) สำหรับรูปภาพทั้งหมดถูกสร้างขึ้นและการจำแนกหมวดหมู่ที่ทราบแน่ชัดของแต่ละภาพเนื้องอกสมองถูกระบุและใช้เป็นฉลากสำหรับรูปภาพเพื่อดำเนินการลำดับถัดไป กลุ่มข้อมูลถูกแบ่งออกเป็นสองกลุ่มอย่างสุ่มเพื่อเป็นกลุ่มอบรมและทดสอบซึ่งประกอบไปด้วยอัตราส่วน 70 และ 30 เปอร์เซ็นต์ ตามลำดับ ดังนั้นกลุ่มข้อมูลเพื่ออบรมจะประกอบด้วยภาพ MRI 108 ภาพ และกลุ่มข้อมูลเพื่อการทดสอบประกอบด้วย MRI 47 ภาพ นอกจากนี้แต่ละกลุ่มข้อมูลประกอบด้วยเนื้องอกสมอง GBM และสมองปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 โฟลว์ชาร์ตการวินิจฉัยเนื้องอกสมองด้วยการเรียนรู้แบบถ่ายโอน

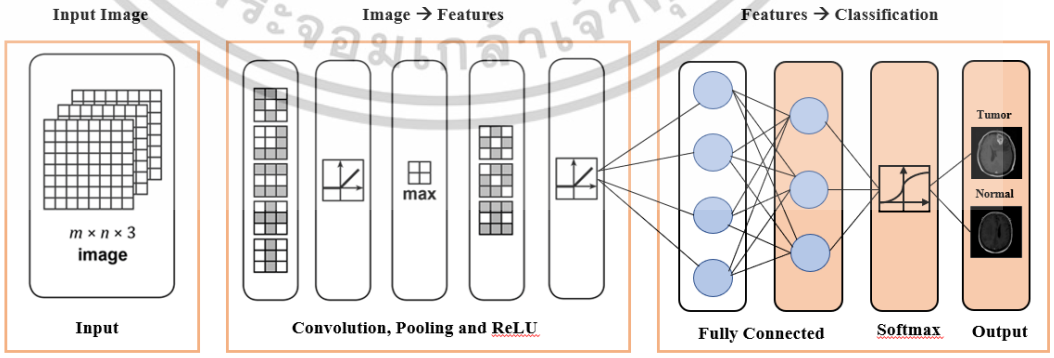
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2 ขั้นตอนการเตรียมกระบวนการโดยสร้างพื้นที่จัดเก็บรูปภาพส่วนขยาย

พื้นที่จัดเก็บข้อมูลรูปภาพส่วนขยาย (Augmented Image Datastores) จะถูกสร้างในขั้นตอนเตรียมกระบวนการ (Preprocessing) เพื่อจัดเตรียมขนาดอินพุทและฟอร์แมทที่เหมาะสมสำหรับโครงข่ายที่อบรมไว้ก่อนแบบ GoogLeNet และ AlexNet โดยสำหรับโครงข่ายแบบ GoogLeNet ต้องการภาพอินพุทขนาด 224 x 224 x 3 และเป็นภาพแบบ Color Channel ในขณะที่โครงข่ายแบบ AlexNet ต้องการภาพอินพุทขนาด 227 x 227 x 3 Color Channel ซึ่งคุณสมบัตินี้ถูกปรับตั้งค่าในส่วนประกอบเริ่มต้นของเลเยอร์โครงข่าย ที่มีชื่อเรียกว่าเลเยอร์อินพุทรูปภาพก่อนที่จะดำเนินการในขั้นตอนต่อไป โดยการทดลองส่วนนี้ใช้โปรแกรม MATLAB R2020R ร่วมกับโมดูลโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกที่ได้รับการอบรมไว้ก่อนแล้วที่มีชื่อเรียกว่า GoogLeNet และ AlexNet เพื่อใช้การจัดหมวดหมู่เนื้องอกสมอง GBM

3.2.3 การจัดหมวดหมู่เนื้องอกสมองโดยใช้ GoogLeNet

ในงานวิจัยนี้โครงข่ายประสาทเทียมแบบคอนโวลูชันที่อบรมไว้แล้วแบบ GoogLeNet ใช้ทำงานสำหรับการจัดหมวดหมู่กลุ่มข้อมูลรูปภาพที่รวบรวมไว้ใหม่ โดย GoogLeNet ถูกอบรมโดยใช้รูปจำนวนล้านรูปภาพและสามารถจำแนกความแตกต่างรูปภาพจำนวนมากถึงหนึ่งพันหมวดหมู่ด้วยกัน [45] โครงข่ายแบบ GoogLeNet ถูกต่อแบบคุณลักษณะกราฟอวัฏจักรระบุทิศทาง (Directed Acyclic Graph, DAG) และนำเอารูปภาพเป็นอินพุทและสร้างส่งออกเอาท์พุทแบบฉลากสำหรับรูปภาพวัตถุ สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชันแสดงได้ดังรูปที่ 3.9 ทั้งนี้การปรับแต่งแบบละเอียดของโครงข่ายด้วยการถ่ายโอนการเรียนรู้นั้นรวดเร็วและง่ายกว่าการอบรมโครงข่ายตั้งแต่จุดเริ่มต้น (Scratch) โดยการเรียนรู้แบบถ่ายโอน ประกอบด้วย เลเยอร์โครงข่ายที่อบรมไว้แล้ว ข้อมูลที่จะใช้สำหรับอบรม (Training Data) และอัลกอริทึมเพิ่มเติม (Algorithm Options)



รูปที่ 3.9 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (CNN)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.3.1 โครงข่ายที่ได้รับการอบรมไว้ก่อนแบบ GoogLeNet

โครงข่ายที่ได้รับการอบรมไว้ก่อนแบบ GoogLeNet ถูกโหลดไปยังพื้นที่ทำงานของ MATLAB และผลลัพธ์การวิเคราะห์เลเยอร์แสดงได้ดังรูปที่ 3.10 ก. โดยที่โครงข่ายถูกต่อแบบคุณลักษณะกราฟอวัฏจักรระบุทิศทาง (Directed Acyclic Graph, DAG) และประกอบด้วย 144 เลเยอร์และต้องการรูปภาพอินพุทขนาด 224 x 224 x 3 ซึ่งเลเยอร์ชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-Connected Layer) และชั้นการจัดหมวดหมู่เอาท์พุทถูกแทนที่สำหรับความเหมาะสมกับการใช้งาน ดังนั้นชั้นเชื่อมโยงแบบสมบูรณ์ได้รับการเปลี่ยนเป็นสองหมวดหมู่นั่นคือเนื้องอกสมองและสมองปกติ ในขณะที่อัลกอริทึมเพิ่มเติมถูกตั้งค่าเป็น Stochastic Gradient Descent with Momentum ด้วยค่า epochs เท่ากับ 20 และอัตราการเรียนรู้เริ่มต้น (Initial Learning Rate) ถูกตั้งค่าเป็น 0.0001 และความก้าวหน้าการอบรมถูกพล็อตกราฟเพื่อการเฝ้าติดตามมอนิเตอร์อีกด้วย จะเห็นได้ว่าส่วนประกอบของการเรียนรู้แบบถ่ายโอนซึ่งประกอบด้วย เลเยอร์โครงข่ายที่อบรมไว้แล้ว ข้อมูลที่ใช้สำหรับอบรม (Training Data) และอัลกอริทึมเพิ่มเติม (Algorithm Options) ได้ถูกจัดเตรียมและเตรียมพร้อมเพื่อการปฏิบัติการอบรมโครงข่ายในขั้นตอนถัดไป

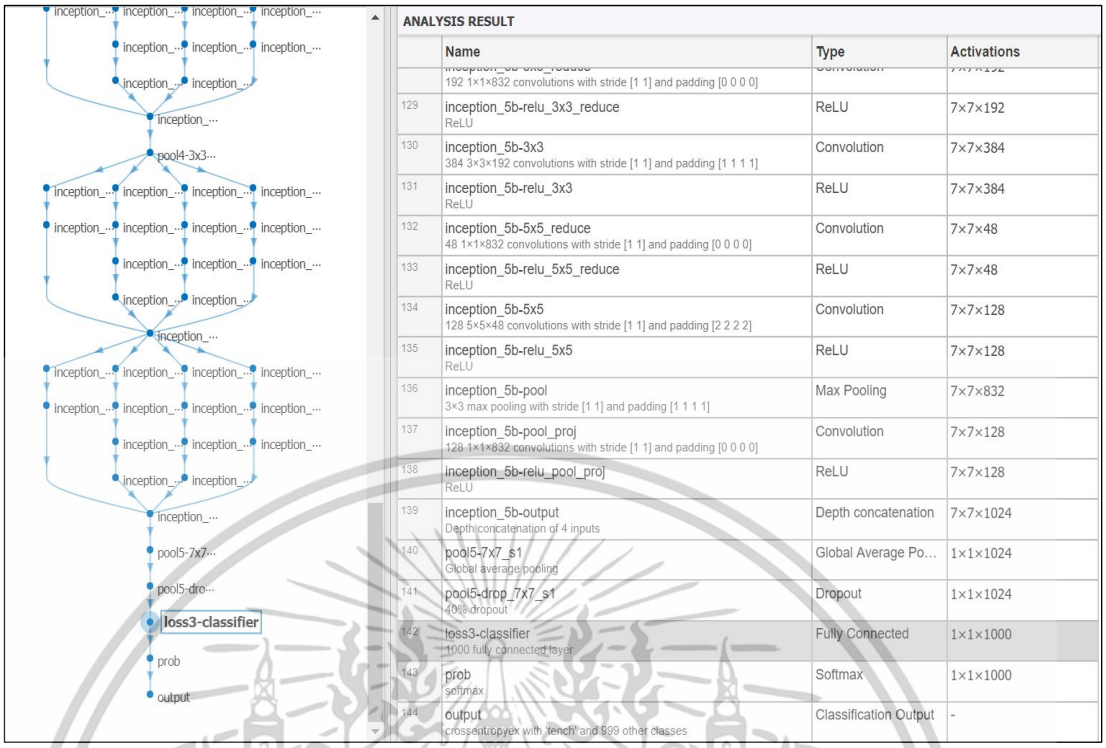
3.2.3.2 อบรมโครงข่ายและการทำนายเอาท์พุท

โครงข่ายที่ได้รับการปรับแต่งจะถูกอบรมในลำดับถัดไปโดยใช้โครงข่ายที่อบรมไว้แล้วแบบ GoogLeNet ด้วยอัลกอริทึมเพิ่มเติม “SGDM” และอัตราการเรียนรู้เริ่มต้น (Initial Learning Rate) ถูกตั้งค่าเป็น 0.0001 และด้วยค่า epochs เท่ากับ 20 ความก้าวหน้าการอบรมและผลลัพธ์การอบรมถูกพล็อตกราฟและแสดงค่าตามรูปที่ 3.10 ข. โดยใช้เวลาในการอบรมประมาณ 11 นาที 52 วินาที ด้วยระบบแบบ CPU อย่างเดียว รวมทั้งค่าความแม่นยำและความสูญเสียของการอบรมถูกนำมาแสดงด้วยเช่นกัน สำหรับในขั้นตอนถัดไปโครงข่ายที่ได้รับการอบรมแล้วจะถูกนำไปใช้ประโยชน์เพื่อการทำนายกลุ่มข้อมูลเพื่อการทดสอบต่อไป โดยกลุ่มข้อมูลทดสอบนี้ถูกแบ่งแยกแบบสุ่มทั้งหมด 47 รูปภาพ ความแม่นยำของการทำนายและเมทริกซ์คอนฟิวชันถูกคำนวณออกมาและพล็อตกราฟสำหรับการปรับปรุงให้ดีขึ้น ผลลัพธ์การทดลองที่ได้รับ คือ ความเที่ยงตรง 80.85 เปอร์เซ็นต์

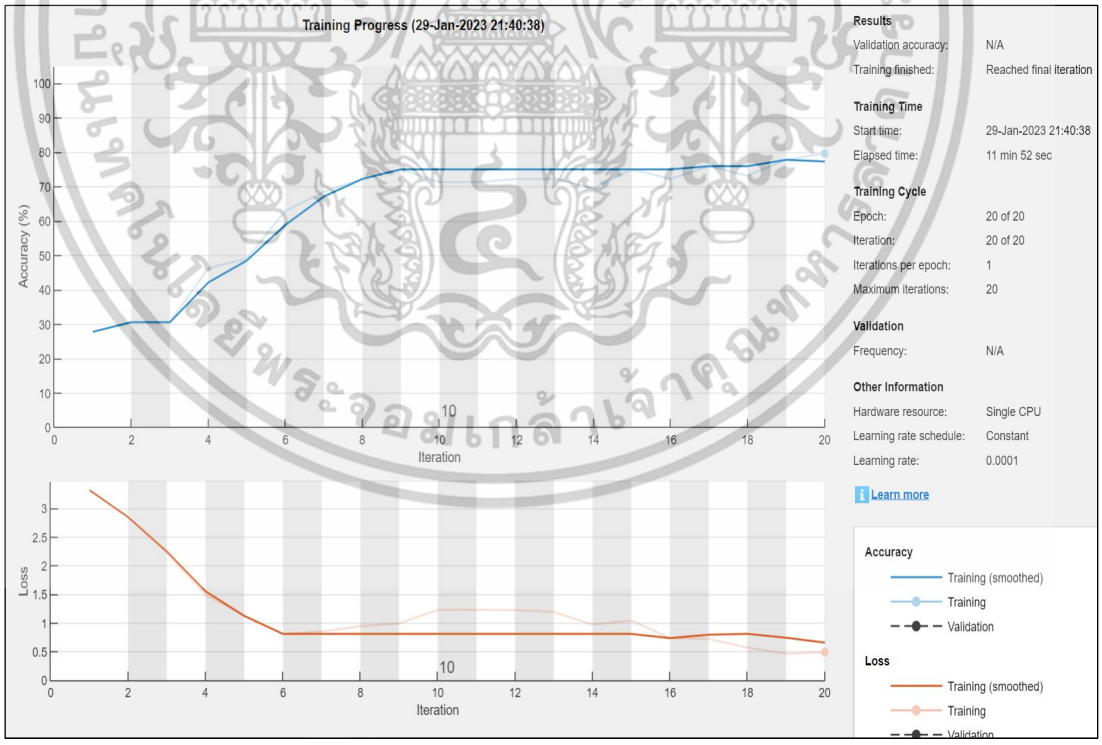
3.2.4 การจัดหมวดหมู่เนื้องอกสมองโดยใช้ AlexNet

โครงข่ายประสาทเทียมคอนโวลูชันที่อบรมไว้ก่อนแบบ AlexNet ใช้ทำงานเพื่อจัดหมวดหมู่คอลเลคชันใหม่ของกลุ่มข้อมูลรูปภาพเพื่อปรับปรุงความแม่นยำเพิ่มเติมจากที่ได้จาก GoogLeNet ซึ่ง AlexNet ได้รับการอบรมโดยใช้จำนวนหนึ่งล้านรูปภาพและสามารถจำแนกความแตกต่างรูปภาพมากถึงหนึ่งพันหมวดหมู่ และโมเดลเรียนรู้ฟีเจอร์จำนวนมากสำหรับรูปภาพที่มีขอบเขตกว้างขวาง [46]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. ผลลัพธ์การวิเคราะห์โครงข่ายของ GoogLeNet



ข. ภาพชาร์ทการอบบรมโครงข่ายของ GoogLeNet

รูปที่ 3.10 ผลลัพธ์การวิเคราะห์และชาร์ทการอบรมของโครงข่าย GoogLeNet

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่โครงข่ายแบบ AlexNet ถูกต่อกันเป็นแบบโครงข่ายอนุกรม ซึ่งโครงข่ายประสาทเทียมสำหรับการเรียนรู้เชิงลึกที่เลเยอร์ถูกจัดเรียงแบบต่อกันไปเรื่อยหนึ่งต่อหนึ่ง และใช้รูปภาพเป็นอินพุทและสร้างฉลากเอาท์พุทสำหรับวัตถุรูปภาพนั้นๆ ประกอบด้วยเลเยอร์อินพุท 1 ชั้นและเลเยอร์เอาท์พุทอีก 1 ชั้นเช่นกัน สถาปัตยกรรมของโครงข่ายประสาทเทียมคอนโวลูชันแบบถ่ายโอนการเรียนรู้ ถูกแสดงดังรูปที่ 3.9 ซึ่งประกอบด้วยรูปภาพอินพุท การแปลงรูปภาพเป็นฟีเจอร์โดยประกอบด้วยคอนโวลูชัน พูลลิง (Pooling) และ เลเยอร์ ReLU หลายอันต่อกัน และสุดท้ายส่วนการแปลงฟีเจอร์เป็นการจัดหมวดหมู่ (Features to Classification) ที่ซึ่ง 3 เลเยอร์สุดท้ายจะถูกแทนที่เพื่อสอดคล้องกับการประยุกต์ใช้งาน ขั้นตอนสำหรับการจัดหมวดหมู่เนื้องอกสมองโดยใช้ AlexNet นั้นแตกต่างเล็กน้อยจากการใช้โครงข่ายที่อบรมไว้แล้วแบบ GoogLeNet ยกตัวอย่างเช่นขนาดภาพอินพุทจะถูกเปลี่ยนเป็น 227 x 227 x 3 และจำนวนของเลเยอร์สำหรับ AlexNet ประกอบด้วยจำนวน 25 เลเยอร์เท่านั้นซึ่งนั่นส่งผลต่อการใช้เวลาสำหรับการอบรมโมเดลที่น้อยกว่า

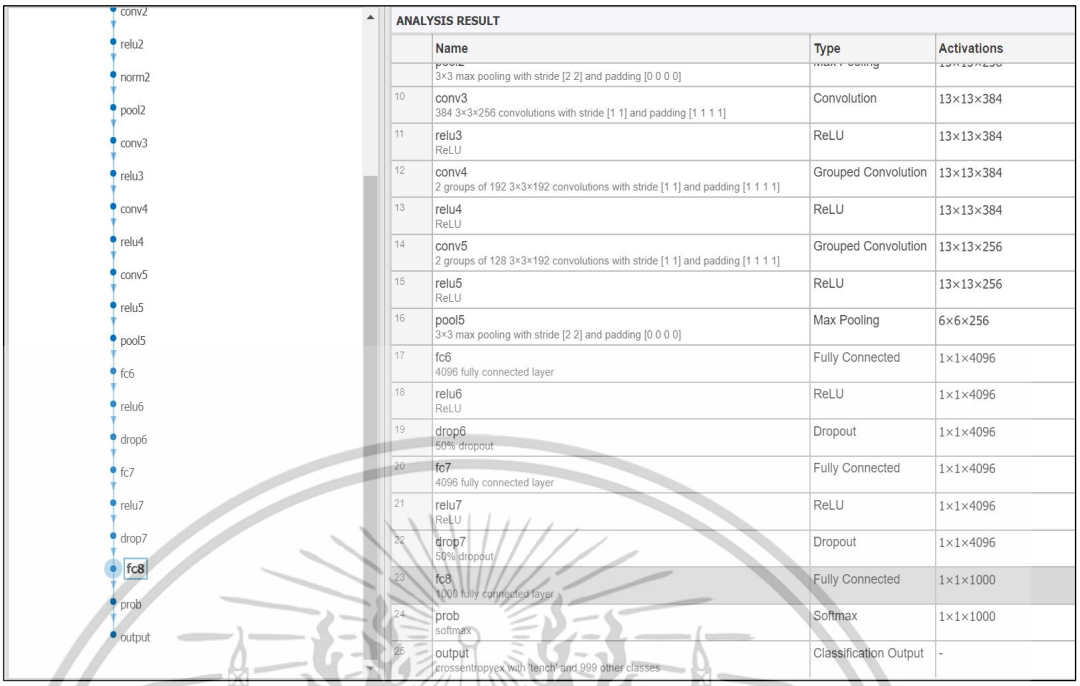
3.2.4.1 เลเยอร์โครงข่ายที่ได้รับการอบรมแล้วแบบ AlexNet

โครงข่ายที่อบรมไว้แล้วแบบ AlexNet ถูกโหลดเข้าสู่พื้นที่ทำงานโปรแกรม MATLAB และผลลัพธ์การวิเคราะห์โครงข่ายถูกแสดงดังรูปที่ 3.11 ก. โครงข่ายถูกต่อในรูปแบบอนุกรมและประกอบไปด้วย 25 เลเยอร์และต้องการขนาดของภาพอินพุท 227 x 227 และ 3 Color Channels เลเยอร์ชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-Connected Layer) และชั้นการจัดหมวดหมู่เอาท์พุทถูกแทนที่และชั้นเชื่อมโยงแบบสมบูรณ์ถูกแก้ไขเปลี่ยนแปลงเป็นสองหมวดหมู่นั่นคือตามฉลากรูปภาพเนื้องอกสมองและสมองปกติ อัลกอริทึมเพิ่มเติมถูกปรับแต่งเป็น Stochastic Gradient Descent with Momentum ด้วยค่า epochs เท่ากับ 20 และค่าอัตราการเรียนรู้เริ่มต้นถูกตั้งค่าเป็น 0.0001 ซึ่งเหมือนกันกับโครงข่ายที่อบรมไว้แล้วแบบ GoogLeNet

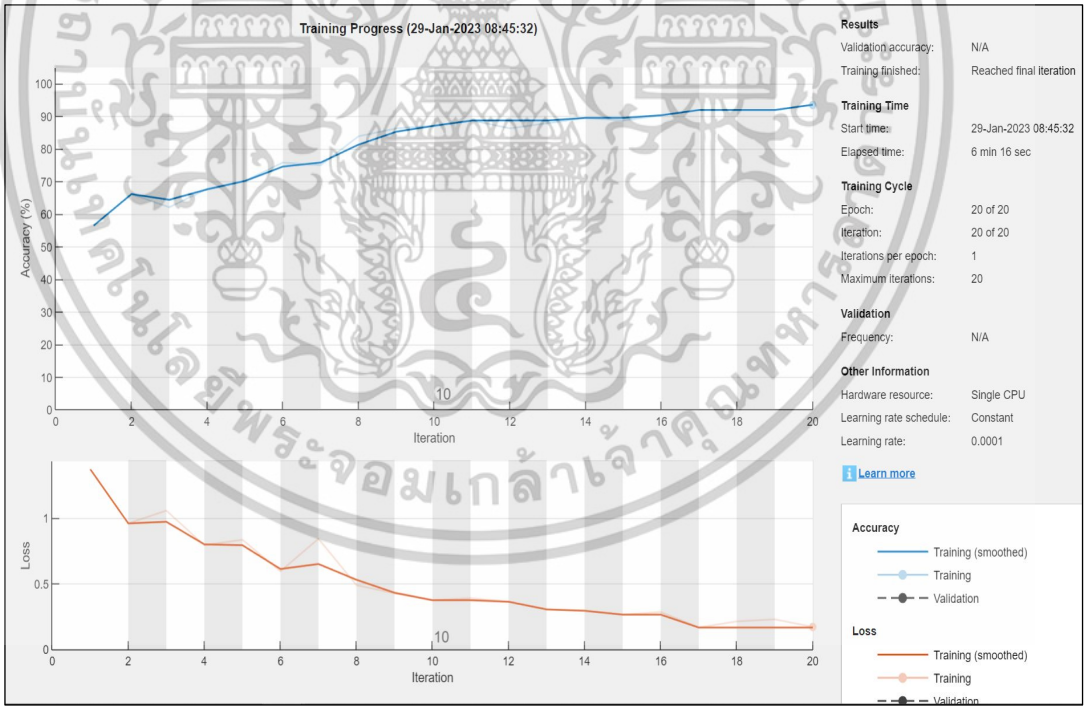
3.2.4.2 การอบรมโครงข่ายและทำนายเอาท์พุท

โครงข่ายที่ได้รับการดัดแปลงถูกอบรมโดยใช้โครงข่ายที่อบรมไว้แล้วแบบ AlexNet ด้วยอัลกอริทึมเพิ่มเติมแบบ “SGDM” และอัตราการเรียนรู้เริ่มแรกเท่ากับ 0.0001 ด้วยค่า epochs เท่ากับ 20 ผลลัพธ์การอบรมถูกพล็อตกราฟและแสดงดังรูปที่ 3.11 ข. โดยใช้เวลาในการอบรม 6 นาที 16 วินาที โดยประมาณด้วยการใช้ CPU อย่างเดียว สังเกตได้ว่าเวลาที่ใช้ในการอบรมโครงข่ายแบบ AlexNet นั้นน้อยกว่า GoogLeNet เนื่องจากจำนวนเลเยอร์ที่น้อยกว่าเพียง 25 เลเยอร์เท่านั้น และหลังจากนั้นโครงข่ายที่ได้รับการอบรมแล้วถูกนำไปใช้ประโยชน์สำหรับการทำนายผลของกลุ่มข้อมูลเพื่อการทดสอบ ซึ่งประกอบไปด้วยรูปภาพที่ถูกแบ่งแบบสุ่มจำนวน 47 รูปภาพ ความแม่นยำของการทำนายผลและเมทริกซ์ Confusion ถูกคำนวนออกมาและพล็อตกราฟเพื่อใช้ในการปรับปรุงเพิ่มเติม ผลลัพธ์การทดลองได้รับความแม่นยำถึง 93.62 เปอร์เซ็นต์ สำหรับ AlexNet ซึ่งแสดงให้เห็นว่าได้รับการปรับปรุงอย่างมีนัยสำคัญ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. ผลลัพธ์การวิเคราะห์โครงข่ายของ AlexNet



ข. ภาพชาร์ทการอบบรมโครงข่ายของ AlexNet

รูปที่ 3.11 การวิเคราะห์เลเยอร์และชาร์ทการอบรมโครงข่าย AlexNet

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

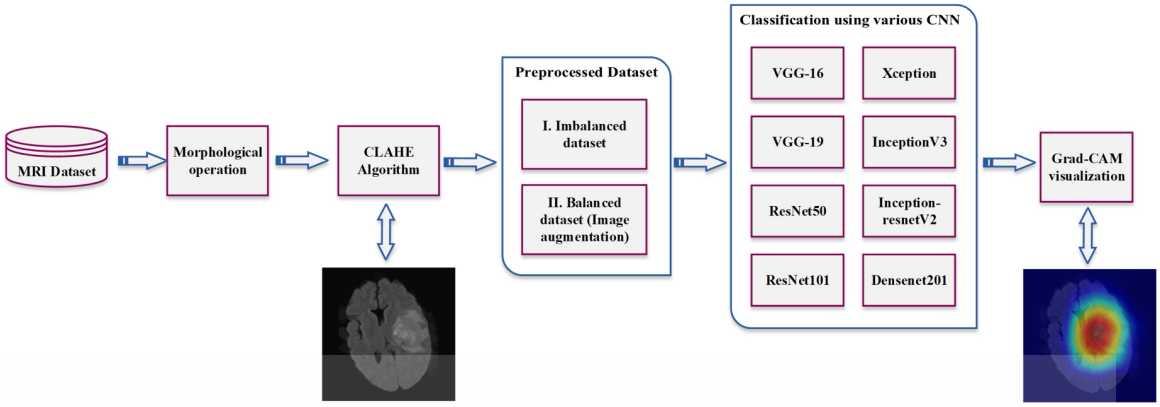
บทที่ 4
การวินิจฉัยเนื้องอกสมองกลิโอมาที่มีกลุ่มข้อมูลขนาดใหญ่ด้วยโครงข่ายประสาทเทียมคอนโวลูชัน CNNs

ในบทนี้กล่าวถึงการจัดหมวดหมู่เนื้องอกสมองที่มีกลุ่มข้อมูลขนาดใหญ่ด้วยโครงข่ายประสาทเทียมคอนโวลูชันที่ทันสมัยหลากหลายแบบ ซึ่งได้รับกลุ่มข้อมูลเพิ่มขึ้นจากสองแหล่งฐานข้อมูลเพื่อลดปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งมากเกินไป และนำเสนอการเตรียมกระบวนการรูปภาพก่อนนำไปจัดหมวดหมู่เพื่อเพิ่มความแม่นยำในการทำนายผล กระบวนวิธีการทดลองสามแบบถูกนำมาใช้ กล่าวคือ อบรมกลุ่มข้อมูลซึ่งไม่สมดุลกันด้วยโครงข่ายประสาทเทียมคอนโวลูชัน ดำเนินการเสริมภาพเพื่อทำให้เกิดความสมดุลกันขึ้นในกลุ่มข้อมูลอบรมของโครงข่ายประสาทเทียมคอนโวลูชัน และเพื่อประเมินผลลัพธ์เพิ่มเติมได้ศึกษาเกี่ยวกับการเรียนรู้แบบผสมผสาน ท้ายที่สุดนำเสนอการสร้างภาพการมองเห็นพื้นที่ซึ่งถูกแยกแยะเพื่อยืนยันผลการทำนายอีกครั้ง โดยรายละเอียดกล่าวได้ดังต่อไปนี้

4.1 กระบวนวิธีการทดลอง

ในส่วนนี้นำเสนอการจัดหมวดหมู่เนื้องอกสมองกลิโอมาโดยใช้การเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network, CNNs) และขั้นตอนกระบวนการแสดงดังรูปบล็อกไดอะแกรมรูปที่ 4.1 เริ่มแรกกลุ่มข้อมูล MRI ที่แปลงเป็นภาพ JPEG และเก็บในฐานข้อมูลส่วนบุคคล (Local Database) ก่อนจะนำไปสู่ขั้นตอนเตรียมกระบวนการรูปภาพด้วยกระบวนการสัณฐานวิทยา (Morphological Operation) สำหรับแยกแยะส่วนกะโหลกและสิ่งแปลกปลอมออกจากพื้นที่สมองเพื่อได้รับภาพส่วนพื้นที่สมองเท่านั้นเพื่อดำเนินการขั้นถัดไป จากนั้นนำภาพที่ได้เพื่อปรับคอนทราสให้สมดุลและทำเป็นมาตรฐานเดียวกันโดยใช้หลักการ Contrast Limited Adaptive Histogram Equalization (CLAHE) และกลุ่มข้อมูลรูปภาพที่ผ่านกระบวนการที่กล่าวมาทั้งหมดที่ได้รับเกิดความไม่สมดุลกันอย่างมากระหว่างสองคลาส และเพื่อจัดการกับปัญหานี้ได้ตั้งค่าแบบจำลองการทดสอบดังต่อไปนี้ แบบจำลองที่หนึ่งใช้กลุ่มข้อมูลที่ไม่สมดุลเพื่อใช้อบรมโดย CNNs หลากหลายแบบ และแบบจำลองที่สองใช้หลักการเสริมรูปภาพ (Image Augmentation) เพื่อเพิ่มจำนวนรูปภาพคลาสที่เป็นส่วนน้อย (Minority Class) ซึ่งในที่นี้คือสมองปกติ เพื่อทำให้กลุ่มข้อมูลเกิดความสมดุลกันขึ้น และนำไปอบรมโดยใช้ CNNs หลากหลายแบบ เมื่อกลุ่มข้อมูลได้รับการอบรมโดยโครงข่ายประสาทเทียมคอนโวลูชันแล้ว โครงข่ายที่ได้รับการอบรมจะถูกนำไปใช้เพื่อทำนายกลุ่มข้อมูลทดสอบแบบไบนารี ว่าเป็นภาพเนื้องอกหรือสมองปกติ ท้ายที่สุดนำอัลกอริทึม Gradient-weighted Class Activation Mapping เพื่อแสดงให้เห็นพื้นที่รูปภาพที่แยกแยะ (Discriminative Image Region) เพื่อให้เข้าใจและแปลความหมายภาพที่ทำนายได้ง่ายขึ้น รวมทั้งเพื่อให้รู้ว่าส่วนไหนของภาพที่มีผลต่อการทำนายผลว่าเป็นเนื้องอกหรือสมองปกติโดยใช้แผนที่ความร้อน ซึ่งในแผนที่ความร้อนสีแดงหมายถึงส่วนของภาพที่มีผลกระทบต่อการทำนายผลเนื้องอก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 บล็อกไดอะแกรมการจัดหมวดหมู่เนื้องอกสมองกลิโอมาด้วย CNNs

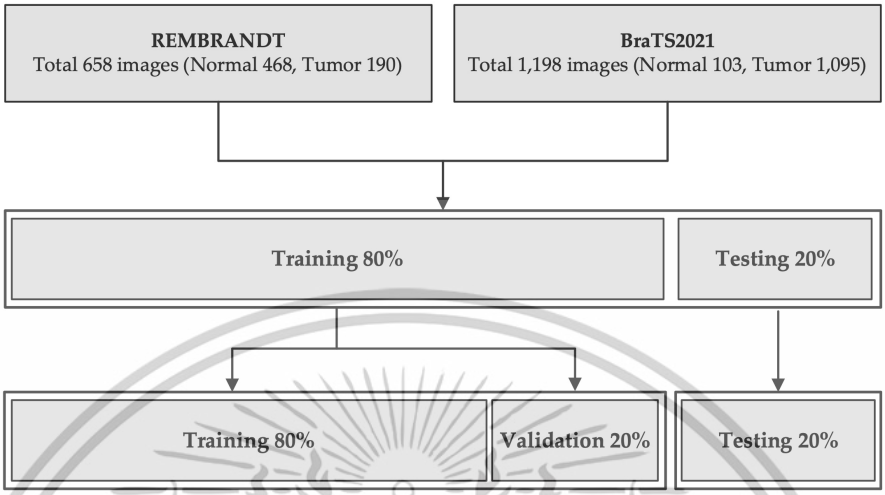
4.2 การออกแบบตั้งค่าการทดลอง

กลุ่มข้อมูลเนื้องอกสมองเพื่อใช้ในการทดลองในส่วนนี้ได้รับจากฐานข้อมูลสองแหล่งที่แตกต่างกัน โดยกลุ่มข้อมูล REMBRANDT ประกอบด้วยภาพ MRI ของผู้ป่วย 130 ท่านรวมทั้งข้อมูลวิเคราะห์ทางด้านคลินิกและอยู่ในรูปแบบฟอร์แมท DICOM ภาพตัดเฉือนในรูปแบบ FLAIR ของเนื้องอกสมองตำแหน่งที่ต่างกันและสมองปกติเพื่อใช้ในการจัดหมวดหมู่ต่อไป [3]

ภาพถ่าย MRI 658 ภาพจากฐานข้อมูล REMBRANDT ซึ่งประกอบด้วยเนื้องอกสมองเกรด 3 (Astrocytoma) และเกรด 4 (GBM) ซึ่งเป็นระดับรุนแรง จำนวน 190 รูปภาพและสมองปกติอีก 468 ภาพได้นำมารวมกันในกลุ่มข้อมูล ในส่วนฐานข้อมูลที่สองได้รับภาพ MRI จากกลุ่มข้อมูล BraTS2021 ซึ่งประกอบด้วยภาพสแกน MRI ของเนื้องอกสมอง GBM เกรด 4 ระดับรุนแรงจากสถาบันการแพทย์หลายแห่งและอยู่ในรูปแบบฟอร์แมทชนิด NifTI [4], [5], [6] ภาพถ่าย MRI ที่อยู่ในรูปแบบขั้นตอน T_{2FLAIR} จากฐานข้อมูล BraTS2021 ถูกแปลงเป็นฟอร์แมท JPEG ทั้งหมด 1,198 รูปภาพ ซึ่งประกอบด้วยเนื้องอกสมอง 1,095 ภาพและสมองปกติ 103 ภาพ เพื่อถูกรวมในกลุ่มข้อมูลเดียวกันจะได้รูปภาพ MRI รวมทั้งหมดจากสองฐานข้อมูลเท่ากับ 1,856 ภาพ ซึ่งประกอบด้วยเนื้องอกสมอง 1,285 ภาพและสมองปกติ 571 ภาพ จำนวนรูปภาพ MRI ที่ถูกแปลงเป็นฟอร์แมท JPEG ที่ได้จากแต่ละฐานข้อมูลแสดงดังในตารางที่ 4.1 จะถูกใช้สำหรับการจัดหมวดหมู่โดยใช้ CNNs หลากหลายแบบ จากนั้นกลุ่มข้อมูลถูกแบ่งเป็นกลุ่มข้อมูลอบรม (Training Dataset) และกลุ่มข้อมูลทดสอบ (Testing Dataset) ด้วยอัตราส่วน 80 และ 20 เปอร์เซ็นต์ ตามลำดับ นอกจากนั้นภาพเนื้องอกและสมองปกติจากทั้งสองฐานข้อมูลถูกจัดสรรและเฉลี่ยเท่ากันทั้งกลุ่มข้อมูลเพื่ออบรมและทดสอบอีกด้วย ดังนั้นแล้วจะได้รับกลุ่มข้อมูลอบรม 1,484 ภาพ ประกอบด้วยภาพเนื้องอก 1,028 ภาพและภาพสมองปกติ 456 ภาพ ในส่วนกลุ่มข้อมูลทดสอบทั้งหมด 372 ภาพ ซึ่งประกอบด้วยเนื้องอก 257 ภาพและสมองปกติ 115 ภาพ ทั้งนี้กลุ่มข้อมูลอบรมถูกแบ่งเป็นกลุ่มข้อมูลเพื่ออบรม (Training) และกลุ่มข้อมูลทำให้สมบูรณ์ (Validation) ด้วยอัตราส่วน 80 และ 20 เปอร์เซ็นต์ ตามลำดับอีกด้วย ก่อนที่จะนำเข้าสู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กระบวนการอบรมโดยใช้ CNNs หลากหลายแบบ การรวบรวมและจัดสรรรูปภาพสำหรับกลุ่มข้อมูล แสดงได้ดังรูปที่ 4.2



รูปที่ 4.2 กลุ่มข้อมูลที่รวบรวมสำหรับการจัดหมวดหมู่เนื้องอกสมองด้วย CNNs

ตารางที่ 4.1 รูปภาพ MRI ที่ถูกแปลงเป็นฟอร์แมท JPEG ที่ได้จากแต่ละฐานข้อมูล

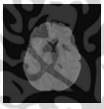
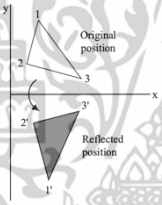
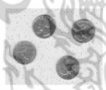
Dataset	MRI (original)		Training (80%)		Testing (20%)	
	Normal	Tumor	Normal	Tumor	Normal	Tumor
REMBRANDT	468	190	**374**	152	94	38
BraTS2021	103	1095	**82**	876	21	219
Total	571	1285	**456 (1:2.25)**	1028	115	257

กลุ่มข้อมูลทดสอบที่ได้รับเกิดความไม่สมดุลอย่างมาก และเพื่อจัดการกับปัญหานี้วิธีการเสริมจำนวนรูปภาพได้นำมาใช้ให้เกิดประโยชน์เนื่องจากเป็นวิธีการที่ไม่ยุ่งยากและตรงไปตรงมาที่สุดเพื่อสร้างกลุ่มข้อมูลที่สมดุลกัน โดยได้เพิ่มรูปภาพที่เป็นคลาสส่วนน้อย (Minority) กล่าวคือรูปภาพสมองปกติถูกเสริมเพิ่มเติมโดยใช้การหมุน 90 องศาและการสะท้อนรอบแกนเอ็กซ์ ซึ่งภาพสมองปกติจากฐานข้อมูล REMBRANDT ที่ถูกแบ่งเป็นกลุ่มข้อมูลอบรม (Training Dataset) จำนวน 374 ภาพจะถูกเสริมภาพถ่ายโดยใช้การหมุนภาพ 90 องศา (Rotation) อย่างเดียวทำให้ได้ภาพจำนวน 748 รูปภาพ ในขณะที่ภาพสมองปกติของฐานข้อมูล BraTS2021 ที่ถูกจัดสรรเป็นกลุ่มข้อมูลอบรมจำนวน 82 ภาพ ซึ่งมีจำนวนน้อยที่สุดจะถูกเสริมภาพโดยใช้เทคนิคการหมุนภาพ 90 องศา (Rotation) และการสะท้อนของภาพ (Reflection) เพื่อเพิ่มปริมาณภาพสำหรับกลุ่มข้อมูลอบรม และได้รับภาพเพิ่มเป็น 246 รูปภาพ นอกจากนั้นได้เลือกภาพจากทั้งสองกลุ่มข้อมูลเพื่อเสริมภาพด้วยวิธีสร้างสัญญานรบกวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Synthetic Noise) และเบลอภาพ (Synthetic Blur) รวมเป็น 34 ภาพ รายละเอียดการเสริมภาพกับกลุ่มข้อมูลส่วนน้อยแสดงดังตารางที่ 4.2 โดยหลักการเสริมรูปภาพนี้ทำเฉพาะกับกลุ่มข้อมูลอบรมเท่านั้นทำให้ได้รับภาพรวมทั้งหมดของสมองปกติสำหรับกลุ่มข้อมูลอบรม 1,028 รูปภาพ ในขณะที่ภาพเนื้องอกสมองสำหรับกลุ่มข้อมูลอบรม 1,028 รูปภาพซึ่งสร้างความสมดุลให้เกิดขึ้น และจะได้รับจำนวนรูปภาพสำหรับแต่ละวิธีทดลองตามที่แสดงในตารางที่ 4.3

ตารางที่ 4.2 เทคนิคการเสริมภาพสำหรับรูปภาพสมองปกติซึ่งเป็นคลาสส่วนน้อย

Augmentation	Characteristic	Augmented MRI image	Dataset applicable	
			REMBRANDT	BraTS2021
Normal brain for training (original)	-		374	82
Rotation 90 degree			374	82
Reflection around x axis			-	82
Synthetic noise			10	7
Synthetic blur			10	7

ตารางที่ 4.3 การผสมผสานรูปภาพสำหรับกลุ่มข้อมูลของแต่ละวิธีการ

Approach	Dataset	Normal	Glioma	Total
Approach 1 Imbalanced dataset	Training	456	1,028	1,484
	Testing	115	257	372
Approach 2 Balanced dataset	Training	1,028	1,028	2,056
	Testing	115	257	372

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

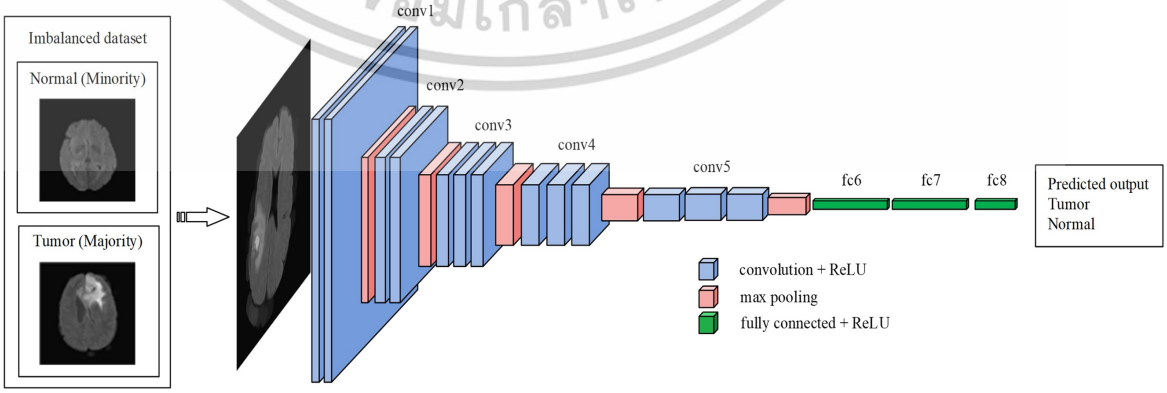
ในงานวิจัยส่วนนี้การทดลองทั้งหมดเริ่มตั้งแต่การเตรียมกระบวนการ การอบรมกลุ่มข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันจนถึงการจัดหมวดหมู่แบบไบนารีของเนื้องอกสมองและการตรวจสอบหลังการทำนายผล เพื่อยืนยันความมั่นใจและโปร่งใสในการแยกแยะพื้นที่รูปภาพ ใช้โปรแกรม MATLAB R2023a ร่วมกับโมดูลเสริมภายในโปรแกรม (Toolbox) ซึ่งประกอบด้วย Image Processing, Machine Learning และ Deep Learning Toolbox ดังจะกล่าวรายละเอียดต่อไป

