



การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึก

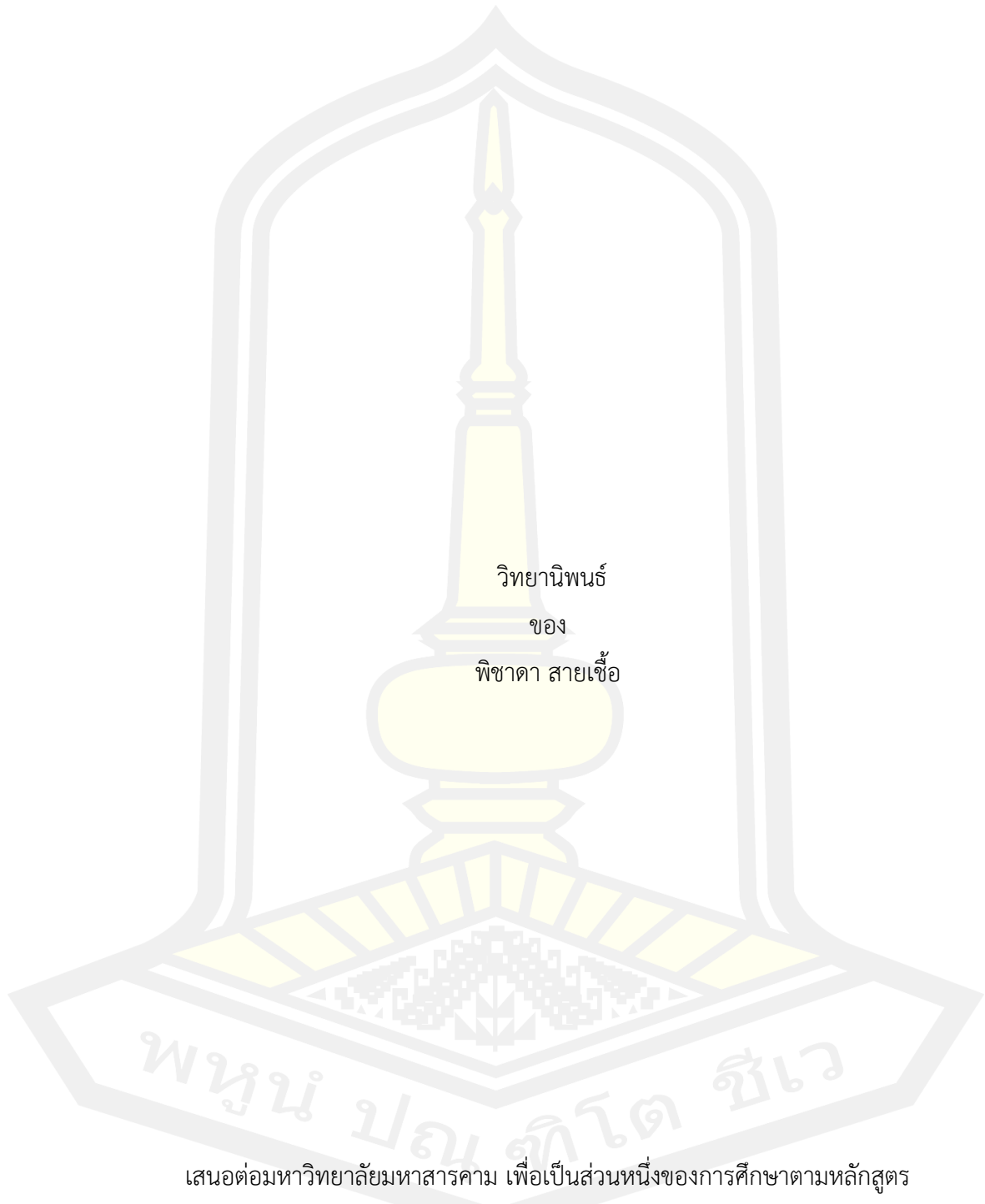
วิทยานิพนธ์
ของ
พิชาดา สายเชื้อ

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

มิถุนายน 2565

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึก

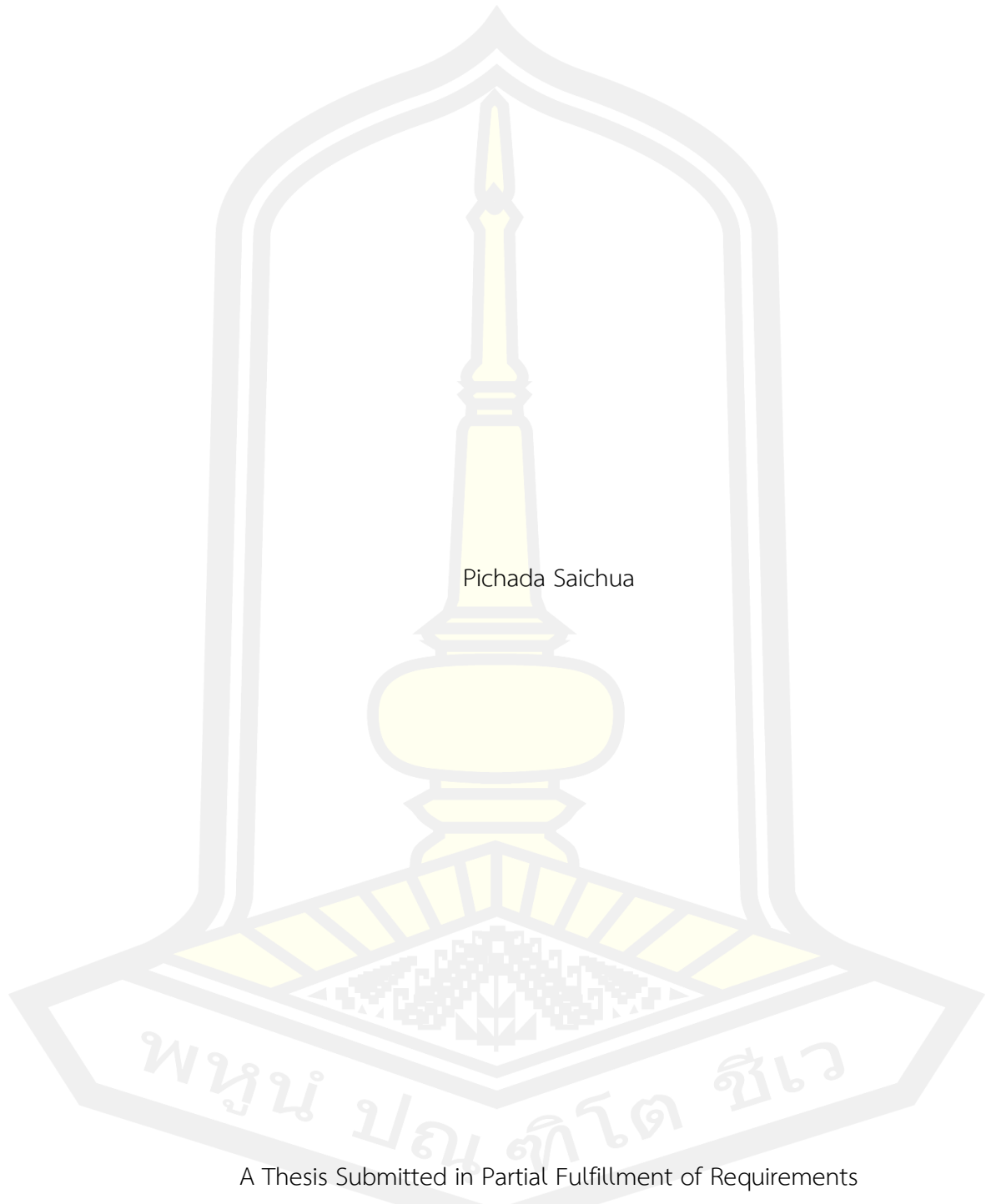


เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

มิถุนายน 2565

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Classification of Diabetic Retinopathy Images using Deep Learning