4.2.1 กระบวนวิธีการที่ 1 การอบรมโดยใช้กลุ่มข้อมูลไม่สมดุลกัน

วิธีการแรกที่นำเสนอในส่วนนี้คือการใช้โครงข่ายประสาทเทียมที่หลากหลายเพื่ออบรมกลุ่มข้อมูลที่ไม่สมดุลกัน เป็นการนำรูปภาพ MRI ที่ได้หลังผ่านขั้นตอนเตรียมกระบวนการ แต่ยังไม่ได้มีการเสริมภาพดังที่กล่าวมาแล้ว โดยรูปภาพทั้งหมดเพื่อใช้อบรม 1,484 รูปภาพ ประกอบด้วยสมองปกติ 456 ภาพและเนื้องอกสมอง 1,028 ภาพ ดังแสดงในตารางที่ 4.3 ซึ่งจะถูกใช้เป็นกลุ่มข้อมูลอบรมสำหรับโครงข่ายประสาทเทียมหลายหลายแบบ ซึ่งขนาดของรูปภาพอินพุทแตกต่างกันดังต่อไปนี้ โดยกลุ่มข้อมูล REMBRANDT ขนาดรูปภาพ 1178 x 1177 พิกเซลและกลุ่มข้อมูล BraTS2021 ขนาดรูปภาพ 1152 x 1152 พิกเซล ภาพที่มีขนาดพิกเซลใหญ่ก็จะใช้เวลาในการประมวลผลเพิ่มขึ้นตามไปด้วยเนื่องจากหลักการของคอนโวลูชันจะทำงานในระดับพิกเซลเพื่อสกัดฟีเจอร์ออกมาอย่างอัตโนมัติ (Feature Map Extraction) ในทางตรงกันข้ามถ้าภาพขนาดเล็กลงความละเอียดของภาพก็จะน้อยทำให้การสกัดฟีเจอร์และทำนายผลลดความแม่นยำลงได้เช่นกัน จากนั้นลาเบล (Label) ของแต่ละรูปภาพถูกจัดเตรียมไว้ในรูปแบบตาราง ซึ่งจะระบุคลาสที่ถูกต้องของรูปภาพนั้นๆ กล่าวคือ เนื้องอก (Tumor) หรือสมองปกติ (Normal) ในขั้นตอนการอบรมนั้นความเหมือนหรือตรงกันของลาเบลและรูปภาพที่ใช้อบรมสำคัญอย่างยิ่ง เพราะจะเป็นตัวกำหนดความแม่นยำของฟีเจอร์ที่ถูกสกัดออกมา และการจัดหมวดหมู่อีกด้วย กลุ่มข้อมูลอบรมจะถูกแบ่งอีกครั้งเป็นกลุ่มข้อมูลอบรม (Training) และกลุ่มข้อมูลที่ทำให้สมบูรณ์ (Validation) ด้วยอัตราส่วน 80 และ 20 เปอร์เซ็นต์ ตามลำดับ การเรียนรู้แบบถ่ายโอน นั้นมีองค์ประกอบสำคัญสามอย่างคือ กลุ่มข้อมูลที่จะใช้เพื่ออบรม (Training Dataset) เลเยอร์โครงข่าย (Network Layer) และอัลกอริทึมเพิ่มเติม (Algorithm Options) ในงานวิจัยนี้ใช้ เลเยอร์โครงข่ายที่อบรมไว้ก่อนแล้ว โดยการดัดแปลงสามเลเยอร์สุดท้าย นั่นคือ Fully Connected, Softmax และ Classification Output Layer โดยเลเยอร์ Fully Connected จะดัดแปลงขนาดเอาท์พุท (Output Size) จากหนึ่งพันหมวดหมู่เป็นสองหมวดหมู่นั่นคือ สมองปกติและเนื้องอกตามการประยุกต์ใช้งาน โดยการนำมาใช้โครงข่ายที่อบรมไว้ก่อนแล้วนั้นจะช่วยให้ไม่ต้องเริ่มอบรมโครงข่ายตั้งแต่เริ่มต้น (From Scratch) ซึ่งจะช่วยลดเวลาลงได้ บล็อกไดอะแกรมสำหรับโครงข่ายการอบรมโดยใช้การเรียนรู้แบบถ่ายโอนแสดงดังรูปที่ 3.9 สังเกตว่าประกอบด้วยสามส่วนคือการจัดเตรียมภาพอินพุท การแปลงภาพเป็นฟีเจอร์แมพ และท้ายสุดการจัดหมวดหมู่ฟีเจอร์โดยใช้ Fully Connected Layer

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึมเพิ่มเติม (Algorithm Option) ที่นำมาใช้งานประกอบด้วย Stochastic Gradient Descent with Momentum อัตราการเรียนรู้เริ่มต้นเท่ากับ 0.0001 และ epochs เท่ากับ 20 เพื่อใช้ในการอบรมโดยใช้โครงข่ายแบบ CNNs ซึ่งในงานวิจัยนี้ใช้ VGG16, VGG19, ResNet50, ResNet101, Xception, InceptionV3, InceptionResNetV2 และ Densenet201 สุดท้ายแล้วจะได้รับโครงข่ายที่อบรมแล้วเพื่อนำไปใช้ทำนายกลุ่มข้อมูลทดสอบ (Testing Dataset) จำนวน 372 รูปภาพต่อไป และในที่นี้กราฟความความคืบหน้าการอบรม (Training Progress) และชาร์ตคอนฟิวชัน (Confusion Chart) ถูกพล็อตไว้เพื่อจุดประสงค์ในการมอนิเตอร์และปรับปรุงต่อไปอีกด้วย บล็อกไดอะแกรมของการอบรมโครงข่ายของกลุ่มข้อมูลอบรมแบบไม่สมดุลกันแสดงดังรูปที่ 4.3 จากบล็อกไดอะแกรมภาพที่ผ่านการเตรียมกระบวนการถูกป้อนไปยังโครงข่ายประสาทเทียมที่ได้รับการอบรมไว้ก่อนแล้วหลากหลายแบบเพื่ออบรมโมเดลโดยใช้อัลกอริทึมออพชันดังที่กล่าวมา นอกจากนั้นโมเดลที่ได้รับการอบรมโดยการเรียนรู้แบบถ่ายโอนแต่ละแบบในที่นี้เรียกว่า Trained Model จะถูกบันทึกไว้เพื่อนำไปใช้จัดหมวดหมู่ (Prediction) ในอนาคตอีกด้วย โดยจะอยู่ในรูปไฟล์ trainedmodel.mat หรือ MATLAB data เมื่อจะเรียกใช้งานจะทำได้โดยใช้คำสั่ง load ได้โดยไม่ต้องอบรมโมเดลใหม่ ซึ่ง Trained Model นี้จะเป็นการนำกลับมาใช้ใหม่ให้เกิดประโยชน์ ทั้งการทดสอบการทำนายกลุ่มข้อมูลทดสอบเพื่อยืนยันความถูกต้อง การทำนายรูปภาพ MRI แบบเฉพาะเจาะจงแต่ละภาพ (Particular MRI Image) เพื่อการวินิจฉัยโดยใช้คอมพิวเตอร์ช่วย (Computer-Aided Diagnosis) รวมทั้งการยืนยันความถูกต้องเพื่อเพิ่มความเชื่อมั่นว่าการวินิจฉัยโดยคอมพิวเตอร์นั้นแม่นยำจริงโดยใช้อิลกอริทึม Gradient Weight Class Activation Mapping เพื่อเป็นการเน้นย้ำว่าการทำนายผลเนื้องอกสมองนั้นส่วนไหนของภาพที่มีผลกระทบต่อการตัดสินใจและทำนายของคอมพิวเตอร์มากที่สุด โดยเฉพาะอย่างยิ่งภาพตัดเฉือนส่วนที่ขนานกับดวงตานั้นคอมพิวเตอร์อาจจะตรวจจับดวงตาเป็นเนื้องอกได้ และอัลกอริทึม Grad-CAM จะช่วยเป็นเครื่องมือในการตรวจสอบอีกขั้นได้เช่นกัน



รูปที่ 4.3 บล็อกไดอะแกรมสำหรับการอบรมกลุ่มข้อมูลซึ่งไม่สมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

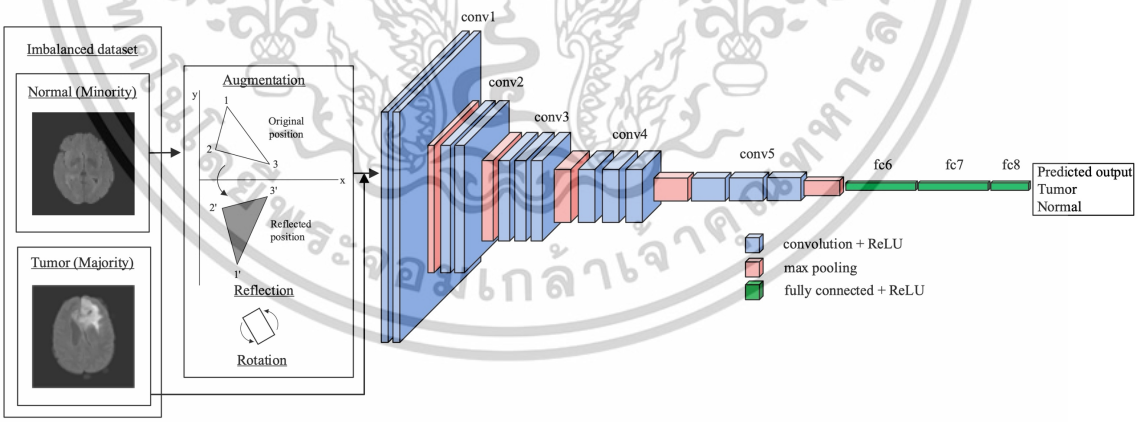
เพื่อความเข้าใจยิ่งขึ้น ขั้นตอนการวินิจฉัยเนื้องอกสมองกลิโอมาโดยใช้โครงข่ายประสาทเทียมคอนโวลูชันกับกลุ่มข้อมูลซึ่งไม่สมดุลกันตั้งแต่เริ่มต้นจนวินิจฉัยเนื้องอกสมองแสดงในรูปโฟลว์ชาร์ตดังรูปที่ 4.4

รูปที่ 4.4 โฟลว์ชาร์ตการอบรมโครงข่ายโดยใช้กลุ่มข้อมูลซึ่งไม่สมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 กระบวนวิธีการที่ 2 การอบรมโดยใช้กลุ่มข้อมูลที่สมดุลกัน

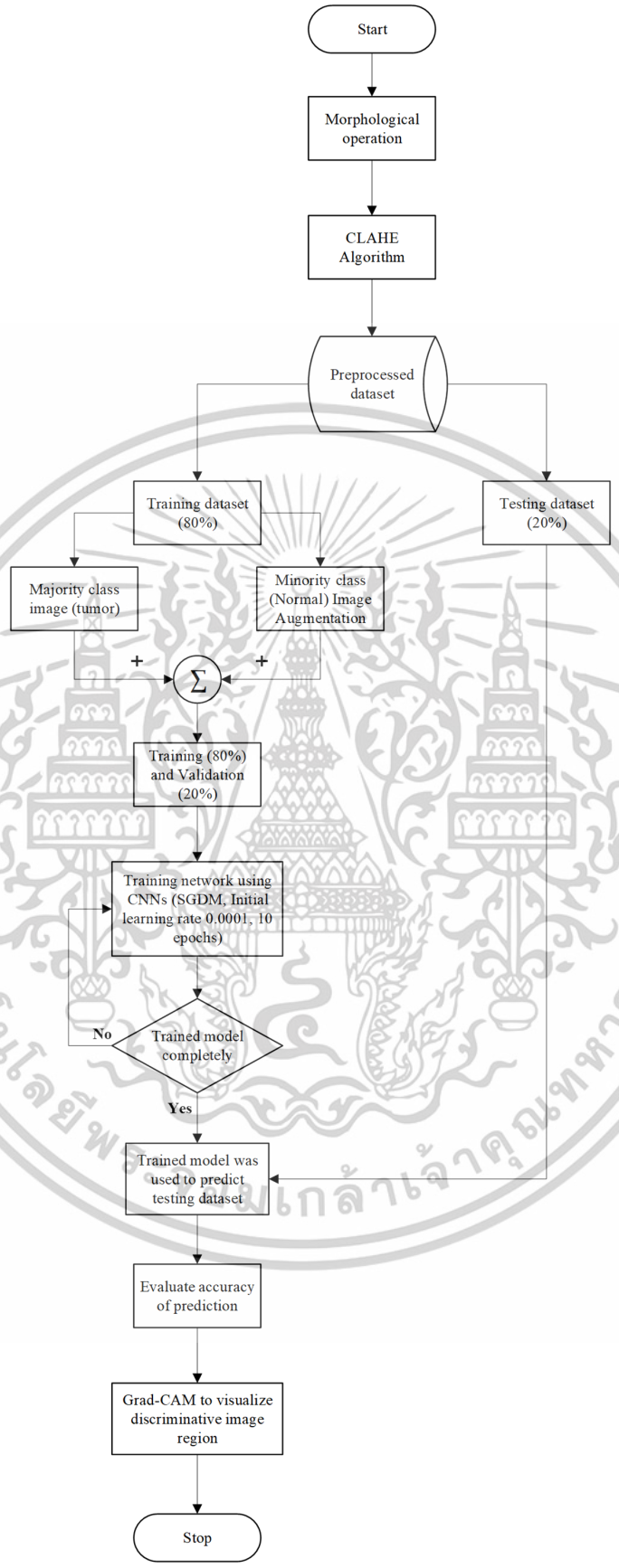
กระบวนวิธีการที่สองนำเสนอการเสริมรูปภาพ (Image Augmentation) เพื่อเพิ่มจำนวนรูปภาพคลาสส่วนน้อย (Minority Class) ทำให้กลุ่มข้อมูลสมดุลกัน และปรับปรุงความแม่นยำอย่างมีนัยสำคัญ ซึ่งเป็นวิธีที่ไม่ยุ่งยากและมีประสิทธิภาพ และจากงานวิจัยที่ผ่านมาแสดงให้เห็นว่าการกระจายตัวของข้อมูลมีผลกระทบต่อความแม่นยำการทำนายของโครงข่ายประสาทเทียมแบบคอนโวลูชันทำให้ได้รับผลการทำนายที่ดีถ้าใช้กลุ่มข้อมูลอบรมที่สมดุลกัน (Balanced Dataset) [42], [43] ดังนั้นจึงนำเสนอวิธีการนี้เพื่อเสริมภาพสมองปกติซึ่งเป็นคลาสส่วนน้อยในกลุ่มข้อมูลอบรมและไม่สมดุลกับภาพเนื้องอกสมองอย่างมาก และท้ายที่สุดหลังจากเสริมภาพได้รับสมองปกติ 1,028 รูปภาพ และเนื้องอกสมองกลิโอมา 1,028 ภาพ เพื่อใช้ประโยชน์ในการอบรมโครงข่ายต่อไป โดยเลเยอร์โครงข่ายและอัลกอริทึมเพิ่มเติมพารามิเตอร์เหล่านี้ยังคงเหมือนกับวิธีการที่หนึ่ง เมื่อครบทั้งสามองค์ประกอบคือ Training Dataset, Network Layer, และ Algorithm Options จึงทำการอบรมโครงข่ายโดยใช้โครงข่ายที่อบรมไว้แล้วแต่ละแบบ รวมทั้งกราฟความก้าวหน้าการอบรมถูกพล็อตไว้เพื่อติดตามมอนิเตอร์และปรับแต่งละเอียด (Fine Tuning) ในอนาคตต่อไป ในส่วนกลุ่มข้อมูลทดสอบยังมีจำนวนเท่ากับวิธีการแรก นั่นคือ 372 รูปภาพซึ่งประกอบด้วยสมองปกติ 115 ภาพและเนื้องอก 257 ภาพ โครงข่ายที่ได้รับการอบรมแล้วถูกใช้เพื่อทำนายผลกลุ่มข้อมูลทดสอบและความแม่นยำในการทำนายระหว่างรูปภาพและตารางความจริงของรูปภาพได้รับการประเมินผล รวมทั้ง Confusion Chart ได้ถูกพล็อตไว้เพื่อวิเคราะห์พารามิเตอร์ทางด้านสมรรถนะการทำนายผลอื่นๆ อีกด้วย บล็อกไดอะแกรมสำหรับการอบรมกลุ่มข้อมูลแบบสมดุลกันถูกแสดงดังรูปที่ 4.5



รูปที่ 4.5 บล็อกไดอะแกรมการอบรมกลุ่มข้อมูลซึ่งสมดุลกันโดยการเสริมภาพ

กระบวนการวิธีที่ 2 การอบรมและวินิจฉัยเนื้องอกสมองด้วยกลุ่มข้อมูลอบรม (Training Dataset) ซึ่งถูกทำให้สมดุลกันโดยใช้หลักการเสริมรูปภาพส่วนน้อย (Augmentation) สามารถอธิบายเพิ่มเติมเพื่อความเข้าใจยิ่งขึ้นตามโฟลว์ชาร์ตรูปที่ 4.6 ส่วนรายละเอียดแต่ละขั้นตอนอธิบายคร่าวๆ ได้ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 โฟลว์ชาร์ตการอบรมโครงข่ายโดยใช้กลุ่มข้อมูลซึ่งสมดุลกันด้วยการเสริมรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2.1 Morphological Operation Preprocessing

เมื่อได้รับภาพที่แปลงเป็น JPEG ฟอร์แมทแล้วภาพเหล่านั้นต้องผ่านการเตรียมกระบวนการก่อนที่จะส่งเข้าไปอบรมโดย CNNs เพื่อความเหมือนกันของทุกภาพและเป็นมาตรฐานเดียวกัน ซึ่งขั้นแรกของการเตรียมกระบวนการคือ ปฏิบัติการทางสัณฐานวิทยา (Morphological Operation) เป็นการแยกส่วนกะโหลกและสิ่งแปลกปลอมออกจากส่วนสมองเพื่อง่ายต่อการทำงานขั้นถัดไป ประกอบด้วยสี่ขั้นตอนตามภาพที่ 4.7 คือ ปฏิบัติการสัณฐานวิทยาแบบเปิด (Opening Morphological) การใช้ฟังก์ชันลบภาพ (Subtract) เพื่อลบส่วนที่ไม่ได้ใช้งานออก และท้ายที่สุดจะได้พื้นที่รูปภาพสมองอย่างเดียว จากรูปที่ 4.7 ก. ภาพที่ได้รับหลังจากผ่านกระบวนการสัณฐานวิทยาแบบเปิด (Open Image) รูปที่ 4.7 ข. คือ ภาพได้รับจากภาพต้นแบบ (Original Image) ลบด้วยภาพที่ผ่านการเปิด (Opened Image) รูปที่ 4.7 ค. คือภาพที่ได้จากกะโหลกที่มีส่วนแปลกปนที่ได้จากรูป 4.7 ข. ลบด้วยภาพที่ผ่านกระบวนการแบบเปิดในรูป 4.7 ก. ส่วนรูปที่ 4.7 ง. คือภาพที่เป็นส่วนพื้นที่สมองอย่างเดียวที่จะนำไปดำเนินการต่อไป

ก. ภาพที่ผ่านการ Opening ข. ภาพต้นแบบลบด้วยภาพ Opening

ค. รูปภาพส่วนกะโหลก ง. ภาพพื้นที่สมองที่แยกออกมา

รูปที่ 4.7 ขั้นตอนการปฏิบัติการทางสัณฐานวิทยา

สามารถอธิบายการใช้โปรแกรม MATLAB ในการทำ Morphological เพื่อนำส่วนกะโหลก (Skeleton) และสิ่งแปลกปลอม (Artifact) ออกจากส่วนพื้นที่สมอง ซึ่งจะมีทั้งหมด 4 ขั้นตอนคร่าวๆ ดังต่อไปนี้ [10], [47]

1. Opening Morphological Operation คือ การใช้ Opening เพื่อแยกกะโหลกกับส่วนพื้นที่สมองออกจากกัน ขั้นแรกอ่านหรือเรียกรูปภาพที่จะดำเนินการเข้ามายังพื้นที่ทำงานใน MATLAB workspace

```
I=imread("0001.jpg");
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กำหนดโครงสร้างอิลิเมนท์เป็นแบบดิสค์ขนาดรัศมี 25 พิกเซล การกัดกร่อนรูปภาพ (Erodes) จากนั้นขยายรูปภาพที่ถูกกัดกร่อนออก (Dilates) โดยใช้โครงสร้างอิลิเมนท์เท่ากัน

```
SEdisk=strel('disk',25);
```

นำส่วนเล็กๆ ที่มีรัศมีน้อยกว่า 25 พิกเซลออกโดยการ Opening ด้วยโครงสร้างอิลิเมนท์แบบดิสค์ รวมทั้งแสดงรูปภาพที่ได้รับการ Opening แล้ว

```
background=imopen(I,SEdisk);
imshow(background)
```

การ Opening ของเซต A ด้วยโครงสร้างอิลิเมนท์ B แสดงได้ดังสมการ $A \circ B = (A \ominus B) \oplus B$ ซึ่งเป็นการ Erosion ของ A โดย B และจากนั้นตามด้วยการ Dilation ของผลลัพธ์โดย B นั่นเอง

2. Subtract Opened Image From Original หรือการลบภาพออกซึ่งในที่นี้จะทำทีละขั้นตอน ขั้นแรกลบภาพที่ได้จากการ Opening ขั้นตอนแรกออกจากภาพต้นแบบ จะได้ภาพกะโหลกและส่วนสมองมาด้วยเช่นกัน ดังรูปที่ 4.7 ข.

```
Isub=imsubtract(I,background);
```

3. Subtract Opened Image From Isub เพื่อให้ได้ส่วนกะโหลกเท่านั้น โดยการนำรูปภาพที่ได้จากการ opening ในขั้นตอนที่ 1 ลบออกจากภาพกะโหลกและพื้นที่สมองบางส่วนที่ได้รับในขั้นตอนที่ 2 ทำให้ได้รับภาพกะโหลกอย่างเดียวตามรูป 4.7 ค.

```
Isub1=imsubtract(Isub,background);
```

4. Subtract Skeleton from Original Image ขั้นตอนสุดท้ายเพื่อให้ได้พื้นที่สมองอย่างเดียว นำภาพกะโหลกที่ได้จากขั้นตอนที่ 3 มาลบออกจากภาพต้นแบบจะได้รับภาพที่ต้องการซึ่งจะเป็นภาพสมองเท่านั้นดังรูปที่ 4.7 ง.

```
Isub2=imsubtract(I,Isub1);
```

สามารถเขียนรูปภาพเพื่อบันทึกเก็บไว้ได้โดยใช้คำสั่ง Imwrite รวมทั้งเรียกดูภาพก่อนและหลังปฏิบัติการสัณฐานวิทยาโดยใช้คำสั่ง Imshowpair จากนั้นเพื่อพล็อตฮิสโตแกรมของภาพคำสั่ง Imhist จะถูกใช้เพื่อดำเนินการ

```
imwrite(Isub2,'0001m.jpg')
imshowpair (I,Isub2,'montage');
figure;
imhist(Isub2);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2.2 Contrast Limited Adaptive Histogram Equalization

รูปภาพที่ผ่านกระบวนการ Morphological Operation แล้วจะต้องนำมาปรับคอนทราสให้เป็นมาตรฐานเดียวกันโดยใช้อัลกอริทึม CLAHE [11], [12] รายละเอียดทฤษฎีดังที่กล่าวในหัวข้อ 2.3.1.3 โดยขั้นตอนและส่วนประกอบแสดงดังรูปที่ 4.8 ข้อดีหลักของอัลกอริทึม CLAHE คือความต้องการการคำนวณและการนำไปใช้งานที่ไม่ยุ่งยาก ซึ่งใช้พารามิเตอร์ตัวเดียวเท่านั้นคือ Clip Limit รวมทั้งได้รับผลลัพธ์ที่ยอดเยี่ยมกับรูปภาพส่วนใหญ่

ปัญหาสัญญาณรบกวนใน Adaptive Histogram Equalization (AHE) สามารถลดได้โดยการจำกัด Contrast Enhancement โดยเฉพาะอย่างยิ่งในพื้นที่เนื้อเดียวกัน (Homogeneous) ซึ่งพื้นที่ลักษณะนี้จะเกิดพีคที่สูงในฮิสโตแกรมเพราะจำนวนพิกเซลจะมีระดับสีเทาในพื้นที่เดียวกันเยอะ ด้วยการใช้ CLAHE ระดับสีเทาจะถูกจำกัดไว้โดยอนุญาตเพียงแค่พิกเซลที่ไม่เกินค่ากำหนดสูงสุดเท่านั้น หลังจากการตัดฮิสโตแกรมพิกเซลที่ถูกตัดออกจะถูกกระจายเท่าๆ กันไปยังฮิสโตแกรมทั้งหมด เพื่อทำให้ค่าฮิสโตแกรมโดยรวมยังเท่าเดิม อธิบายตัวอย่างการใช้โปรแกรม MATLAB ในการเตรียมกระบวนการอัลกอริทึม CLAHE ได้ดังต่อไปนี้

ขั้นแรกอ่านรูปภาพที่จะปรับสมดุลคอนทราส และเปลี่ยนรูปภาพเป็นชนิดเกรย์สเกล

```
I1 = imread('0002.jpg');
I2=im2gray(I1)
```

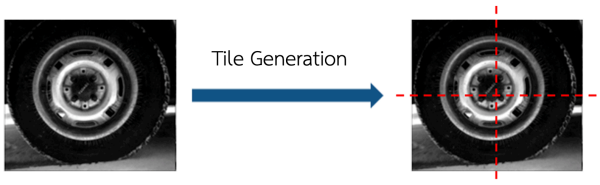
จากนั้นใช้คำสั่ง Adapthisteq และกำหนด Clip Limit เป็น 0.00001 เพื่อปรับสมดุลคอนทราสให้เป็นมาตรฐานเดียวกัน และเขียนเพื่อเก็บรูปภาพเพื่อดำเนินการขั้นถัดไป

```
I3 = adapthisteq (I2,'clipLimit',0.00001,'Distribution','rayleigh');
imwrite(I3,'0002C.jpg');
```

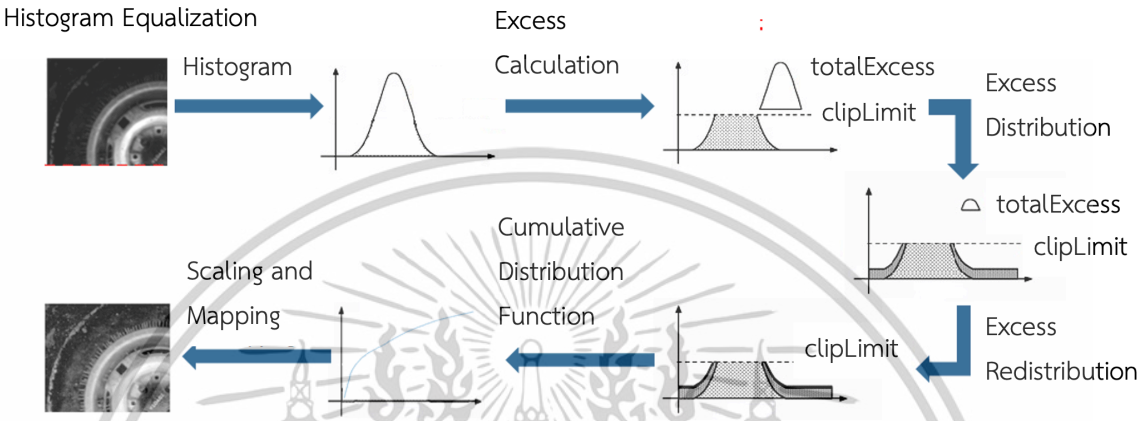
จากนั้นภาพที่ได้ต้องแปลงจากชนิดเกรย์สเกลกลับเป็นชนิด RGB เพื่อนำไปใช้งานอบรมโดยใช้การเรียนรู้แบบถ่ายโอน ในขั้นตอนต่อไป รวมทั้งแสดงรูปฮิสโตแกรมโดยใช้คำสั่ง imhist และแสดงค่าเทรชโฮลด์ของภาพ

```
imggray = imread('0002C.jpg');
rgbImage = cat(3, imggray, imggray, imggray);
imwrite(rgbImage,'0002RGB.jpg')
imshow(rgbImage)
figure;
imhist(rgbImage);
T = graythresh(rgbImage)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. Tile Generation



ข. Histogram Equalization

ค. Bilinear Interpolation

รูปที่ 4.8 ขั้นตอนของอัลกอริทึม CLAHE

4.2.2.3 การเสริมภาพเพิ่มจำนวนกลุ่มข้อมูลอบรม (Augmentation)

สามารถใช้คำสั่งในโปรแกรม MATLAB เพื่อทำการเสริมรูปภาพ (Image Augmentation) ซึ่งในงานนี้ใช้การหมุนภาพ 90 องศา (Rotation) และการสะท้อน (Reflection) [47] ดังที่กล่าวมาแล้วในบทที่ 2 ตัวอย่างโปรแกรมที่ใช้เสริมภาพด้วยการหมุนภาพอธิบายได้ดังนี้

1. กำหนดรูปแบบการหมุนของภาพอินพุทแบบสุ่ม -90 ถึง 90 องศา

```
tform = randomAffine2d(Rotation=[-90 90]);
outputView = affineOutputView(size(imOriginal),tform);
imAugmented = imwarp(imOriginal,tform,OutputView=outputView);
imshow(imAugmented)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การสร้างรูปแบบภาพสะท้อนที่หมุนภาพอินพุทรอบแกน x และ y แบบสุ่ม

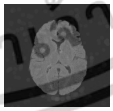
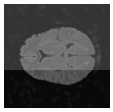
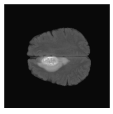
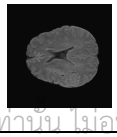
```
tform = randomAffine2d(XReflection=true,YReflection=true);
outputView = affineOutputView(size(imOriginal),tform);
imAugmented = imwarp(imOriginal,tform,OutputView=outputView);
imshow(imAugmented)
```

4.2.2.4 การอบรมโดยใช้โครงข่ายประสาทเทียมคอนโวลูชัน

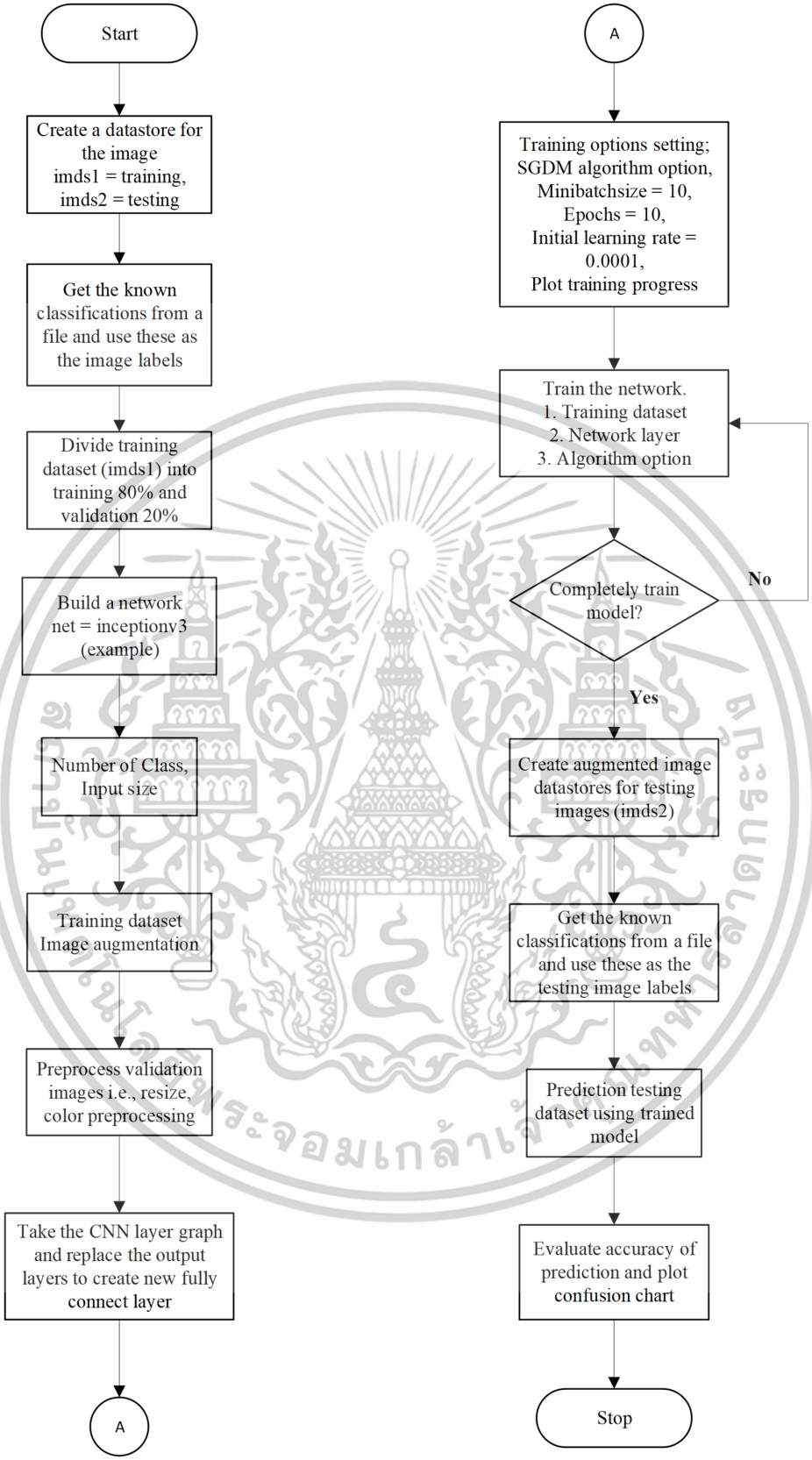
หลังจากได้รูปภาพที่ผ่านการเตรียมกระบวนการแล้วจะต้องนำมาอบรมโดยใช้โครงข่ายประสาทเทียมคอนโวลูชัน ในขณะเดียวกัน Ground Truth Label ของกลุ่มข้อมูลก็จะถูกจัดเตรียมให้ตรงกับความจริงของภาพ กล่าวคือสมองปกติหรือเนื้องอก เพื่อนำไปใช้ในการอบรมโมเดลโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันต่อไปดังตารางที่ 4.4 กล่าวอย่างเข้าใจง่ายๆ คือ นำภาพสมองและฉลากความจริงป้อนเข้าไปยังโครงข่ายจากนั้นฟีเจอร์แมพจะถูกสกัดเพื่อจำแนกความแตกต่างระหว่างสมองปกติและเนื้องอก

ในส่วนของโปรแกรมขั้นตอนการอบรมโมเดลโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันนั้นอธิบายได้ด้วยโฟลว์ชาร์ตรูปที่ 4.9 โดยรูปภาพที่ผ่านขั้นตอนเตรียมกระบวนการ (Preprocessing) แล้วจะถูกเก็บไว้ในฐานข้อมูลส่วนบุคคล จากนั้นดำเนินตามขั้นตอนตามโฟลว์ชาร์ตเพื่ออบรมโครงข่ายและทำนายผลกลุ่มข้อมูลทดสอบ (Testing Dataset)

ตารางที่ 4.4 ตัวอย่าง Ground Truth Label กลุ่มข้อมูลเพื่อใช้อบรมโครงข่าย

Image No.	Ground Truth Label	JPEG image
00004.jpg	Tumor	
00195.jpg	Normal	
000195L.jpg	Normal	
00916.jpg	Tumor	
01111.jpg	Normal	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.9 โฟลว์ชาร์ตขั้นตอนการอบรมโมเดลโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน

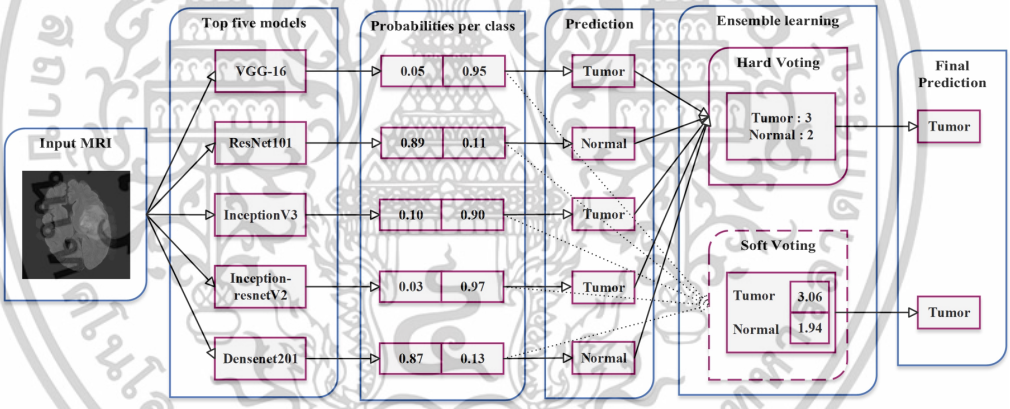
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.3 กระบวนวิธีการที่ 3 การจัดหมวดหมู่โดยใช้ Majority Voting

เมื่อไม่นานมานี้ หลายงานการศึกษาแนะนำการเรียนรู้แบบรวมหลายโมเดลที่แตกต่างและมีอิสระจากกันเข้าด้วยกัน (Ensemble Learning) เพื่อเพิ่มประสิทธิภาพของโมเดล รวมทั้งปรับปรุงความแม่นยำของการจัดหมวดหมู่รูปภาพทางการแพทย์ [43], [44] ในการศึกษานี้ Ensemble Hard และ Soft Majority Voting ได้นำมาใช้เพื่อวิเคราะห์สืบค้นประสิทธิภาพของการจัดหมวดหมู่เนื้องอกสมอง โมเดลที่ได้ผลลัพธ์ที่มีประสิทธิภาพสูงสุดห้าอันดับแรกได้นำมาใช้เพื่อการโหวต รายละเอียดของวิธีการโหวตแบบเสียงส่วนมากอธิบายได้ดังนี้

การโหวตเสียงส่วนมากแบบฮาร์ด (Majority Hard Voting) คือ ผลลัพธ์การรวมฉลากผลการทำนายหมวดหมู่จากทุกๆ อุปกรณ์จัดหมวดหมู่ที่ได้ทำนายภาพอินพุท และท้ายที่สุดผลรวมจำนวนสูงสุดจะเป็นผลโหวตของฉลากที่ได้รับ แสดงได้ดังรูปที่ 4.10

การโหวตเสียงส่วนมากแบบซอฟท์ (Majority Soft Voting) คือ ผลลัพธ์การรวมค่าความน่าจะเป็นต่อคลาสที่ถูกสร้างโดยทุกๆ อุปกรณ์จัดหมวดหมู่สำหรับภาพอินพุท และผลลัพธ์สุดท้ายของลาเบลที่ทำนายได้ มาจากค่าความน่าจะเป็นรวมสูงสุด แสดงได้ดังรูปที่ 4.10



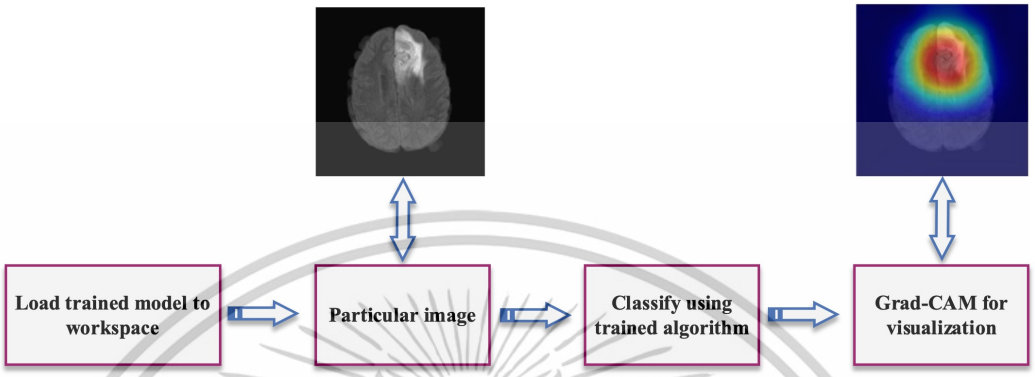
รูปที่ 4.10 Ensemble Learning โดยใช้การโหวต Majority Voting

4.2.4 การสร้างภาพการมองเห็นพื้นที่ซึ่งถูกแยกแยะโดยใช้ GRAD-CAM

งานวิจัยนี้ได้ทำการศึกษาและนำอัลกอริทึม Grad-CAM มาใช้งานจริงเพื่อมองเห็นพื้นที่รูปภาพซึ่งถูกแยกแยะ (Discriminative Image Region) ของรูปภาพ MRI ที่ถูกทำนายผลแล้วเพื่อการตีความหมายและความน่าเชื่อถือได้ของผลลัพธ์ทำนายดังกล่าว โปรแกรมซึ่งสร้างและดัดแปลงเพื่อใช้มองเห็นถูกแสดงดังบล็อกไดอะแกรมรูปที่ 4.11 อธิบายเป็นขั้นตอนได้ดังนี้โดยโมเดลที่ได้รับการอบรมแล้วของแต่ละอัลกอริทึม CNNs ถูกอัพโหลดไปยังพื้นที่ทำงาน (Workspace) จากนั้นรูปภาพแบบเฉพาะเจาะจงที่ต้องการแสดงการมองเห็นและตรวจสอบถูกอ่านมายังพื้นที่ทำงานด้วยเพื่อดำเนินการขั้นถัดไป รูปภาพที่เฉพาะเจาะจงนี้ถูกจัดหมวดหมู่โดยใช้อัลกอริทึมโมเดลโครงข่ายที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อบรมไว้แล้วเพื่อทำนายวินิจฉัยภาพสมองปกติหรือเนื้องอกกลิโอมา ค่าความน่าจะเป็นถูกระบุในภาพอีกด้วย ท้ายที่สุดผลลัพธ์การทำนายได้รับการตีความหมายโดยใช้อัลกอริทึม Grad-CAM รวมทั้งรูปภาพแผนที่ความร้อนได้ถูกสร้างขึ้นสำหรับแสดงการมองเห็น ซึ่งได้แสดงโค้ดเทียม (Pseudo code) สำหรับการมองเห็นรูปภาพที่เฉพาะเจาะจงแสดงได้ดังตารางที่ 4.5



รูปที่ 4.11 บล็อกไดอะแกรมเพื่อการมองเห็นรูปภาพการทำนายโดยใช้อัลกอริทึม Grad-CAM

ตารางที่ 4.5 Pseudo code สำหรับการจัดหมวดหมู่และมองเห็นรูปภาพแบบเฉพาะเจาะจง

Pseudo code for visualization of the image

Begin

Require: CNN trained model using brain tumor training dataset with transfer learning.

1. Load CNN trained model to workspace.
2. Verify the compatible of input image size // i.e., the required input image is 299 x 299 for inceptionV3.
3. Read input image RGB type which requires to predict and visualize into the workspace.
4. Resize the input image to equal with the required input image of network.
5. Classify the input image to indicate tumor or normal brain using the trained model network.
6. Compute the Grad-CAM map for visualization the predicted image :

$$L^{c}_{Grad-CAM} = ReLU\left(\sum_{k} \alpha^{c}_{k} A^{k}\right)$$

7. Display the discriminative image region of classified image.

End

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การนำมาใช้งานจริงของโปรแกรมซึ่งถูกสร้างขึ้นนี้จะช่วยเหลือเจ้าหน้าที่ทางการแพทย์ในการวินิจฉัยเนื้องอกสมองกลิโอมาโดยแยกแยะรูปภาพทีละภาพได้ รูปถ่ายทางการแพทย์ MRI ซึ่งได้รับการเตรียมกระบวนการตามที่อธิบายก่อนหน้าจะถูกโหลดเข้ามายังพื้นที่ทำงานของโปรแกรม MATLAB จากนั้นโมเดล CNN ที่ถูกอบรมไว้แล้วจะทำนายภาพดังกล่าวเพื่อระบุเป็นเนื้องอกหรือสมองปกติ นอกเหนือจากนั้นภาพที่ได้รับการทำนายจะได้รับการยืนยันอีกครั้งโดยใช้อัลกอริทึม Grad-CAM เพื่อสร้างพื้นที่รูปภาพซึ่งได้รับการแยกแยะ

4.3 การประเมินผลสมรรถนะ

ผลลัพธ์การทดลองได้รับการประเมินโดยใช้คุณลักษณะของการทดลองวินิจฉัยที่มีชื่อว่าความไว (Sensitivity) ความเฉพาะเจาะจง (Specificity) และความแม่นยำ (Accuracy) ซึ่งถูกใช้อย่างกว้างขวางทางสถิติเพื่ออธิบายผลทดสอบวินิจฉัย [48] หลากหลายเทอมตัวแปรถูกใช้สำหรับจำกัดความของความไว ความเฉพาะเจาะจง และความแม่นยำ เหล่านั้นคือ True Positive (TP), True Negative (TN), False Negative (FN) และ False Positive (FP) ในงานวิจัยนี้ถ้าเนื้องอกถูกพิสูจน์ว่ามีอยู่จริงในผู้ป่วยและผลลัพธ์การทำนายระบุได้ว่าพบว่ามีโรคอยู่จริง ผลลัพธ์ของการทดสอบวินิจฉัยถูกพิจารณาเป็น True Positive ในทำนองเดียวกันถ้าเนื้องอกถูกพิสูจน์ว่าไม่มีอยู่จริงในผู้ป่วย ผลลัพธ์การทำนายระบุว่าไม่พบเนื้องอกในผู้ป่วยด้วยเหมือนกัน ผลลัพธ์การทดสอบจะได้รับการพิจารณาเป็น True Negative

True Positive และ Negative แนะนำความเหมือนกันของผลลัพธ์ระหว่างการทดสอบวินิจฉัยและสถานะที่ถูกพิสูจน์แล้ว ในทางตรงกันข้ามถ้าเนื้องอกถูกพิสูจน์ว่ามีอยู่ในผู้ป่วยซึ่งแท้ที่จริงแล้วไม่มีโรคแต่อย่างใด ผลลัพธ์ทดสอบถูกพิจารณาเป็น False Positive ในทำนองเดียวกันถ้าผลลัพธ์การทำนายถูกพิสูจน์และระบุว่าไม่พบเนื้องอกในผู้ป่วยผู้ที่ซึ่งเป็นโรคเนื้องอกสมองอย่างแน่นอน ผลลัพธ์การทดสอบถูกพิจารณาเป็น False Negative โดย False Positive และ Negative นี้จะระบุว่าผลลัพธ์การทดสอบตรงข้ามกับสภาวะจริง ผลการทำนายจะถูกแสดงใน Confusion Chart ดังที่กล่าวไปแล้ว และค่าตัวแปรต่างๆ ถูกระบุเพื่อทำความเข้าใจได้ง่ายขึ้นดังแสดงในรูปที่ 4.12

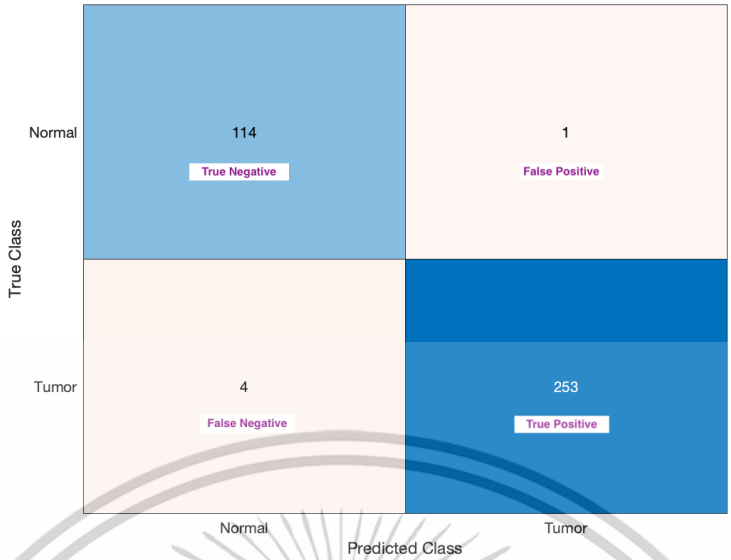
ความไว ความเฉพาะเจาะจง และความแม่นยำถูกอธิบายในเทอมของ TP, TN, FN และ FP ดังต่อไปนี้

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{4.1}$$

$$Specif icity = \frac{TN}{(TN + FP)} \tag{4.2}$$

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)} \tag{4.3}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 ผลการทำนายที่แสดงในรูปแบบ Confusion Chart

ค่าตัวเลขของความไว (Sensitivity) เป็นตัวแทนของความน่าจะเป็นของการทดสอบวินิจฉัยซึ่งระบุผู้ป่วยซึ่งพบว่ามีโรค การทดสอบด้วยค่าความไวที่สูงมีแนวโน้มที่จะเก็บรวบรวมความเป็นไปได้ทั้งหมดของสภาวะที่เป็นผลบวกโดยปราศจากความผิดพลาดต่อผู้ป่วยรายใดรายหนึ่ง ดังนั้นการทดลองด้วยความไวที่สูงบ่อยครั้งถูกใช้เพื่อคัดกรองผู้ป่วย

ค่าตัวเลขของความเฉพาะเจาะจง (Specificity) เป็นตัวแทนความน่าจะเป็นของผลวินิจฉัยที่ระบุผู้ป่วยซึ่งไม่พบว่ามีโรคแต่อย่างใด การทดสอบด้วยค่าความเฉพาะเจาะจงสูงนั้นเมื่อนำไปวินิจฉัยกับผู้ป่วยที่ปราศจากโรคอย่างแน่นอน มีโอกาสสูงที่ผู้ป่วยรายนั้นถูกระบุเป็นลบได้อย่างถูกต้อง ผลลัพธ์การทำนายผลที่ได้รับคือ True Positive Rate (TPR) ที่ซึ่งเป็นความสัมพันธ์กันระหว่างความไวและ False Positive Rate (FPR) ซึ่งมีค่าเท่ากับ (1 – ความเฉพาะเจาะจง) การผสมผสานกันของความเป็นไปได้ทั้งหมดของ TPR และ FPR จะก่อให้เกิดพื้นที่ Receiver Operating Characteristic (ROC) เพิ่มเติมจากนั้นพื้นที่ด้านล่าง ROC Curve (AUC) วัดความแม่นยำของผลวินิจฉัยได้อีกด้วย พื้นที่มากกว่าระบุได้ว่าผลวินิจฉัยที่แม่นยำกว่า พื้นที่ด้านล่าง ROC curve สามารถวัดค่าได้โดยใช้สมการดังต่อไปนี้ โดยที่ t = (1 – Specificity) และ ROC (t) คือความไว (Sensitivity)

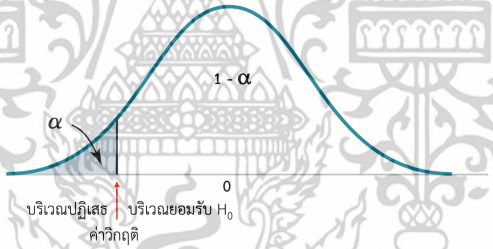
$$AUC = \int_0^1 ROC(t)\,\mathrm{d}t \tag{4.4}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.1 การวิเคราะห์เชิงสถิติโดยใช้ p-value

ในบางกรณีปัญหาทางด้านวิศวกรรมต้องการตัดสินใจว่าข้อความสองอย่างอันไหนถูกต้อง ข้อความดังกล่าวเรียกว่า สมมติฐาน (Hypotheses) และขั้นตอนตัดสินใจถูกเรียกว่าการทดสอบสมมติฐาน (Hypothesis Testing) ซึ่งส่วนประกอบของการทดสอบสมมติฐานคือ สมมติฐานหลัก (Null Hypothesis, H_0) คือการกล่าวอ้างว่าสมมติเบื้องต้นนั้นถูกต้อง และสมมติฐานรอง (Alternative Hypothesis, H_1) คือการระบุว่าขัดแย้งกับสมมติฐานหลัก ซึ่งในงานวิจัยนี้ใช้ One-Sided Alternative Hypotheses ทางเดียวที่จะรายงานผลลัพธ์การทดสอบสมมติฐานคือระบุว่าสมมติฐานหลักถูกปฏิเสธหรือไม่ (Rejecting H_0) ที่ค่าระดับนัยสำคัญที่ระบุนั้น (α, Significant Level) เรียกการทดสอบว่า ค่าระดับความสำคัญแบบคงที่ (Fixed Significance Level) เพื่อลดความยุ่งยาก วิธีการ p-value ถูกนำมาใช้อย่างกว้างขวางเพื่อทดสอบสมมติฐาน คำจำกัดความ p-value คือ ระดับความสำคัญที่น้อยสุดที่จะนำไปสู่การกำจัดสมมติฐานหลัก (Rejecting H_0) โดยใช้ข้อมูลที่ให้มา

การทดสอบทางด้านเดียวด้านซ้าย (Left-Tailed Test) เป็นการทดสอบสมมติฐาน ที่สมมติฐานรองเป็นเครื่องหมายน้อยกว่า ($<$) โดยมีเขตวิกฤติอยู่ที่ปลายหางด้านซ้ายของการแจกแจง กล่าวคือ $H_0 : \theta \geq \theta_0$ และ $H_1 : \theta < \theta_0$



รูปที่ 4.13 วิธีคำนวณ p-value สำหรับ t-test การทดสอบทางเดียวด้านซ้าย

ค่าสถิติทดสอบ คือ ค่าที่คำนวณมาจากข้อมูลตัวอย่าง และใช้ในการตัดสินใจเกี่ยวกับการปฏิเสธสมมติฐานหลัก

ระดับนัยสำคัญ (α) คือความน่าจะเป็นที่ค่าสถิติทดสอบจะมีค่าอยู่ในเขตวิกฤติ เมื่อสมมติฐานหลักเป็นจริง ถ้าค่าสถิติทดสอบที่คำนวณได้มีค่าอยู่ในเขตวิกฤติจะตัดสินใจปฏิเสธสมมติฐานหลัก

ค่าวิกฤติ คือ ค่าใดๆ ที่ใช้ในการแบ่งเขตวิกฤติและเขตยอมรับ ค่าวิกฤตินี้ขึ้นอยู่กับลักษณะของสมมติฐานหลัก การแจกแจงค่าจากตัวอย่างของค่าสถิติที่ใช้และระดับนัยสำคัญ

ในการตัดสินใจและสรุปผล จะใช้ค่า p-value ซึ่งเป็นค่าความน่าจะเป็นที่แปลงมาจากค่าสถิติทดสอบ โดยมีเกณฑ์ดังนี้

ปฏิเสธ H_0 ถ้าค่า p-value มีค่าน้อยกว่า ระดับนัยสำคัญ

ยอมรับ H_0 ถ้าค่า p-value มีค่ามากกว่าหรือเท่ากับ ระดับนัยสำคัญ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนการทดสอบสมมติฐานมี 4 ขั้นตอนดังนี้

1. กำหนดสมมติฐานหลักและสมมติฐานรอง
2. กำหนดระดับนัยสำคัญ (α) เช่น 0.05, 0.01
3. กำหนดตัวสถิติทดสอบ และคำนวณหาค่า p-value
4. สรุปผล โดยพิจารณาค่า p-value ดังนี้

   ถ้า p-value $\geq \alpha$ ยอมรับ H_0

   ถ้า p-value $< \alpha$ ปฏิเสธ H_0

Test Statistic สามารถคำนวณโดยใช้สมการดังนี้

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \tag{4.5}$$

$\bar{x}$ คือ Sample Mean

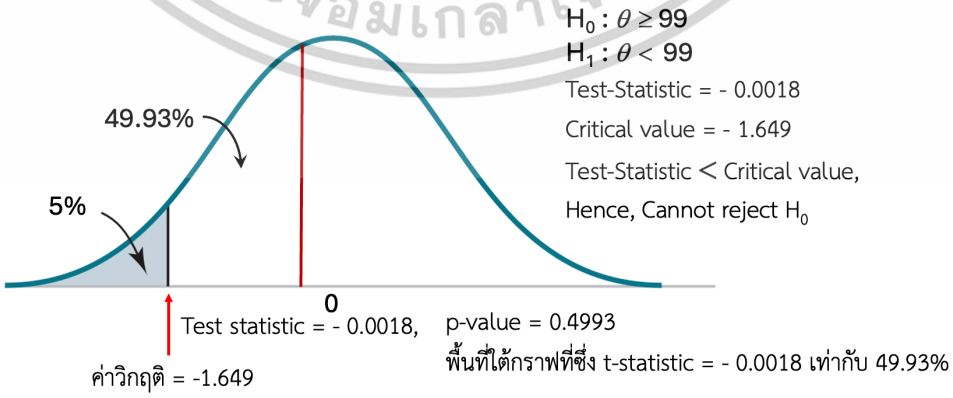
μ คือ Hypothesized Mean

s คือ Sample Standard Deviation

n คือ จำนวนตัวอย่าง

ตัวอย่างจากการทำนายเนื้องอกสมองด้วย InceptionV3 ซึ่งกลุ่มข้อมูลทดสอบทั้งหมด 372 รูปภาพ และจะทดสอบผลการทำนายซึ่งมีค่าความน่าจะเป็นมากกว่าหรือเท่ากับ 0.99 จะหาค่า t-statistic โดยใช้โปรแกรม MATLAB ได้ดังนี้

ขนาดกลุ่มข้อมูล:	n = size (x, 1) = 372
หาค่า mean:	xbar = mean(x) = 0.99
หาค่าเบี่ยงเบนมาตรฐาน:	stdev = std(x) = 0.0503
คำนวณค่า t-statistic:	tstat = (xbar-0.99)/(stdev/sqrt(n)) = - 0.0018
คำนวณค่าวิกฤติ:	critical = tinv (0.05, n-1) = - 1.649
คำนวณค่า p-value:	pvalue = tcdf(tstat,n-1) = 0.4993



รูปที่ 4.14 ผลการคำนวณ p-value สำหรับ InceptionV3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
การทดลองและการวิเคราะห์ผลการทดลอง

งานวิจัยนี้นำเสนอระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค เพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์ในการวินิจฉัยโรคเนื้องอกสมองกลิโอมา GBM เกรด 4 ซึ่งเป็นชนิดร้ายแรงที่สุด จากภาพถ่ายทางการแพทย์ จุดมุ่งหมายเพื่อพัฒนาระบบวินิจฉัยโรคอย่างรวดเร็วหรือเน้นย้ำผลการวินิจฉัยอีกครั้ง เพื่อให้การรักษาโรคอย่างทันท่วงทีด้วยวิธีการที่เหมาะสมและจะทำให้เพิ่มโอกาสรอดชีวิตของผู้ป่วยได้ โดยนำเสนอสองกระบวนวิธีการทดลองสำหรับการวินิจฉัยโรคเนื้องอกสมอง กล่าวคือ การจัดหมวดหมู่เนื้องอกสมองกลิโอมาที่มีกลุ่มข้อมูลจำนวนน้อยและไม่สมดุลกันของกลุ่มข้อมูลอบรมระหว่างเนื้องอกและสมองปกติ โดยใช้วิธีการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน และเปรียบเทียบกับการใช้การเรียนรู้แบบถ่ายโอน ส่วนกระบวนวิธีการที่สองนำเสนอการทดลองกับกลุ่มข้อมูลที่มีจำนวนมากขึ้นและใช้การเรียนรู้เชิงลึกแบบโครงข่ายประสาทเทียมคอนโวลูชันในการจัดหมวดหมู่ และนำเสนอการเรียนรู้แบบผสมผสานเพื่อศึกษาผลลัพธ์เพิ่มเติม ผลการทดลองที่ได้รับในแต่ละกระบวนวิธีการกล่าวได้ดังต่อไปนี้

5.1 ผลการทดลองการวินิจฉัยเนื้องอกสมองที่มีกลุ่มข้อมูลขนาดเล็ก

5.1.1 ผลการทดลองการวินิจฉัยเนื้องอกสมองด้วยซัพพอร์ตเวกเตอร์แมชชีน

โมเดลที่ได้รับการอบรมแล้วถูกโหลดมายังพื้นที่ทำงานของโปรแกรม MATLAB จากนั้นกลุ่มข้อมูลทดสอบทั้งหมด 32 ฟีเจอร์ที่เป็นตัวแทนของรูปภาพเนื้องอกหรือสมองปกติ และแสดงในรูปแบบตารางถูกอ่านมายังพื้นที่ทำงาน (Workspace) เพื่อทำนายว่าภาพไหนเป็นเนื้องอกหรือสมองปกติ โดยใช้อัลกอริทึมการเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีน ดังที่กล่าวในหัวข้อ 3.1 และท้ายที่สุดได้รับผลลัพธ์การทำนายเอาท์พุท จากนั้นผลลัพธ์การทำนายถูกนำมาวัดผลเปรียบเทียบกับ Ground Truth Label คุณลักษณะของการทดสอบวินิจฉัยสามารถวัดผลได้ โดยคุณสมบัติเหล่านั้นถูกเรียกว่าความไว (Sensitivity) ความเฉพาะเจาะจง (Specificity) และความแม่นยำ (Accuracy) ถูกใช้อย่างกว้างขวางทางสถิติเพื่ออธิบายการทดสอบวินิจฉัย [48] ค่าพารามิเตอร์ใช้เพื่อคำนวณคุณลักษณะของการทดสอบวินิจฉัยประกอบด้วย

True Positive คือ มีเนื้องอกอยู่และถูกตรวจจับได้ถูกต้อง

True Negative คือ สมองปกติและระบุได้ถูกต้อง

False Positive คือ สมองปกติแต่ผลลัพธ์ระบุได้เป็นเนื้องอกซึ่งไม่ถูกต้อง

False Negative คือ มีเนื้องอกอยู่แต่ผลลัพธ์ไม่สามารถตรวจจับได้และระบุเป็นสมองปกติ

ซึ่งรายละเอียดอธิบายในหัวข้อ 4.3 ผลลัพธ์ค่าพารามิเตอร์และคุณลักษณะเพื่อวัดผลสมรรถนะของการทำนายถูกแสดงในตารางที่ 5.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.1 ผลลัพธ์การวินิจฉัยเนื้องอกด้วยซัพพอร์ตเวกเตอร์แมชชีน

True Positive	True Negative	False Positive	False Negative
19	12	0	1
Evaluated result			
Accuracy	Sensitivity	Specificity	
96.875%	95%	100%	

ผลลัพธ์การทดลองที่ได้รับถูกนำไปเปรียบเทียบกับงานวิจัยที่เกี่ยวข้องที่ผ่านมาและพบว่าความแม่นยำได้รับการปรับปรุงให้ได้ผลลัพธ์ที่ดีขึ้นคือ 96.875 เปอร์เซ็นต์ และผลลัพธ์การเปรียบเทียบได้แสดงในตารางที่ 5.3 ข้อดีของการจัดหมวดหมู่โดยใช้การเรียนรู้ของเครื่องจักร คืออุปกรณ์คอมพิวเตอร์ที่ใช้ในการคำนวณไม่จำเป็นต้องมีคุณสมบัติที่สูงเมื่อเปรียบเทียบกับการจัดหมวดหมู่ด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน เนื่องจากค่าฟีเจอร์ของกลุ่มข้อมูลทดสอบถูกสกัดออกมาและจัดเตรียมในรูปแบบตาราง ก่อนการอบรมด้วยอัลกอริทึมการเรียนรู้ของเครื่องจักร

5.1.2 ผลการทดลองวินิจฉัยเนื้องอกสมองด้วยการเรียนรู้แบบถ่ายโอน

กลุ่มข้อมูลทดสอบถูกแบ่งออกแบบสุ่มซึ่งคิดเป็น 30 เปอร์เซ็นต์ ที่ประกอบด้วย 47 รูปภาพ ดังที่กล่าวในหัวข้อ 3.2 ที่ผ่านมา คุณลักษณะของการทดสอบวินิจฉัยโรคสามารถวัดผลได้ คุณสมบัติเหล่านี้มีชื่อเรียกว่าความไว (Sensitivity) ความเฉพาะเจาะจง (Specificity) และความแม่นยำ (Accuracy) และ Confusion Chart สำหรับแต่ละโครงข่ายที่ได้รับการอบรมและทดสอบแล้วถูกแสดงดังรูปที่ 5.1 กลุ่มข้อมูลทดสอบจำนวน 47 รูปภาพถูกจัดหมวดหมู่โดยใช้โครงข่ายที่ได้รับการอบรมไว้แล้ว ผลลัพธ์การจัดหมวดหมู่และคุณลักษณะที่ใช้วัดประสิทธิผลของการทำนายผลถูกแสดงได้ดังตารางที่ 5.2 จะเห็นได้ว่า GoogLeNet ได้รับผล False Positive 7 รูปภาพ, False Negative 2 รูปภาพ และความแม่นยำโดยรวมเท่ากับ 80.85 เปอร์เซ็นต์ ในขณะที่ AlexNet ได้รับผลลัพธ์ False Positive 2 รูปภาพ, False Negative 1 รูปภาพ และความแม่นยำโดยรวมเท่ากับ 93.62 เปอร์เซ็นต์ ผลลัพธ์การทดลองสำหรับแต่ละโครงข่ายที่ได้รับการอบรมไว้ก่อนแล้วถูกนำเสนอในกราฟผลลัพธ์การทำนายผลรูปที่ 5.2 ที่แสดงให้เห็นว่าความไว (Sensitivity) ของ GoogLeNet และ AlexNet เท่ากับ 93.94 และ 96.97 เปอร์เซ็นต์ ตามลำดับ ในส่วนความเฉพาะเจาะจง (Specificity) ของแต่ละโครงข่ายคือ 50 และ 85.71 เปอร์เซ็นต์ ตามลำดับ

ผลลัพธ์ความแม่นยำของกระบวนวิธีการที่นำเสนอถูกนำไปเปรียบเทียบกับงานอื่นที่เกี่ยวข้องในลักษณะเดียวกันที่ผ่านมา แสดงได้ดังตารางที่ 5.3 และพบว่าได้รับสมรรถนะที่ดี และสามารถพัฒนาต่อยอดเพื่อนำไปใช้งานจริงเพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์สำหรับการวินิจฉัยโรคเนื้องอกสมอง GBM ด้วยระบบคอมพิวเตอร์ต่อไปในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.2 ผลลัพธ์การทำนายเนื้องอกสมองด้วยการเรียนรู้แบบถ่ายโอน

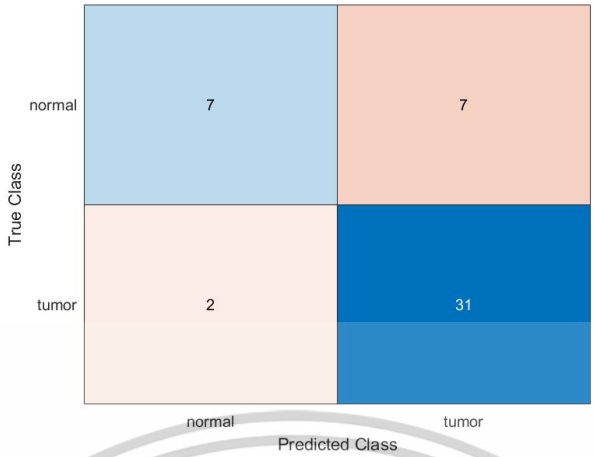
Pretrained Network	True Positive	True Negative	False Positive	False Negative
GoogLeNet	31	7	7	2
AlexNet	32	12	2	1

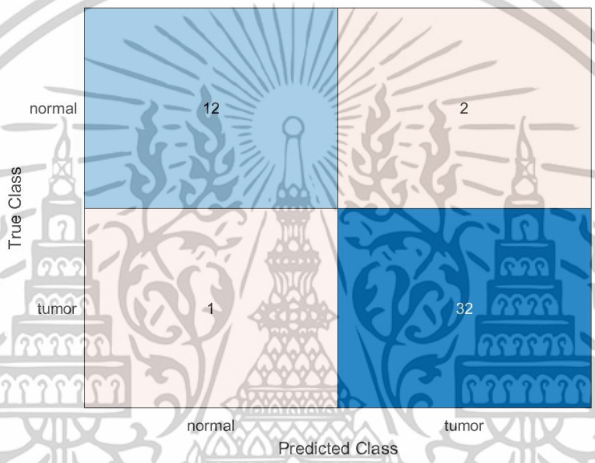
Evaluated result			
Pretrained Network	Accuracy (%)	Sensitivity (%)	Specificity (%)
GoogLeNet	80.85	93.94	50
AlexNet	93.62	96.97	85.71

ตารางที่ 5.3 ผลลัพธ์การเปรียบเทียบกับงานวิจัยที่ผ่านมาสำหรับกลุ่มข้อมูลขนาดเล็ก

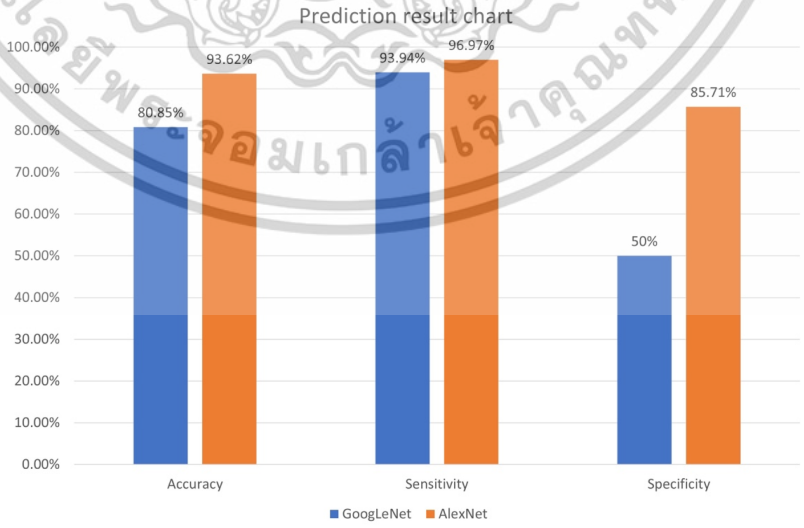
Reference	Methods	Accuracy (%)
Zaw et al. (2019)	Maximum entropy threshold and Naïve Bayes classifier	94
Dubey et al. (2018)	Wavelet transform and Neural network	80
Gaikwad et al. (2021)	CNN with VGG-16 model	99.03
Hossain et al. (2019)	Fuzzy C-means and SVM	92.42
	Convolution Neural Network	97.87
Dipu et al. (2021)	YOLOV5	85.95
	FastAi	95.78
Arora & Sharma (2021)	VGG16	90.54
	ResNet	81.92
Proposed methodology	Support Vector Machine	96.875
Proposed methodology	GoogLeNet	80.85
	AlexNet	93.62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้





รูปที่ 5.1 ผลการทำนายที่แสดงในรูปแบบ Confusion Chart



รูปที่ 5.2 ประสิทธิผลของการทำนายโดยใช้ GoogLeNet และ AlexNet

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ผลการทดลองการวินิจฉัยเนื้องอกสมองที่กลุ่มข้อมูลมีจำนวนมากขึ้นโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน

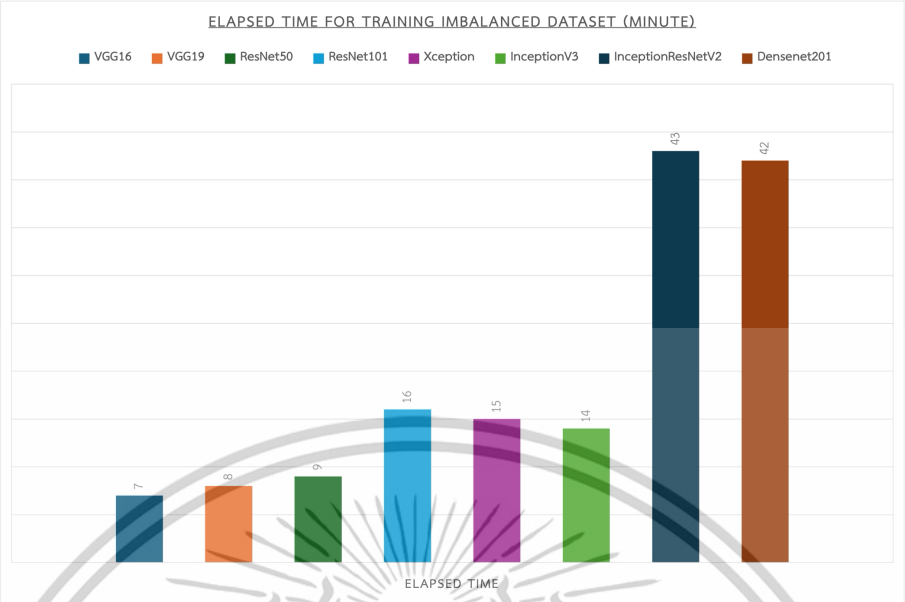
5.2.1 ผลการทดลองวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน

ในงานวิจัยส่วนนี้ได้ประยุกต์ใช้งานจริงของกระบวนวิธีการสองแบบที่ต่างกันสำหรับการจัดหมวดหมู่เนื้องอกสมองกลิโอมาโดยใช้การเรียนรู้แบบถ่ายโอนหลากหลายแบบซึ่งอยู่บนพื้นฐานของ CNNs ตามที่อธิบายก่อนหน้านี้ กระบวนวิธีการทดลองแบบที่ 1 รูปภาพ MRI ซึ่งผ่านขั้นตอนเตรียมกระบวนการถูกใช้เพื่ออบรมและทดสอบโดยใช้ CNNs ที่หลากหลาย 8 แบบด้วยกัน ดังที่กล่าวในหัวข้อ 4.2.1 ที่ผ่านมา ในงานวิจัยส่วนนี้ใช้คอมพิวเตอร์ส่วนบุคคลเพื่อการอบรมกลุ่มข้อมูลทดสอบโดยรายละเอียดดังนี้ CPU Intel Core i7, 3.0 GHz, RAM 32 GB, and NVIDIA GeForce RTX 3070, Graphics Processing Unit (GPU) และโปรแกรมที่ใช้ในการอบรมและทดสอบกลุ่มข้อมูลคือ MATLAB R2023a การกำหนดพารามิเตอร์ที่สำคัญดังนี้ อัลกอริทึมออพชันแบบ Stochastic Gradient Descent with Momentum ด้วยอัตราการเรียนรู้เริ่มต้นเท่ากับ 0.0001, Max epochs 20 และข้อมูลสรุปผลลัพธ์การอบรม Validation Accuracy และเวลาที่ใช้ในการอบรมและทดสอบกลุ่มข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบต่างๆ แสดงดังตารางที่ 5.4 และแผนภูมิรูปภาพเปรียบเทียบระยะเวลาการอบรมด้วยโครงข่ายประสาทเทียมคอนโวลูชันแบบต่างๆ แสดงดังรูปที่ 5.3

ตารางที่ 5.4 ผลลัพธ์การอบรมกลุ่มข้อมูลด้วยวิธีการที่ 1 กลุ่มข้อมูลที่ไม่สมดุลกัน

CNNs Transfer Learning	Validation Accuracy (%)	Training Elapsed Time	Testing Inference Time (sec.)
VGG16	81.48	7 min. 53 sec.	4.578
VGG19	85.19	8 min. 43 sec.	4.764
ResNet50	79.80	9 min. 33 sec.	5.221
ResNet101	77.10	16 min. 42 sec.	6.312
Xception	74.41	15 min. 10 sec.	6.658
InceptionV3	79.46	14 min. 29 sec.	5.924
InceptionResNetV2	83.50	43 min. 45 sec.	9.445
Densenet201	78.79	42 min. 51 sec.	9.859

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

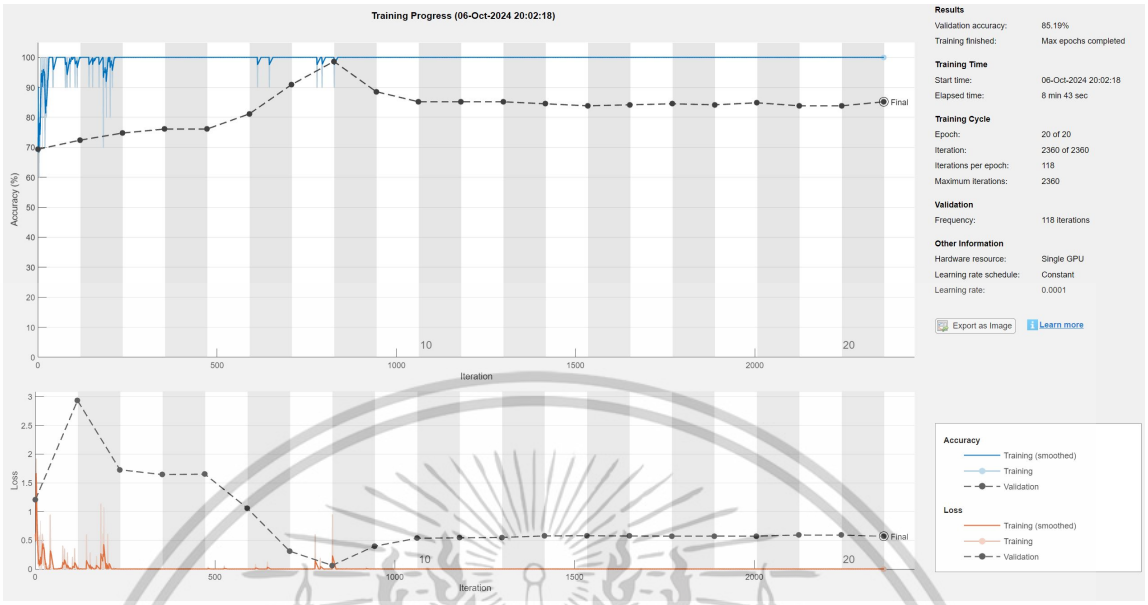


รูปที่ 5.3 เปรียบเทียบเวลาที่ใช้ในการอบรมซึ่งใช้กลุ่มข้อมูลไม่สมดุลกัน

ในระหว่างกลุ่มข้อมูลได้รับการอบรมโดยโครงข่ายประสาทเทียมนั้นความก้าวหน้าการอบรม (Training Progress) ได้ถูกพล็อตไว้ดังตัวอย่างรูปที่ 5.4 แสดงตัวอย่างความก้าวหน้าการอบรมที่ได้จากโครงข่ายแบบ VGG19 และ InceptionResNetV2 ตามลำดับ เมื่อโครงข่ายได้รับการอบรมแล้วนั้นจะนำไปทำนายวินิจฉัยกลุ่มข้อมูลทดสอบ (Testing Dataset) และท้ายที่สุดผลลัพธ์การทำนายที่ได้รับการวัดผลความแม่นยำและคุณภาพของการทดลองวินิจฉัย ผลลัพธ์การทดลองระบุว่า VGG19 คือเครื่องมือจัดหมวดหมู่ที่ดีสุดสำหรับการจัดหมวดหมู่กลุ่มข้อมูลซึ่งไม่สมดุลกัน โดยได้รับความแม่นยำ 95.70 เปอร์เซ็นต์ ความไว 99.22 เปอร์เซ็นต์ ความเฉพาะเจาะจง 87.83 เปอร์เซ็นต์ และพื้นที่ใต้กราฟ (AUC) 0.9978 ตามลำดับ โดยใช้เวลาอบรมกลุ่มข้อมูล 1,484 รูปภาพ ทั้งหมด 8 นาที 43 วินาที และเวลาทำนายผลกลุ่มข้อมูลทดสอบ 372 รูปภาพ เท่ากับ 4.764 วินาที ตัวอย่างของกราฟเส้นโค้ง ROC ที่ถูกพล็อตซึ่งได้รับจากความสัมพันธ์ของ True Positive Rate และ False Positive Rate ซึ่งผลลัพธ์ได้จากวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน (Imbalanced Dataset) ถูกแสดงดังรูปที่ 5.5

ในส่วนพารามิเตอร์ของความไว ความเฉพาะเจาะจง ความแม่นยำ และพื้นที่ใต้กราฟ สำหรับอุปกรณ์จัดหมวดหมู่ CNNs อื่นๆ ถูกระบุไว้ในตารางที่ 5.5 และผลลัพธ์ที่ได้จากโครงข่ายการเรียนรู้แบบถ่ายโอนแต่ละแบบได้แสดงเปรียบเทียบในรูปแบบแผนภูมิเพื่อความเข้าใจง่ายดังที่แสดงในรูปที่ 5.6 จากข้อมูลผลลัพธ์แสดงให้เห็นว่าค่าความเฉพาะเจาะจงยังต้องการการปรับปรุงให้ดีขึ้น และโมเดลที่ได้รับการอบรมแล้วนี้เมื่อนำไปทดสอบกับภาพสมองปกติความแม่นยำจะลดลงด้วยเช่นกัน ซึ่งสอดคล้องกับค่าความเฉพาะเจาะจงที่มีค่าไม่สูงนัก เมื่อเปรียบเทียบกับพารามิเตอร์อื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

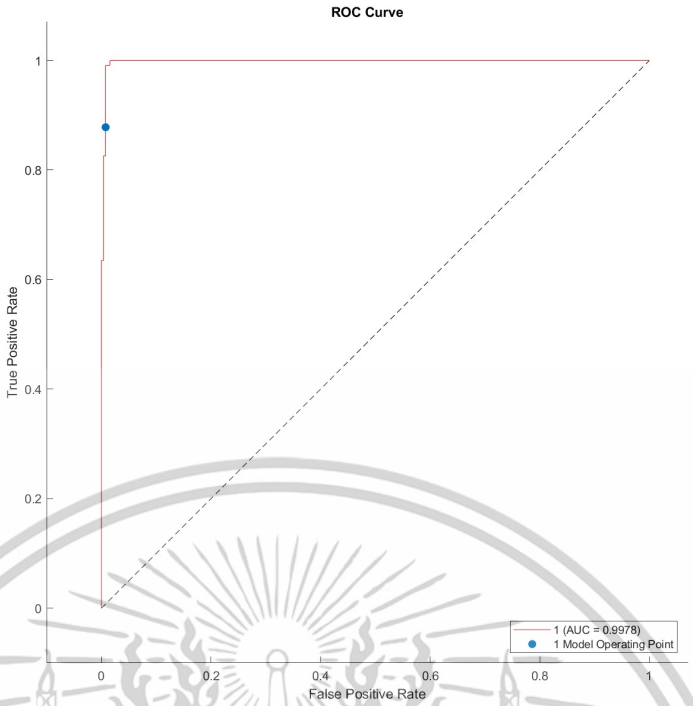


ก. ความก้าวหน้าการอบรมของ VGG19 ที่ใช้วิธีการทดลองที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน

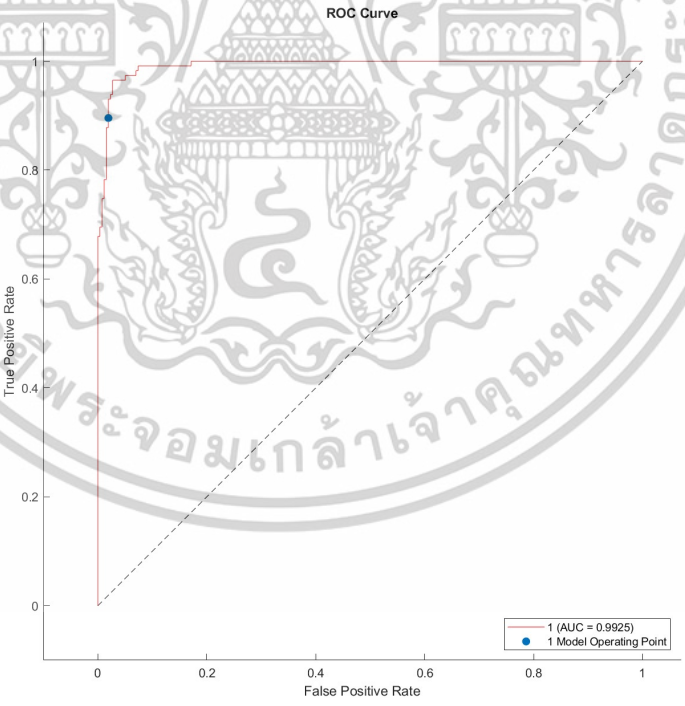
ข. ความก้าวหน้าการอบรมของ InceptionResNetV2 ที่ใช้วิธีการทดลองที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน

รูปที่ 5.4 ตัวอย่างความก้าวหน้าการอบรมที่พล็อตไว้สำหรับวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. กราฟเส้นโค้ง ROC ที่ถูกพล็อตสำหรับ VGG19 ที่ใช้วิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน



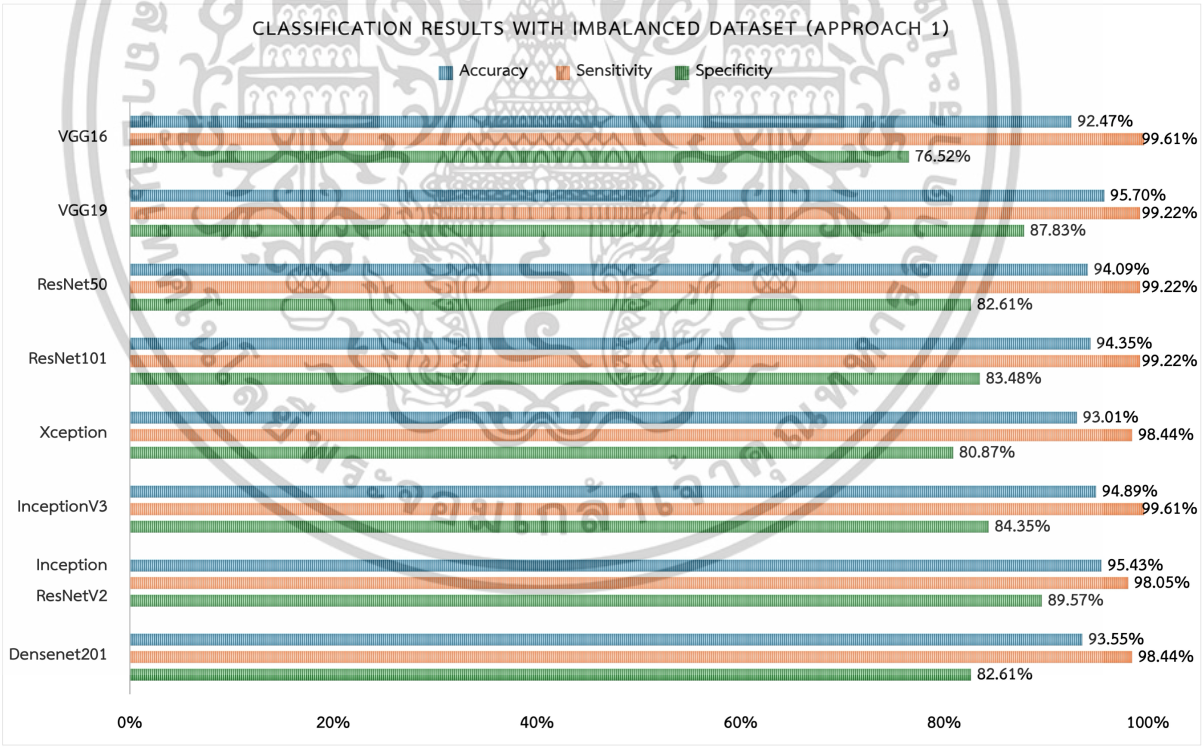
ข. กราฟเส้นโค้ง ROC ที่ถูกพล็อตสำหรับ InceptionResNetV2 ที่ใช้วิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน

รูปที่ 5.5 ตัวอย่างกราฟเส้นโค้ง ROC ที่ถูกพล็อตไว้สำหรับวิธีการที่ 1 กลุ่มข้อมูลซึ่งไม่สมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.5 ผลลัพธ์สมรรถนะการจัดหมวดหมู่ของกระบวนวิธีการที่ 1 กลุ่มข้อมูลที่ไม่สมดุล

CNNs	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
VGG16	92.47	99.61	76.52	0.9961
VGG19	**95.70**	**99.22**	**87.83**	**0.9978**
ResNet50	94.09	99.22	82.61	0.9957
ResNet101	94.35	99.22	83.48	0.9948
Xception	93.01	98.44	80.87	0.9774
InceptionV3	94.89	99.61	84.35	0.9973
InceptionResNetV2	95.43	98.05	89.57	0.9925
Densenet201	93.55	98.44	82.61	0.9912



รูปที่ 5.6 แผนภูมิรูปภาพเปรียบเทียบผลลัพธ์การทำนายเมื่อใช้กับกลุ่มข้อมูลซึ่งไม่สมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

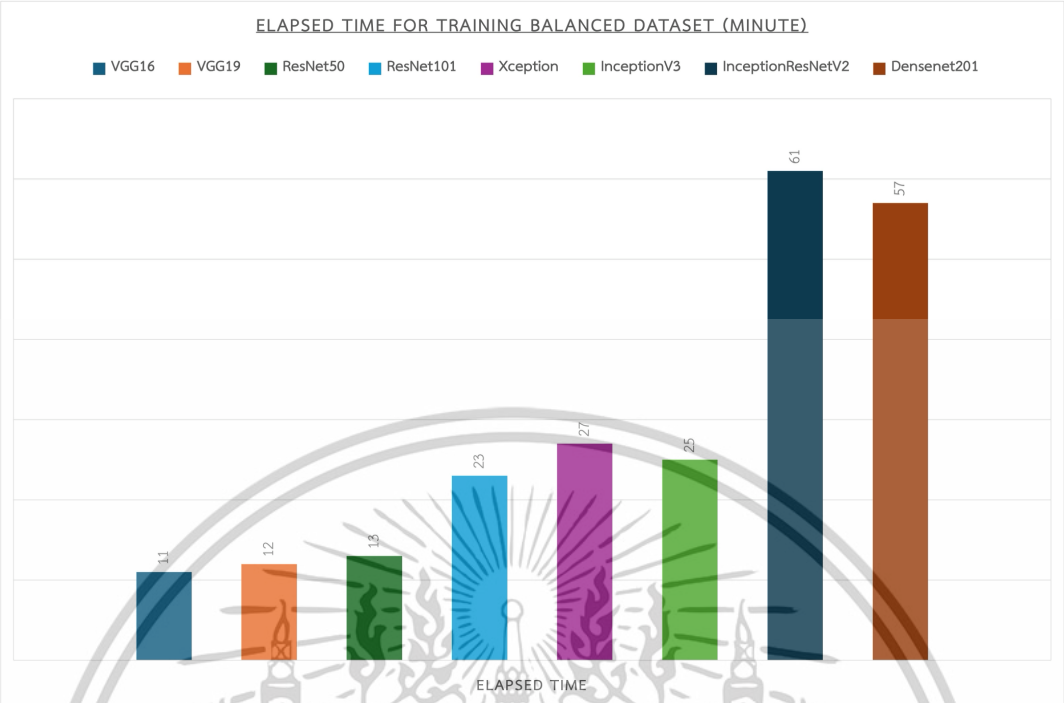
5.2.2 ผลการทดลองวิธีการที่ 2 กลุ่มข้อมูลที่ซึ่งสมดุลกัน

ในส่วนกระบวนการวิธีที่ 2 นั้น ความแม่นยำและคุณภาพของการทดสอบวินิจฉัยได้รับการปรับปรุงให้ดีขึ้นอย่างมีนัยสำคัญโดยใช้หลักการเสริมภาพเพื่อเพิ่มจำนวนของรูปภาพในคลาสส่วนน้อย ตามที่กล่าวมาเบื้องต้นก่อนหน้านี้ในหัวข้อ 4.2.2 กระบวนวิธีการเสริมภาพถูกนำมาใช้เพื่อจัดการปัญหาความไม่สมดุลของกลุ่มข้อมูลและสร้างสมดุลให้แก่กลุ่มข้อมูลอบรม ผลลัพธ์การทดลองแสดงให้เห็นว่าทุกค่าพารามิเตอร์ของสมรรถะได้รับการปรับปรุงอย่างมีนัยสำคัญจนเกือบสมบูรณ์แบบสำหรับการจัดหมวดหมู่แบบไบนารีของเนื้องอกสมองกลิโอมา ได้นำมาใช้งานจริงของ CNNs หลากหลายแบบเพื่ออบรมกลุ่มข้อมูลพร้อมทั้งใช้อัลกอริทึมออพชันดังที่กล่าวมาแล้วก่อนหน้านี้ เพื่อให้ได้รับโมเดลที่ได้ผ่านการอบรมแล้ว ความก้าวหน้าการอบรมของอัลกอริทึมแต่ละตัวถูกพล็อตไว้ความก้าวหน้าของการอบรมระบุผลลัพธ์ของ Validation Accuracy และเวลาการอบรม ตัวอย่างเช่น InceptionV3 ได้รับผลลัพธ์ Validation 99.51 เปอร์เซ็นต์ ในขณะที่ InceptionResNetV2 ได้รับผลลัพธ์ 99.76 เปอร์เซ็นต์ ในส่วนอัลกอริทึมอื่นผลลัพธ์ Validation Accuracy และเวลาที่ใช้อบรมและทดสอบโครงข่ายแสดงดังตารางที่ 5.6 รูปภาพเปรียบเทียบเวลาที่ใช้ในการอบรมสำหรับโครงข่ายแต่ละแบบแสดงดังรูปที่ 5.7 จะเห็นได้ว่า InceptionResNetV2 และ Densenet201 ใช้เวลาอบรมโครงข่ายมากกว่าแบบอื่น ๆ เนื่องจากจำนวนพารามิเตอร์และสถาปัตยกรรมโครงข่ายที่แตกต่างกัน และทั้งนี้ในส่วนกราฟความก้าวหน้าการอบรม (Training Progress) สำหรับการเรียนรู้แบบถ่ายโอนทั้งหมด 8 แบบที่พล็อตไว้ถูกแสดงดังรูปที่ 5.8 ก. ถึง 5.8 ฌ.

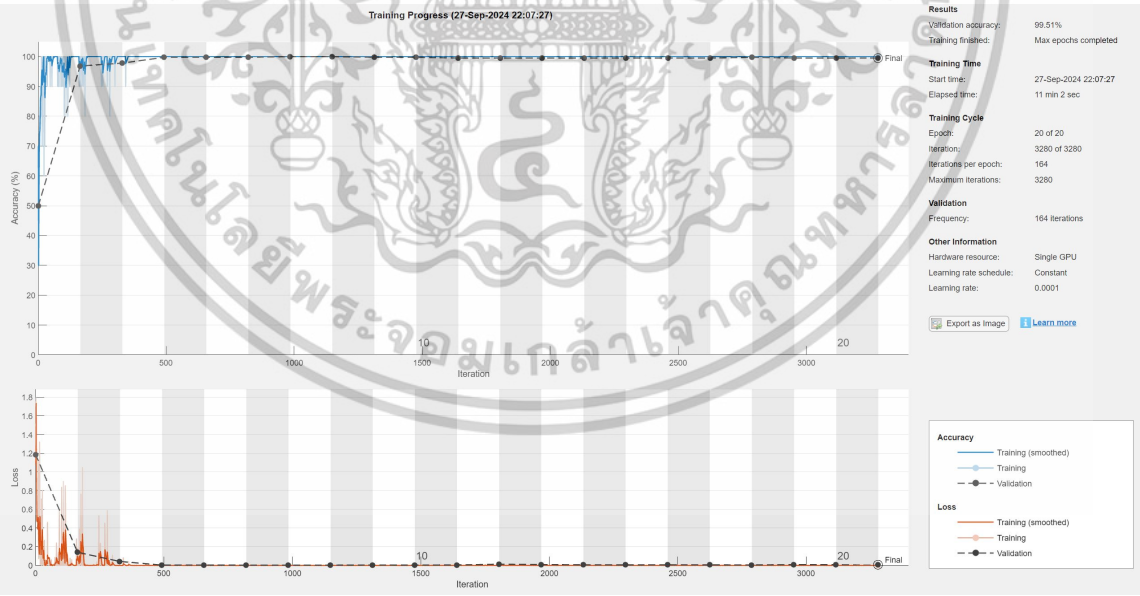
ตารางที่ 5.6 ผลลัพธ์การอบรมกลุ่มข้อมูลของกระบวนวิธีการที่ 2 กลุ่มข้อมูลที่สมดุลกัน

CNNs Transfer Learning	Validation Accuracy (%)	Training Elapsed Time	Testing Inference Time (sec.)
VGG16	99.51	11 min. 2 sec.	5.955
VGG19	98.79	12 min. 7 sec.	4.884
ResNet50	99.03	12 min. 50 sec.	5.412
ResNet101	99.03	23 min. 21 sec.	6.62
Xception	96.60	27 min. 19 sec.	6.625
InceptionV3	99.51	25 min. 15 sec.	5.878
InceptionResNetV2	99.76	61 min. 3 sec.	9.315
Densenet201	99.76	57 min. 25 sec.	9.413

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

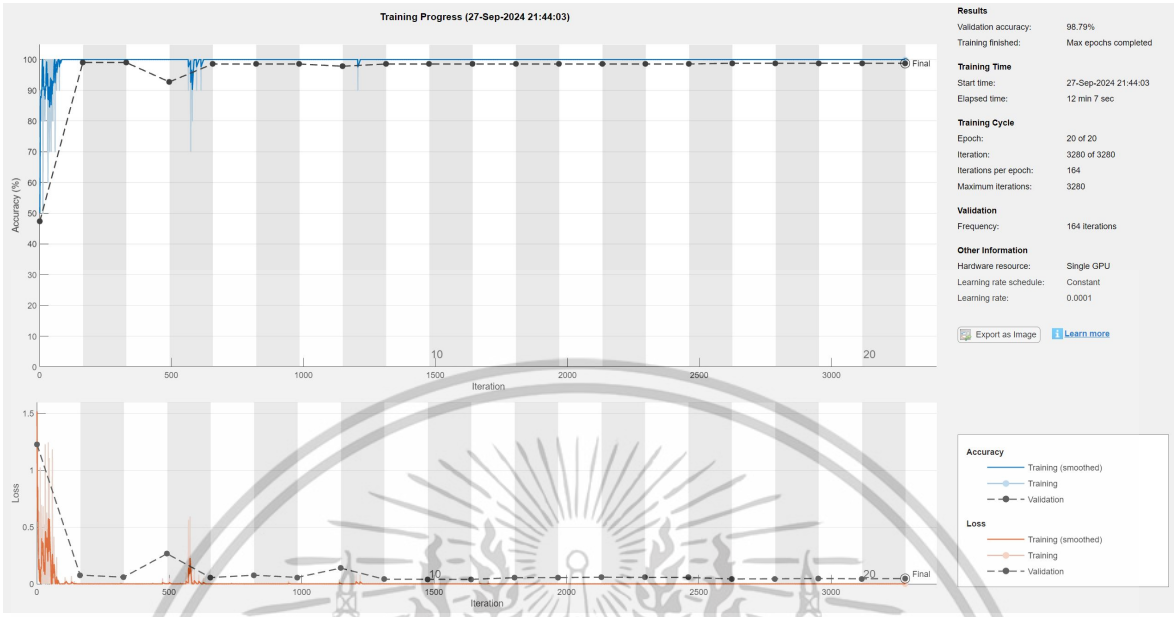


รูปที่ 5.7 เปรียบเทียบเวลาที่ใช้ในการอบรมซึ่งใช้กลุ่มข้อมูลที่สมดุลกัน

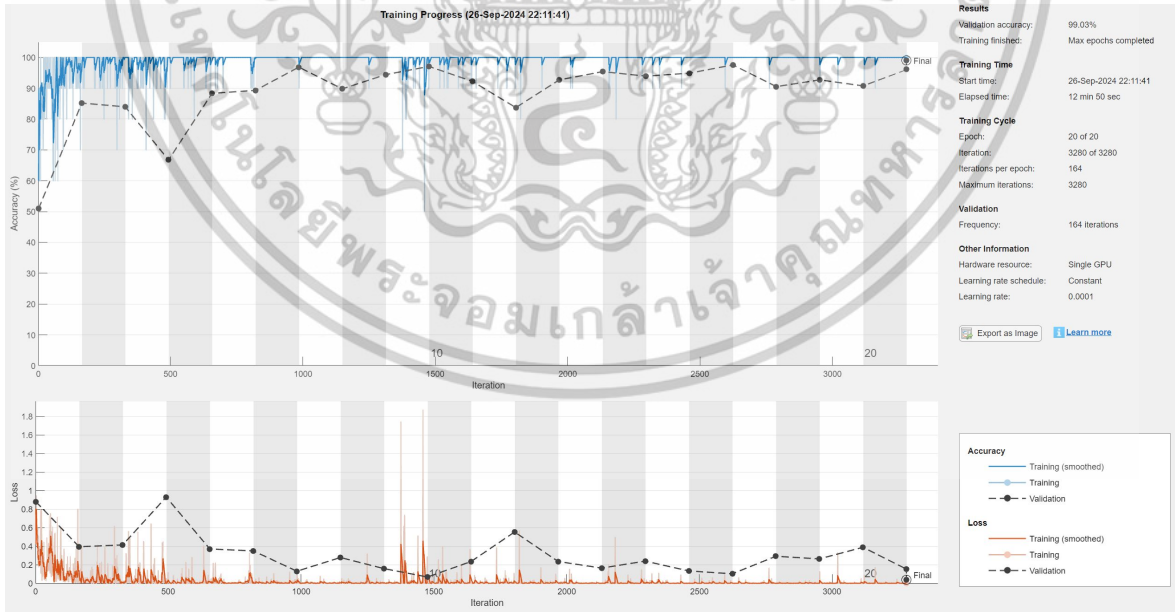


5.8 ก. ความก้าวหน้าการอบรมของ VGG16 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

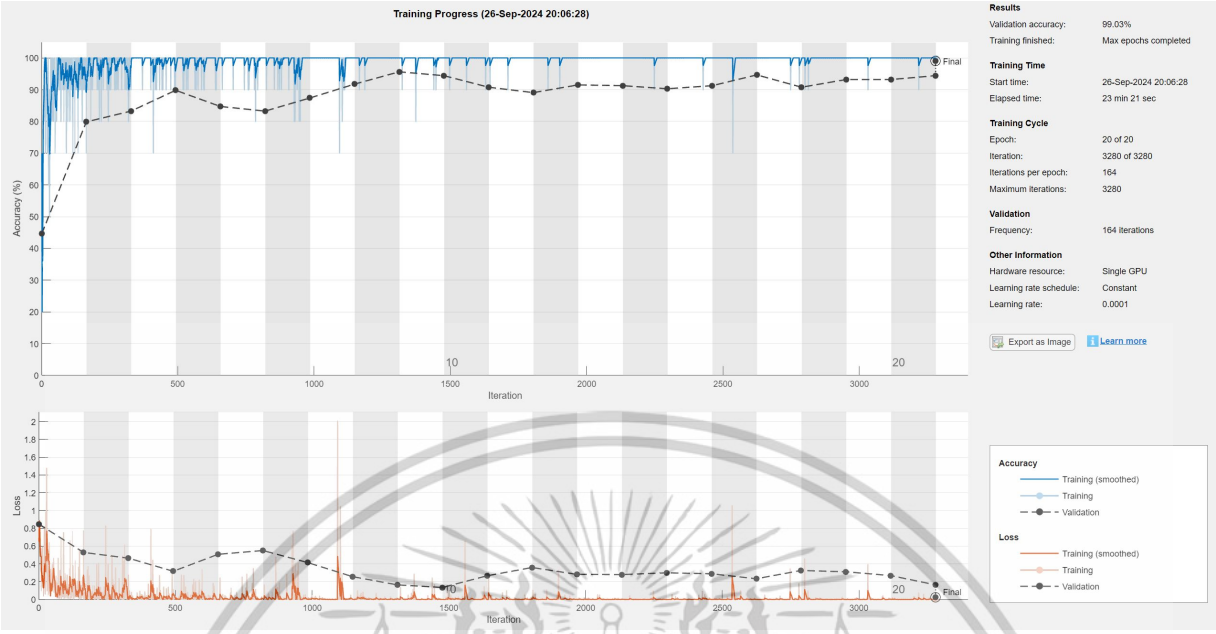


5.8 ข. ความก้าวหน้าการอบรมของ VGG19 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

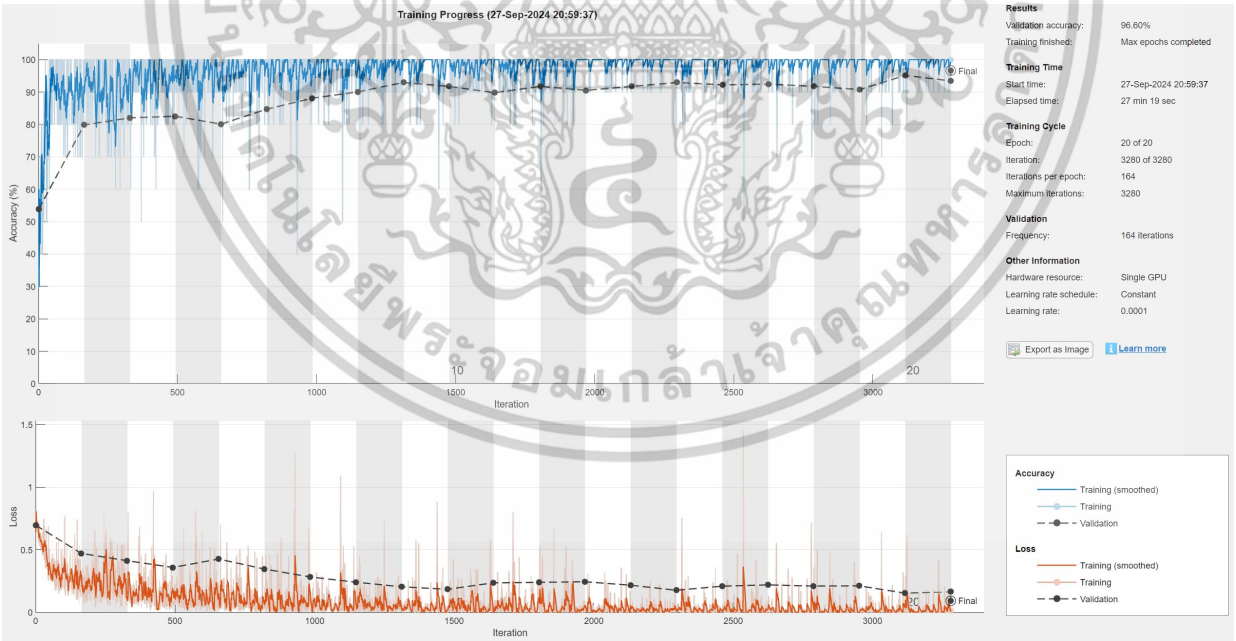


5.8 ค. ความก้าวหน้าการอบรมของ ResNet50 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

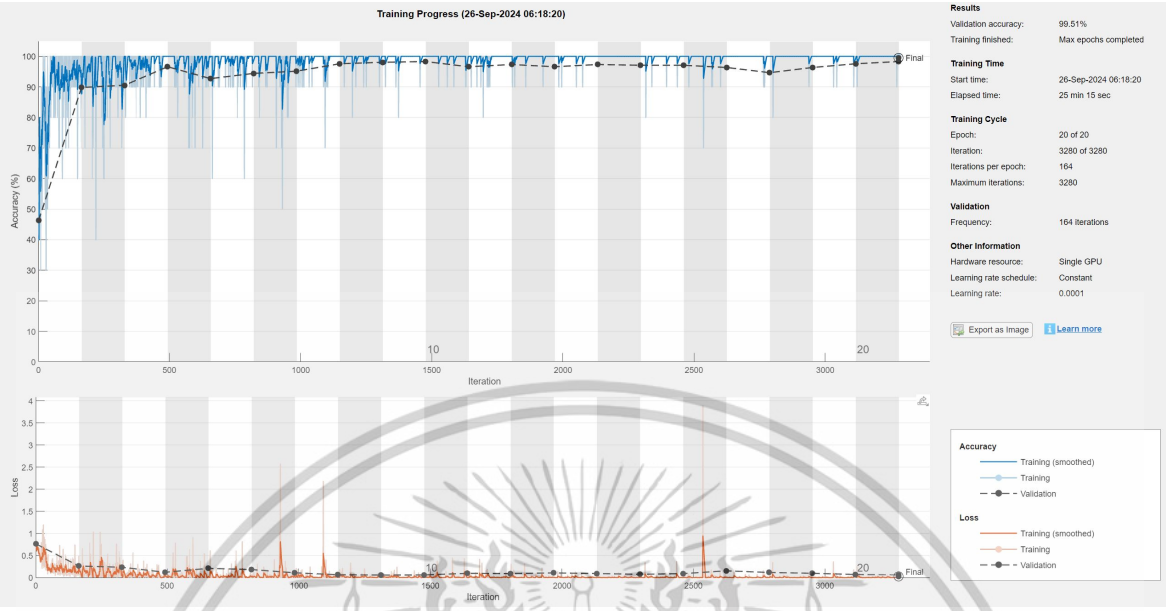


5.8 ง. ความก้าวหน้าการอบรมของ ResNet101 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

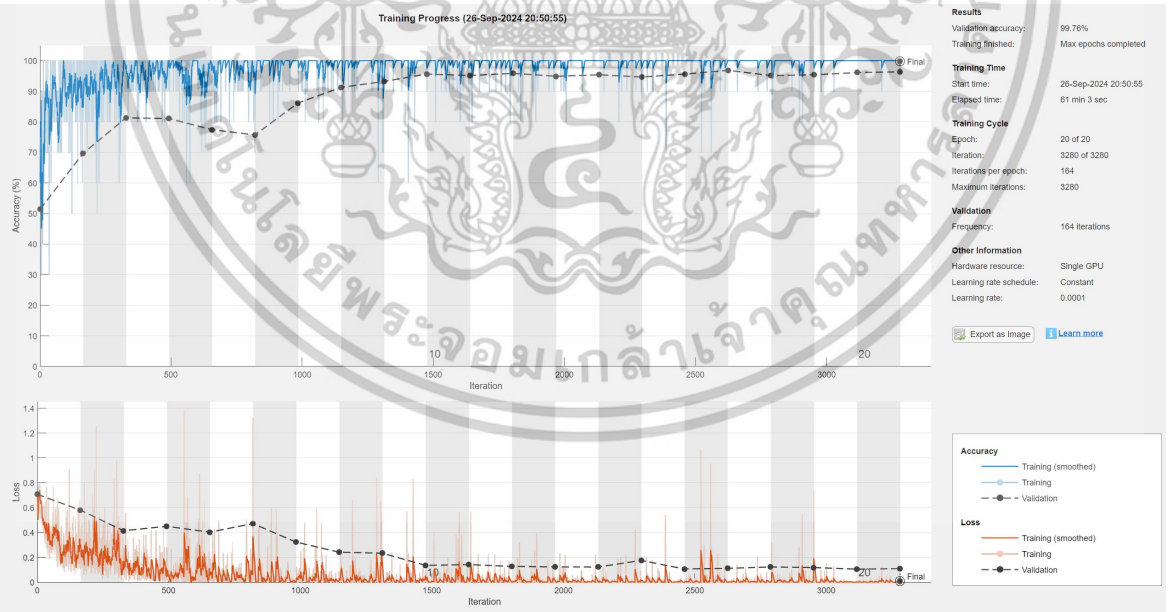


5.8 จ. ความก้าวหน้าการอบรมของ Xception ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

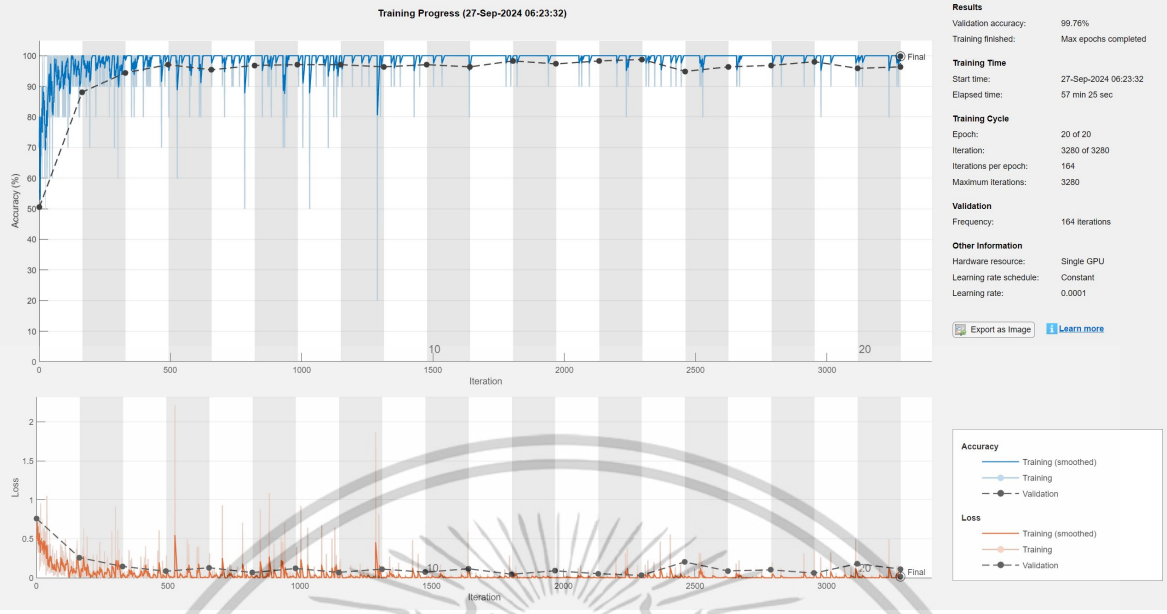


5.8 ฉ. ความก้าวหน้าการอบรมของ InceptionV3 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน



5.8 ช. ความก้าวหน้าการอบรมของ InceptionResNetV2 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



5.8 ฌ. ความก้าวหน้าการอบรมของ Densenet201 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกัน

รูปที่ 5.8 ความก้าวหน้าการอบรมสำหรับวิธีการที่ 2 กลุ่มข้อมูลซึ่งสมดุลกัน

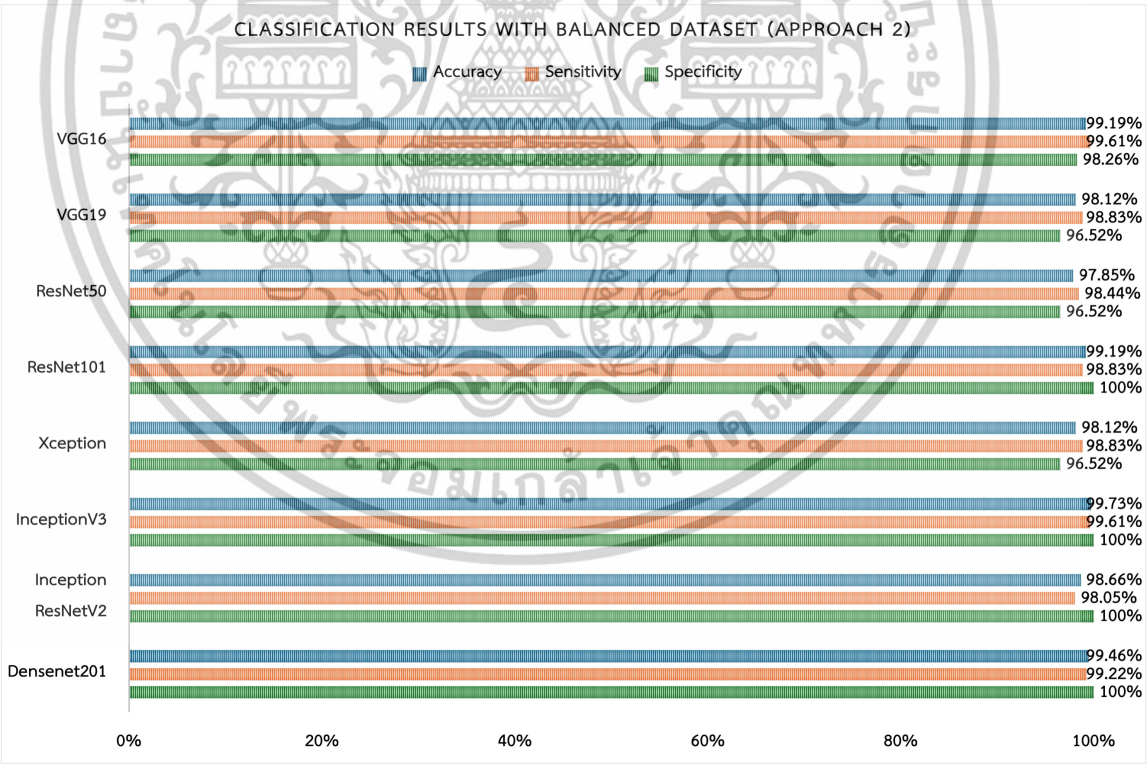
จากนั้นโมเดลที่ได้รับการอบรมถูกนำไปใช้เพื่อจัดหมวดหมู่เนื้องอกกลิโอมาและคอนฟิวชันชาร์ต (Confusion Chart) ได้ถูกพล็อตไว้เพื่ออ้างอิงความแม่นยำและพารามิเตอร์ในการประเมินผลอื่น และจากนั้นผลลัพธ์สมรรถนะได้รับการคำนวณซึ่งเครื่องมือจัดหมวดหมู่ CNNs ที่ได้รับผลลัพธ์ที่ดีสุดคือ InceptionV3 ซึ่งได้รับความแม่นยำเท่ากับ 99.73 เปอร์เซ็นต์ ความไวเท่ากับ 99.61 เปอร์เซ็นต์ ความเฉพาะเจาะจงเท่ากับ 100 เปอร์เซ็นต์ และพื้นที่ใต้กราฟ AUC เท่ากับ 1 ตามลำดับ โดยใช้เวลาอบรมกลุ่มข้อมูล 2,056 รูปภาพ ทั้งหมด 25 นาที 15 วินาที และเวลาทำนายผลกลุ่มข้อมูลทดสอบ 372 รูปภาพ เท่ากับ 5.878 วินาที และเครื่องมือจัดหมวดหมู่แบบ CNNs อื่นๆ แสดงผลลัพธ์การทดลองดังตารางที่ 5.7 พร้อมกันนี้ได้แสดงแผนภูมิเปรียบเทียบสมรรถนะการวินิจฉัยของแต่ละโครงข่ายดังที่แสดงในรูปที่ 5.9 ซึ่งจะเห็นได้ว่าทุกพารามิเตอร์ได้รับการปรับปรุงอย่างมีนัยสำคัญเมื่อใช้กลุ่มข้อมูลการอบรมที่สมดุลกัน โดยเฉพาะอย่างยิ่งค่าความเฉพาะเจาะจง ดังนั้นโมเดลที่ได้รับการอบรมแล้วนี้เมื่อนำไปทดสอบภาพเนื้องอกหรือสมองปกติผลลัพธ์จะแม่นยำขึ้นมาก

ซึ่งในที่นี้ได้แสดงภาพผลลัพธ์คอนฟิวชันชาร์ตที่ได้รับจากการทำนายโดยแต่ละโครงข่ายประสาทเทียม ดังแสดงในรูปที่ 5.10 เป็นตัวอย่าง Confusion Chart ของผลลัพธ์การทำนายได้จาก InceptionV3 และ Densenet201 ในส่วนความสัมพันธ์ระหว่าง True Positive Rate (TPR) และ False Positive Rate (FPR) ถูกนำไปใช้เพื่อพล็อตเส้นโค้ง ROC รูปที่ 5.11 แสดงตัวอย่างเส้นโค้ง ROC ที่ได้รับจากการอบรมโครงข่ายแบบ InceptionV3 และ Densenet201 ซึ่งจะนำไปหาค่า AUC ต่อไปดังที่กล่าวในสมการที่ (4.4) โดยค่า AUC ที่ได้เท่ากับ 1 และ 0.9999 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

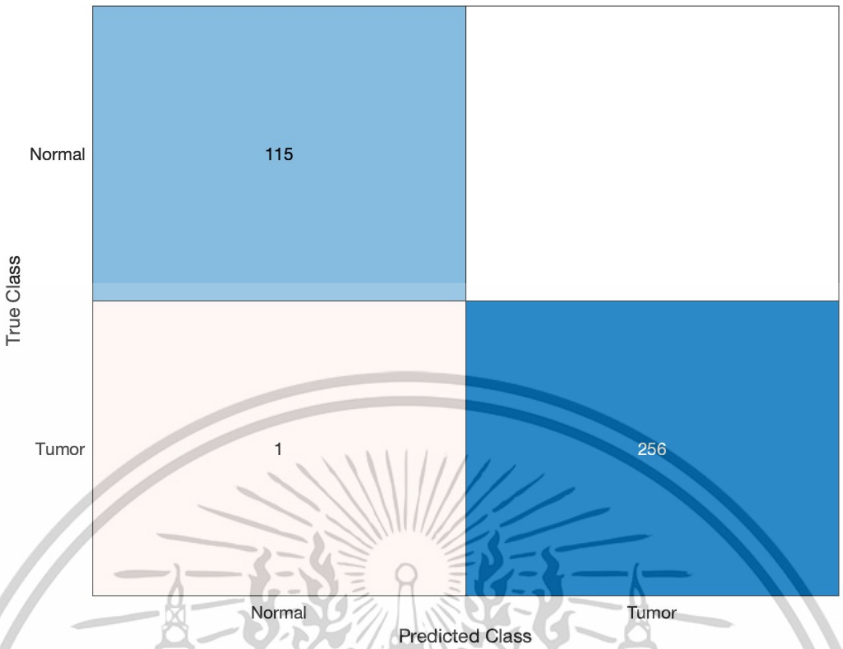
ตารางที่ 5.7 ผลลัพธ์สมรรถนะการจัดหมวดหมู่ด้วยวิธีการที่ 2 กลุ่มข้อมูลที่สมดุลกัน

CNNs	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
VGG16	99.19	99.61	98.26	0.9993
VGG19	98.12	98.83	96.52	0.9994
ResNet50	97.85	98.44	96.52	0.9982
ResNet101	99.19	98.83	100	0.9999
Xception	98.12	98.83	96.52	0.9980
InceptionV3	**99.73**	**99.61**	**100**	**1.000**
InceptionResNetV2	98.66	98.05	100	0.9989
Densenet201	99.46	99.22	100	0.9999

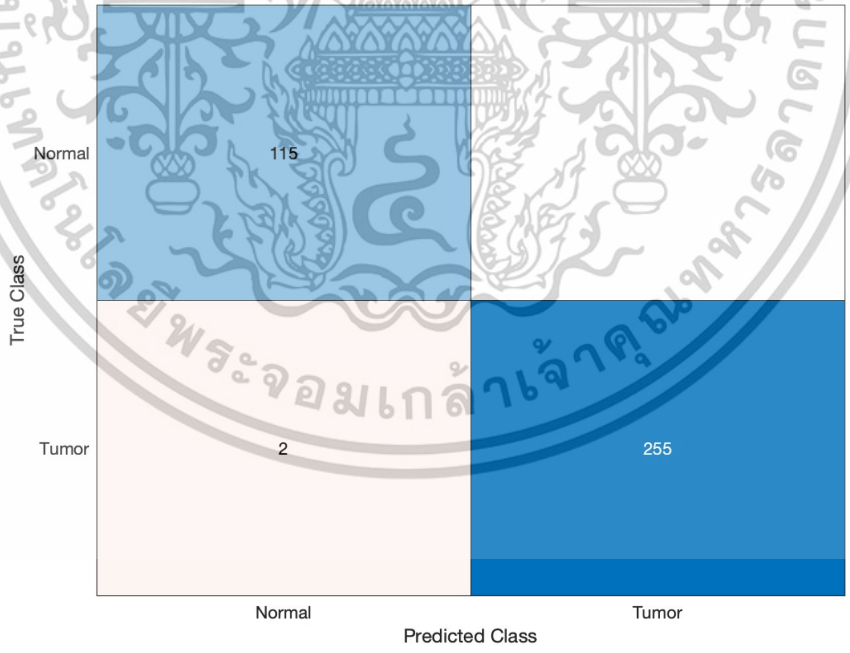


รูปที่ 5.9 แผนภูมิรูปภาพเปรียบเทียบผลลัพธ์การทำนายเมื่อใช้กับกลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



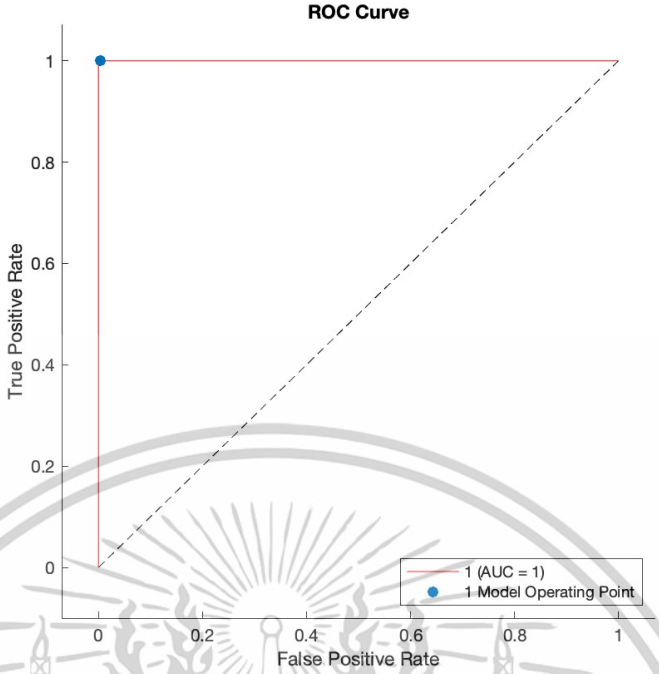
ก. Confusion Chart สำหรับ InceptionV3 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกันเพื่ออบรมโครงข่าย



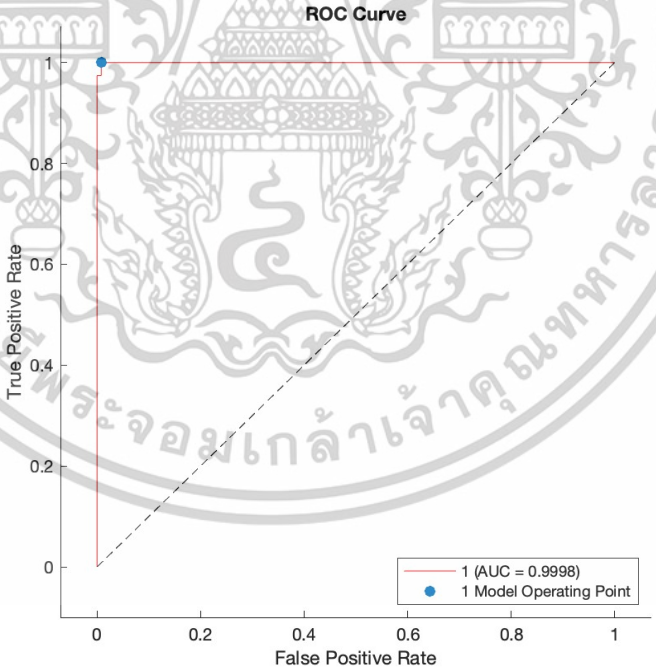
ข. Confusion Chart สำหรับ Densenet201 ที่ใช้กลุ่มข้อมูลซึ่งสมดุลกันเพื่ออบรมโครงข่าย

รูปที่ 5.10 Confusion Chart ของการทำนายผลโดยใช้กลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. กราฟเส้นโค้ง ROC สำหรับ InceptionV3 ที่ใช้วิธีการทดลองที่ 2 กลุ่มข้อมูลซึ่งสมดุลกัน



ข. กราฟเส้นโค้ง ROC สำหรับ Densenet201 ที่ใช้วิธีการทดลองที่ 2 กลุ่มข้อมูลซึ่งสมดุลกัน

รูปที่ 5.11 ตัวอย่างพล็อตกราฟเส้นโค้ง ROC สำหรับวิธีการที่ 2 กลุ่มข้อมูลซึ่งสมดุลกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2.3 ผลการทดลองวิธีการที่ 3 การเรียนรู้แบบ Ensemble Learning

เพื่อสืบค้นประสิทธิภาพของการจัดหมวดหมู่ การเรียนรู้แบบผสมผสาน (Ensemble Learning) ด้วยวิธีการโหวตเสียงส่วนมาก (Majority Voting) ทั้งแบบฮาร์ดและซอฟท์ ที่ซึ่งเป็นวิธีการแบบตรงไปตรงมาถูกนำมาใช้งาน ด้วยวิธีการนี้ไม่ต้องการการอบรมโครงข่ายใหม่อีกครั้ง อุปกรณ์จัดหมวดหมู่ที่มีประสิทธิภาพสูงสุดห้าอุปกรณ์ถูกใช้เพื่อการโหวตเสียงส่วนมาก ดังนั้น VGG16, InceptionV3, InceptionResNetV2, ResNet101 และ Densenet201 ซึ่งใช้กลุ่มข้อมูลที่สมดุลกัน จากวิธีการที่สองที่กล่าวมาถูกนำมาใช้ ท้ายที่สุดได้รับผลลัพธ์ประเมินผลตามตารางที่ 5.8

ตารางที่ 5.8 ผลลัพธ์สมรรถนะการจัดหมวดหมู่ด้วยวิธีการที่ 3 การเรียนรู้แบบผสมผสาน

Ensemble Learning	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Majority Hard Voting	99.46	99.22	100	n/a
Majority Soft Voting	99.46	99.22	100	n/a

5.2.4 ผลการวิเคราะห์เชิงสถิติการทำนายผลด้วย p-value

ถึงแม้ว่าได้นำเสนอการประเมินสมรรถนะการทำนายผลโดยใช้ Accuracy, Sensitivity และ Specificity นั้นเพียงพอแล้วที่จะวัดผลตัดสินใจและเปรียบเทียบแต่ละอุปกรณ์จัดหมวดหมู่ที่แตกต่างกัน มากกว่านั้นแล้ว ได้นำเสนอการวิเคราะห์เชิงปริมาณโดยใช้หลักสถิติที่เหมาะสมนั่นคือ p-value โดยค่านี้เป็นตัวแทนของความน่าจะเป็นของตัวอย่างที่ได้รับการสังเกตที่ทำให้สมมติฐานหลักเป็นจริง โดยสมมติฐานหลัก (Null Hypothesis, H_0) จะถูกกำจัด (Rejected) ถ้า p-value น้อยกว่าระดับนัยสำคัญ [50] ในงานวิจัยนี้ใช้การทดสอบทางด้านเดียวฝั่งซ้าย (Left-Tailed Test) เป็นการทดสอบสมมติฐาน โดยกำหนดให้ $H_0 : \theta \geq 99$ นั่นหมายความว่าความน่าจะเป็นของการทำนายผลต้องมากกว่าหรือเท่ากับ 0.99 รวมทั้งกำหนดให้ระดับนัยสำคัญสำหรับการทดสอบเท่ากับ 5 เปอร์เซ็นต์ ซึ่งนำไปสู่ค่าวิกฤติที่คำนวณได้เท่ากับ -1.649 ตามที่กล่าวรายละเอียดแล้วก่อนหน้านี้ในหัวข้อ 4.3.1 ผลลัพธ์ทดสอบทางสถิติสำหรับอุปกรณ์จัดหมวดหมู่ที่ได้รับสมรรถะดีสุดห้าอุปกรณ์นำเสนอได้ดังตารางที่ 5.9 ในส่วน Rejected H_0 นั้นถ้าระบุเป็นศูนย์ จะไม่สามารถ Reject สมมติฐานหลักได้ ผลลัพธ์ที่ได้แสดงให้เห็นว่าอุปกรณ์จัดหมวดหมู่ที่นำมาพิจารณาทั้งหมดมีค่า p-value มากกว่าระดับนัยสำคัญ ดังนั้นไม่สามารถ Reject สมมติฐานหลักได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

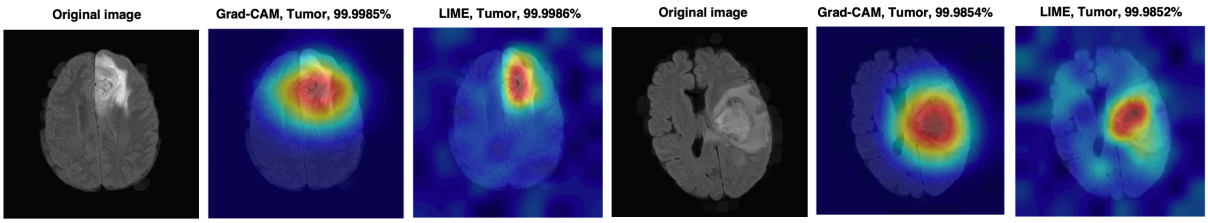
ตารางที่ 5.9 การวิเคราะห์เชิงสถิติสำหรับการทำนายโดยใช้ p-value

CNNs	Test Statistic	Confident Interval	p-value	Reject H_0
VGG16	0.0068	0.9972	0.5027	0
ResNet101	-0.1454	0.9947	0.4422	0
InceptionV3	-0.0018	0.9943	0.4993	0
InceptionResNetV2	-1.5617	0.9905	0.0596	0
Densenet201	0.2697	0.9966	0.6062	0

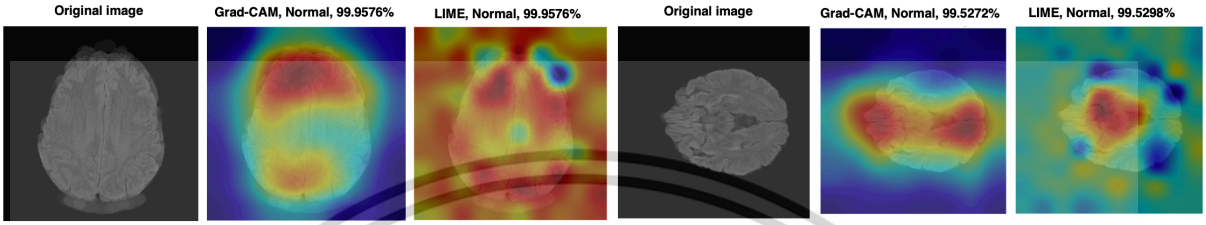
5.2.5 ผลการทดลองการมองเห็นพื้นที่รูปภาพซึ่งถูกแยกแยะ

ท้ายที่สุด Grad-CAM สำหรับการทำให้มองเห็นพื้นที่รูปภาพซึ่งถูกแยกแยะ (Discriminative Image Region) และการแปลความหมายถูกดำเนินการเพื่อให้เข้าใจว่าส่วนไหนของรูปภาพที่มีส่วนต่อการตัดสินใจทำนายผลและจัดหมวดหมู่มากที่สุด แผนที่ความร้อนแสดงให้เห็นสีแดงสำหรับพื้นที่ที่สัมพันธ์กับส่วนเนื้องอกที่สุด การมองเห็นได้ของพื้นที่ที่ระบุสำหรับความโปร่งใสและสะท้อนให้เห็นว่าการทำนายโดยเครื่องจัดหมวดหมู่นั้นมีความน่าเชื่อถือและยอมรับได้มากน้อยเพียงใด ในส่วนนี้ได้นำเสนอการออกแบบโปรแกรมที่ปรับปรุงขึ้นเพื่อนำมาใช้งานจริงของอัลกอริทึม Grad-CAM เพื่อมองเห็นรูปภาพแต่ละภาพอย่างเฉพาะเจาะจงสำหรับการมอนิเตอร์และตัดสินใจเพิ่มเติมสำหรับผลลัพธ์การทำนายตามที่อธิบายในบล็อกไดอะแกรมรูปที่ 4.11 และโค้ดเทียมในตารางที่ 4.5 อัลกอริทึม Grad-CAM โดยใช้ InceptionV3 ร่วมกับกระบวนวิธีการกลุ่มข้อมูลสมดุลกันถูกดำเนินการเพื่อทำให้มองเห็นพื้นที่รูปภาพซึ่งถูกแยกแยะของสมองปกติและเนื้องอก นอกเหนือจากนั้นแล้วได้วิเคราะห์เพิ่มเติมสำหรับการมองเห็นและแปลความหมายการทำนายโดยใช้อัลกอริทึม Local Interpretable Model-agnostic Explanation (LIME) ร่วมกับ InceptionV3 ด้วยกระบวนวิธีการนี้ทำให้แน่ใจว่าอุปกรณ์จัดหมวดหมู่ระบุตำแหน่งเนื้องอกสมองได้อย่างแม่นยำ ดังที่แสดงในรูป 5.12 ขั้นตอนกระบวนการนี้ทำได้เหมือนกับอัลกอริทึม Grad-CAM ที่ได้อธิบายก่อนหน้านี้ แต่ทว่าแผนที่ความร้อนถูกสร้างขึ้นโดยใช้อัลกอริทึม LIME แทน [51] เพื่อเปรียบเทียบให้เห็นพื้นที่รูปภาพที่มีผลต่อการทำนายของอัลกอริทึมได้ชัดเจนยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ก. พื้นที่รูปภาพซึ่งถูกแยกแยะของเนื้องอกสมองกลิโอมา



ข. พื้นที่รูปภาพซึ่งถูกแยกแยะของสมองปกติ

ก. รูปภาพเนื้องอกสมองซึ่งถูกแยกแยะ ข. รูปภาพสมองปกติซึ่งได้รับการทำนายและแยกแยะพื้นที่

รูปที่ 5.12 อัลกอริทึม Grad-CAM และ LIME โดยใช้ InceptionV3 กับกลุ่มข้อมูลสมดุลกัน

5.2.6 ผลลัพธ์การทดลองเปรียบเทียบกับงานเกี่ยวข้องที่ผ่านมา

ผลลัพธ์การทดลองถูกประเมินผลและเปรียบเทียบกับงานที่เกี่ยวข้องในลักษณะเดียวกันที่ผ่านมาและพบว่ากระบวนวิธีการทดลองที่นำเสนอได้รับผลลัพธ์การทดลองที่ดีมากดังแสดงในตารางที่ 5.10 ในงานที่ผ่านมานักวิจัยหลายท่านรายงานความแม่นยำที่มีสมรรถนะสูง เช่น Zaw และคณะ นำเสนอความแม่นยำของการจัดหมวดหมู่ถึง 94 เปอร์เซ็นต์ และจุดเด่นสำคัญของการศึกษานี้คือ อัลกอริทึมสามารถตรวจจับเนื้องอกได้ทุกตำแหน่งพื้นที่ โดยเฉพาะอย่างยิ่งเนื้องอกที่อยู่ระนาบเดียวกับสายตาซึ่งบ่อยครั้งถูกระบุเป็นดวงตาซึ่งไม่ถูกต้อง [21] Ahmed Wasif Reza และคณะ นำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบหลายคลาสโดยใช้ VGG16 ที่ถูกดัดแปลงและแสดงผลความแม่นยำ 99.5 เปอร์เซ็นต์ [33] S. Asif และคณะ นำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบไบนารี และเครื่องมือจัดหมวดหมู่ที่ดีสุดคือ Xception ที่ใช้งานร่วมกับ ADAM optimizer ด้วยผลลัพธ์ความแม่นยำ 99.67 เปอร์เซ็นต์ ความไว 99.68 เปอร์เซ็นต์ และความเฉพาะเจาะจง 99.66 เปอร์เซ็นต์ ตามลำดับ [36] S. Mohsen และคณะ นำเสนอการจัดหมวดหมู่เนื้องอกสมองแบบไบนารีโดยใช้ VGG19 และ ResNext101_32x8d ที่ได้รับการดัดแปลง ซึ่งได้รับความแม่นยำ 99.89 เปอร์เซ็นต์ และ 100 เปอร์เซ็นต์ ตามลำดับ [37] ถึงแม้ว่างานที่เกี่ยวข้องที่ผ่านมาหลากหลายงานรายงานการจัดหมวดหมู่เนื้องอกสมองที่ได้ผลลัพธ์ความแม่นยำสูง อย่างไรก็ตาม ความน่าเชื่อถือและไว้วางใจได้ของเครื่องจัดหมวดหมู่ถือเป็นสิ่งสำคัญอย่างยิ่งยวด รวมทั้งปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งจนเกินไปของเครื่องมือจัดหมวดหมู่เป็นอีกข้อจำกัดที่ต้องได้รับการปรับปรุงแก้ไข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นกระบวนวิธีการที่นำเสนอถูกนำมาใช้งานจริงกับกลุ่มข้อมูลสองกลุ่มที่แตกต่างกัน รวมทั้งจำนวนของรูปภาพที่เพิ่มจำนวนขึ้นเป็นการป้องกันปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งเกินไป กระบวนวิธีการที่นำเสนอได้รับผลลัพธ์สมรรถะสูง InceptionV3 ซึ่งใช้งานกับกลุ่มข้อมูลสมดุลกันคือเครื่องมือจัดหมวดหมู่ที่ดีสุดด้วยผลลัพธ์ความแม่นยำ (Accuracy) 99.73 เปอร์เซ็นต์ ความไว (Sensitivity) 99.61 เปอร์เซ็นต์ ความเฉพาะเจาะจง (Specificity) 100 เปอร์เซ็นต์ และพื้นที่ใต้กราฟ (AUC) เท่ากับ 1 ตามลำดับ นอกเหนือจากนั้นได้นำมาใช้งานจริงของอัลกอริทึม Grad-CAM สำหรับการมองเห็นของพื้นที่รูปภาพซึ่งถูกแยกแยะที่แสดงให้เห็นความน่าเชื่อถือได้ (Trustworthiness) ของเครื่องมือจัดหมวดหมู่อีกด้วย

ตารางที่ 5.10 การเปรียบเทียบผลลัพธ์การทดลองกับงานในลักษณะเดียวกันที่ผ่านมา

Authors	Classification Methods	Dataset	Accuracy
Zaw et al. [21]	Naïve Bayes classifier	REMBRANDT Tumor 90 Normal 24	Accuracy 94% Sen. 81.25% Spec. 100%
Hossain et al. [24]	Fuzzy C-means with SVM and CNN	BraTS dataset, Tumor 187 Normal 30	SVM, Accuracy 92.42% Sensitivity 98.3% Specificity 42.8% CNN, Accuracy 97.87%
Dipu et al. [25]	YOLOV5 and FastAi	BraTS2018, 1,992 images	FastAi, Accuracy 95.78% Precision 96.70% Recall 95.65% YOLOv5, Accuracy 85.95% Precision 92.67% Recall 84.33%
Arora & Sharma [27]	VGG16 and ResNet	BraTS2015, 3,764 images	SVM 50.51% VGG16 90.54% InceptionV3 63.94% ResNet 81.92%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.10(ต่อ) การเปรียบเทียบผลลัพธ์การทดลองกับงานในลักษณะเดียวกันที่ผ่านมา

Authors	Classification Methods	Dataset	Accuracy
Dubey et al. [28]	Neural Network classifier	Real MRI images from Hospital, Meningioma 20, Glioma 15	Accuracy 95%
Gaikwad et al. [29]	VGG-16 transfer learning	MRI on Kaggle, Tumor 155 Normal 98	Accuracy 99.03%
Kumar Baranwal et al. [30]	Multiclass classify using SVM and customized CNN	Nanfang and Tianjing Medical University Hospital, Meningioma : 708 Glioma : 1,426 Pituitary : 930	Concluded that CNN was the best classifier, Accuracy, Meningioma: 94.12% Glioma : 94.3% Pituitary : 98.85%
Hareem Kibriya et al. [31]	Multiclass using ResNet18 and GoogLeNet with SVM	Public available Figshare dataset, Meningioma : 708 Glioma : 1,426 Pituitary : 930	GoogLeNet+SVM Accuracy 97.6% Precision 97.3% Recall 97.3% ResNet18 + SVM Accuracy 98% Precision 98.3% Recall 98%
Aboussaleh et al. [32]	Customized CNN classifier	BraTS2017, High-Grade Glioma (HGG) 210 Low-Grade Glioma (LGG) 75 With data augmentation increased to 1100 labeled images, 550 for each label.	Accuracy for training (HGG+ LGG) 98%, accuracy for validation (HGG+ LGG) 91%, and segmentation with post processing (HGG+ LGG) obtained dice scores 82.35%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.10(ต่อ) การเปรียบเทียบผลลัพธ์การทดลองกับงานในลักษณะเดียวกันที่ผ่านมา

Authors	Classification Methods	Dataset	Accuracy
Ahmed Wasif Reza et al. [33]	Multiclass classify using modified VGG16	Public available Kaggle and Figshare Glioma : 2,547 Meningioma : 2,582 Pituitary : 2,658 No tumor : 2,396	Glioma : 98.98% Meningioma: 99.13% Pituitary : 99.95% No tumor : 99.81% Overall : 99.5%
Ayesha Younis et al. [34]	CNN, VGG16 and Ensemble	MRI images for Brain Tumor Detection Tumor 155 Normal 98	CNN 96% VGG16 98.5% Ensemble 98.14%
S. Asif et al. [36]	Xception with ADAM optimizer	Dataset (large) Tumor 1500 Normal 1500 Dataset (small) Tumor 155 Normal 98	MRI large dataset with Xception and ADAM optimizer was better than the other, Accuracy 99.67% Sen. 99.68% Spec. 99.66%
S. Mohsen et al. [37]	VGG19 ResNext101_32×8d	Brain Tumor dataset on Kaggle, Glioma : 900 Pituitary : 900	VGG19 Accuracy 99.98% Precision 99.90% Recall 99.89% ResNext101 Accuracy, Precision, and Recall 100%
Proposed method	Preprocessing techniques with various 8 CNNs. InceptionV3 with balanced dataset is the best classifier.	REMBRANDT, BraTS2021 Normal : 1028 Glioma : 1028	Acc 99.73% Sen 99.61% Spe 100% AUC 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6
สรุปผลการทดลอง ข้อเสนอแนะ และแนวทางพัฒนา

ในบทท้ายนี้กล่าวถึงสรุปผลการทดลองที่ได้รับจากแต่ละกระบวนวิธีการ ข้อดีและข้อเสียของแต่ละอุปกรณ์การจัดหมวดหมู่ซึ่งเหมาะกับงานจัดหมวดหมู่รูปภาพทางการแพทย์ต่างกันออกไป โดยขึ้นอยู่กับขนาดกลุ่มข้อมูลอบรมและทดสอบ ทรัพยากรคอมพิวเตอร์ในการประมวลผล รวมทั้งกล่าวถึงข้อเสนอแนะและแนวทางพัฒนาในอนาคต เพื่อให้การวินิจฉัยเนื้องอกในสมองด้วยระบบคอมพิวเตอร์วิชันได้รับผลลัพธ์ที่ดีและเหมาะสมในการนำไปใช้งานจริงได้ดียิ่งขึ้น

6.1 สรุปผลการทดลอง

งานวิจัยนี้นำเสนอระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรค ในการจัดหมวดหมู่เนื้องอกสมองกลิโอมาชนิด GBM เกรด 4 ซึ่งร้ายแรงที่สุด โดยแบ่งเป็นสองหมวดหมู่หลักๆ คือการใช้กลุ่มข้อมูลขนาดไม่ใหญ่มากและใช้การเรียนรู้ของเครื่องจักรแบบซัพพอร์ตเวกเตอร์แมชชีนเป็นเครื่องมือในการจัดหมวดหมู่ ซึ่งการแบ่งออกเป็นส่วนและสกัดฟีเจอร์อาศัยการทำด้วยมือ (Handcrafted) และเหมาะกับกรณีที่มีทรัพยากรคอมพิวเตอร์ที่จำกัด ผลลัพธ์การทำนายความแม่นยำได้รับการปรับปรุงให้ดีขึ้น คือ 96.875 เปอร์เซ็นต์ ดีกว่างานที่ผ่านมาที่เกี่ยวข้องกัน จากนั้นเพื่อเปรียบเทียบกับวิธีการอื่น การเรียนรู้เชิงลึกแบบถ่ายโอนได้นำเสนอการใช้ GoogLeNet และ AlexNet เพื่อช่วยในการจัดหมวดหมู่ซึ่งฟีเจอร์แมพจะถูกสกัดออกมาและจัดหมวดหมู่เนื้องอกสมองอย่างอัตโนมัติ แต่วิธีการนี้ต้องอาศัยอุปกรณ์คอมพิวเตอร์ที่มีคุณสมบัติค่อนข้างสูงเนื่องจากการสกัดฟีเจอร์ต้องพิจารณาในแต่ละพิกเซลของภาพ ผลลัพธ์การทำนายที่ได้คือความแม่นยำโดยรวมเท่ากับ 80.85 เปอร์เซ็นต์ และ 93.62 เปอร์เซ็นต์ตามลำดับ

และเพื่อศึกษาพัฒนาและปรับปรุงให้ดีขึ้น ได้นำเสนอการจัดหมวดหมู่เนื้องอกสมองกลิโอมาแบบอัตโนมัติโดยใช้การเรียนรู้เชิงลึกแบบ CNNs ซึ่งใช้กับกลุ่มข้อมูลที่มากขึ้นในการอบรมเพื่อแก้ปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งเกินไป โดยรูปภาพ MRI จากฐานข้อมูลที่แตกต่างกันสองแหล่งถูกนำมารวมกันและใช้ประโยชน์เพื่องานจัดหมวดหมู่แบบไบนารีของเนื้องอกสมองชนิด GBM เกรด 4 ขั้นตอนเตรียมกระบวนการรูปภาพด้วยกระบวนการสัณฐานวิทยา (Morphological Operation) สำหรับแยกแยะส่วนกะโหลกและสิ่งแปลกปลอมออกจากพื้นที่สมองเพื่อได้รับภาพส่วนพื้นที่สมองเท่านั้น จากนั้นนำภาพที่ได้เพื่อปรับคอนทราสให้สมดุลและทำเป็นมาตรฐานเดียวกันโดยใช้หลักการ Contrast Limited Adaptive Histogram Equalization และกลุ่มข้อมูลรูปภาพที่ได้รับเกิดความไม่สมดุลอย่างมากระหว่างสองคลาสคือเนื้องอกและสมองปกติซึ่งเป็นคลาสส่วนน้อย ดังนั้นเพื่อจัดการกับปัญหานี้ได้ตั้งค่าแบบจำลองการทดสอบดังต่อไปนี้ แบบจำลองที่หนึ่งใช้กลุ่มข้อมูลที่ไม่สมดุลเพื่อใช้อบรมโดย CNNs หลากหลายแบบ และแบบจำลองที่สองใช้หลักการเสริมรูปภาพ (Image Augmentation) เพื่อเพิ่มจำนวนรูปภาพคลาสที่เป็นส่วนน้อย (Minority Class) ซึ่งในที่นี้คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมองปกติ เพื่อทำให้กลุ่มข้อมูลเกิดความสมดุลกันขึ้น และนำไปอบรมโดยใช้ CNNs หลากหลายแบบ ซึ่งประกอบด้วยการเรียนรู้แบบถ่ายโอน 8 แบบถูกนำมาใช้เพื่อจัดหมวดหมู่เนื้องอกสมองแบบไบนารี ผลลัพธ์การทดลองที่ได้รับแสดงให้เห็นว่าการจัดการทำให้กลุ่มข้อมูลสมดุลกันทำให้ความแม่นยำได้รับการปรับปรุงให้ดีขึ้นอย่างมีนัยสำคัญ สมรรถนะและคุณภาพของผลลัพธ์การทำนายได้รับการประเมิน และพบว่าดีกว่างานวิจัยที่ผ่านมา โดยเครื่องมือจัดหมวดหมู่ที่ดีสุดคือ InceptionV3 ซึ่งใช้กับกลุ่มข้อมูลสมดุลกัน และได้รับความแม่นยำ 99.73 เปอร์เซ็นต์ ความไว 99.61 เปอร์เซ็นต์ ความเฉพาะเจาะจง 100 เปอร์เซ็นต์ และพื้นที่ใต้กราฟ AUC เท่ากับ 1 ตามลำดับ โดยใช้เวลาอบรมกลุ่มข้อมูล 2,056 รูปภาพ ทั้งหมด 25 นาที 15 วินาที และเวลาทำนายผลกลุ่มข้อมูลทดสอบ 372 รูปภาพ เท่ากับ 5.878 วินาที และเพื่อการสืบค้นปรับปรุงสมรรถะการทำนายเพิ่มเติมได้นำเสนอหลักการเรียนรู้แบบผสมผสาน (Ensemble Learning) บนหลักการโหวตเสียงส่วนมาก (Majority Hard and Soft Voting) และที่มากกว่านั้นอัลกอริทึม Grad-CAM และ LIME ได้นำมาใช้งานเพื่อการแสดงให้เห็นพื้นที่รูปภาพที่แยกแยะระหว่างก้อนเนื้องอกและพื้นที่ส่วนสมองเพื่อสร้างความน่าเชื่อถือได้ของผลลัพธ์การทำนายโดย CNNs วิธีการที่นำเสนอทำนายเนื้องอกสมองกลิโอมาได้อย่างแม่นยำ และในอนาคตข้างหน้าเครื่องมือที่มีประสิทธิภาพ และจำแนกเนื้องอกสมองได้ตั้งแต่เกรด 1 ถึง เกรด 4 จะถูกพัฒนาอย่างต่อเนื่อง รวมทั้งจะพิจารณาทดสอบผลกระทบต่อการปฏิบัติทางการแพทย์ เพื่อดูว่าผลจากงานวิจัยมีความหมายอย่างไรต่อการดูแลรักษาผู้ป่วยในทางปฏิบัติ เพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์ในการวินิจฉัยเนื้องอกสมองกลิโอมาโดยใช้การประมวลผลภาพต่อไป

6.2 ข้อเสนอแนะ และแนวทางพัฒนาในอนาคต

เพื่อพัฒนาระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคให้แมนยำและนำไปใช้งานได้จริง ดังนั้นกล่าวถึงข้อเสนอแนะเกี่ยวกับงานในอนาคตดังนี้

1. การเพิ่มจำนวนรูปภาพสำหรับกลุ่มข้อมูลอบรมและทดสอบเพื่อความทนทานของอัลกอริทึมและแก้ปัญหาความเหมาะสมกับสิ่งใดสิ่งหนึ่งเกินไป รวมทั้งปรับปรุงวิธีการเตรียมกระบวนการเพื่อลดขั้นตอนและเพิ่มประสิทธิภาพของผลลัพธ์การทดสอบ

2. การเลือกใช้อัลกอริทึมเพิ่มเติมสำหรับการอบรมโครงข่ายประสาทเทียมแบบอื่นนอกเหนือจาก Stochastic Gradient Descent with Momentum เช่น Root Mean Square Propagation (RMSProp) หรือ Adaptive Moment Estimation (Adam) รวมทั้งปรับค่า epochs ค่าอื่นๆ เพื่อดูผลลัพธ์ ซึ่งรวมถึงการปรับแต่งอย่างละเอียด (Fine Tuning) ของพารามิเตอร์อื่นๆ

3. การศึกษาเพิ่มเติมเกี่ยวกับการนำหลักการ Encoder-Decoder มาใช้จริง เช่น Semantic Segmentation โดยใช้ DeeplabV3+ เพื่อลดขั้นตอนการเตรียม Label และ Ground Truth Table ซึ่งเป็นการลดขั้นตอนการทำงานเตรียมฉลากให้ตรงกับรูปภาพจริง ซึ่งอาจจะเกิดความผิดพลาดจากมนุษย์ได้ (Human Error)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. เพื่อความสะดวกและทำงานได้ง่ายขึ้น ควรมีการพัฒนาระบบ Graphic User Interface (GUI) ในการทำนายผลลัพธ์จากรูปภาพ MRI ต่อไป

5. ในอนาคตควรศึกษาโมเดลรูปภาพทางการแพทย์แบบสามมิติ (Transformer-Based Models) ซึ่งจะช่วยสืบหาเพื่อปรับปรุงประสิทธิภาพในการวินิจฉัยเนื้องอกสมองให้ดียิ่งขึ้น

6. พัฒนาระบบคอมพิวเตอร์เพื่อช่วยวินิจฉัยโรคเนื้องอกสมองกลิโอมาที่สามารถตรวจจับเนื้องอกสมองตั้งแต่เกรด 1 ถึงเกรด 4 เพื่อให้วินิจฉัยโรคได้อย่างถูกต้องและรวดเร็วในระยะเริ่มต้นของเนื้องอกในสมอง และทำการรักษาอย่างทันทีเพื่อเพิ่มโอกาสรอดชีวิตของผู้ป่วย

7. พิจารณาทดสอบผลกระทบต่อการปฏิบัติทางการแพทย์ (Clinical Implications) เพื่อดูว่าผลจากงานวิจัยมีความหมายอย่างไรต่อการดูแลรักษาผู้ป่วยในทางปฏิบัติ และนำระบบที่พัฒนาขึ้นไปทดลองใช้งานจริงเพื่อช่วยเหลือเจ้าหน้าที่ทางการแพทย์วินิจฉัยโรคเนื้องอกสมองต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] S. Bauer, R. Wiest, L. P. Nolte, and M. Rayes, “A survey of MRI-based medical image analysis for brain tumor studies,” Phys. Med. Biol., vol. 58, no. 13, pp. 97–129, 2013.

[2] D. N. Louis, A Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, ‘‘The 2016 world health organization classification of tumors of the central nervous system: A summary,’’ Acta Neuropathol., vol. 131, no. 6, pp. 803–820, Jun. 2016.

[3] The Cancer Imaging Archive, 2017, REMBRANDT https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT

[4] U. Baid, et al., “The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification,” arXiv:2107.02314, 2021.

[5] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS),” IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 1993-2024, 2015, DOI: 10.1109/TMI.2014.2377694.

[6] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., “Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features, Nature Scientific Data,” 4:170117 (2017) DOI: 10.1038/sdata.2017.117.

[7] R. Gonzalez and R. Woods, Digital Image Processing, 4th ed., Pearson Education, 2018.

[8] C. Cortes and V. Vapnik, “Support-vector networks,” Machine learning 20, 3, 273-297, (1995).

[9] การประมวลผลภาพดิจิทัลด้วย MATLAB, รศ.ดร.ชูชาติ ปิณฑวิรุจน์

[10] R. Gonzalez, R. Woods and S. Eddins, Digital Image Processing using MATLAB, Pearson Education.

[11] S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B.T.H. Romeny, J.B. Zimmerman, and K. Zuiderveld, “Adaptive Histogram Equalization and Its Variations,” in Proc. Computer Vision, Graphics, And Image Processing, vol. 39, 1987, pp. 355-368.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[12] K. Zuiderveld, “Contrast limited adaptive histogram equalization,” In Graphics Gems IV; Academic Press Professional, Inc.: Cambridge, MA, USA, 1994; pp. 474–485.

[13] Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, The MIT press, 2017.

[14] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proc. 3rd Int. Conf. Learn. Represent., San Diego, CA, USA, May 2015.

[15] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

[16] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017, pp. 1251–1258.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016, pp. 2818–2826.

[18] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017, Volume 4, p. 12.

[19] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, “Learning Deep Features for Discriminative Localization. In Proceedings of the CVPR,” Las Vegas, NV, USA, 26 June–1 July 2016.

[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, “Grad-cam: Visual explanations from deep networks via gradient-based localization,” In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017, pp. 618–626.

[21] H. T. Zaw, N. Maneerat, and K. Y. Win, “Brain tumor detection based on Naïve Bayes Classification,” in Proc. 5th Int. Conf. on Engineering, Applied Sciences and Technology (ICEAST), Luang Prabang, Laos, 2019, pp. 1-4.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[22] P. Pilaoon, N. Maneerat, K. Yajima, R. Varakulsiripunth, and K. Homamoto, “Brain Tumor Classification using Supervised Support Vector Machine,” in Proc. 9th Int. Conf. on Engineering, Applied Sciences and Technology (ICEAST), Vientiane, Laos, 2023, pp. 79-83.

[23] P. Pilaoon, N. Maneerat, A. Nakthewan, R. Varakulsiripunth, and K. Hamamoto, “Brain Tumor Classification using Pretrained Deep Convolutional Neural Network,” in Proc. 9th Int. Conf. on Engineering, Applied Sciences and Technology (ICEAST), Vientiane, Laos, 2023, pp. 84-88.

[24] T. Hossain, F. Shishir , M. Ashraf, A. Nasim, and F. Shah, “Brain Tumor Detection Using Convolutional Neural Network,” in Proc. 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 (ICASERT 2019), IEEE.

[25] N.M. Dipu, S.A. Shohan, and K.M.A. Salam, “Deep Learning Based Brain Tumor Detection and Classification,” in Proc. International Conference on Intelligent Technologies (CONIT), Karnataka, India. June 25-27, 2021, IEEE.

[26] S. Pereira, A. Pinto, V. Alves, and C.A. Silva, “Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images,” IEEE Transactions on medical imaging, vol. 35, No. 5, May 2016, pp. 1240-1251.

[27] S. Arora and M. Sharma, “Deep Learning for Brain Tumor Classification from MRI Images,” in Proc. 6th International Conference on Image Information Processing (ICIIP), 2021, pp. 409-412.

[28] Y.K. Dubey, M.M. Mushrif, and K. Pisar, “Brain Tumor Type Detection Using Texture Features in MR Images,” IEEE, 2018.

[29] S. Gaikwad, S. Patel, and A. Shetty, “Brain Tumor Detection: An Application Based on Machine Learning,” in Proc. 2nd Int. Conf. for Emerging Technology (INCET), IEEE, India, 2021.

[30] S.K. Baranwal, K. Jaiswal, K. Vaibhav, A. Kumar, and R. Srikantaswamy, “Performance analysis of Brain Tumor Image Classification using CNN and SVM,” in Proc. of 2nd Int. Conf. on Inventive Research in Computing Applications (ICIRCA), IEEE, 2020, pp. 537-542.

[31] H. Kibriya, M. Masood, M. Nawaz, R. Rafique, and S. Rehman, “Multiclass Brain Tumor Classification Using Convolutional Neural Network and Support Vector Machine,” in Proc. of Int. Conf. on Computing (MAJICC), IEEE, 2021.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[32] I. Aboussaleh, J. Riffi, A.M. Mahraz, and H. Tairi, “Brain Tumor Segmentation Based on Deep Learning's Feature Representation,” Journal of Imaging, vol.7, 2021.

[33] A.W. Reza, M.S. Hossain, M.A. Wardiful, M. Farzana, S. Ahmad, F. Alam, R.N. Nandi, and N. Siddique, “A CNN-Based Strategy to Classify MRI-Based Brain Tumors Using Deep Convolutional Network,” Applied Sciences, vol. 13, p. 312, 2023.

[34] A. Younis, L. Qiang, C.O. Nyatega, M.J. Adamu, and H.B. Kawuwa, “Brain Tumor Analysis Using Deep Learning and VGG-16 Ensembling Learning Approaches,” Applied Sciences, vol. 12, p. 7282, 2022.

[35] H. Hu, X. Li, W. Yao, and Z. Yao, “Brain Tumor Diagnose Applying CNN through MRI,” in Proc. of 2nd Int. Conf. on Artificial Intelligence and Computer Engineering (ICAICE), IEEE, 2021, pp. 430-434.

[36] S. Asif, W. Yi, Q. U. Ain, J. Hou, T. Yi, and J. Si, “Improving Effectiveness of Different Deep Transfer Learning-Based Models for Detecting Brain Tumors From MR Images,” IEEE access, vol. 10, pp. 34716 – 34730, April 5, 2022.

[37] S. Mohsen, A.M. Ali, E. M. El-Rabaie, A. Elkaseer, S.G. Scholz, and A.M. Ali Hassan, “Brain Tumor Classification Using Hybrid Single Image Super-Resolution Technique With ResNext101_32 x 8d and VGG19 Pre-Trained Models,” IEEE access, vol. 11, pp. 55582 – 55595, June 8, 2023.

[38] K. Win, N. Maneerat, S. Choomchuay, S. Sreng, and K. Hamamoto, “Suitable Supervised Machine Learning Techniques for Malignant Mesothelioma Diagnosis,” The 2018 Biomedical Engineering International Conference (BMEiCON-2018), IEEE.

[39] K. Win, N. Maneerat, K. Hamamoto, and S. Sreng, “Hybrid Learning of Hand-Crafted and Deep Activated Features Using Particle Swarm Optimization and Optimized Support Vector Machine for Tuberculosis Screening,” Appl. Sci. 2020.

[40] S. Sreng, N. Maneerat, D. Isarakorn, K. Hamamoto, and R. Panjaphongse, “Primary Screening of Diabetic Retinopathy Based on Integrating Morphological Operation and Support Vector Machine,” 2017 Bioinformatics, Medical Imaging and Neuroscience (ICIIBMS), pp. 250-254, IEEE, 2017.

[41] S. Sreng, N. Maneerat, K. Win, K. Hamamoto, and R. Panjaphongse, “Classification of Cotton Wool Spots Using Principal Components Analysis and Support Vector Machine,” The 2018 Biomedical Engineering International Conference (BMEiCON), IEEE, 2018.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[42] I. Valova, C. Harris, T. Mai, and N. Gueorguieva, “Optimization of Convolutional Neural Networks for Imbalanced Set Classification,” KES, Elsevier B.V., 2020, pp. 660-669.

[43] K. Win, N. Maneerat, S. Sreng, and K. Hamamoto, “Ensemble Deep Learning for the Detection of COVID-19 in Unbalanced Chest X-ray Dataset,” Appl. Sci., vol. 11, p. 10528, 2021.

[44] S. Tang, C. Wang, J. Nie, N. Kumar, Y. Zhang, Z. Xiong, and A. Barnawi, “EDL-COVID: Ensemble Deep Learning for COVID-19 Case Detection From Chest X-Ray Images,” IEEE Transactions on industrial informatics, vol. 17, no. 9, 2021, pp.6539-6549.

[45] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, Vincent Vanhoucke, and Andrew Rabinovich, “Going deeper with convolutions”, 2014.

[46] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks”, In Advances in Neural Information Processing Systems 25, pp. 1106–1114, 2012.

[47] MATLAB webpage https://www.mathworks.com (online training and resource).

[48] W. Zhu, N. Zeng, and N. Wang, Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. NESUG proceedings: health care and life sciences, November 14-19, 2010, Baltimore, Maryland.

[49] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 2261–2269.

[50] D. C. Montgomery and G. C. Runger, “Applied Statistics and Probability for Engineers,” 7th ed., Wiley, 2018, pp. 194 -243.

[51] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why Should I Trust you? Explaining the Prediction of Any Classifier,” in Proc. of the 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, 2016, pp. 1135-1144.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก
บทความที่ได้รับการตีพิมพ์และเผยแพร่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Brain Tumor Classification using Supervised Support Vector Machine

Pongsak Pilaoon [1] , Noppadol Maneerat [1, *] , Kuniaki Yajima [2] , Ruttikorn Varakulsiripunth [3] , Kazuhiko Hamamoto [4]

[1] *School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520 Thailand*
[2] *National Institute of Technology (KOSEN), Sendai College, 4-16-1, Ayashi-Chuou, Aoba-ku, Sendai, Miyagi 989-3128, Japan*
[3] *Thai-Nichi Institute of Technology, Bangkok, 10250 Thailand*
[4] *Department of Information Media Technology, School of Information Science and Technology, Tokai University, 4-1-1 Kitakaname, Hiratsuka-shi, Kanagawa, 259-1292, Japan*
E-mail: [1]ppilaoon@gmail.com, [2]yajima@sendai-nct.ac.jp, [3]ruttikorn@tni.ac.th, [4]hama@keyaki.cc.u-tokai.ac.jp
[*] Correspondence: noppadol.ma@kmitl.ac.th; Tel.: +66-87-911-3939

2023 9th International Conference on Engineering, Applied Sciences, and Technology (ICEAST) | 979-8-3503-2204-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICEAST58324.2023.10156975

***Abstract*—A Glioblastoma (GBM) is a malignant brain tumor earlier detection and diagnosis will increase survival opportunities. This research has developed for binary classification of GBM brain tumors by supervised machine learning from magnetic resonance imaging (MRI). DICOM medical images have been converted into JPEG files and morphological operation has been implemented to separate the brain region from the skull image for preparation and easier for tumor segmentation in preprocessing stage. The global thresholding segmentation has been proposed to segment the brain tumor from the artifact and then the features have been extracted by gray level coefficient matrix feature extraction (GLCM). In this research, a support vector machine has been conducted for binary classification and finally, GBM grade-4 brain tumor is distinguished from normal brain images. The dataset comprises 155 MRI images 80% has been assigned for training and another 20% will be the testing dataset. The experimental output prediction result is 96.875 % accuracy, 95 % sensitivity, and 100% specificity. The performance of classification has been improved and shown better results when compared with previous research work.**

Keywords—GBM Brain Tumor, Morphological Operation, GLCM feature extraction, Support Vector Machine Classification

I. INTRODUCTION

Astrocytoma tumors arise from a star-shaped glial cell in the brain and spinal cord namely astrocytes that perform many functions, including provision nutrients to the nervous tissue, regulation of cerebral blood flow and repair process of the brain and spinal cord following infection and traumatic injuries. Astrocytoma tumors have been classified into four grades, and grade 4 referred to Glioblastoma Multiforme (GBM) that denotes the most malignant tumor and occurs often in frontal and temporal lobes of the brain. The early detection of GBM and treatment immediately affect with overall survival [1].

Magnetic Resonance Imaging (MRI) is an important diagnostic tool for GBM. Usually, images are acquired both before and after the administration of intravenous contrast. If the image shows much contrast enhancement is an indication of a higher grade of astrocytoma. Low-grade tumors usually do not show much contrast enhancement, while GBMs display strong contrast enhancement. After a brain tumor is detected on an MRI scan, a doctor obtains the location and size of the tumor for further treatment [2][3].

While brain tumor diagnoses from MRI scans by humans are prone to the error happening, the digital image processing technique has been conducted to help the doctor to diagnose the GBM brain tumor from MRI scans by computer system instead. The image processing technique has been widely used and proved to be an effective diagnostic tool in various medical applications.

The brain tumor detection and classification has been proposed by various researchers. Hein Tun Zaw et al. shown the segmentation tumor regions based on maximum entropy threshold and naïve bayes was used for classification the brain tumor from REMBRANDT database, the main advantage reported that can detect tumor at temporal lobe which align with the eye level [4]. Anandgaonkar and Sable presented segmentation based on Fuzzy C-means clustering algorithm to divide medical image in number of clusters, the tumor was extracted using thresholding and classified type of tumor based on its area [5]. Sujan et al. shown the brain tumor blob detection by Otsu's method thresholding and morphological operation was performed to reduce noise and improve accuracy of detection result [6]. Dubey et al. presented segmentation using stationary wavelet transform (SWT), watershed algorithm for GLCM feature extraction after that used neural network for classification the brain tumor from local dataset. The algorithm could successfully classify two type of brain tumor Meningioma and glioma [7]. Hossain et al. presented Fuzzy C-means clustering algorithm for tumor segmentation followed by traditional machine learning classifiers and compare with convolution neural network using brain tumor images from BRATS dataset [8]. Gaikwad et al. presented brain tumorous and non-tumorous classification using CNN with transfer learning approach VGG-16 model in python language that shown outperform result but use time consuming for training a model [9]

The purpose of this research is continuous development and improvement of the accuracy from past research work by Hein Tun Zaw et al. [4] that can detect the tumor in the temporal lobe, but this research improved to detection of GBM brain tumors in all regions. The research implemented the image processing technique and supervised machine learning to classify MRI medical images from the REMBRANDT database whether it is a GBM brain tumor or not. In preprocessing stage morphological operation is proposed to separate the brain region from the skull. Global thresholding is used for the segmentation of tumor blob and a threshold value is carefully adjusted to detect the tumor region. The features of brain tumors are extracted using GLCM feature extraction and a Support Vector Machine is used to classify the extracted feature in tabulated form finally, the prediction result is justified whether a tumor or not. The obtained experimental result is 96.88% accuracy.

979-8-3503-2204-0/23/$31.00 ©2023 IEEE

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The rest of the research is organized as follows: the proposed method is described in section II, the experimental results are described in section III and finally, conclusion is described in section IV.

II. METHODOLOGY

The proposed method for detection GBM brain tumor is shown in figure 1. The procedure to distinguish between brain tumor and normal consist with preprocessing, Global thresholding segmentation, GLCM feature extraction and classification by support vector machine classifier. The advantage of using machine learning is suitable for big data diagnose instead of visual inspection by human visualize.

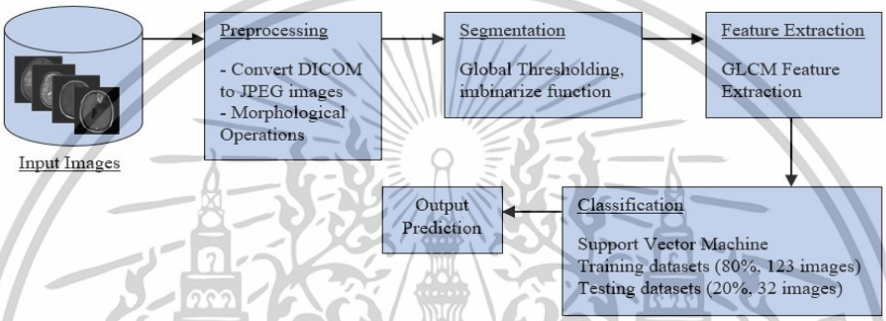
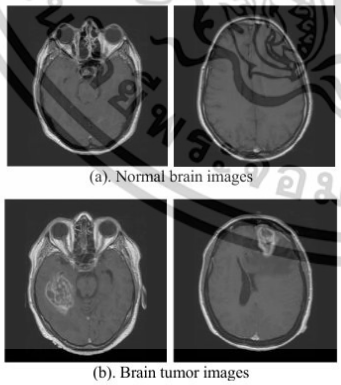


Fig. 1 Block diagram of brain tumor detection based on Support Vector Machine classification

A. Image Database

All MRI images use in this research are from The Repository of Molecular Brain Neoplasia Data (REMBRANDT). It contains the MRI images from 130 patients with clinical data analysis in DICOM format [10]. These dataset images were converted into JPEG file and stored in local database. A converted images totally are 155 images contains with GBMs brain tumor in different location 111 images and normal brain 44 images were used in this work. The brain tumor and normal images is shown in figure 2.



(a). Normal brain images

(b). Brain tumor images

Fig. 2 Example of MRI brain images converted to JPEG format

B. Preprocessing

This stage the MRI images were converted into JPEG format and resolution of images were adjusted to 630 x 630 pixels. The morphological operation and pixel subtraction were used for removal the skeletons from brain region. First step the image was converted to grayscale and the opening morphological operation with disk structuring element was performed for removing small objects and thin lines from an image while preserving the shape and size of larger objects in the image. A subtract image background function was used in the next sequence to remove the skull from the brain region, the remaining brain region only would be better for the segmentation stage. This research uses MATLAB R2020b program for the segmentation and classification of GBM brain tumors.

B.1. Morphological operation

The opening morphological operation is an erosion followed by a dilation by using the same structuring element for both operations. This research opening morphological with disk structuring element was used in preprocessing stage. The opening of set A by structuring element B is defined as,

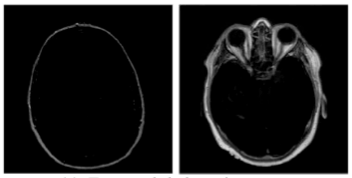
$$A \circ B = (A \ominus B) \oplus B \quad (1)$$

The opening A by B is an erosion of A by B followed by a dilation of the result by B [11][12]. The skeleton was removed by opening and remaining brain region only. However, some pixel of the brain was removed as well. Therefore, the pixel subtract function was used to separate the brain region from the skeleton.

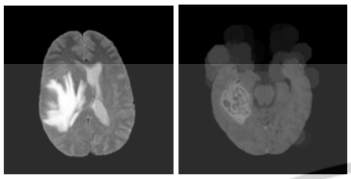
B.2. Image subtraction

The pixel subtraction was used to remove the skeleton from brain regions for the readiness segmentation stage. First, the opened image was subtracted from the original image to obtain the skull image and artifacts brain region and then the opened image was subtracted from the skull with the artifact brain in the first step to obtain the skeleton only. Finally, the separated skeleton was subtracted from the original image and obtained the brain region only which is ready for further segmentation as shown in figure 3.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a). Extracted skeleton images



(b). Brain region images

Fig. 3 Example of pre-processed brain images

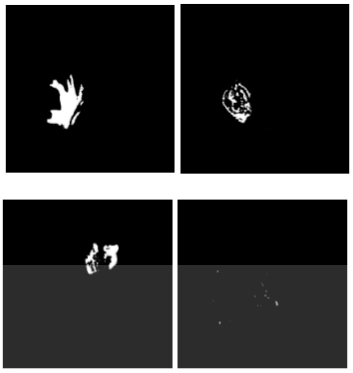


Fig. 5 The segmented tumor and normal brain images

C. Segmentation

A global thresholding was used for segmentation in this research to distinguish the brain tumor region from background for further processing. The threshold is selected by visual inspection of the image histogram, the solution to extract the objects form the background is to select a suitable threshold value. Any point *(x, y)* for which *f (x, y)* greater than *T* is called an object point and otherwise the point is called a background point [11][12]. The threshold image *g (x, y)* is defined as,

$$g(x, y) = 1 \text{ if } f(x, y) \geq T \quad (2)$$
$$= 0 \text{ if } f(x, y) < T$$

The histogram of preprocessed brain image is shown in figure 4, from visual inspection found that the background point is threshold less than 85 and otherwise is a tumor object. The threshold is normalized to the range [0, 1], therefore the proper threshold is 0.33 which is derived from 85 divided by 255.

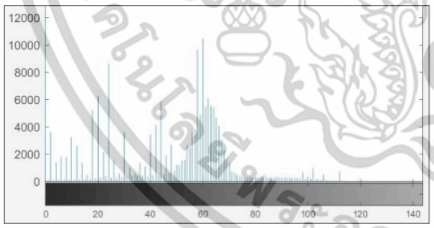


Fig. 4 Preprocessed brain image histogram

The threshold was carefully adjusted until found that produces a good result of brain tumor image was observed. To segment the image the threshold is used in the function *imbinarize* and the result of segmentation is shown in figure 5. The segmented image of both tumor and normal brain image readiness for extraction of the feature in the next step.

D. Gray Level Co-occurrence Matrix Feature extraction

Feature extraction is a process to provide the represent data for binary classification in machine learning classifier. GLCM is a statistical method of examining texture that consider the spatial relationship of pixels and the GLCM feature extraction was used to extract a feature in this research. Firstly, the segmented grayscale image was converted to GLCMs matrix using *graycomatrix* function and then the several statistics were derived from them using *graycoprops* function, the statistics provide information about the texture of an image. The statistics namely contrast, correlation, energy and homogeneity are defined as,

Contrast measure of intensity contrast between a pixel and its neighbor over the whole image. The contrast value is 0 for a constant image.

Correlation measure of how correlated a pixel is to its neighbor over the whole image. Range of value is [1,-1], corresponding to perfect positive and negative correlations. The correlation value is not a number (NaN) for a constant image.

Energy measure of uniformity in the range [0,1], Uniformity is 1 for a constant image.

Homogeneity measures the closeness of the distribution of elements in GLCM to the GLCM diagonal. Range of value is [0,1] and homogeneity value is 1 for a diagonal GLCM.

All GLCM features were extracted and summarized in tabulated format is shown in table 1. The binary class of feature was specified to 1 for brain tumor image and 0 for normal brain image and further processing to classify by supervised machine learning.

TABLE I. EXAMPLE OF EXTRACTED FEATURE USING GLCM

Contrast	Correlation	Energy	Homogeneity	Class
0.0725	0.9064	0.9827	0.9987	1
0.064	0.9463	0.9743	0.9988	1
0.0265	0.9528	0.988	0.9995	1
0.0027	0.5412	0.9998	0.9999	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

E. SVM Classification

Support Vector Machine is a supervised machine learning that has been widely used for binary classification in medical image diagnostic research [13] – [16]. SVM classifies data by finding the best hyperplane to separate data points of one class from those of another class. The best hyperplane means the largest margin between two classes and the support vectors are the points that closest with hyperplane [17]. The definition is illustrated in figure 6, positive indicates data points of type 1 and negative indicates data points of type -1.

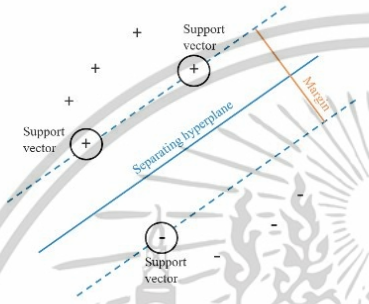


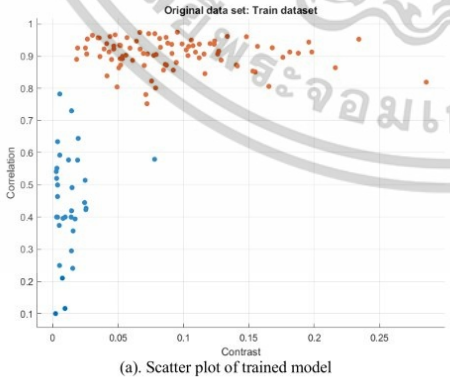
Fig. 6 Support Vector Machine definition

Given a dataset $D = \{(x_i, y_i), i = 1 - n\}$ and $y_i = \in \{-1, 1\}$, Support Vector Machine is defined as;

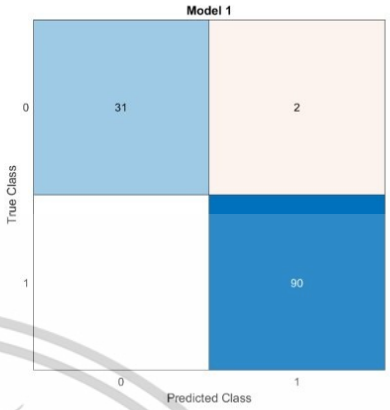
$$\min_{w,b,\zeta} \frac{1}{2}\|w\|^2 + C\sum_{i}^{n}\zeta_i \qquad (3)$$

Subject to $y_i(w \cdot x_i + b) \geq 1 - \zeta_i, \zeta_i \geq 0, i = 1 - n$,
where ζ_i is slack variables and C is regularization parameter.

The extracted features in tabulated format were imported into SVM classifier, the features dataset comprise of brain tumor 111 features and normal brain 44 features. Firstly, the features dataset was randomly arranged for preparation and separated for training and testing by the ratio 80 : 20 respectively. The training dataset is 123 features were fed into linear SVM classifier and obtained the scatter plot and confusion matrix as shown in figure 7.



(a). Scatter plot of trained model



(b). Confusion matrix of trained model

Fig. 7 Scatter plot and confusion matrix of training dataset

After that the trained model was utilized for testing the remaining 32 features dataset to predict and classify that the images are normal or brain tumor.

III. Experimental Results

The testing dataset 32 features were imported to the workspace for prediction whether it is tumor or normal brain image using trained model and finally, the output prediction result obtained. The attributes of the diagnostic tests can be measured. These attributes namely sensitivity, specificity and accuracy have been widely used statistics to describe a diagnostic test [18]. The parameters used for calculation the attributes compose with following,

True positive is tumor exist and was correctly detected.
True negative is normal brain and correctly identified.
False positive is normal brain but identified as tumor.
False negative is tumor exist but can't detected.
And the evaluated result is calculated as follows,

Sensitivity = TP / (TP + FN)

Specificity = TN / (TN + FP)

Accuracy = (TN + TP) / (TN+TP+FN+FP)

The parameters and attributes to measure the performance of prediction is shown in table 2.

TABLE II. Prediction Result Obtained

True Positive	True Negative	False Positive	False Negative
19	12	0	1
Evaluated result			
Accuracy	**Sensitivity**	**Specificity**	
96.875%	95%	100%	

The obtained experimental result has been compared with the past research and found that the accuracy has been improved with better results 96.875 %. The comparison result is shown in table 3.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE III. COMPARISON RESULT

Reference	Methods	Accuracy
Dubey et al. (2018)	Wavelet transform and Neural network	80% (Glioma)
Zaw et al. (2019)	Maximum entropy threshold and Naïve Bayes classifier	94%
Hossain et al. (2019)	Fuzzy C-means and SVM,	92.42%
	Convolution Neural Network	97.87%
Gaikwad et al. (2021)	CNN with VGG-16 model	99.03%
Proposed methodology	Support Vector Machine	96.875%

IV. CONCLUSION

In this research, the automated diagnosis of Glioblastoma (GBM) brain tumor has been proposed to assist the medical staff such as neurosurgeons or radiologists to an early detect brain tumor. MRI images from the REMBRANDT database have been used for an experiment, the morphological operation used to remove the skeleton from the brain area. The tumor blob was segmented using global thresholding and the GLCM features were extracted and arranged in a tabulated format. The extracted features are imported to the Support Vector Machine classifier and finally, output prediction has been generated whether the tumor or normal brain images. The experimental result of the proposed methodology obtains an accuracy of 96.875%, a sensitivity of 95%, and a specificity of 100% respectively, the result better than past research work that uses a similar database. The proposed methodology's benefit is suitable for use with big data of medical image diagnosis and less time-consuming for training the model when compared with CNN deep learning method. Therefore, In the future work we will try to implement the methodology with the large dataset for testing the robustness and against overfitting problem.

REFERENCES

[1] https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Glioblastoma-Multiforme

[2] E. Holland, "Glioblastoma multiforme: the terminator," Proceedings of the National Academy of Sciences, 97(12), pp.6242-6244.

[3] K. Urbańska, J. Sokołowska, M. Szmidt, and P. Sysa, 2014. Glioblastoma overview. Contemporary oncology, *18*(5), p.307.

[4] H. Zaw, N. Maneerat and K. Win, "Brain tumor detection based on Naïve Bayes Classification," 2019, IEEE.

[5] G. Anandgaonkar and G. Sable, "Brain tumor detection and identification from T1 post contrast MR images using cluster based segmentation," International Journal of Science and Research, vol. 3, pp. 814-817, April 2014.

[6] M. Sujan, N. Alam, S. Abdullah and M. Jahirul, "A Segmentation based Automated System for Brain Tumor Detection," International Journal of Computer Applications, vol. 153, pp. 41-49, November 2016.

[7] Y. Dubey, M. Mushrif and K. Pisar, "Brain tumor type detection using texture feature in MR images," 2018, IEEE.

[8] T. Hossain, F. Shishir , M. Ashraf, A. Nasim and F. Shah, "Brain Tumor Detection Using Convolutional Neural Network," 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 (ICASERT 2019), IEEE.

[9] S. Gaikwad, S. Patel and A. Shetty, "Brain tumor detection: An application based on machine learning," 2021 2nd International Conference for Emerging Technology (INCET), IEEE.

[10] The Cancer Imaging Archive, 2017. REMBRANDT https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT

[11] R. Gonzalez and R. Woods, Digital Image Processing, 4th ed., Pearson Education, 2018.

[12] R. Gonzalez, R. Woods and S. Eddins, Digital Image Processing using MATLAB, Pearson Education.

[13] K. Win, N. Maneerat, S. Choomchuay, S. Sreng and K. Hamamoto, "Suitable Supervised Machine Learning Techniques for Malignant Mesothelioma Diagnosis," The 2018 Biomedical Engineering International Conference (BMEiCON-2018), IEEE.

[14] K. Win, N. Maneerat, K. Hamamoto and S. Sreng, "Hybrid Learning of Hand-Crafted and Deep-Activated Features Using Particle Swarm Optimization and Optimized Support Vector Machine for Tuberculosis Screening," Appl. Sci. 2020.

[15] S. Sreng, N. Maneerat, D. Isarakorn, K. Hamamoto and R.Panjaphongse, "Primary Screening of Diabetic Retinopathy Based on Integrating Morphological Operation and Support Vector Machine," 2017 Bioinformatics, Medical Imaging and Neuroscience (ICIIBMS 2017) (pp. 250-254). IEEE.

[16] S. Sreng, N. Maneerat, K. Win, K. Hamamoto and R.Panjaphongse, "Classification of Cotton Wool Spots Using Principal Components Analysis and Support Vector Machine," The 2018 Biomedical Engineering International Conference (BMEiCON-2018), IEEE.

[17] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning 20, 3, 273-297, (1995).

[18] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy,associated confidence interval and ROC analysis with practical SAS implementations," NESUG proceedings: health care and life sciences, November 14-19, 2010, Baltimore, Maryland.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Brain Tumor Classification using Pretrained Deep Convolutional Neural Network

Pongsak Pilaoon [1] , Noppadol Maneerat [1, *] , Athasart Nakthewan [1] , Ruttikorn Varakulsiripunth [2] , Kazuhiko Hamamoto [3]

[1] *School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520 Thailand;*
ppilaoon@gmail.com (P.P.); athasart.na@kmitl.ac.th (A.N.)
[2] *Thai-Nichi Institute of Technology, Bangkok, 10250 Thailand;*
ruttikorn@tni.ac.th
[3] *Department of Information Media Technology, School of Information Science and Technology, Tokai University, 4-1-1 Kitakaname, Hiratsuka-shi, Kanagawa, 259-1292, Japan;*
hama@keyaki.cc.u-tokai.ac.jp
[*] Correspondence: noppadol.ma@kmitl.ac.th; Tel.: +66-87-911-3939

2023 9th International Conference on Engineering, Applied Sciences, and Technology (ICEAST) | 979-8-3503-2204-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICEAST58324.2023.10157725

***Abstract*—In this research deep learning convolution neural network (CNN) has been implemented for binary classification GBM brain tumor. The dataset from REMBRANDT database that comprise of 155 MRI images has been utilized in this research. The transfer learning by pretrained network namely GoogleNet and AlexNet have been conducted to classify the GBM brain tumor form normal brain. The advantage of classification by transfer learning is the manual segmentation and feature extraction were replaced by automatic procedure that reduce the error from human handcraft. The prediction result by using Googlenet pretrained network shown accuracy result 80.85% and Alexnet pretrained network obtained accuracy 93.62%. The GBM brain tumor classification by deep learning pretrained network obtained good result of accuracy and can be implemented in practical to help the medical staff for earlier diagnosis for further treatment and increasing survivor rate of patients. The manual adjustment for segmentation and feature extraction are improved by automatic classification using deep learning pretrained network is main advantage of this research. The future work we will try to implement with increasing images dataset to improve accuracy, robustness testing and prevent overfitting problem.**

***Keywords*—GBM Brain Tumor, Deep Learning, Transfer Learning, Pretrained deep convolutional neural Network**

I. Introduction

Glioblastoma Multiforme (GBM) is a grade 4 astrocytoma tumor that denotes the most malignant brain tumor. The early detection of GBM and treatment immediately affect overall survival [1].

Magnetic Resonance Imaging (MRI) is an important tool for GBM diagnosis. If the image shows much contrast enhancement indicating a higher grade of astrocytoma, the GBMs display strong contrast enhancement. After a brain tumor is detected on an MRI scan, a doctor will immediately treat it to increase the survival rate [2][3]. The disease diagnosis from medical images using the computer system can reduce the error from human sight and recently the segmentation and classification the medical images has used the digital image processing technique and machine learning for assistant.

Although the digital image processing segmentation with machine learning classification technique has been widely used and proved to be an effective diagnostic tool in various medical applications. And the GBM brain tumor diagnosis using machine learning classification can reduce the error by human visual inspection with superior accuracy result, but the input and parameter of the segmentation and feature extraction stage still need to be manually fed and adjusted by human. In recently the deep learning technique has been implemented in various medical applications and obtains very good results. The main advantage of deep learning is the segmentation and extraction the feature will be automatically performed by deep convolutional neural network.

The Segmentation and Classification of Brain tumors have been introduced in various research. Hein Tun Zaw et al. presented segmentation of the medical image from the REMBRANDT database based on the maximum entropy threshold method to enhance the brain image and obtain the tumor region. And then the morphology features and intensity features are extracted to utilize for classification based on Naïve Bayes supervised machine learning. This method can detect the tumor in different regions including the middle region that aligns with eye level which is a significant advantage [4]. Hossain et al. presented Fuzzy C-means clustering algorithm for tumor segmentation followed by traditional machine learning classifiers and compare with convolution neural network using brain tumor images from BRATS dataset [5]. N.M. Dipu et al. presented brain tumor classification using YOLOv5 and FastAI deep learning with database BRATS 2018 that contains 1,992 MRI images [6]. Arora and Sharma presented brain tumor classification using various deep learning VGG16, inception V3 and ResNet with BRATS 2015 that contains 3,764 MRI images [7]

The purpose of this research is improvement and looking for new methodology to classify the GBM brain tumor by reference to past research work presented by Hein Tun Zaw et al. [4] which the main advantage is can detect the tumor in the temporal lobe, but this research is improved for detection GBM brain tumors in all regions and furthermore the segmentation and feature extraction were fully automated by deep learning. The deep learning has been implemented in various research work to diagnose the disease from medical images [5]-[10]. In this research the pretrained deep convolutional neural network namely GoogLeNet and AlexNet have been implemented to classify the MRI medical images from the REMBRANDT database whether it is a GBM brain tumor or not. The whole provided MRI images are 155 images are arranged in datastores and the image labels is provided and mentioned that it is brain tumor or normal brain. The dataset was divided into training and testing dataset with ratio 70% and 30% respectively. For readiness to training by each pretrained network the input image size requires to preprocessing stage first. The GoogLeNet pretrained network requires input images size 224 x 224 x3, meanwhile the AlexNet pretrained network requires input images size 227 x 227 x3, 3 is RGB color channel.

979-8-3503-2204-0/23/$31.00 ©2023 IEEE

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The rest of the research is organized as follows: the methodology is described in section II that comprise of brain tumor classification using GoogLeNet and AlexNet pretrained network, the experimental results are described in section III and finally, conclusion is described in section IV.

II. METHODOLOGY

The methodology for binary classification GBM brain tumor using transfer learning is shown in figure 1. The procedure consists of input MRI images that have been converted from DICOM medical image to JPEG file for further proceeding. The datastore and image labels were created that comprise of 155 images dataset including with GBM brain tumor and normal brain condition. The images were randomly divided into training and testing dataset and then for preprocessing stage the datastores were performed augmentation to properly resize input image for each pretrained network. The transfer learning namely GoogLeNet and AlexNet have been implemented to classify the brain tumor, the transfer learning components that comprise with network layer, training data and algorithm option were properly set prior to the training network stage. And then training data was trained by pretrained deep convolutional neural network and training progress was plotted for monitoring purpose. The trained network has been utilized for prediction the divided testing dataset in final stage and the confusion matrix was plotted and analyzed accuracy as well.

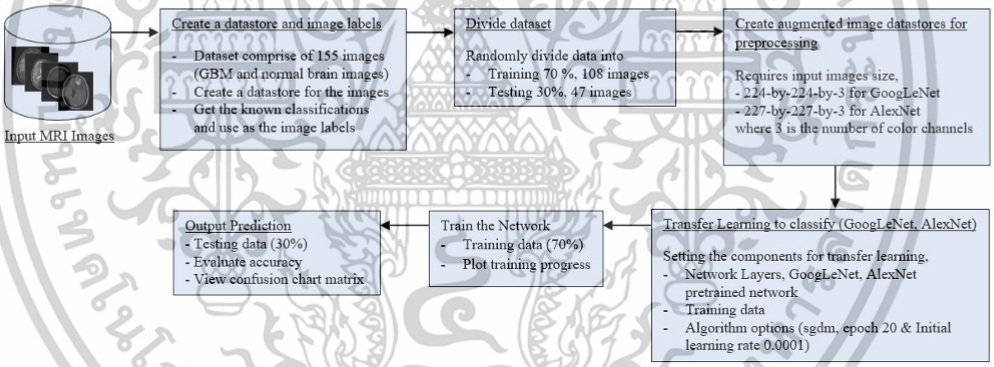


Fig. 1 Block diagram of brain tumor classification based on Deep Learning Network

A. Image Database

All MRI images used in this research are from The Repository of Molecular Brain Neoplasia Data (REMBRANDT). It contains the MRI images from 130 patients with clinical data analysis in DICOM format [11]. These dataset images were converted into JPEG file and stored in local database. A converted images totally are 155 images contains with GBMs brain tumor in different location 111 images and normal brain 44 images were used in this work. The brain tumor and normal images are shown in figure 2. The datastore for the whole images was created and the known classification of brain tumor type was set and used as the image labels for further processing. The dataset was randomly divided into 2 groups for training and testing dataset that comprise of 70% and 30% respectively. Therefore, the training dataset comprise of 108 MRI images and testing dataset comprise of 47 MRI images and each dataset combined with GBM brain tumor and normal brain condition.

B. Preprocessing by create augemented image datastores

The augmented image datastores were created in preprocessing stage to provide the suitable input size and format for the GoogLeNet and AlexNet pretrained network. For the GoogLeNet network requires input images size 224 x 224 and 3 color channels. The AlexNet pretrained network requires input images size 227 x 227 and 3 color channels, this property was set in the first element of the network layer namely image input layer prior next further proceeding. This research uses MATLAB R2020b program with the pretrained deep convolutional neural network module namely GoogLeNet and AlexNet for the classification of GBM brain tumors.

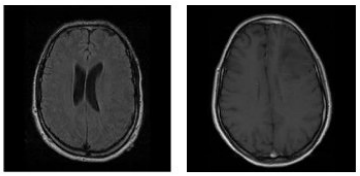
C. Brain tumor classification using GoogLeNet

In this research the pretrained GoogLeNet convolutional neural network has performed for classification a new collection of images dataset. GoogLeNet has been trained on million images and can classify image into 1000 object categories [12]. The GoogLeNet network is connected in Directed Acyclic Graph (DAG) concept and takes an image as input and outputs a label for the object image. The Convolutional Neural Network Architecture is shown in figure 3. A fine tuning a network with transfer learning is faster and easier than training a network from scratch. Transfer learning comprise of pretrained network layer, training data and algorithm options.

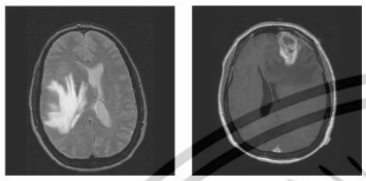
C.1. GoogLeNet pretrained network layer

The GoogLeNet pretrained network was loaded to MATLAB workspace and the layer analysis result is shown in figure 4a, the network is connected in Directed Acyclic Graph (DAG) concept and comprise with 144 layers and the input images requires size is 224 x 224 x 3. The fully connected layer and classification output layer were replaced for suitable with our application. Therefore, the fully connected layer has been changed into 2 categories are tumor and normal brain image labels. The algorithm options were set to stochastic gradient descent with momentum (SGDM) with 20 epochs and the initial learning rate was set to 0.0001 and training progress was plotted for monitoring as well. The component of transfer learning comprises of pretrained network layer, training data and algorithm options were arranged and readiness to perform training the network in next step.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a). Normal brain images



(b). GBM Brain tumor images

Fig. 2 Example of MRI brain images converted to JPEG format

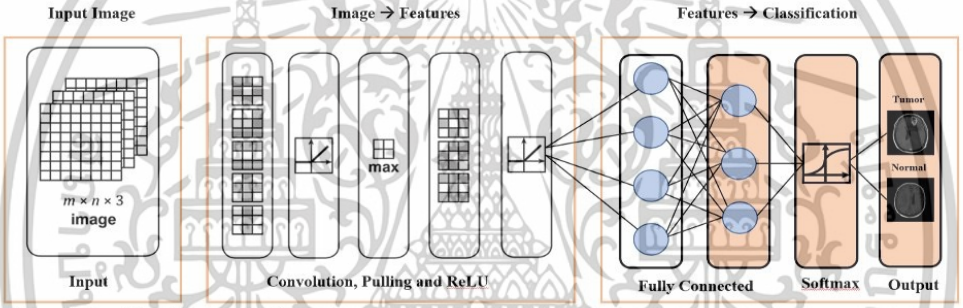
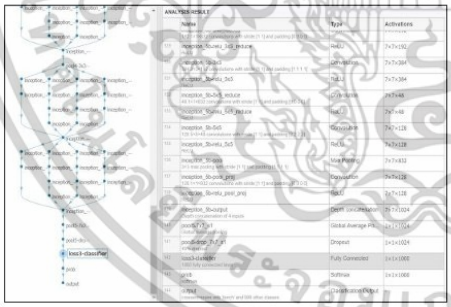
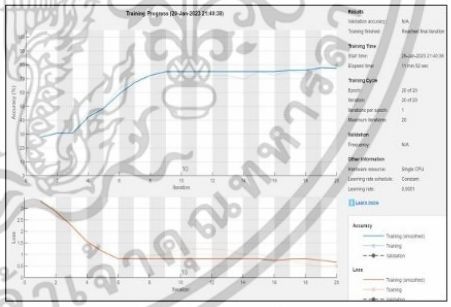


Fig. 3 Convolutional Neural Network Architecture



(a). GoogLeNet analysis result for network layer



(b). Training result chart of GoogLeNet

Fig. 4 GoogLeNet analysis result for network layer and training chart

C.2. Train the network and output prediction

The modified network has been trained using pretrained GoogLeNet network with training option "SGDM" and initial learning rate 0.0001 with 20 epochs. The training progress and training result has been plotted and shown as per figure 4b, the training time is 11.52 minute approximately with single CPU, the accuracy and loss of training is shown as well. For the next step the trained network was utilized for prediction the testing dataset that contains with the randomly divided 47 images. The accuracy of prediction and confusion matrix has been calculated and plotted for improvement. The experimental result obtained accuracy 80.85% for GoogLeNet.

D. Brain tumor classification using AlexNet

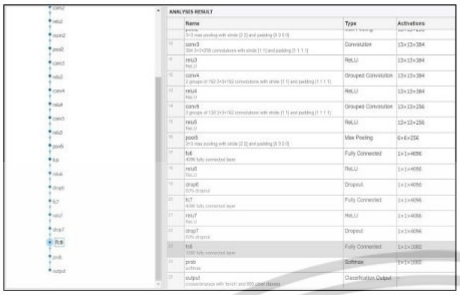
The pretrained AlexNet convolutional neural network has performed for classification a new collection of images dataset for further improve the accuracy. AlexNet has been trained on million images and can classify image into 1000 object categories, the model has learned many features for a wide range of images [13]. The AlexNet network is connected in series network that a neural network for deep learning with layer is arranged one after the other and takes an image as input and outputs a label for the object image. It contains single input layer and single output layer. The Convolutional Neural Network Architecture is shown in figure 3 that comprise of input image, image to features that contains with many convolution, pooling and ReLU layer and features to classification portion which the last 3 layers will be replaced to corresponding with our application. The procedure for classification brain tumor using AlexNet slightly different form GoogLeNet pretrained network such as the input image size will be changed to 227x227x3 and the number of layers for AlexNet contains only 25 layers that lead to less time consume for training the network.

D.1. AlexNet pretrained network layer

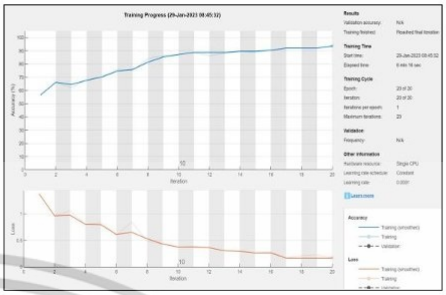
The AlexNet pretrained network was loaded to MATLAB workspace and the layer analysis result is shown in figure 5a, the network is connected in series network and comprise with 25 layers and the input images requires size is 227 x 227 and 3 color channels. The fully connected layer and classification output layer were replaced, and the fully connected layer has been changed into 2 categories are tumor and normal brain image labels. The algorithm options were set to stochastic gradient descent with momentum with 20 epochs and the

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

initial learning rate was set to 0.0001 similar with GoogLeNet pretrained network.



(a). AlexNet analysis result for network layer



(b). Training result chart of AlexNet

Fig. 5 AlexNet analysis result for network layer and training chart

D.2. Train the network and output prediction

The modified network has been trained using pretrained AlexNet network with training option "SGDM" and initial learning rate 0.0001 with 20 epochs. The training result has been plotted and shown as per figure 5b, the training time is 6.16 minute approximately with single CPU. The training time for AlexNet consume less than GoogLeNet due to the number layer is 25 layers only. And then the trained network was utilized for prediction the testing dataset that contains with the randomly divided 47 images. The accuracy of prediction and confusion matrix has been calculated and plotted for further improvement. The experimental result obtained accuracy 93.62% for AlexNet shown that it has been significantly improved.

III. Experimental Results

The testing dataset is randomly divided into 30% that comprise of 47 images. The attributes of the diagnostic tests can be measured. These attributes namely sensitivity, specificity and accuracy have been widely used statistics to describe a diagnostic test [14]. The parameters used for calculation the attributes compose with following,
True positive is tumor exist and was correctly detected.
True negative is normal brain and correctly identified.
False positive is normal brain but identified as tumor.
False negative is tumor exist but can't detected.
And the evaluated result is calculated as follows,

Sensitivity = TP / (TP + FN)

Specificity = TN / (TN + FP)

Accuracy = (TN + TP) / (TN+TP+FN+FP)

The confusion matrix for each pretrained network is shown in figure 6, the testing dataset 47 images was classified using the trained network. The classification result and attributes to measure effectiveness of prediction is shown in table 1. The GoogLeNet obtained false positive 7 images, false negative 2 images and overall accuracy 80.85%. The AlexNet obtained false positive 2 images, false negative 1 image and overall accuracy 93.62%. The experimental result for each pretrained network was presented in prediction result chart figure 7 that shown the sensitivity of GoogLeNet and AlexNet 93.94% and 96.97% respectively, the specificity of each pretrained network are 50% and 85.71% respectively.

TABLE I. Prediction Result Obtained

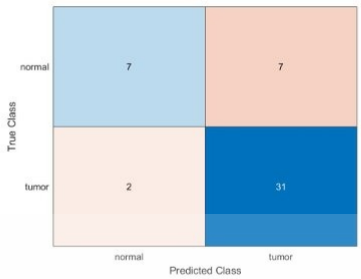
Pretrained Network	True Positive	True Negative	False Positive	False Negative
GoogleNet	31	7	7	2
AlexNet	32	12	2	1
Evaluated result				
Pretrained Network	**Accuracy**	**Sensitivity**	**Specificity**	
GoogleNet	80.85%	93.94%	50%	
AlexNet	93.62%	96.97%	85.71%	

TABLE II. Comparison Result

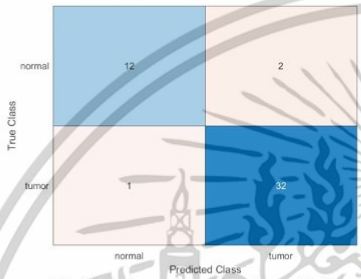
Reference	Methods	Accuracy
Zaw et al. (2019)	Maximum entropy threshold and Naïve Bayes classifier	94%
Hossain et al. (2019)	Fuzzy C-means and SVM	92.42%
	Convolution Neural Network	97.87%
Dipu et al. (2021)	YOLOV5	85.95%
	FastAi	95.78%
Arora & Sharma (2021)	VGG16	90.54%
	ResNet	81.92%
Proposed methodology	GoogLeNet	80.85%
	AlexNet	93.62%

The proposed methodology has been compared to another past related work and shown in table 2. And found that the obtained accuracy of proposed method is good performance and can be implemented in practical work to assist the medical staff for diagnosis the GBM brain tumor.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a). Confusion matrix of GoogLeNet prediction



(b). Confusion matrix of AlexNet prediction

Fig. 6 Confusion matrix of classification by GoogLeNet and AlexNet

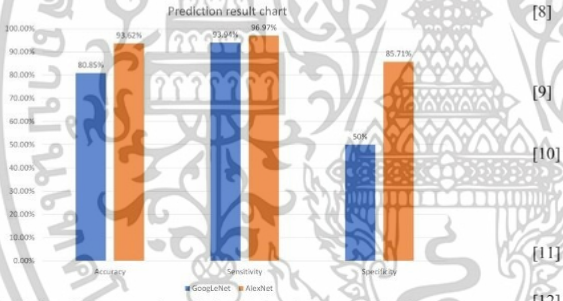


Fig. 7 The effectiveness of prediction by GoogLeNet and AlexNet

IV. Conclusion

In this research, MRI images from the REMBRANDT database have been used to study the classification of GBM brain tumor using the pretrained deep convolutional neural network. The pretrained network namely GoogLeNet and AlexNet were conducted to classify the GBM brain tumor and normal brain. The MRI images were randomly divided into training and testing datastore in the ratio of 70% and 30% respectively and then procedures of classification were performed as explained in block diagram. The deep learning conceptual was implemented to classify GBM brain tumor instead the manual segmentation and feature extraction that prone to error happening. The deep convolutional neural network has trained the training dataset using algorithm option and then the trained network was used for classification the testing dataset. The prediction result obtained good accuracy with 80.85% for GoogLeNet and 93.62% for AlexNet pretrained network that can be implemented in practical work to assist the medical staff for diagnosis the GBM brain tumor. In future work, we will try to implement with the large dataset for robust testing and against overfitting classification problems.

References

[1] https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Glioblastoma-Multiforme

[2] E. Holland, "Glioblastoma multiforme: the terminator," Proceedings of the National Academy of Sciences, 97(12), pp.6242-6244.

[3] K. Urbańska, J. Sokołowska, M. Szmidt, and P. Sysa, 2014. Glioblastoma overview. Contemporary oncology, *18*(5), p.307.

[4] H. Zaw, N. Maneerat and K. Win, "Brain tumor detection based on Naïve Bayes Classification," 2019, IEEE.

[5] T. Hossain, F. Shishir , M. Ashraf, A. Nasim and F. Shah, "Brain Tumor Detection Using Convolutional Neural Network," 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 (ICASERT 2019), IEEE.

[6] N.M. Dipu, S.A. Shohan and K.M.A. Salam, "Deep Learning Based Brain Tumor Detection and Classification", 2021 International Conference on Intelligent Technologies (CONIT), Karnataka, India. June 25-27, 2021, IEEE

[7] S. Arora and M. Sharma, "Deep Learning for Brain Tumor Classification from MRI Images", 2021 Sixth International Conference on Image Information Processing (ICIIP), pp.409-412.

[8] S. Pereira, A. Pinto, V. Alves and C.A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016, pp.1240-1251.

[9] K. Win, N. Maneerat, S. Sreng and K. Hamamoto, "Ensemble Deep Learning for the Detection of COVID-19 in Unbalanced Chest X-ray Dataset", Appl. Sci. 2021, 11, 10528

[10] S. Tang, C. Wang, J. Nie, N. Kumar, Y. Zhang, Z. Xiong and A. Barnawi, "EDL-COVID: Ensemble Deep Learning for COVID-19 Case Detection From Chest X-Ray Images", IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 17, NO. 9, SEPTEMBER 2021, pp.6539-6549

[11] The Cancer Imaging Archive, 2017. REMBRANDT https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT

[12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, Vincent Vanhoucke and Andrew Rabinovich, "Going deeper with convolutions", 2014.

[13] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks", In Advances in Neural Information Processing Systems 25, pp. 1106–1114, 2012.

[14] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy,associated confidence interval and ROC analysis with practical SAS implementations," NESUG proceedings: health care and life sciences, November 14-19, 2010, Baltimore, Maryland.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Glioma Brain Tumor Classification using Transfer Learning

Pongsak Pilaoon [1] , Athasart Narkthewan [2*], Noppanat Wadlom [3] , Ruttikorn Varakulsiripunth [4] , Kazuhiko Hamamoto [5] , Noppadol Maneerat [2]

[1] School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520 Thailand;
[2] Department of Engineering, King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, 86160 Thailand;
[3] School of Computing, Korea Advanced Institute of Science and Technology, Daejeon, 34141 South Korea;
[4] Thai-Nichi Institute of Technology, Bangkok, 10250 Thailand;
[5] Department of Information Media Technology, School of Information Science and Technology, Tokai University, 4-1-1 Kitakaname, Hiratsuka-shi, Kanagawa, 259-1292, Japan;

E-mail: [1] ppilaoon@gmail.com, [3] noppanat_w@kaist.ac.kr, [4] ruttikorn@tni.ac.th, [5] hama@keyaki.cc.u-tokai.ac.jp, [1] noppadol.ma@kmitl.ac.th
[*] Correspondence: athasart.na@kmitl.ac.th; Tel.: +66-81-632-6779

2024 10th International Conference on Engineering, Applied Sciences, and Technology (ICEAST) | 979-8-3503-8594-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICEAST61342.2024.10553834

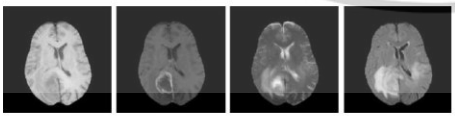
***Abstract*—In this research the glioma brain tumor binary classification using transfer learning was introduced. The MRI image from 2 datasets comprised with REMBRANDT and BraTS2021 with increasing number of MRI images was proposed to prevent overfitting problem. MRI images were converted to JPEG format and heavily imbalanced with normal brain image is minority class. Morphological operation was used to remove skeletons and artifacts from brain region. We have introduced Contrast Limited Adaptive Histogram Equalization to preprocess and enhance contrast before classify using various CNNs. To handle imbalanced dataset problem, we proposed image augmentation to increase the number of images and obtain balanced dataset. The various CNNs transfer learning was implemented to classify glioma brain tumor. Finally, the best classifier is InceptionV3 with balanced dataset that obtained accuracy 99.19%, sensitivity 98.83%, and specificity 100% respectively, better than our past research work.**

Keywords—Glioma Brain Tumor, morphological operation, CLAHE, Transfer Learning, Convolution Neural Network

I. Introduction

Glioma brain tumor has been classified into four grades by referring to aggressiveness and disease severity. Gliomas brain tumor are the most primary brain tumors in adults and deem as 70% of adult malignant primary brain tumors that arise from glial cells. The high-grade gliomas are grade III and IV that are malignant tumors. The glioma grade IV is also known as glioblastoma multiforme (GBM) which has been the most malignant primary brain tumor and very rapid growth [1-2]. Average survival time for GBM patient is one year approximately. Earlier detection and immediate treatment will increase survival rate. Basically, the glioma brain tumor diagnosis uses MRI image to visual inspection by medical staff that will lead to error from human sight. Hence, this study has proposed the computer vision technique to assist medical staff in diagnosis of the disease.

The magnetic resonance imaging (MRI) has been standard technique for visual inspection and diagnosis disease which is a non-invasive and widely used in clinicals.



(a) T_1 (b) T_{1C} (c) T_2 (d) T_{2FLAIR}

Fig. 1 MRI image's axial slice sequences of a high-grade glioma.

Currently, different MRI sequences are employed for diagnosis of brain tumor includes T_1-weighted MRI (T_1), T_1-weighted MRI with contrast enhancement (T_{1c}), T_2-weighted MRI (T_2), T_2-weighted MRI with fluid attenuated inversion recovery (T_{2FLAIR}) as showed in figure 1.

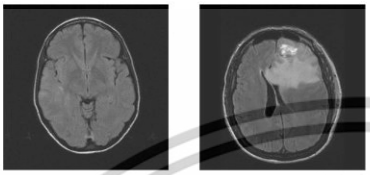
In recently, various research work has presented brain tumor diagnosis with good performance result. Hein Tun Zaw et al. presented segmentation brain tumor medical image from REMBRANDT database using maximum entropy threshold method to obtain tumor region. The morphology and intensity features were extracted for binary classification based on Naïve Bayes machine learning [3]. And for our previous study for binary classification brain tumor using traditional machine learning Support Vector Machine. The medical images form REMBRANDT dataset was used for global threshold segmentation and GLCM feature extraction and then feed in the tabulated dataset to SVM machine learning algorithm. The obtained result from our studied showed a good result of accuracy [4], Hossain et al. presented Brain tumor segmentation using Fuzzy C-means clustering algorithm and followed by using traditional machine learning classifier to compare the performance with CNN [5]. N.M. Dipu et al. presented brain tumor classification using YOLOv5 and FastAI deep learning with BraTS2018 database that contains 1,992 MRI images and obtained good accuracy result [6]. Arora & Sharma presented the binary classification brain tumor using various traditional machine learning and transfer learning network i.e., VGG16, InceptionV3 and ResNet [7]. Dubey et al. presented brain tumor detection using Stationary wavelet transform (SWT) and watershed algorithm for segmentation, GLCM for extract the feature and finally, neural network classification was implemented [8]. Gaikward et al. presented binary classification of brain tumor, the preprocessing and augmentation was conducted for dataset images before classify using transfer learning VGG16 [9].

II. Dataset description

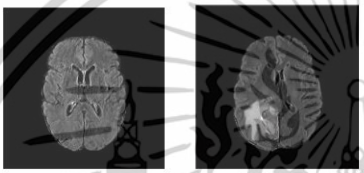
In this study, the dataset of brain tumor MRI images from 2 different sources were utilized. First dataset we used MRI images form Repository of Molecular Brain Neoplasia Data (REMBRANDT) that contains the MRI images from 130 patients with clinical data analysis in DICOM format [10]. Second dataset we utilized MRI images from Brain Tumor Segmentation (BraTS) 2021 that contains glioma brain tumor, with pathologically confirmed diagnosis. All MRI image from BraTS2021 is available as NIfTI file format and were exhibited with 4 modalities T_1-weighted MRI (T_1), T_1-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

weighted MRI with contrast enhancement (T_{1c}), T_2-weighted MRI (T_2), T_2-weighted MRI with fluid attenuated inversion recovery (T_{2FLAIR}) and were acquired with different clinical protocols and various scanners from multiple data contributing institutions [11-13]. In this study, the T_2-weighted MRI with fluid attenuated inversion recovery (T_{2FLAIR}) was used for binary classification. All MRI images were converted to JPEG images and illustrated in figure 2 and the image number from each dataset is further explained in table 1.



(a) FLAIR MRI image from REMBRANDT



(b) FLAIR MRI image from BraTS2021

Fig. 2 Example of MRI images from each dataset

TABLE I. THE OBTAINED MRI IMAGES FORM EACH DATABASE

Dataset	Normal Brain	Glioma Tumor
REMBRANDT	468	190
BraTS2021	103	1095

III. METHODOLOGY

In this study, we have proposed brain tumor classification using transfer learning pretrained network and the procedure is illustrated in block diagram figure 3. Initially, the converted MRI image in JPG format was performed image preprocessing with morphological operation for removal the skeleton and artifact from brain image to obtain brain region only. The contrast Limited Adaptive Histogram Equalization (CLAHE) performed for enhancement of the contrast of images.

The obtained dataset was heavily imbalanced. Hence, to handle this constraint the image augmentation was performed to increase the number of MRI images and create the balanced dataset. Finally, the training dataset was trained using various transfer learning with optimized parameter. The trained model was utilized for prediction brain tumor from testing dataset and the performance was evaluated to compare with each classifier.

A. Image preprocessing

In our dataset the MRI images have been taken from various sources. The difference of contrast is a normal practice. The preprocessing stage is necessary to normalize all images before classification and lead to the accuracy is significantly improved. We have implemented preprocessing stage as follows step; Morphological operation was performed to remove the skeleton, noises, and artifact and to obtain brain region only. And then CLAHE algorithm performed for enhancement the contrast of image. Augmentation was performed to increase the number of minority class.

A.1. Morphological operation

The opening morphological operation was performed for removal skeleton from brain region, which is an erosion followed by a dilation using same structuring element that is disk structuring for this study. The opening of set A by structuring element B is defined as:

$$A \circ B = (A \ominus B) \oplus B \quad (1)$$

The opening A by B is an erosion ($\ominus$) A by B and followed by a dilation ($\oplus$) of the obtained result by B [14-15]. Firstly, the skeleton was removed by opening morphological operation with disk structuring, and some pixels of the brain region were removed as well as figure 4a. Therefore, the pixel subtraction function is performed to separate the skeleton from the brain region. The original image was subtracted by the opened image in figure 4a and obtained the skeleton and some brain regions as figure 4b. To obtain the skeleton only the skeleton with some brain regions image that obtained in the previous step was subtracted by the opened image in figure 4a. Finally, the skeleton image was subtracted from the original image and obtained brain region only for further processing, the morphological operation procedure is illustrated in Figure 4.

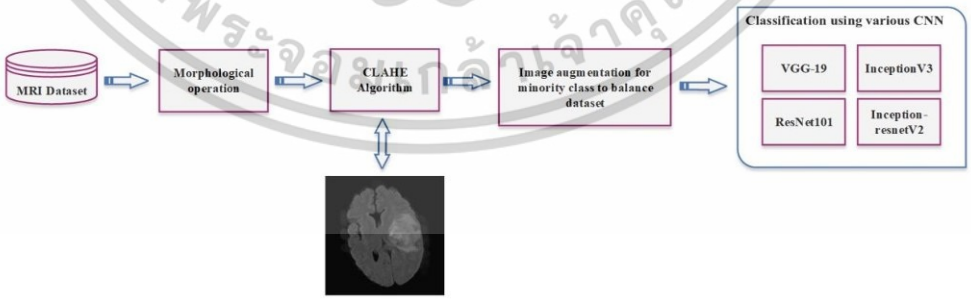


Fig. 3 Overall block diagram of proposed methodology for glioma brain tumor classification

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

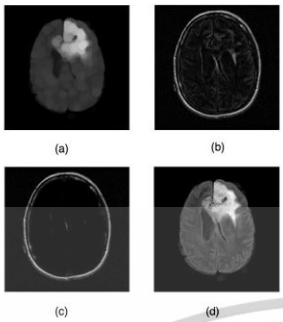


Fig. 4 Morphological operation preprocessing: (a) Opened image; (b) Original was subtracted by opened image; (c) Skeleton with some artifacts image was subtracted by opened image; (d) Separated brain region.

A.2. Contrast Limited Adaptive Histogram Equalization (CLAHE)

The CLAHE algorithm has been developed and improved from Adaptive Histogram Equalization (AHE) for contrast enhancement of medical image [16-17]. This study the dataset has taken from the different sources led to a variation contrast and intensity. Therefore, the contrast enhancement using CLAHE algorithm helps to improve and normalize the image before classification process.

The CLAHE algorithm comprises 3 major parts: tile generation, histogram equalization and bilinear interpolation. The Histogram equalization consists of five steps: histogram computation, excess calculation, excess distribution, excess redistribution and scaling and mapping using a cumulative distribution function (CDF) that is exhibited in equation 2.

$$S_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) \qquad (2)$$

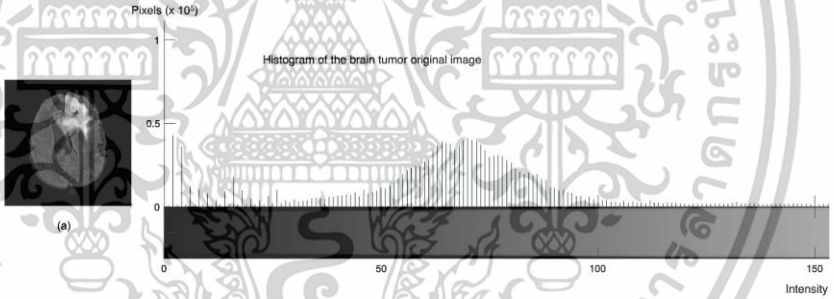
$$= \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \qquad k = 0, 1, 2, \ldots, L-1$$

Where T (r_k) is histogram equalization transformation, MN is the total number of pixels in the image, n_j is the number of pixels that have intensity r_j, L is the number of possible intensity levels in the image.

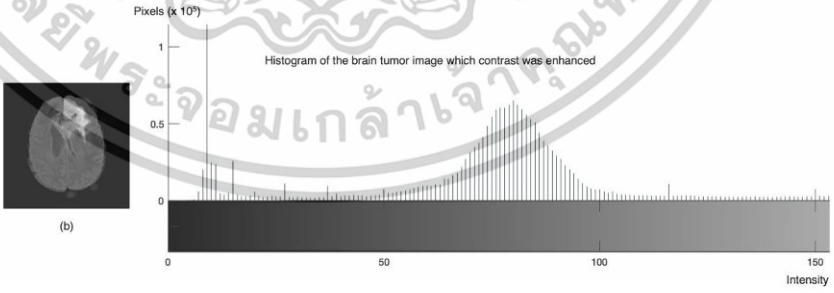
The original and contrasted enhancement image of the brain tumor medical image with the histogram for each image is illustrated in Figure 5.

A.3. Image Augmentation for minority class

In the past research work demonstrated that the distribution of the training data has a significant impact on the performance of CNN [18-19]. Hence, we have implemented image augmentation to increase number of the minority classes images for training model dataset.



(a) Brain tumor original image with histogram.



(b) The enhanced contrast brain image with histogram.

Figure 5. CLAHE algorithm to enhance contrast of brain MRI image.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The simple augmentation rotation and reflection were applied with the minority class, which is a normal brain image for our dataset. MRI medical images normal brain from REMBRANDT were augmented using rotation while normal brain MRI images from BraTS2021 that is the most minority were augmented using rotation and reflection method to increase number of training dataset, figure 6 showed image augmentation for each dataset.

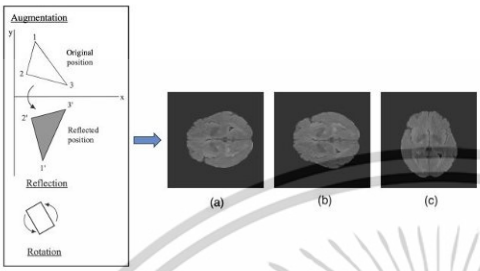


Figure 6. Image augmentation for normal brain minority class: (a) Original image; (b) Reflection around x axis; (c) Rotation 90 degree

B. *Various transfer learning pretrained network for classificaiton glioma brain tumor*

In this study the various pretrained CNNs network have been implemented for classification task. The advantage of transfer learning is the networks have learnt rich feature representations for a wide range of images. The network takes images as input and ground truth label for training, the feature is automatically extracted and finally outputs label is predicted. Fined tuning a network with transfer learning is faster and easier than training a network from scratch. The transfer learning network comprise with pretrained network layer, training dataset and algorithm options.

The convolution operation is typically denoted with an asterisk.

$$s(t) = (x*w)(t) \tag{3}$$

whereas x is often referred to as the input and the second argument was the kernel. The output is some time referred to as the feature map.

If we use a two-dimensional image I as input, we also want to use a two-dimensional kernel K [20].

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n)K(i-m, j-n) \tag{4}$$

In this study, the stochastic gradient descent with momentum (SGDM) is implemented for algorithm option set up and the initial learning rate is set to 0.0001. The SGDM algorithm is described in mathematic equation (5 – 7), where ϵ is the learning rate, $\boldsymbol{x}^{(i)}$ is the input image, $\boldsymbol{y}^{(i)}$ is the label and $\boldsymbol{\theta}$ is the weight. Firstly, compute gradient estimate.

$$g = \frac{1}{m} \nabla_\theta \sum_{i=1}^{m} L(f(x^{(i)}; \theta), y^{(i)}) \tag{5}$$

And then compute the updated velocity.

$$v \leftarrow \alpha v - \epsilon \nabla_\theta \left(\frac{1}{m} \sum_{i=1}^{m} L(f(x^{(i)}; \theta), y^{(i)}) \right) \tag{6}$$

Where $\boldsymbol{\alpha}$ is the momentum parameter.

$$\theta \leftarrow \theta + v \tag{7}$$

The velocity $\boldsymbol{v}$ accumulates the gradient elements g.

In this study, we have implemented 4 pretrained network for classification task. VGG19 [21], ResNet101 [22], InceptionV3 [23] and InceptionResNetV2 [24], the network properties for each network are described in table 2. All input images were converted to RGB type prior utilized for training by the algorithms.

TABLE II. THE CNN TRANSFER LEARNING PROPERTIES USED IN THIS STUDY.

CNNs	Layers	Parameters (x 10^6)	Input Size
VGG19	19	144	224 x 224
ResNet101	101	44.6	224 x 224
InceptionV3	48	23.9	299 x 299
InceptionResNetV2	164	55.9	299 x 299

IV. EXPERIMENTAL SETUP

The total combination of MRI image from 2 databases contains with 1,856 images comprising of brain tumor 1,285 images and normal brain 571 images were utilized for classification using various CNNs. The dataset was divided into training and testing with ratio 80% and 20% respectively, the brain tumor and normal images were equally distributed into training and testing dataset as well. Therefore, we obtained 1,484 images for training dataset that comprise of 1,028 tumor images and 456 normal brain images while testing dataset contains with 372 images that comprise of 257 tumor images and 115 normal brain images. The training dataset was divided into training and validation dataset with ratio 80% and 20% respectively before utilized for training using various CNNs. The dataset combination and management are illustrated in figure 7.

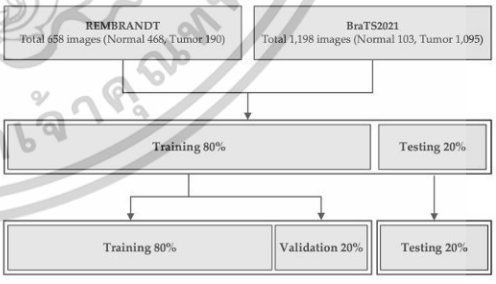


Figure 7. Dataset combination for brain tumor classification

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The dataset is heavily imbalanced. Hence, image augmentation with minority class was performed which is the normal brain image for our dataset. Finally, the balanced dataset is obtained, and the image number as follows; the normal brain 994 images and gliomas tumor 1,028 images for utilization in training network. The block diagram for training balanced dataset is illustrated in figure 8. In this study we performed preprocessing and classification using MATLAB R2023a program with image processing and deep learning toolbox.

TABLE III. EXPERIMENTAL RESULT OBTAINED FROM EACH CNN

CNNs	Accuracy	Sensitivity	Specificity
VGG19	97.58%	99.22%	93.91%
ResNet101	98.12%	98.05%	98.26%
InceptionV3	99.19%	98.83%	100%
InceptionResNetV2	98.66%	98.44%	99.13%

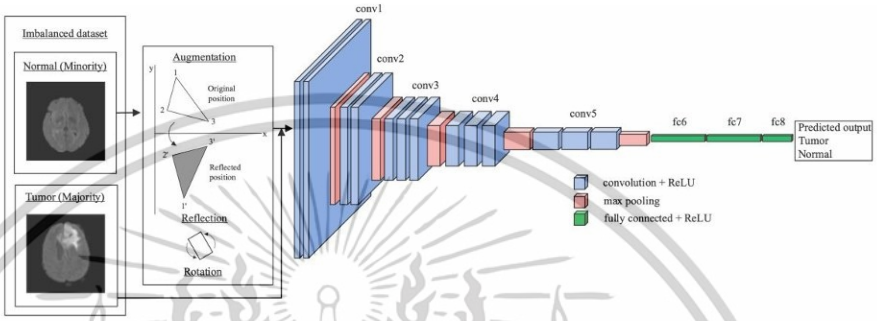


Figure 8. Block diagram for training balanced dataset.

V. EXPERIMENTAL RESULTS

The testing dataset contains 372 images that comprise 257 tumor images and 115 normal brain images. The attributes of the diagnostic tests can be measured. These attributes, namely sensitivity, specificity and accuracy have been widely used in statistics to describe a diagnostic test [25]. The parameters used for calculation the attributes compose with following,
True positive is tumor exist and was correctly detected.
True negative is normal brain and correctly identified.
False positive is normal brain but identified as tumor.
False negative is tumor exist but can't detected.
And the evaluated result is calculated as follows,

$$\text{Sensitivity} = TP / (TP + FN) \quad (8)$$

$$\text{Specificity} = TN / (TN + FP) \quad (9)$$

$$\text{Accuracy} = (TN + TP) / (TN+TP+FN+FP) \quad (10)$$

The obtained result from the proposed methodology was significantly improved when compared with previous work. We have implemented various CNNs for training datasets with algorithm option as explained in previous section. The training progress and confusion chart are plotted. The best classifier was InceptionV3 which obtained an accuracy of 99.19%, sensitivity of 98.83%, and specificity of 100% respectively. Another CNNs classifier showed experimental result in Table 3.

The proposed methodology has been compared to another past related work and shown in table 4. And found that the obtained accuracy of proposed method is almost perfect performance and can be implemented in practical work to assist the medical staff for diagnosis the glioma brain tumor.

TABLE IV. PERFORMANCE COMPARISION WITH PAST RELATED STUDY

Authors	Classification Methods	Accuracy
Zaw et al. [3]	Naïve Bayes classifier	Acc 94 % Sen 81.25% Spe 100%
Pongsak et al. [4]	SVM classifier	Acc 96.875% Sen 95% Spe 100%
Hossain et al. [5]	Fuzzy C-means with SVM and CNN	92.42%, 97.87%
Dipu et al. [6]	YOLOV5 and FastAi	85.95%,95.78%
Arora & Sharma [7]	VGG16 and ResNet	90.54%,81.92%
Dubey et al. [8]	Neural Network classifier	95%
Gaikwad et al. [9]	VGG-16 transfer learning	99.03%
Proposed method	Preprocessing techniques with various CNNs. InceptionV3 with balanced dataset was the best classifier	Acc 99.19% Sen 98.83% Spe 100%

As shown in Table 4, our proposed methodology demonstrated remarkable performance when compared to previous similar works. Zaw et al. presented brain tumor classification using the Naïve Bayes classifier, which achieved an accuracy of 94%, sensitivity of 81.25%, and specificity of 100%. Our previous research work used an SVM classifier for brain tumor classification and achieved an accuracy of 96.875%, sensitivity of 95%, and specificity of 100%. The results of another related work are also presented in Table 4. The InceptionV3 classifier with a balanced dataset performed the best in our methodology, achieving an accuracy of 99.19%, sensitivity of 98.83%, and specificity of 100%.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

VI. Conclusion

In this study, Glioma brain tumor MRI images from two different sources REMBRANDT and BraTS2021 were utilized for binary classification. The FLAIR MRI image was used to preprocess using morphological operation to remove skeleton and artifact to obtain brain region only and CLAHE algorithm to enhance and normalize the contrast of the image before classification using various CNNs. The balanced dataset using image augmentation with a minority class to handle imbalanced dataset problems helps to improve accuracy. The four types of transfer learning pre-trained networks were conducted for training and testing the balanced dataset. Finally, the best classifier is InceptionV3 which has achieved an accuracy of 99.19%, sensitivity 98.83%, and specificity 100% respectively, better than past related work. With the improved methodology and the superior experimental result, we envision that our proposed methodology will be implemented to assist medical staff in diagnosing glioma brain tumors using computer vision techniques in future work.

References

[1] S. Bauer, R. Wiest, L. P. Nolte, and M. Rayes, "A survey of MRI-based medical image analysis for brain tumor studies," Phys. Med. Biol., vol. 58, no. 13, pp. 97–129, 2013.

[2] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: A summary," Acta Neuropathol., vol. 131, no. 6, pp. 803–820, Jun. 2016.

[3] H. T. Zaw, N. Maneerat and K. Y. Win, "Brain tumor detection based on Naïve Bayes Classification," in Proc. 5th Int. Conf. on Engineering, Applied Sciences and Technology (ICEAST), Luang Prabang, Laos, 2019, pp. 1-4.

[4] P. Pilaoon, N. Maneerat, K. Yajima, R. Varakulsiripunth and K. Hamamoto, "Brain Tumor Classification using Supervised Support Vector Machine," in Proc. 9th Int. Conf. on Engineering, Applied Sciences, and Technology (ICEAST), Vientiane, Laos, 2023, pp. 79-83.

[5] T. Hossain, F. S. Shishir, M. Ashraf, M. A. A. Nasim, and F. M. Shah, "Brain Tumor Detection Using Convolutional Neural Network," in Proc. 1st Int. Conf. on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE, 2019.

[6] N. M. Dipu, S. A. Shohan, and K. M. A. Salam, "Deep Learning Based Brain Tumor Detection and Classification," in Proc. Int. Conf. on Intelligent Technologies (CONIT), Karnataka, India, 2021, pp. 1-6.

[7] S. Arora and M. Sharma, "Deep Learning for Brain Tumor Classification from MRI Images," in Proc. 6th Int. Conf. on Image Information Processing (ICIIP), 2021, pp.409-412.

[8] Y. K. Dubey, M. M. Mushrif, and K. Pisar, "Brain Tumor Type Detection Using Texture Features in MR Images," IEEE, 2018.

[9] S. Gaikwad, S. Patel, and A. Shetty, "Brain Tumor Detection: An Application Based on Machine Learning," in Proc. 2nd Int. Conf. for Emerging Technology (INCET), Belgaum, India, 2021, pp. 1-4.

[10] The Cancer Imaging Archive, 2017, REMBRANDT https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT

[11] U. Baid et al., "The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification," 2021, arXiv:2107.02314.

[12] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 1993-2024, 2015, DOI: 10.1109/TMI.2014.2377694.

[13] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features, Nature Scientific Data," 2017, DOI: 10.1038/sdata.2017.117.

[14] R. Gonzalez and R. Woods, "Digital Image Processing," 4th ed., Pearson Education, 2018.

[15] R. Gonzalez, R. Woods, and S. Eddins, "Digital Image Processing using MATLAB," 2nd ed., Pearson Education, 2017.

[16] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive Histogram Equalization and Its Variations," in Proc. Computer Vision, Graphics, And Image Processing, vol. 39, 1987, pp. 355-368.

[17] K. Zuiderveld, "Contrast limited adaptive histogram equalization," In Graphics Gems IV; Academic Press Professional Inc., Cambridge, 1994, pp. 474–485.

[18] I. Valova, C. Harris, T. Mai, and N. Gueorguieva, "Optimization of Convolutional Neural Networks for Imbalanced Set Classification," in Proc. 24th int. conf. KES, 2020, pp. 660-669.

[19] K. Y. Win, N. Maneerat, S. Sreng, and K. Hamamoto, "Ensemble Deep Learning for the Detection of COVID-19 in Unbalanced Chest X-ray Dataset," Appl. Sci., vol. 11, p. 10528, 2021.

[20] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning," The MIT press, 2017, pp. 294 -299.

[21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. 3rd Int. Conf. Learn. Represent., San Diego, CA, USA, May 2015.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[23] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 2818–2826.

[24] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in Proc. 31st AAAI Conf. Artif. Intell., San Francisco, CA, USA, 2017, pp. 4278–4284.

[25] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," NESUG proceedings: health care and life sciences, Baltimore, Maryland, Nov. 2010.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal
10 YEARS
OVER 10 YEARS OF QUALITY OPEN ACCESS PUBLISHING

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 February 2025, accepted 5 March 2025, date of publication 10 March 2025, date of current version 20 March 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3550012

RESEARCH ARTICLE

Glioma Brain Tumor Classification Using Convolution Neural Network and Majority Voting

PONGSAK PILAOON[1], KAZUHIKO HAMAMOTO[2], (Member, IEEE), AND NOPPADOL MANEERAT[1]

[1]School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[2]School of Information and Telecommunication Engineering, Tokai University, Tokyo 108-8619, Japan

Corresponding author: Noppadol Maneerat (noppadol.ma@kmitl.ac.th)

ABSTRACT Glioma brain tumors are malignant diseases for which early detection and instant treatment will increase the survival rate. Several studies have reported the efficiency of deep learning convolutional neural networks (CNN) in diagnosing brain tumors using magnetic resonance imaging (MRI). In this study, we investigated the potential of state-of-the-art classifiers to achieve the highest accuracy in the detection of brain tumors using MRI. For this purpose, we introduced a comparative study of eight state-of-the-art classifiers. The methodology comprised three different approaches: 1) an imbalanced dataset, 2) a balanced dataset using image augmentation, and 3) ensemble learning using the best of the top five models for majority hard and soft voting. The dataset comprised converted MRI data from the repository of molecular brain neoplasia data (REMBRANDT) and brain tumor segmentation 2021 (BraTS) databases. An increasing number of MRI images and datasets has prevented overfitting. Initially, a preprocessing stage morphological operation and contrast-limited adaptive histogram equalization (CLAHE) algorithms were used to remove skeletons and artifacts and optimize the image contrast for readiness classification. The stochastic gradient descent with the momentum algorithm option was used to train the network. The trained model was used to predict the testing dataset, and the results from each pretrained network were evaluated. The experimental results demonstrated that the prediction accuracy of the trained network was significantly improved using a balanced training dataset. The discriminative image region used to interpret the predicted result using the gradient-weighted class activation mapping (Grad-CAM) algorithm was proposed in the final stage for trustworthiness. The experimental results showed that the best approach was inceptionV3 with a balanced dataset. The accuracy, sensitivity, specificity, and area under the curve were 99.73%, 99.61%, 100%, and 1.00, respectively.

INDEX TERMS Glioma brain tumor, CLAHE, morphological operation, convolutional neural network, Grad-CAM, deep learning, majority voting.

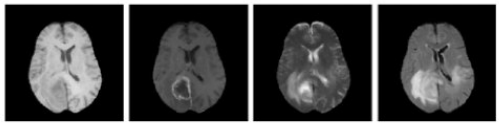
I. INTRODUCTION

With a prevalence of less than 1% in the Western population, brain tumors are not very common; however, they are among the most fatal cancers. A recent study estimated the incidence rate of primary brain tumors in the United States to be approximately 0.025%; one-third of the tumors are malignant and the rest are benign or borderline malignant. Brain tumors can be classified according to their source or aggressiveness. Primary brain tumors arise in the brain, whereas metastatic brain tumors originate and spread from other cancers in the body, such as lung, colorectal, and breast cancer. The World Health Organization (WHO) classifies brain tumors into four grades based on aggressiveness and disease severity.

The associate editor coordinating the review of this manuscript and approving it for publication was Abdullah Iliyasu.

© 2025 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a) T_1 (b) T_{1C} (c) T_2 (d) T_{2FLAIR}

FIGURE 1. MRI image's axial slice of a high-grade glioma.

Gliomas are the most common primary brain tumors in adults, constituting 70% of all malignant primary brain tumors arising from the glial cells. They were classified into four grades. High-grade gliomas are grades III and IV malignant tumors that lead to an end of life. Grade IV gliomas are also known as glioblastoma multiforme (GBM). These are the most rapidly growing malignant primary brain tumors. These tumors infiltrated and spread along the white matter fiber tracts. Abnormal vessels and necrotic cores are observed, and surrounding edema leads to a mass effect on healthy brain tissues [1], [2].

The average survival time of patients with GBM is one year. The treatment options for gliomas include surgery, radiation therapy, chemotherapy, or a combination of these.

Magnetic resonance imaging (MRI) is a standard non-invasive technique that is widely used in clinics for brain tumor diagnosis. To produce different types of tissue contrast, varying excitation and repetition times have been applied to create a versatile tool for imaging different structures of interest. In current clinical methods, different MRI sequences are used for the diagnosis of brain tumors. These sequences or modalities include T_1-weighted MRI (T_1), T_1-weighted MRI with contrast enhancement (T_{1c}), T_2-weighted MRI (T_2), and T_2-weighted MRI with fluid-attenuated inversion recovery (T_{2FLAIR}). Fig. 1 illustrates an axial slice of the four standard sequences for a patient with glioma.

T_1-weighted MRI was used for structural analysis and easy annotation of healthy tissues. On T_1-weighted MRI with contrast enhancement, the tumor borders appeared brighter and necrotic and active tumor regions were clearly distinguished. On T_2-weighted MRI, the edematous region surrounding the tumor appeared bright. T_{2FLAIR}(FLAIR) is a special sequence that helps to separate the edema region from the cerebrospinal fluid (CSF) by increasing the time to echo (TE) and repetition time (TR); thus, the brain tumor remains bright but normal CSF is attenuated and dark. This sequence is very sensitive to pathology and makes differentiation between CSF and an abnormality much easier.

Glioma brain tumor diagnosis by medical staff is time-consuming and should be performed with medical expertise, which is a constraint in upcountry or remote areas. Therefore, this study investigated the potential of state-of-the-art classifiers to achieve the highest accuracy in brain tumor prediction using MRI. We envision the development of accurate and effective tools to assist medical staff in classifying glioma brain tumors using artificial intelligence (AI) and computer vision techniques.

A. LITERATURE REVIEW AND ANALYSIS

Various studies on brain tumor segmentation and classification have recently been conducted using traditional machine and deep-learning algorithms, as described in Table 1. In this table, we enumerate the year of publication, methodology, datasets, performance, contributions, advantages, and disadvantages of each methodology, based on previous related studies.

Traditional machine learning and state-of-the-art deep learning methods have been applied for segmentation and classification in various studies. Traditional machine learning classification requires manual handcrafting for each stage of tumor segmentation and feature extraction before arranging the data in a tabulate for readiness to feed into each machine learning algorithm. Using this methodology, we can visually inspect the segmented tumor and select the appropriate extracted features to represent normal or tumor regions. Dataset training in a tabulate format does not require the high computer specifications. However, this method is not appropriate for applications with large datasets [3], [4], [5]. Advanced deep-learning techniques are being increasingly utilized in classification tasks. A significant advantage of deep learning is its capability to automatically extract features and segment brain tumors, a task in which traditional machine learning methods often fall short. Although convolutional neural networks (CNNs) have shown high accuracy in both segmentation and classification, they are frequently susceptible to overfitting. This problem often stems from the relatively small size of the datasets, which typically consist of images sourced from a single origin. [6], [7], [8], [9], [10], [11], [12], [13], [14]. Previous research has demonstrated impressive results in the segmentation and classification of brain tumors [15], [16], [17], [18], [19]. However, image preprocessing techniques such as morphological operations or intensity normalization may be necessary to enhance the accuracy of brain tumor detection. Furthermore, assessing the reliability of the predictions can provide a means to verify the effectiveness of the algorithm. Our proposed methodology includes preprocessing techniques that utilize morphological operations and the contrast-limited adaptive histogram equalization (CLAHE) algorithm to effectively separate the brain region from the skull and enhance image contrast. To improve accuracy, we implemented brain tumor classification using a variety of eight CNNs deep learning models along with an ensemble majority voting system based on the top five classifiers. Compared to traditional machine learning methods, our proposed approach is more suitable for large datasets of brain tumors. We aimed to prevent overfitting by combining the data from two different sources. Finally, we applied gradient-weighted class activation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

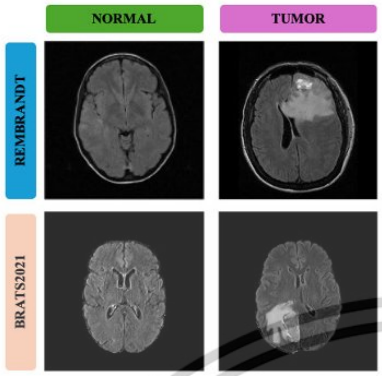


FIGURE 2. Example of MRI images from each dataset.

mapping (Grad-CAM) and local interpretable model-agnostic explanations (LIME) to visualize the regions affected by brain tumors, thereby enhancing the trustworthiness of our predictions.

II. DATASET DESCRIPTION

In this study, an MRI dataset was obtained from two sources. The first dataset used MRI images from the repository of molecular brain neoplasia data (REMBRANDT), which contained MRI images from 130 patients with clinical data analysis using digital imaging and communications in medicine (DICOM) format, and glioma grades III and IV were used for further processing [20]. The images were converted to a joint photographic expert group (JPEG) format with dimensions of $1178 \times 1177 \times 3$. In the second dataset, we utilized MRI images from brain tumor segmentation (BraTS) 2021, which contains brain glioma tumors with a pathologically confirmed diagnosis. All BraTS MRI scans are available in neuroimaging informatics technology initiative (NIfTI) file format and were exhibited with four modalities T_1, T_{1c}, T_2, and T_{2FLAIR} and were acquired with different clinical protocols and various scanners from multiple data contributing institutions [21], [22], [23]. The images were converted to the JPEG format with dimensions of $1152 \times 1152 \times 3$. The MRI image from BraTS was originally intended for the segmentation of the subparts of the brain tumor, that is, the tumor core, enhancing tumor, and complete tumor. In this study, T_{2FLAIR} modality MRI images were used to detect whole or complete tumors. The contrast between the normal brain and the tumor was distinguished, which led to easier classification. The converted JPEG images are shown in Fig. 2 and the image numbers and characteristics of each dataset are listed in Table 2. The number of original MRI images (without augmentation), training, and testing datasets are enumerated, the normal brain minority class is emphasized in bold letters, and the imbalance ratio is specified in this table.

III. METHODOLOGY

In this study, we proposed brain tumor classification using a deep learning CNN, as illustrated in the block diagram in Fig. 3. Initially, the MRI datasets were subjected to image preprocessing in morphological operations to remove the skeleton from the brain region, and artifact removal to retain the brain region only for further processing. Contrast limited adaptive histogram equalization (CLAHE) was used to balance the image contrast.

A heavily imbalanced image dataset is used in this study. To address this problem, we set up the following experimental scenarios: (i). Imbalanced dataset for training using various CNN, and (ii). Image augmentation was performed to increase the number of minority classes in the normal brain in our dataset, and the model was trained using various CNN. Finally, gradient-weighted class activation mapping (Grad-CAM) was performed to visualize the discriminative image regions for ease of understanding and interpretation. (iii). We implemented ensemble majority voting to investigate and improve the performance of brain tumor prediction. The top five prediction models were used for the majority of the hard and soft votes.

A. IMAGE PREPROCESSING

We used datasets from various sources. The difference in contrast is a normal practice, and the images also contain skeletons, noise, and artifacts that need to be removed. Therefore, image preprocessing is a prerequisite before being fed into a CNN for precise classification. We implemented image preprocessing to normalize and enhance the quality of all images as follows. A morphological operation was performed to remove skeleton, noise, and artifacts and obtain only the brain regions. Subsequently, a contrast-limited adaptive histogram equalization (CLAHE) algorithm was used for image contrast enhancement. Image preprocessing is illustrated in Fig. 4.

1) MORPHOLOGICAL OPERATION

The opening morphological operation was performed to remove the skeleton from the brain region, which is an erosion followed by dilation, using the same structuring element as the disk structuring for this study. The opening of set A by structuring element B is defined as follows:

$$A^{\circ}B = (A \Theta B) \oplus B \tag{1}$$

Opening A by B is an erosion (Θ) A by B followed by a dilation ($\oplus$) of the result obtained by B [24], [25]. First, the skeleton was removed using an opening morphological operation, and pixels in the brain region were removed. Therefore, a pixel subtraction function is used to separate the skeleton from the brain region. The open image was subtracted from the original image to obtain the skeleton and brain regions, as shown in Fig. 5b. To obtain only the skeleton, the open image (Fig. 5a) was subtracted from the skeleton with some of the brain region images obtained in the previous step. Finally, the skeleton image was subtracted from the

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 1. Literature review.

Authors	Classification Methods	Dataset (image)	Accuracy	Advantages / Contribution	Disadvantages
Zaw et al. [3] Year 2019	Morphological operation to separate the brain area from the skull. The maximum entropy method was applied to enhance and obtain brain tumor images in the segmentation stage. Many morphology and intensity features were extracted from the segmented tumor area. The extracted features matrix was trained and predicted using the Naïve Bayes classifier.	REMBRANDT Tumor 90 Normal 24	Accuracy 94% Sensitivity 81.25% Specificity 100%	Developing a method to detect the brain tissues that are affected by cancer, especially for grade-4 tumor, Glioblastoma Multiforme (GBM). When compared to other methods, the significant advantage of this method is can properly detect the tumor located in different regions of the brain including the middle region which aligned with eye level that are often detected as an eye.	Segmentation tumor area and feature extraction performed by manual handcraft for each stage. Implementation with small size of the dataset which is susceptibility to overfitting.
Pilaoon et al. [4] Year 2023	This research was further developed from the previous work of Zaw et al. [3] to improve the accuracy. The methodology was composed of, preprocessing with the morphological operation to separate the brain region from the skull image and a global thresholding method was implemented for tumor segmentation. The features were extracted using a gray level co-occurrence matrix (GLCM) and a support vector machine (SVM) classifier was conducted to predict brain tumors.	REMBRANDT Tumor 111 Normal 44	Accuracy 96.875% Sensitivity 95% Specificity 100%	Proposed the automated diagnosis system to early detect GBM brain tumors. The accuracy results were significantly improved and better than the previous work. The traditional machine learning SVM was conducted to classify the represented brain tumor or normal in tabulated format. This methodology is simple and effective for manual segmentation and feature extraction. The normal computer specification is adequate for the machine learning to classify the data based on tabulated format.	Segmentation of tumor area and feature extraction were performed by manual handcraft for each stage which was time-consuming and not appropriate to implement with the large dataset. This research was implemented with the small size of the dataset which is susceptible to overfitting.
Dubey et al. [5] Year 2018	Proposed morphological and thresholding method for automatic skull stripping of the brain image. After that tumors were segmented using stationary wavelet transform (SWT), self-organizing map (SOM), and watershed algorithm. The features were extracted using a gray level co-occurrence matrix (GLCM) algorithm and were fed into feed forward neural network classifier.	Real MRI images from the Hospital, Meningioma 20, and Glioma 15 images	Accuracy 95%	Manual segmentation and feature extraction were clearly explained and applied with real MRI images from the hospital. The two tumor type meningioma and glioma were successfully classified. The proposed system can be useful for the classification of other tumor type such as lymphoma, astrocytoma, and medulloblastoma.	Segmentation of tumor area and feature extraction were performed by manual handcraft for each stage which was time-consuming and not appropriate to implement with the large dataset. This research was implemented with the small size of the dataset which is susceptible to overfitting.
Pilaoon et al. [6] Year 2023	Proposed binary classification of GBM brain tumor using convolutional neural network namely GoogLeNet and AlexNet. The actual image label was created and the stochastic gradient descent with momentum algorithm option with an initial learning rate of 0.0001 was used for the training dataset. Then the trained model was used for the prediction of the testing dataset whether a tumor or a normal brain.	REMBRANDT, Tumor 111 Normal 44	GoogLeNet, Accuracy 80.85% Sensitivity 93.94% Specificity 50% AlexNet, Accuracy 93.62% Sensitivity 96.97% Specificity 85.71%	Manual segmentation and feature extraction were automatically performed by a transfer learning network that reduced the error from human handcraft. This methodology can be implemented in practice to help the medical staff with earlier diagnosis for further treatment and increase the survivor rate of patients.	Brain tumor classification based on CNN, the raw image directly feeds into the network, and features are automatically extracted by the network. Hence, every image pixel needs to be analyzed which leads to high computer specification consumption. It seems that implemented with a small dataset which is susceptible to overfitting.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 1. ***(Continued.)*** **Literature review.**

Authors	Classification Methods	Dataset (image)	Accuracy	Advantages / Contribution	Disadvantages
Hossain et al. [7] Year 2019	Proposed brain tumor segmentation by fuzzy c-means clustering algorithm (FCT) and classification using different traditional machine learning, i.e., K-Nearest Neighbor, Logistic Regression, Multilayer Perceptron, Naïve Bayes, Random Forest, and Support Vector Machine. And CNN was implemented to improve the brain tumor prediction results.	BraTS dataset, Tumor 187 Normal 30	SVM, Accuracy 92.42% Sensitivity 98.3% Specificity 42.8% CNN, Accuracy 97.87%	Comparison of the accuracy of brain tumor prediction using the various traditional machine learning and CNN. In the traditional classifier part, the SVM classifier obtained the highest accuracy. Moreover, for better results, this research implemented CNN which brought good accuracy.	It seems that the dataset is small size which leads to overfitting occurrence. And the obtained accuracy results need to be further improved such as poor specificity parameter result because of the imbalanced dataset.
Dipu et al. [8] Year 2021	This paper proposed two deep learning-based approaches for brain tumor detection and classification using the cutting-edge object detection framework You Only Look Once (YOLOv5) and the deep learning library FastAi. The training directory that contained tumor and normal brain was created and utilized for training using the Google colab environment. Then the trained model was used for predicting brain tumors.	BraTS2018, 1,992 MRI images	FastAi, Accuracy 95.78% Precision 96.70% Recall 95.65% YOLOv5, Accuracy 85.95% Precision 92.67% Recall 84.33%	Comparison between two different deep learning-based approaches with good accuracy result. A large number of datasets were used for training networks with high computer specifications which led to good training and testing inference time. These two models can be applied in real-time brain tumor detection for early diagnosis of brain cancer.	Contained MRI images from BraTS 2018 single dataset only and this system will provide much better results if it were to be trained on a larger dataset for a more extended period. And the preprocessing techniques need to be implemented such as morphological operations or intensity normalization.
Arora & Sharma [9] Year 2021	Brain tumor classification using traditional machine learning. The features were extracted using the gray level co-occurrence matrix (GLCM) algorithm. The benign and malignant brain tumor was classified using SVM and Random Forest classifier. And comparative analysis with CNN transfer learning, i.e., VGG16, InceptionV3, and ResNet transfer learning.	BraTS2015, 3,764 MRI images	SVM 50.51% Random Forest 64.23% VGG16 90.54% InceptionV3 63.94% ResNet 81.92%	Brain tumor classification using different approaches with a large MRI image dataset. The work can be extended to the classification of brain tumors into different levels of malignancy. Transfer learning models are quite helpful in training medical images with very good accuracy.	It seems that overfitting problem occurrence and the obtained accuracy results needs to be further improved, and another parameter need to be evaluated such as specificity and sensitivity.
Gaikwad et al. [10] Year 2021	Preprocessing and augmentation techniques were performed before brain tumor classification using VGG-16 transfer learning. The architecture was based on CNN along with the transfer learning model VGG-16. The proposed system contained a dataset from Kaggle which was then preprocessed followed by augmentation and passed to the architecture.	Brain MRI images on Kaggle, Tumor 155 Normal 98	Accuracy 99.03%	Image augmentation was implemented for increasing and balancing the dataset before training the model which led to good accuracy obtained. The objective of this paper was to build an application of CNN by proposing a medical expert-based approach to detect brain tumor.	This study reported that time-consuming to train the model assuming that computer specification was a constraint. And another parameter needs to be evaluated such as specificity and sensitivity.
Kumar Baranwal et al. [11] Year 2020	Proposed system to classify the brain tumor images into three sub-types: Meningioma, Glioma, and Pituitary using SVM and customized CNN. Raw images are labeled and pre-processed to enhance the performance and robustness of	Nanfang and Tianjing Medical University Hospital, Meningioma : 708 Glioma : 1,426	Concluded that CNN was the best classifier, Accuracy, Meningioma: 94.12% Glioma :	The system for automatic classification of the brain tumor type based on input MRI was created. Multiclass brain tumor classification was proposed with good performance results. Three sub-types of brain tumors: Meningioma, Glioma, and	In this study, the customized CNN classifier with only one network type was proposed. Preprocessing technique to resize images from 512 x 512 to 384 x 384 pixels only. And the other

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 1. ***(Continued.)*** **Literature review.**

Authors	Classification Methods	Dataset (image)	Accuracy	Advantages / Contribution	Disadvantages
	the model. Features were extracted using the gray level co-occurrence matrix (GLCM) algorithm before being fed into linear and polynomial SVM classifiers. Certain hyper-parameters in CNN are tuned to make the model more efficient and finally model is trained and tested.	Pituitary : 930	94.3% Pituitary : 98.85% Recall, Meningioma: 0.91 Glioma : 0.92 Pituitary : 0.99	Pituitary were classified using machine learning and CNN. Many parameters were evaluated for the performance of brain tumor classification, i.e., accuracy, precision, and recall.	preprocessing techniques need to be further implemented such as morphological operations or intensity normalization.
Aboussaleh et al. [12] Year 2021	The proposed system was divided into two phases. Firstly, The MRI image was classified into a tumor or normal by a CNN model with two neurons in the output layer, the ground truth label as tumor or normal was used in this stage. Secondly, the segmentation was applied to the images that contained the tumor using the extracted features from the last convolution layer of the proposed CNN architecture and their gradients. Finally, post-processing was applied to improve the results.	BraTS2017, High-grade glioma (HGG) 210 images Low-grade glioma (LGG) 75 images. With data augmentation the dataset was increased to 1100 labeled images, 550 for each label.	Binary classification accuracy for training (HGG+ LGG) 98%, accuracy for validation (HGG+ LGG) 91%, and segmentation with post processing (HGG+ LGG) obtained dice scores 82.35%	The strength of this approach is the intervention of a specialist to manually locate the tumor pixel by pixel because it is a complex and time-consuming task. This methodology solves these problems by using the extracted features from the CNN architecture and independently of the ground truth detected manually by specialists. The experimental results show good performance and a significant result when compared with the past related works.	This research proposed brain tumor segmentation with good performance results. However, seems that the image dataset before augmentation was a small dataset and taken from a single source only which is prone to overfitting.
Ayesha Younis et al. [13] Year 2022	Presented a binary classification of brain tumors using CNN and could be trained for faster and more convenient training using a pretrained VGG16 model. The ensemble learning was also implemented to improve the performance of brain tumor prediction.	Brain MRI images for Brain Tumor Detection Tumor 155 Normal 98	CNN 96% VGG16 98.5% Ensemble 98.14%	Various types of deep learning and ensemble learning were used for brain tumor classification. The proposed model integrated deep learning and transfer learning models to achieve a remarkable accuracy rate. The study successfully yielded better outcomes while requiring less computing time.	The preprocessing techniques need to be further implemented such as morphological operations or intensity normalization. The dataset is quite small and tends to have overfitting problems.
Hanming Hu et al. [14] Year 2021	Brain tumor detection using various CNNs to test their efficiency, including VGG16/19, AlexNet, GoogLeNet, and ResNet, and implemented a YOLO to snip the tumor's exact location. Preprocessing involves rotation, flipping, and gamma correction and is also manually labeled to train a YOLO model for detecting the tumor location. Then with the labels, two different graphic enhancement techniques were applied for cropping and color differential.	Brain Tumor Classification (MRI) dataset	ResNet50 76.904% GoogLeNet 75.38% AlexNet 74.365% VGG16 77.157% VGG19 76.396% Average 76.105%	Various CNNs were implemented for the detection of brain tumors and a web-based graphic user interface was introduced. The development of artificial intelligence in health care, and deep learning has much room to grow in the future and can further contribute to the complex medical problems of the future.	The obtained accuracy needs to be improved, and another parameter needs to be evaluated such as specificity and sensitivity. It seems that this study was implemented with a small dataset that led to overfitting occurrence. Another preprocessing technique needs to be applied such as a morphological operation to separate the skull from the brain region before feeding into the classifiers.
Pereira et al. [15] Year 2016	This paper was inspired by the work of Simonyan and Zisserman who developed	BraTS2013, BraTS2015 challenge	Dice similarity coefficient (DSC) for the	Proposed a novel CNN-based method for segmentation of brain tumors in MRI images	The proposed method was inspired by the VGG architecture. However, a

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 1. ***(Continued.)*** **Literature review.**

Authors	Classification Methods	Dataset (image)	Accuracy	Advantages / Contribution	Disadvantages
	VGG transfer learning but this work investigate the potential of using deep architectures with small convolutional kernels for the segmentation of gliomas. The CNN with small 3x3 kernels was proposed for automatic brain tumor segmentation. And also investigated the use of intensity normalization as a preprocessing stage that aims to address data heterogeneity caused by multi-site and scanner acquisitions of MRI images.	dataset	complete, core and enhancing regions respectively. BraTS2013 0.88, 0.83, and 0.77 BraTS2015 0.78, 0.65, and 0.75	with a small 3x3 kernels CNN. The intensity normalization preprocessing technique influenced to achievement of good segmentation. In the BraTS challenge 2013, the proposed method was ranked in the first position in the DSC metric in the complete, core, and enhancing regions segmentation. And able to reduce the computation time approximately ten-fold.	significant issue with this type of network is the vanishing gradient problem. Many architectures have been developed to address this issue, resulting in improved performance, such as ResNet, Inception, and DenseNet.
Hareem Kibriya et al. [16] Year 2021	Presented multiclass brain tumor classification using deep learning and machine learning techniques. First, the MRI images were classified via CNN models, i.e., ResNet18 and GoogLeNet. The extracted features from the CNN models are also classified using SVM. MRI Images were preprocessed based on the min-max normalization technique and augmented against imbalanced dataset issues.	Public available Figshare dataset, Meningioma : 708 Glioma : 1,426 Pituitary : 930	GoogLeNet + SVM Accuracy 97.6% Precision 97.3% Recall 97.3% ResNet18 + SVM Accuracy 98% Precision 98.3% Recall 98%	Image preprocessing and augmentation were performed to normalize and balance the dataset before being fed into CNN transfer learning. SVM-classified deep feature vectors from the dedicated CNNs were applied and it was observed that the proposed system's accuracy increased when SVM was used instead of the softmax layer of CNN.	Another preprocessing technique needs to be applied such as a morphological operation to separate the skull from the brain region before feeding into the classifiers. The visualization and trustworthiness of classification need to be further investigated.
Ahmed Wasif Reza et al. [17] Year 2023	Propose the multiclass brain tumor classification using modified VGG16 architecture with 3 different classes. Preprocessing stage by resizing, orienting, and coloring before training. The proposed model is composed of 21 layers, where the first 20 layers use the rectified linear unit (ReLU) function, and the last one is the SoftMax function. And used dropout layers to reduce overfitting, the model applied 20% dropout which gave the most accurate values.	Public available Kaggle and Figshare Glioma : 2,547 Meningioma : 2,582 Pituitary : 2,658 No tumor : 2,396	Glioma : 98.98% Meningioma: 99.13% Pituitary : 99.95% No tumor : 99.81% Overall : 99.5%	The proposed model was implemented with a large number of real datasets without augmentation and obtained remarkable results when compared to other related work. Using a modified VGG-16 architecture on 10,153 MRI images with 3 different classes, the network performs significantly well.	The significant limitation of this model is that it is comparatively slow as it contains high numbers of parameters. Additionally, this model cannot detect the exact location of the tumor.
S. Asif et al. [18] Year 2022	The MRI images were preprocessed (cropping, resizing, splitting, and normalization) and then data augmentation was used to increase the size of the dataset. The various deep transfer learning (Xception, NasNet Large, DenseNet121, and InceptionResNetV2) was implemented for detecting brain tumors with different optimizers, i.e., adaptive moment estimation (ADAM), root mean squared propagation (RMSprop), and stochastic gradient descent (SGD).	Brain Tumor Detection (large) Tumor 1500 Normal 1500 Brain MRI Images Dataset (small) Tumor 155 Normal 98	MRI large dataset with Xception and ADAM optimizer was better than the other, Accuracy 99.67% Sensitivity 99.68% Specificity 99.66%	Demonstrated the effectiveness of brain tumor classification using various CNN and different optimizers and comparison results between small and large datasets for robustness testing. The MRI images were improved during the preprocessing phase and various techniques such as data augmentation, three different optimizers (ADAM, SGD, and RMSprop), and the L2 regularizer were used to improve performance.	Another preprocessing technique needs to be applied such as a morphological operation to separate the skull from the brain region before feeding into the classifiers. The transparency and trustworthiness of prediction shall be further implemented such as Grad-CAM for discriminative image regions.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 1. *(Continued.)* **Literature review.**

Authors	Classification Methods	Dataset (image)	Accuracy	Advantages / Contribution	Disadvantages
S. Mohsen et al. [19] Year 2023	Presented glioma and pituitary brain tumors classified using ResNext101_32×8d and VGG19. A single-image super-resolution (SISR) technique was applied to the MRI images to improve their resolution before classification. The SISR consists of two phases: generator and discriminator. The ResNext101_32 ×8d model consists of 344 layers including 104 batch normalization layers, 104 conv-layers, 100 ReLU layers, 33 bottleneck layers, a single max-pooling layer, a single adaptive average layer, and one linear layer.	Brain Tumor dataset on Kaggle, Glioma : 900 Pituitary : 900	VGG19 Accuracy 99.98% Precision 99.90% Recall 99.89% ResNext101 Accuracy, Precision, and Recall was all 100%	The SISR technique based on a generative adversarial network (GAN) algorithm was implemented to improve the resolution and led to obtaining the impressive accuracy result. The Models could assist specialist doctors by providing a fast identification of patients with brain tumors, which makes these models useful tools for rapid screening and providing support for medical diagnoses.	The preprocessing MRI images shall be implemented to separate the brain region from the skeleton. The detected tumor with discriminative image region methodology shall be considered.
Our proposed method	Preprocessing techniques with various 8 CNNs and ensemble majority voting. Finally, InceptionV3 with a balanced dataset is the best classifier.	REMBRANDT and BraTS2021 dataset, Tumor 1285 Normal 571	Accuracy 99.73% Sensitivity 99.61% Specificity 100% Area under the ROC curve (AUC) 1.00	The performance was significantly improved and implemented Grad-CAM algorithm for trustworthiness.	In future work, the transformer-based models for the 3D medical images approach will be investigated to improve the performance.

TABLE 2. Converted MRI image to JPEG format from each dataset.

Dataset	MRI image (original)		Training (80%)		Testing (20%)		Remark
	Normal	Glioma	Normal	Glioma	Normal	Glioma	
REMBRANDT, Dimension 1178 x 1177 x 3	468	190	**374**	152	94	38	MRI images from two institutes, Thomas Jefferson University and Henry Ford Hospital. The tumor grades were justified by information from 3 radiologists and many factors were considered such as tumor location, enhancement quality, lesion size, and etc. In this study, we used malignant glioma brain tumor grades III and IV for proceeding.
BraTS2021, Dimension 1152 x 1152 x 3	103	1095	**82**	876	21	219	MRI image from this dataset was originally intended for segmentation tasks, but in this study, we intend to use FLAIR image type for classification of the whole tumor or complete tumor that contains tumor core and enhancing tumor inside.
Total	571	1285	**456 (1:2.25)**	1028	115	257	Normal brain is a minority class with ratio 1:2.25

original image and the brain region was obtained only for further processing. The morphological operation procedure is illustrated in Fig. 5.

MRI images from the REMBRANDT database were subjected to morphological preprocessing to obtain only the brain regions, as per the aforementioned steps. MRI images from the BraTS2021 database were similarly performed to remove some artifacts, and brain regions were obtained only because MRI images from BraTS2021 were provided without a skeleton.

2) CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALIZATION (CLAHE)

The CLAHE algorithm was developed and improved from adaptive histogram equalization (AHE) for contrast enhancement of medical images [26], [27]. In this study, the dataset

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

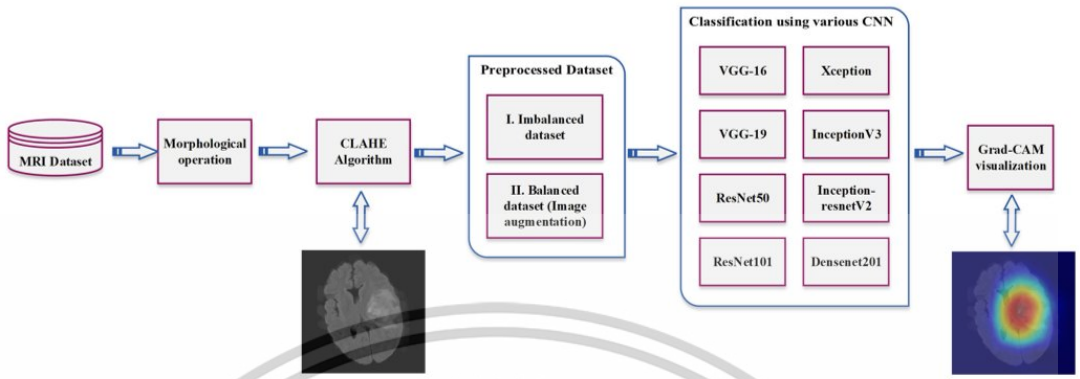


FIGURE 3. Overall block diagram of proposed methodology.

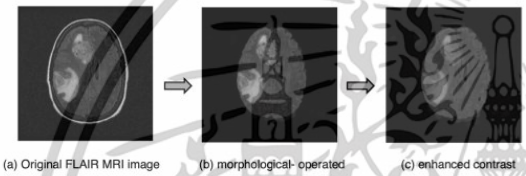


FIGURE 4. Image preprocessing for removal the artifacts and contrast enhancement.

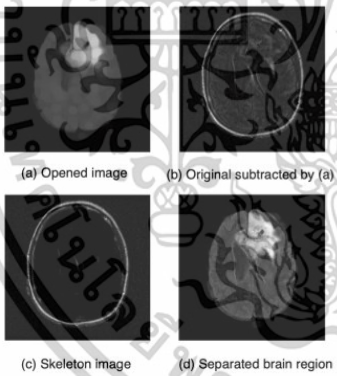


FIGURE 5. Morphological operation preprocessing.

was obtained from different sources, which led to variations in the contrast and intensity. Therefore, contrast enhancement using the CLAHE algorithm improved and normalized the image before the classification process. The CLAHE algorithm comprises of three major parts: tile generation, histogram equalization, and bilinear interpolation. Histogram equalization comprises five steps: histogram computation, excess calculation, excess distribution, excess redistribution, and scaling and mapping using a cumulative distribution function (CDF), as shown in Equation 2.

$$S_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \quad k = 0, 1, 2, \ldots, L-1 \tag{2}$$

where $T(r_k)$ is the histogram equalization transformation, MN is the total number of pixels in the image, n_j is the number of pixels with intensity r_j, and L is the number of possible intensity levels in the image.

The original and contrasted enhancement images of the brain tumor medical image, with a histogram for each image, are illustrated in Fig. 6.

3) IMAGE AUGMENTATION FOR BALANCING THE TRAINING DATASET

In this study, the image dataset was heavily imbalanced, and previous studies demonstrated that the distribution of the training data has a significant impact on the performance of the CNN. Hence, we implemented image augmentation to increase the number of minority class images in the training dataset [28], [29], [30]. Simple augmentation rotation, reflection, synthetic noise, and synthetic blur were applied to the minority class, which was a normal brain image for our dataset. Medical MRI images of the normal brain from REMBRANDT were augmented using rotation, whereas normal brain MRI images from BraTS2021, which was the minority, were augmented using the rotation and reflection method to increase the number of training datasets.

The selected images from both datasets were further subjected to synthetic noise and blur augmentation to equalize the normal and brain tumors of the training dataset, the details of which are listed in Table 3. In this study, image augmentation was performed to increase and balance the training dataset, which is adequate for the procedure.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

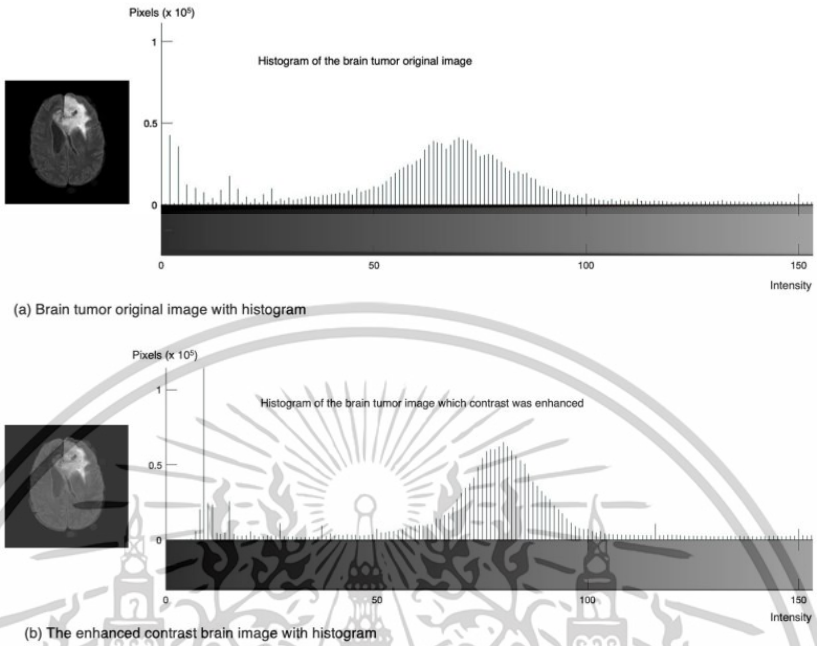


FIGURE 6. CLAHE algorithm to enhance contrast of brain MRI image.

B. VARIOUS DEEP LEARNING PRETRAINED CNNS FOR CLASSIFICATION BRAIN TUMOR

In this study, various pretrained CNNs were implemented for classification tasks. The advantage of transfer learning is that networks can learn rich feature representations from a wide range of images. The network takes images as input and ground truth labels for training, the feature is automatically extracted, and the output label is predicted. Fine-tuning a network using transfer learning is faster and easier than training from scratch. The transfer learning network comprises a pretrained network layer, training dataset, and algorithm options.

The convolution operation is denoted by an asterisk (*).

$$s(t) = (x^* w)(t) \tag{3}$$

where x is often referred to as the input and the second argument is the kernel. The output is the some-time, which is referred to as a feature map.

If we use a two-dimensional image I as the input, we also want to use two-dimensional kernel K.

$$\begin{aligned} S(i,j) &= (I * K)(i,j) \\ &= \sum_m \sum_n I(m,n) K(i-m, j-n) \end{aligned} \tag{4}$$

In this study, stochastic gradient descent with momentum (SGDM) with an initial learning rate of 0.0001 was set up as the algorithm option [31]. The stochastic gradient descent with momentum (SGDM) is described in equations (5), (6), and (7), where $\in$ is the learning rate, $\boldsymbol{x}^{(\mathbf{i})}$ is the input image, $\boldsymbol{y}^{(\mathbf{i})}$ is the label, and $\boldsymbol{\theta}$ is the weight. Firstly, compute gradient estimate.

$$g = \frac{1}{m} \nabla_\theta \sum_{i=1}^{m} L\left(f\left(x^{(i)}; \theta\right), y^{(i)}\right) \tag{5}$$

And then compute the updated velocity.

$$v \leftarrow \alpha v - \in \nabla_\theta \left(\frac{1}{m} \sum_{i=1}^{m} L\left(f\left(x^{(i)}; \theta\right), y^{(i)}\right) \right) \tag{6}$$

where α is the momentum parameter. And the updated weight is computed.

$$\theta \leftarrow \theta + v \tag{7}$$

The velocity $\boldsymbol{v}$ accumulates the gradient elements g, $(\boldsymbol{v} = \alpha v - \in g)$.

We implemented state-of-the-art pretrained network layers, viz., VGG16, VGG19 [32], ResNet50, ResNet101 [33], Xception [34], InceptionV3 [35], InceptionResNetV2 [36], and Densenet201 [37]. The properties of each network are listed in Table 4. All input images were converted to red, green, and blue (RGB) type before being used for training by the algorithms.

In this study, we investigated brain tumor prediction using state-of-the-art transfer learning methods. Recently, several

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 3. Image augmentation for normal brain minority class.

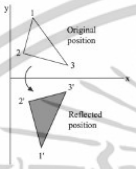
Augmentation	Characteristic	Augmented MRI image	Dataset applicable	
			REMBRANDT	BraTS2021
Normal brain for training (original)	-		374	82
Rotation 90 degree			374	82
Reflection around x axis			-	82
Synthetic noise			10	7
Synthetic blur			10	7

TABLE 4. The CNN transfer learning properties used in this study.

CNNs	Layers	Parameters (x 10^6)	Input Size
VGG16	16	138	224 x 224
VGG19	19	144	224 x 224
ResNet50	50	25.6	224 x 224
ResNet101	101	44.6	224 x 224
Xception	71	22.9	299 x 299
InceptionV3	48	23.9	299 x 299
InceptionResNetV2	164	55.9	299 x 299
Densenet201	201	8.1	224 x 224

CNNs transfer learning methods have been introduced and demonstrated superior results in medical image classification. We intended to use CNNs, which are the latest architectures, or obtained very good accuracy in a previous related work. The details of each CNN implemented in this study are as follows:

1) VISUAL GEOMETRY GROUP (VGG16, VGG19)

The Visual Geometry Group is a section of the Science and Engineering Department of Oxford University. VGG has been applied in facial recognition and image classification, including VGG16 and VGG19. The researcher intended to understand how the depth of the CNN network layer affects classification accuracy. To avoid a large number of parameters, small CNN kernels (3×3) are applied to each layer. The CNN comprises an input layer, output layer, and many hidden layers. VGG16 is a CNN that has been accepted for its excellent performance in computer vision applications. This characteristic is determined by the depth of the layer, which comprises 16 layers, i.e., convolutional 13 layers, max pooling 5 layers, and fully connected 3 layers. A significant improvement in VGG is that the convolution kernel is reduced to 3×3 and the convolution layer is increased. VGG19 is improved from the VGG model and comprises 19 layers and a structure similar to VGG16, but the difference is that the depth of the layer is increased to improve the classification accuracy [32].

2) RESIDUAL NETWORK (RESNET50, RESNET101)

The researcher understood that if we increase the depth of the network layer, the accuracy would improve accordingly. However, when the layer of the network is increased to a limit point, its performance is reduced, which is a bottleneck for the VGG model because we cannot create the network layer as per the requirement because it will lose some

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

performance. A significant problem that ResNet can resolve is the vanishing gradient that occurs when the network layer has a depth greater than the limit by creating a shortcut in the network, which leads to a residual network structure [33].

3) XCEPTION

The Xception convolutional neural network was developed by Google researchers and modified from InceptionV3 to improve classification performance. It was assumed that the mapping of cross-channel correlations and spatial correlations in the feature map of the CNN could be separated (entirely decoupled) because of this assumption as a concentrated version of inception. Thus, Xception is abbreviated from extreme inception based on depthwise separable convolution layers (a depthwise convolution followed by a pointwise convolution). In conclusion, the Xception architecture is a linear connection of depth-wise separable convolution layers combined with a residual network layer to troubleshoot the vanishing problem [34].

4) INCEPTIONV3

InceptionV3 was developed using InceptionV1 (GoogLeNet). The concept is to use multiple filters of different sizes within the same layer, instead of simply increasing the layer depth. The layer is connected in parallel, which leads to an increase in network width. The performance of InceptionV3 was better than that of the previous version, and the architecture was modified as follows:

To reduce the convolution size, a spatial filter size of 5×5 for InceptionV1 was replaced by two convolutions with a filter size of 3×3 to reduce the computation time and cost and improve the classification performance.

Spatial factorization, in terms of asymmetric convolution, was implemented to reduce computational cost. Hence, the symmetric convolution size $n \times n$ was replaced by a convolution of $1 \times n$, followed by $n \times 1$. Although the network layer size of InceptionV3 was increased to 42 layers, the computational cost was higher than GoogLeNet by only 2.5 times and the performance was better than that of VGG [35].

5) INCEPTIONRESNETV2

InceptionResNetV2 was further developed and improved using the inception architecture, which is a combination of a residual network and a precise classification network, as mentioned in the previous section. InceptionResNetV2 was developed from the latest version of InceptionV4 and combined with a residual network to exploit the advantages of each model to improve performance. The architecture comprises layers such as the stem, inception-ResNet-A, −B, and −C, and reductions A and B [36].

6) DENSELY CONNECTED CONVOLUTIONAL NETWORKS (DENSENET201)

DenseNet is a feedforward CNN architecture that connects each layer to every other layer. For each layer, the feature maps of all the preceding layers were used as inputs, and their feature maps were used as inputs for all subsequent layers. DenseNet offers several advantages. It alleviates the vanishing gradient problem, strengthens the feature propagation, encourages feature reuse, and substantially reduces the number of parameters. DenseNet achieves significant improvements over the state-of-the-art on most of them, while requiring fewer parameters and less computation to achieve a high performance [37].

C. GRADIENT-WEIGHTED CLASS ACTIVATION MAPPING (GRAD-CAM) FOR VISUALIZATION

In this study, we implemented a gradient-weighted class activation map (Grad-CAM) to visualize the predicted results of the classifier in the final stage and interpret more transparency by producing visual explanations. The visualization algorithm can discriminate between classes more accurately, reveal the trustworthiness of a classifier, and identify bias in datasets [38], [39]. Grad-CAM uses the gradient information flowing into the last convolutional layer of the CNN so that it can be used to visualize any activation in a deep network. To obtain a class-discriminative localization map, Grad-CAM. Initially, the class score for the output class c in the class activation maps (CAM) model y^c is given by.

$$y^c = \sum_k \alpha_k^c \frac{1}{Z} \sum_i \sum_j A_{ij}^k \tag{8}$$

where α_k^c is the weight corresponding to class c for the k feature map, A_{ij}^k is the pixel at location in the k feature map, and Z is the total number of pixels in the feature map. Grad-CAM is a generalization that overcomes the limitation of class activation maps (CAM), which is the overhead of learning weights for linear mapping. The Grad-CAM weight was calculated as follows:

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{9}$$

The rectified linear unit (ReLU) function was applied to the weighted combination of feature maps to retain only positive correlations in the final activation map. Finally, the class discriminative localization map Grad-CAM was obtained.

$$L_{Grad-CAM}^c = ReLU\left(\sum_k \alpha_k^c A^k\right) \tag{10}$$

IV. EXPERIMENTAL SETUP

A. EXPERIMENTAL AND TRAINING SETUP

The brain tumor dataset used in this study was obtained from two databases. The REMBRANDT database contains MRI images from 130 patients with clinical data analysis in DICOM format. Brain tumors in different locations and normal brains in the axial FLAIR modality were obtained for classification [20]. The 658 MRI images from REMBRANDT comprised grade III and IV brain tumor 190 images and normal brain 468 images were combined

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

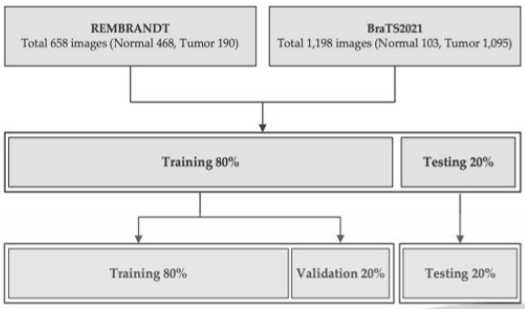


FIGURE 7. Dataset combination for brain tumor classification.

into a dataset. In the second database, MRI images were obtained from BraTS2021, which contains multi-institutional MRI scans in NifTI format [21], [22], [23]. The MRI image T_{2FLAIR} modality in the BraTS2021 database was converted to JPEG format comprising 1,198 images, including 1,095 brain tumors, and 103 normal brain images were combined in the dataset. The total combination of MRI images from the two databases contained 1,856 images comprising of brain tumor 1,285 images and normal brain 571 images were utilized for classification using various CNNs. The dataset was divided into training and testing datasets with ratios of 80% and 20%, respectively, and the brain tumor and normal images were equally distributed into the training and testing datasets. Therefore, we obtained 1,484 images for the training dataset, comprising 1,028 tumor images and 456 normal brain images, whereas the testing dataset contained 372 images, comprising 257 tumor images and 115 normal brain images. The training dataset was divided into training and validation datasets with ratios of 80% and 20%, respectively, before being used for training with the various CNNs. The dataset combination and management are shown in Fig. 7. The training dataset was heavily imbalanced. To address this issue, image augmentation is implemented to increase the number of minority classes. Therefore, the normal brain images were augmented using 90-degree rotation, reflection around the x-axis, synthetic noise, and synthetic blur for the training dataset only. The number of images obtained using each approach is listed in Table 5. In this study, we performed preprocessing and classification using MATLAB R2023a with image processing and a deep learning toolbox.

1) APPROACH 1 TRAINING ALGORITHM USING IMBALANCED DATASET

A dataset comprising 1,484 images was utilized for training using various deep-learning CNNs. A label was provided for each image, and the ground truth table specified the actual class of the brain image, that is, tumor or normal. The consistency of the brain image, image labeling, and ground-truth table was considered because it led to precise feature extraction and classification. The training dataset was divided into training and validation datasets with ratios of 80% and 20%, respectively. Transfer learning using CNNs involves training data, network layers, and algorithms. A pretrained network layer with the modified last 3 layers was used in our study. A block diagram for training the network using transfer learning is shown in Fig. 8.

TABLE 5. Dataset combination for each approach.

Approach	Dataset	Normal	Glioma	Total
Approach 1 Imbalanced dataset	Training	456	1,028	1,484
	Testing	115	257	372
Approach 2 Balanced dataset	Training	1,028	1,028	2,056
	Testing	115	257	372

The algorithm option was aligned as follows: stochastic gradient descent with momentum (SGDM), initial learning rate 0.0001, and 20 epochs for training. The network was trained using these parameters. Finally, we obtained a trained network that was utilized to predict the 372 images of the testing dataset. Training progress and confusion charts were plotted for monitoring purposes.

2) APPROACH 2 TRAINING ALGORITHM USING BALANCED DATASET

Image augmentation to increase the number of images of minority classes, balance the training datasets, and significantly improve the accuracy is a simple and effective methodology. Therefore, we performed augmentation of the normal image in the training dataset because it is heavily imbalanced as explanation in Table 3. Finally, we obtained normal brain 1,028 images and glioma brain tumor 1,028 images for utilization in the training network. The network layer and algorithm option parameters were maintained as in approach 1. Training progress was plotted for monitoring purposes. The testing dataset was maintained similar with a previous approach, that is 372 images comprising 115 normal brain images and brain tumor 257 images. The trained network was utilized for the prediction testing dataset, and the accuracy of the prediction between the predicted result and ground truth table was evaluated. A block diagram for training the balanced dataset is illustrated in Fig. 9.

3) ENSEMBLE LEARNING USING MAJORITY VOTING

Recently, several related studies have introduced ensemble learning to further improve the accuracy of medical image classification [29], [40], [41]. In this study, ensemble hard and soft majority voting were conducted to investigate the performance of brain tumor classification. The top five performance models from the previous section included the majority of the hard and soft votes. The details of the majority voting approach are as follows.

Majority hard voting: Summing the per-class predicted labels from every classifier for the input image and obtaining

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

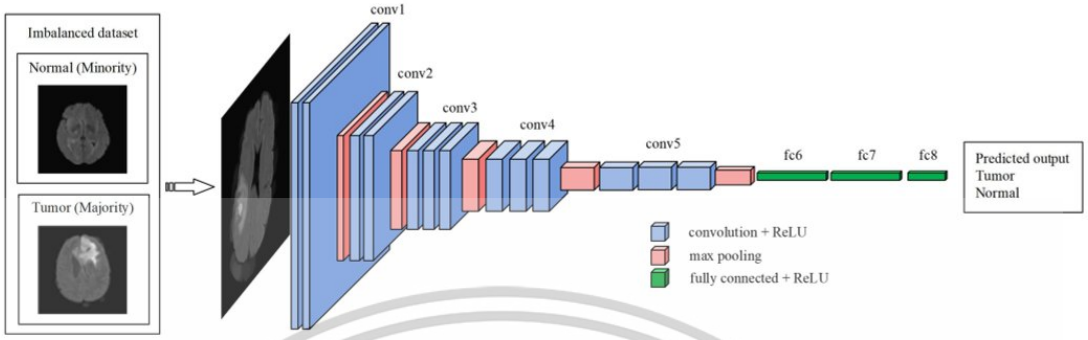


FIGURE 8. Block diagram for training imbalanced dataset.

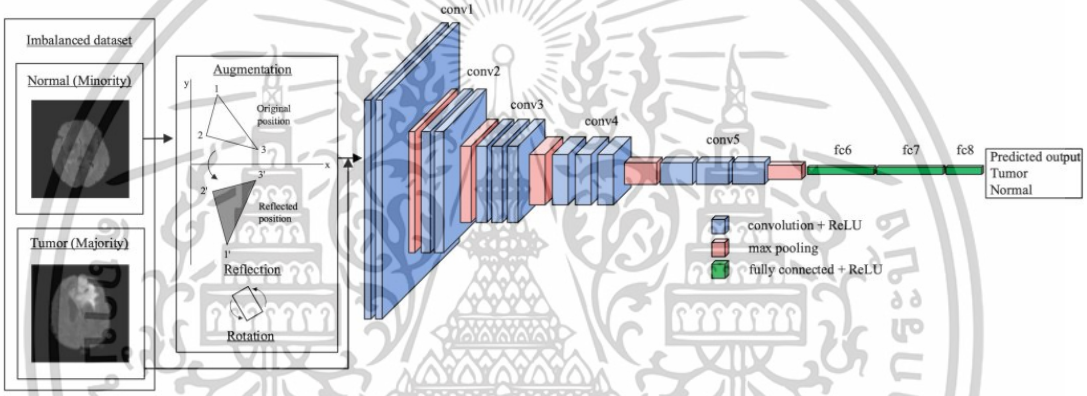


FIGURE 9. Block diagram for training balanced dataset.

the final label of the class with the highest number of voting classes, as shown in Fig. 10.

Majority soft voting: Summing the per-class probability values generated by every classifier for the input image, and obtaining the final label of the class with the greatest probability sum, as shown in Fig. 10.

4) VISUALIZATION USING GRAD-CAM

In this study, Grad-CAM was used to visualize the discriminative regions of the predicted MRI images. A customized program was established according to the block diagram shown in Fig. 11. The trained model from each CNNs algorithm was loaded into the workspace, and a particular image that required visualization and inspection was read into the workspace for further processing. A particular image was classified using the trained model algorithm to predict normal or glioma brain tumors and the probability was indicated. Finally, the predicted results were interpreted using the Grad-CAM algorithm and a heatmap image was generated for visualization. The pseudocode for the visualization of a particular image is presented in Table 6.

The implementation of this customized program will assist the medical staff in diagnosing glioma brain tumors using individual MRI medical images. The medical MRI image, which was pre-processed as explained in the previous section, was loaded into the workspace. The CNN-trained model predicts an individual image to indicate a tumor or a normal tissue. The predicted images were verified by using the Grad-CAM algorithm.

B. PERFORMANCE EVALUATION

The experimental results were evaluated using the following diagnostic test attributes: sensitivity, specificity, and accuracy, which are widely used to describe diagnostic tests [42]. Several terms were used to describe sensitivity, specificity, and accuracy. These included true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP). In this study, if a tumor was proven to be present in

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

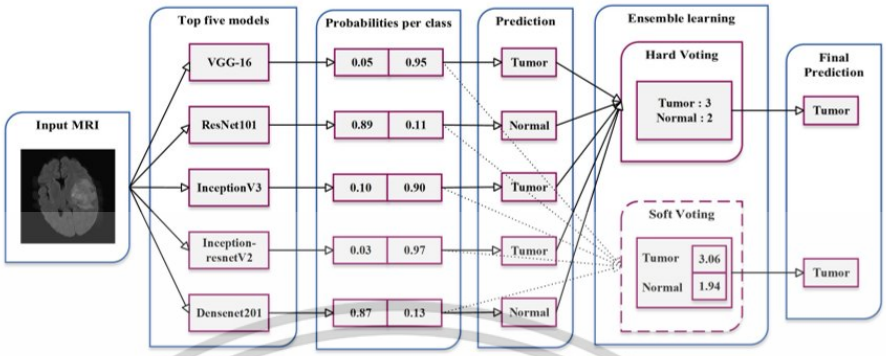


FIGURE 10. Ensemble learning using majority voting.

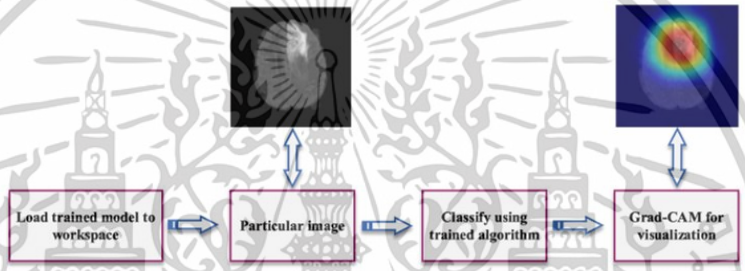


FIGURE 11. Block diagram for visualization using Grad-CAM.

TABLE 6. Pseudo code for classification and visualization of the particular image.

Pseudo-code for image visualization

Begin

Require: CNN trained model using a brain tumor training dataset with transfer learning.

1. Load the CNN trained model to the workspace.
2. Verify the compatibility of the input image size, i.e., the required input image is 299 x 299 for inceptionV3.
3. Read the input image RGB type, which requires prediction and visualization into the workspace.
4. The input image was resized to equal the required input image of the network.
5. The input image is classified to indicate the tumor or normal brain using the trained model network.
6. Compute the Grad-CAM map to visualize the predicted image :

$$L^{c}_{Grad-CAM} = ReLU\left(\sum_{k} \alpha^{c}_{k} A^{k}\right)$$

7. Displays the discriminative image region of the classified image.

End

a patient, the predicted result indicated the presence of the disease, and the result of the diagnostic test was considered true positive. Similarly, if a tumor was proven to be absent in a patient, the predicted result indicated that the tumor was absent, and the test result was considered true negative.

True positive and negative results suggest a consistent result between the diagnostic test and the proven condition. In contrast, if a tumor was proven to be present in a patient who had no such disease, the test result was considered a false-positive. Similarly, if a predicted result indicated that a tumor was absent in a patient with a brain tumor, the test result was considered a false-negative. False positives and negatives indicated that the test results were opposite to the actual conditions.

Sensitivity, specificity, and accuracy are described in terms of *TP*, *TN*, *FN*, and *FP*.

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{11}$$

$$Specificity = \frac{TN}{(TN + FP)} \tag{12}$$

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)} \tag{13}$$

The numerical value of sensitivity represents the probability of a diagnostic test identifying patients with the disease. A test with high sensitivity tends to capture all possible positive conditions without missing anyone. Hence, high-sensitivity tests are often used to screen for disease.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

TABLE 7. Classification results (Approach 1: imbalanced dataset).

CNNs	Accuracy	Sensitivity	Specificity	AUC
VGG16	92.47%	99.61%	76.52%	0.9961
VGG19	**95.70%**	**99.22%**	**87.83%**	**0.9978**
ResNet50	94.09%	99.22%	82.61%	0.9957
ResNet101	94.35%	99.22%	83.48%	0.9948
Xception	93.01%	98.44%	80.87%	0.9774
InceptionV3	94.89%	99.61%	84.35%	0.9973
InceptionResNetV2	95.43%	98.05%	89.57%	0.9925
Densenet201	93.55%	98.44%	82.61%	0.9912

TABLE 8. Training and testing inference time for imbalanced dataset.

CNNs	Training elapsed time	Testing inference time
VGG16	7 min. 53 sec.	4.578 seconds
VGG19	8 min. 43 sec.	4.764 seconds
ResNet50	9 min. 33 sec.	5.221 seconds
ResNet101	16 min. 42 sec.	6.312 seconds
Xception	15 min. 10 sec.	6.658 seconds
InceptionV3	14 min. 29 sec.	5.924 seconds
InceptionResNetV2	43 min. 45 sec.	9.445 seconds
Densenet201	42 min. 51 sec.	9.859 seconds

The numerical value of the specificity represents the probability of a diagnostic test identifying patients without such diseases. A test with high specificity when conducting a diagnostic test in a patient without a certain disease has a high probability of being identified as negative.

The obtained prediction result is the true positive rate (TPR), which is equivalent to the sensitivity, and the false positive rate (FPR), which is equivalent to (1 − specificity). The all-possible combination of TPR and FPR comprises a Receiver Operating Characteristic (ROC) space. The area under the ROC curve (AUC) was used to measure the accuracy of the diagnostic test. A larger area indicated a more accurate diagnostic test. The AUC of the ROC curve can be measured using the following equation, where t = (1 − specificity) and ROC (t) is the sensitivity:

$$AUC = \int_0^1 ROC(t)dt \tag{14}$$

V. RESULTS AND DISCUSSION

In this study, we implemented three different approaches for classifying glioma brain tumors using various CNNs transfer-learning methods and ensemble majority voting, as previously described. In Approach 1, preprocessed MRI images without augmentation were utilized for training and testing using eight CNNs. Finally, the prediction results were used to measure the accuracy and quality of diagnostic tests. The experimental results indicated that VGG19 was the best classifier, with an accuracy of 95.70%, sensitivity of 99.22%, specificity of 87.83%, and AUC of 0.9978. The sensitivity, specificity, accuracy, and AUC of each CNNs classifier are presented in Table 7. The training and testing inference times for each CNNs are presented in Table 8. An example of the ROC curve plotted against the relative true-positive and false-positive rates for the imbalanced dataset is shown in Fig. 12.

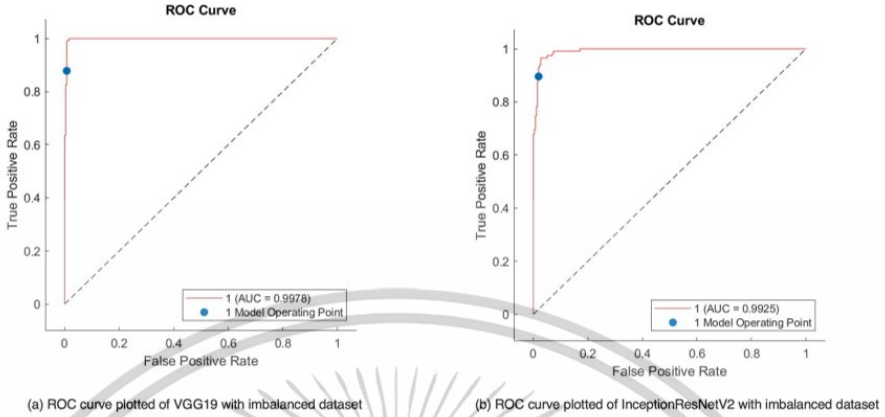
In Approach 2, the accuracy and quality of the diagnostic test were significantly improved by using image augmentation to increase the number of images of minority classes. As described in the previous section, an image augmentation methodology and balanced training dataset were implemented to address the imbalanced dataset problem. The experimental results showed that all the performance parameters were significantly improved for the almost perfect binary classification of glioma brain tumors. As previously explained, we implemented various CNNs for the training dataset using the algorithm option to obtain the trained model. The training progress is plotted to monitor the accuracy and loss of the validation dataset. Fig. 13 shows the training progress plotted for the best classifier that was InceptionV3. The training progress indicated the validation accuracy. InceptionV3 had a validation accuracy of 99.51%. The trained model was then used to classify glioma brain tumors.

The best CNN classifier was InceptionV3, which achieved an accuracy of 99.73%, a sensitivity of 99.61%, a specificity of 100%, and an AUC of 1. Another CNNs classifier and ensemble majority voting showed the experimental results listed in Table 9. The training and testing inference times for each CNNs are presented in Table 10. The relative between the true positive rate (TPR) and false positive rate (FPR) was used to plot the ROC curve. An example of an ROC curve for a balanced dataset is shown in Fig. 14.

Notably, the InceptionV3 classifier with a balanced dataset obtained the best performance, and Densenet201 obtained very good performance with an accuracy of 99.46%, sensitivity of 99.22%, and specificity of 100%. However, InceptionV3 requires a training time of 25 min. 15 sec. and a testing inference time of 5.878 s, which was less than that of the Densenet201 classifier, which used a training time of 57 min. 25 sec. and a testing inference time of 9.413 s. This is because of the differences in network layers and architectures. In this study, we used a laptop with a CPU Intel Core i7, 3.0 GHz, RAM 32 GB, and an NVIDIA GeForce RTX 3070 graphics processing unit (GPU) for training and testing the brain tumor CNNs.

In Approach 3, to further investigate the efficiency of classification, ensemble learning using majority hard and soft voting, which is a straightforward methodology, was conducted, and the network was not required for training

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a) ROC curve plotted of VGG19 with imbalanced dataset (b) ROC curve plotted of InceptionResNetV2 with imbalanced dataset

FIGURE 12. ROC curve plotted of classifiers with imbalanced dataset.

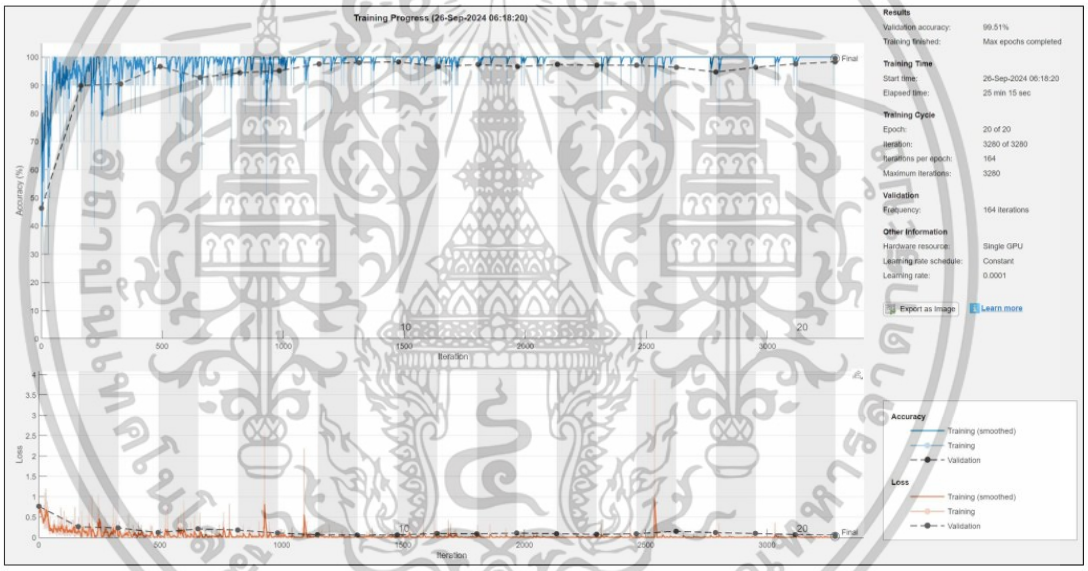


FIGURE 13. Training progress of InceptionV3 with a balanced dataset.

again. The top five classifiers with the best performances were used for majority voting.

Therefore, VGG16, InceptionV3, InceptionResNetV2, ResNet101, and DenseNet201 with a balanced dataset from Approach 2 were utilized. In the hard-voting approach, the predicted label from each classifier was used for summing and voting. In the soft voting approach, the probability per class from each classifier is used for summing and voting, as shown in Fig. 10. Finally, we obtained the evaluated performance from majority voting with an accuracy of 99.46%, a sensitivity of 99.22%, and a specificity of 100%, as shown in Table 9.

Although we present the performance evaluation in Table 9, it is adequate to justify and compare the performances of different classifiers. Furthermore, we propose a quantitative analysis with appropriate statistical testing p-values to evaluate the testing results from each classifier. The p-value represents the probability of observing the sample given that the hypothesis is true. The null hypothesis (H_0) was rejected if the p-value was lower than the significance

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

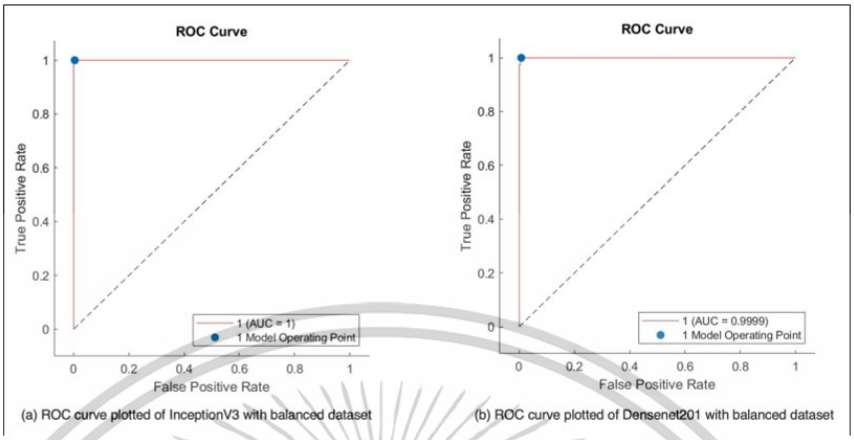


FIGURE 14. ROC curve plotted of classifiers with balanced dataset.

TABLE 9. Classification results (Approach 2: balanced dataset).

CNNs	Accuracy	Sensitivity	Specificity	AUC
VGG16	99.19%	99.61%	98.26%	0.9993
VGG19	98.12%	98.83%	96.52%	0.9994
ResNet50	97.85%	98.44%	96.52%	0.9982
ResNet101	99.19%	98.83%	100%	0.9999
Xception	98.12%	98.83%	96.52%	0.9980
InceptionV3	**99.73%**	**99.61%**	**100%**	**1.000**
InceptionResNetV2	98.66%	98.05%	100%	0.9989
Densenet201	99.46%	99.22%	100%	0.9999
Hard voting	99.46%	99.22%	100%	-
Soft voting	99.46%	99.22%	100%	-

TABLE 10. Training and testing inference time for balanced dataset.

CNNs	Training elapsed time	Testing inference time
VGG16	11 min. 2 sec.	5.955 seconds
VGG19	12 min. 7 sec.	4.884 seconds
ResNet50	12 min. 50 sec.	5.412 seconds
ResNet101	23 min. 21 sec.	6.62 seconds
Xception	27 min. 19 sec.	6.625 seconds
InceptionV3	25 min. 15 sec.	5.878 seconds
InceptionResNetV2	61 min. 3 sec.	9.315 seconds
Densenet201	57 min. 25 sec.	9.413 seconds

level [43]. We used a one-sided test with $H_0 \geq 99$, meaning the probability of prediction result must be greater than or equal to 0.99 for all actual classes, with a significance level of 5% for testing, which led to a critical value of -1.649. The statistical test results for the top five performance classifiers are presented in Table 11. When rejecting H_0, if indicated as zero, we cannot reject the null hypothesis. The results indicate that all proposed classifiers have p-values greater than the significance level. Therefore, the null hypothesis is not rejected.

TABLE 11. Statistical testing of classifiers for quantitative analysis.

CNNs	Test statistic	Confidence interval	p-value	Reject H_0
VGG16	0.0068	0.9972	0.5027	0
ResNet101	-0.1454	0.9947	0.4422	0
InceptionV3	-0.0018	0.9943	0.4993	0
InceptionResNetV2	-1.5617	0.9905	0.0596	0
Densenet201	0.2697	0.9966	0.6062	0

Finally, Grad-CAM was used to visualize the discriminative image region, and interpretability was performed to determine the parts of the image that were the most important for classification. The heatmap shows red color for the region that was most relative to the tumor. The visualization of the indicative region for transparency reflects the reliability and trustworthiness of the classifier.

In this study, we conducted customized programming to implement the Grad-CAM algorithm to visualize particular images for monitoring and to further justify the predicted result, as described in the block diagram in Fig. 11 and the pseudocode in Table 6.

As previously explained, inceptionV3 with a balanced dataset, was the best classifier in our methodology. Hence, the Grad-CAM algorithm using inceptionV3 with a balanced dataset approach was used to visualize the discriminative

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

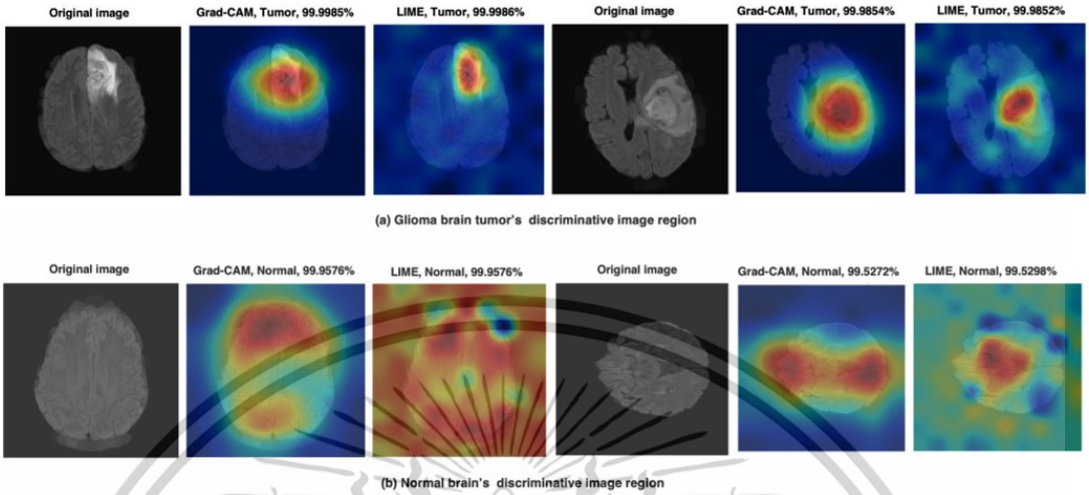


FIGURE 15. Comparison between Grad-CAM and LIME for visualization using InceptionV3.

image regions of the normal brain and tumor; the results are shown in Fig. 15.

We also analyzed the visualization and interpretability of our proposed classifier using the local interpretable model-agnostic explanation (LIME) methodology with InceptionV3, as shown in Fig. 15. This approach ensures that our classifier accurately identifies the discriminative regions of the brain tumors. The visualization process is similar to that of the Grad-CAM algorithm, as explained by the pseudocode in Table 6. However, the map was generated using the LIME algorithm [44].

The experimental results obtained in this study were evaluated and compared with those obtained in previous studies. Our proposed methodology yielded fascinating experimental results. Many previous studies have reported high accuracy in brain tumor classification. However, the trustworthiness and reliability of the classifier are significant and the overfitting problem is a constraint that should be investigated further.

Therefore, the proposed methodology was implemented using two different datasets and the number of images was increased to prevent overfitting. The proposed methodology obtained high performance, and InceptionV3 with a balanced dataset was the best classifier, with an accuracy of 99.73%, sensitivity of 99.61%, specificity of 100%, and AUC of 1.00. Moreover, we implemented the Grad-CAM algorithm to visualize the discriminative image region, which indicated the trustworthiness of the classifier.

VI. CONCLUSION

In this study, we introduce an automated classification method for glioma brain tumors using deep learning CNNs. MRI images from two different databases were combined and used for the binary classification. Preprocessing with a morphological operation to remove the skeleton and artifacts from the brain region and the CLAHE algorithm to enhance the image contrast were performed. The obtained image dataset was highly imbalanced between glioma tumor and normal brain images, which is the minority class. Therefore, two approaches were implemented to prepare the training dataset: imbalanced and balanced MRI.

An image-augmentation methodology was used to increase the number of minority classes to balance the training dataset. Eight state-of-the-art deep learning CNNs were implemented for the binary classification of glioma brain tumors. The experimental results indicated that the accuracy was significantly improved with an approachable balanced dataset. The top five performance classifier was used for ensemble learning using majority hard and soft voting for further investigation as well. The performance and quality of the experimental results were evaluated, and it was found that better than previous related work, the best classifier for our studied was inceptionV3 with a balanced dataset with an accuracy of 99.73%, sensitivity of 99.61%, specificity of 100%, and area under the ROC curve of 1.00. An appropriate statistical analysis was implemented to evaluate the results from different classifiers, and it was found that the top five performance classifiers selected met the criteria for null hypothesis testing. Additionally, Grad-CAM was implemented to visualize the discriminative image region between the tumor blob and brain region to ensure the trustworthiness of the results predicted by the CNNs classifier. The proposed methodology precisely classifies glioma brain tumors and we envision that our proposed

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

algorithm will be further improved to assist medical experts in diagnosing brain tumors. In future work, transformer-based models for 3D medical imaging will be investigated to improve the performance of brain tumor classification.

REFERENCES

[1] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Phys. Med. Biol.*, vol. 58, no. 13, p. R97, Jul. 2013.

[2] D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: A summary," *Acta Neuropathologica*, vol. 131, no. 6, pp. 803–820, Jun. 2016.

[3] H. T. Zaw, N. Maneerat, and K. Y. Win, "Brain tumor detection based on Naïve Bayes classification," in *Proc. 5th Int. Conf. Eng., Appl. Sci. Technol. (ICEAST)*, Luang Prabang, Laos, Jul. 2019, pp. 1–4.

[4] P. Pilaoon, N. Maneerat, K. Yajima, R. Varakulsiripunth, and K. Hamamoto, "Brain tumor classification using supervised support vector machine," in *Proc. 9th Int. Conf. Eng., Appl. Sci., Technol. (ICEAST)*, Vientiane, Laos, Jun. 2023, pp. 79–83.

[5] Y. K. Dubey, M. M. Mushrif, and K. Pisar, "Brain tumor type detection using texture features in MR images," in *Proc. IEEE Region Humanitarian Technol. Conf.*, Dec. 2018, pp. 1–4.

[6] P. Pilaoon, N. Maneerat, A. Nakthewan, R. Varakulsiripunth, and K. Hamamoto, "Brain tumor classification using pretrained deep convolutional neural network," in *Proc. 9th Int. Conf. Eng., Appl. Sci., Technol. (ICEAST)*, Vientiane, Laos, Jun. 2023, pp. 84–88.

[7] T. Hossain, F. S. Shishir, M. Ashraf, M. A. A. Nasim, and F. M. Shah, "Brain tumor detection using convolutional neural network," in *Proc. 1st Int. Conf. Adv. Sci., Eng. Robot. Technol. (ICASERT)*, May 2019, pp. 1–6.

[8] N. M. Dipu, S. A. Shohan, and K. M. A. Salam, "Deep learning based brain tumor detection and classification," in *Proc. Int. Conf. Intell. Technol. (CONIT)*, Karnataka, India, Jun. 2021, pp. 1–6.

[9] S. Arora and M. Sharma, "Deep learning for brain tumor classification from MRI images," in *Proc. 6th Int. Conf. Image Inf. Process. (ICIIP)*, vol. 6, Nov. 2021, pp. 409–412.

[10] S. Gaikwad, S. Patel, and A. Shetty, "Brain tumor detection: An application based on machine learning," in *Proc. 2nd Int. Conf. Emerg. Technol. (INCET)*, Belgaum, India, May 2021, pp. 1–4.

[11] S. K. Baranwal, K. Jaiswal, K. Vaibhav, A. Kumar, and R. Srikantaswamy, "Performance analysis of brain tumour image classification using CNN and SVM," in *Proc. 2nd Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, Jul. 2020, pp. 537–542.

[12] I. Aboussaleh, J. Riffi, A. M. Mahraz, and H. Tairi, "Brain tumor segmentation based on deep learning's feature representation," *J. Imag.*, vol. 7, no. 12, p. 269, Dec. 2021.

[13] A. Younis, L. Qiang, C. O. Nyatega, M. J. Adamu, and H. B. Kawuwa, "Brain tumor analysis using deep learning and VGG-16 ensembling learning approaches," *Appl. Sci.*, vol. 12, no. 14, p. 7282, Jul. 2022.

[14] H. Hu, X. Li, W. Yao, and Z. Yao, "Brain tumor diagnose applying CNN through MRI," in *Proc. 2nd Int. Conf. Artif. Intell. Comput. Eng. (ICAICE)*, Nov. 2021, pp. 430–434.

[15] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016.

[16] H. Kibriya, M. Masood, M. Nawaz, R. Rafique, and S. Rehman, "Multiclass brain tumor classification using convolutional neural network and support vector machine," in *Proc. Mohammad Ali Jinnah Univ. Int. Conf. Comput. (MAJICC)*, Jul. 2021, pp. 1–4.

[17] A. W. Reza, M. S. Hossain, M. A. Wardiful, M. Farzana, S. Ahmad, F. Alam, R. N. Nandi, and N. Siddique, "A CNN-based strategy to classify MRI-based brain tumors using deep convolutional network," *Appl. Sci.*, vol. 13, no. 1, p. 312, Dec. 2022.

[18] S. Asif, W. Yi, Q. U. Ain, J. Hou, T. Yi, and J. Si, "Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from MR images," *IEEE Access*, vol. 10, pp. 34716–34730, 2022.

[19] S. Mohsen, A. M. Ali, E.-S.-M. El-Rabaie, A. ElKaseer, S. G. Scholz, and A. M. A. Hassan, "Brain tumor classification using hybrid single image super-resolution technique with ResNext101_32× 8d and VGG19 pre-trained models," *IEEE Access*, vol. 11, pp. 55582–55595, 2023.

[20] REMBRANDT. (2017). *The Cancer Imaging Archive*. [Online]. Available: https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT

[21] U. Baid et al., "The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification," 2021, *arXiv:2107.02314*.

[22] B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.

[23] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, no. 1, pp. 1–13, Sep. 2017, doi: 10.1038/sdata.2017.117.

[24] R. Gonzalez and R. Woods, *Digital Image Processing*, 4th ed., London, U.K.: Pearson, 2018.

[25] R. Gonzalez, R. Woods, and S. Eddins, *Digital Image Processing Using MATLAB*, 2nd ed., London, U.K.: Pearson, 2017.

[26] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, Sep. 1987.

[27] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, vol. 1994. Cambridge, U.K.: Academic, 1994, pp. 474–485.

[28] I. Valova, C. G. Harris, T. Mai, and N. Gueorguieva, "Optimization of convolutional neural networks for imbalanced set classification," in *Proc. 24th Int. Conf. KES*, Jan. 2020, pp. 660–669.

[29] K. Y. Win, N. Maneerat, S. Sreng, and K. Hamamoto, "Ensemble deep learning for the detection of COVID-19 in unbalanced chest X-ray dataset," *Appl. Sci.*, vol. 11, no. 22, p. 10528, Nov. 2021.

[30] J. Nalepa, M. Marcinkiewicz, and M. Kawulok, "Data augmentation for brain-tumor segmentation: A review," *Frontiers Comput. Neurosci.*, vol. 13, Dec. 2019, Art. no. 83.

[31] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2017, pp. 294–299.

[32] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, May 2015, pp. 1–8.

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[34] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.

[35] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[36] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, San Francisco, CA, USA, Feb. 2017, pp. 4278–4284.

[37] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[38] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2921–2929.

[39] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.

[40] M. Ben Jabra, A. Koubaa, B. Benjdira, A. Ammar, and H. Hamam, "COVID-19 diagnosis in chest X-rays using deep learning and majority voting," *Appl. Sci.*, vol. 11, no. 6, p. 2884, Mar. 2021.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[41] B. Machura, D. Kucharski, O. Bozek, B. Eksner, B. Kokoszka, T. Pekala, M. Radom, M. Strzelczak, L. Zarudzki, B. Gutiérrez-Becker, A. Krason, J. Tessier, and J. Nalepa, "Deep learning ensembles for detecting brain metastases in longitudinal multi-modal MRI studies," *Computerized Med. Imag. Graph.*, vol. 116, Sep. 2024, Art. no. 102401.

[42] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," in *Proc. NESUG Health Care Life Sci.*, Baltimore, MD, USA, Nov. 2010, pp. 1–9.

[43] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 7th ed., Hoboken, NJ, USA: Wiley, 2018, pp. 194–243.

[44] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' explaining the prediction of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 1135–1144.

PONGSAK PILAOON received the B.Eng. degree in instrumentation engineering and the M.Eng. degree in automation engineering from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, in 2003 and 2017, respectively, where he is currently pursuing the Ph.D. degree with the School of Engineering. He has authored two research articles. His research interests include medical image classification, deep learning, and machine learning.

KAZUHIKO HAMAMOTO (Member, IEEE) received the B.Eng., M.Eng., and D.Eng. degrees from Tokyo University of Agriculture and Technology, in 1989, 1991, and 1994, respectively. He is currently a Professor with the Department of Information Media Technology, School of Information Science and Technology, Tokai University, Japan. He published around 60 journal/transaction articles and more than 90 international conference papers. His research interests include medical information, human interface design, and virtual reality. He is a member of many national societies in Japan.

NOPPADOL MANEERAT received the B.Sc. degree in physics from Srinakarinwirot University (SWU), Phitsanulok, Thailand, in 1990, and the M.Eng. and D.Eng. degrees in electrical engineering from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, in 1997 and 2006, respectively. Currently, he is with the Department of Instrumentation and Control Engineering, School of Engineering, KMITL. His research interests include digital image processing, artificial intelligence, and electronic devices.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล นาย พงษ์ศักดิ์ พิลาอุ่น

วัน เดือน ปีเกิด 13 กุมภาพันธ์ 2522

ที่อยู่ หมู่บ้านบริทราเนีย บางนา-สุวรรณภูมิ ต.บางบ่อ อ.บางบ่อ
จ.สมุทรปราการ 10560

ประวัติการศึกษา:

2545 วิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมการวัดคุม คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

2560 วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมอัตโนมัติ คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ผลงานวิจัย:

[1] P. Pilaoon, N. Maneerat, K. Yajima, R. Varakulsiripunth and K. Hamamoto, "Brain Tumor Classification using Supervised Support Vector Machine," in *Proc. 9th Int. Conf. on Engineering, Applied Sciences, and Technology (ICEAST)*, Vientiane, Laos, 2023, pp. 79-83.

[2] P. Pilaoon, N. Maneerat, A. Nakthewan, R. Varakulsiripunth and K. Hamamoto, "Brain Tumor Classification using Pretrained Deep Convolutional Neural Network," in *Proc. 9th Int. Conf. on Engineering, Applied Sciences, and Technology (ICEAST)*, Vientiane, Laos, 2023, pp. 84-88.

[3] P. Pilaoon, A. Nakthewan, N. Wadlom, R. Varakulsiripunth, K. Hamamoto and N. Maneerat, "Glioma Brain Tumor Classification using Transfer Learning," in *Proc. 10th Int. Conf. on Engineering, Applied Sciences, and Technology (ICEAST)*, Luang Prabang, Laos, 2024, pp. 17-22.

[4] P. Pilaoon, K. Hamamoto and N. Maneerat, "Glioma Brain Tumor Classification using Convolution Neural Network and Majority Voting," *IEEE Access*, Vol. 13, pp. 46228-46248, 2025.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้