Pichada Saichua

A Thesis Submitted in Partial Fulfillment of Requirements
for Master of Science (Information Technology)

June 2022

Copyright of Mahasarakham University



คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนางสาวพิชาดา สายเชื้อ
แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ

(รศ. ดร. สิทธิชัย บุขหมั่น)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผศ. ดร. โสฬาริก สุรินตะ)

..... กรรมการ

(ผศ. ดร. สานิต แสงประดิษฐ์)

..... กรรมการ

(ผศ. ดร. แกมกาญจน์ สมประเสริฐศรี)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

.....
(ผศ. ศศิธร แก้วมัน)

คณบดีคณะวิทยาการสารสนเทศ

.....
(รศ. ดร. กริสน์ ชัยมูล)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึก		
ผู้วิจัย	พิชิตดา สายเชื้อ		
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. โอฬาริก สุรินตะ		
ปริญญา	วิทยาศาสตรมหาบัณฑิต	สาขาวิชา	เทคโนโลยีสารสนเทศ
มหาวิทยาลัย	มหาวิทยาลัยมหาสารคาม	ปีที่พิมพ์	2565

บทคัดย่อ

งานวิจัยฉบับนี้ได้ทำการศึกษาและเปรียบเทียบประสิทธิภาพการเรียนรู้เชิงลึกสำหรับการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาโดยใช้ CNN ร่วมกับเทคนิค Ensemble CNN และ Fusion CNN ซึ่งในการทดลองนั้นผู้วิจัยได้เลือกสถาปัตยกรรม CNN ที่แตกต่างกัน 5 รูปแบบ ได้แก่ InceptionV3, ResNet50, ResNet50V2, Xception และ DenseNet121 เพื่อสร้างเป็นแบบจำลองสำหรับการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา โดยในขั้นตอนการเรียนรู้เพื่อสร้างแบบจำลองมีการนำเทคนิค Data Augmentation มาร่วมทดลองเพื่อเพิ่มมิติหรือมุมมองของชุดข้อมูลรูปภาพจอประสาทตาจึงทำให้ความสามารถในการจำแนกประเภทจอประสาทตาของแบบทุกแบบจำลองมีประสิทธิภาพมากยิ่งขึ้น โดยการทดลองที่กล่าวมาเป็นการทดลองเพื่อค้นหาแบบจำลอง CNN ที่มีประสิทธิภาพพร้อมทั้งอัตราความถูกต้องในการจำแนกประเภทจอประสาทตาสูงสุด เพื่อนำไปทำการทดลองร่วมกับเทคนิค Ensemble CNN และ Fusion CNN ซึ่งจากผลลัพธ์การทดลองพบว่าเมื่อนำแบบจำลอง Xception และ InceptionV3 รวมเข้าด้วยกันด้วยเทคนิค Fusion CNN ที่ Dense จำนวน 2 เลเยอร์ โดยมีขนาดของ Dense เป็น 1024 และค่าของ Dropout ที่ 0.4 นั้นสามารถให้อัตราความถูกต้องมากถึงร้อยละ 86.30 และเมื่อรวมแบบจำลอง Xception และ InceptionV3 ด้วยเทคนิค Ensemble CNN ซึ่งเป็นการนำค่าความน่าจะเป็นที่ได้จากการพยากรณ์ของแต่ละแบบจำลองไปทำการคำนวณด้วยวิธี Weighted Average และ Unweighted Average เพื่อเลือกผลลัพธ์หรือคำตอบที่มีค่ามากที่สุด โดยจากผลการศึกษาพบว่าเมื่อทำการรวมแบบจำลองด้วยเทคนิค Fusion CNN นั้นสามารถเพิ่มประสิทธิภาพแบบจำลองได้สูงสุดและมีอัตราความถูกต้องแตกต่างจากการรวมแบบจำลองด้วยเทคนิค Ensemble CNN เพียงเล็กน้อย

คำสำคัญ : การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา, การเรียนรู้เชิงลึก, โครงข่ายประสาทเทียมคอนโวลูชัน, เทคนิคการเพิ่มข้อมูลรูปภาพ

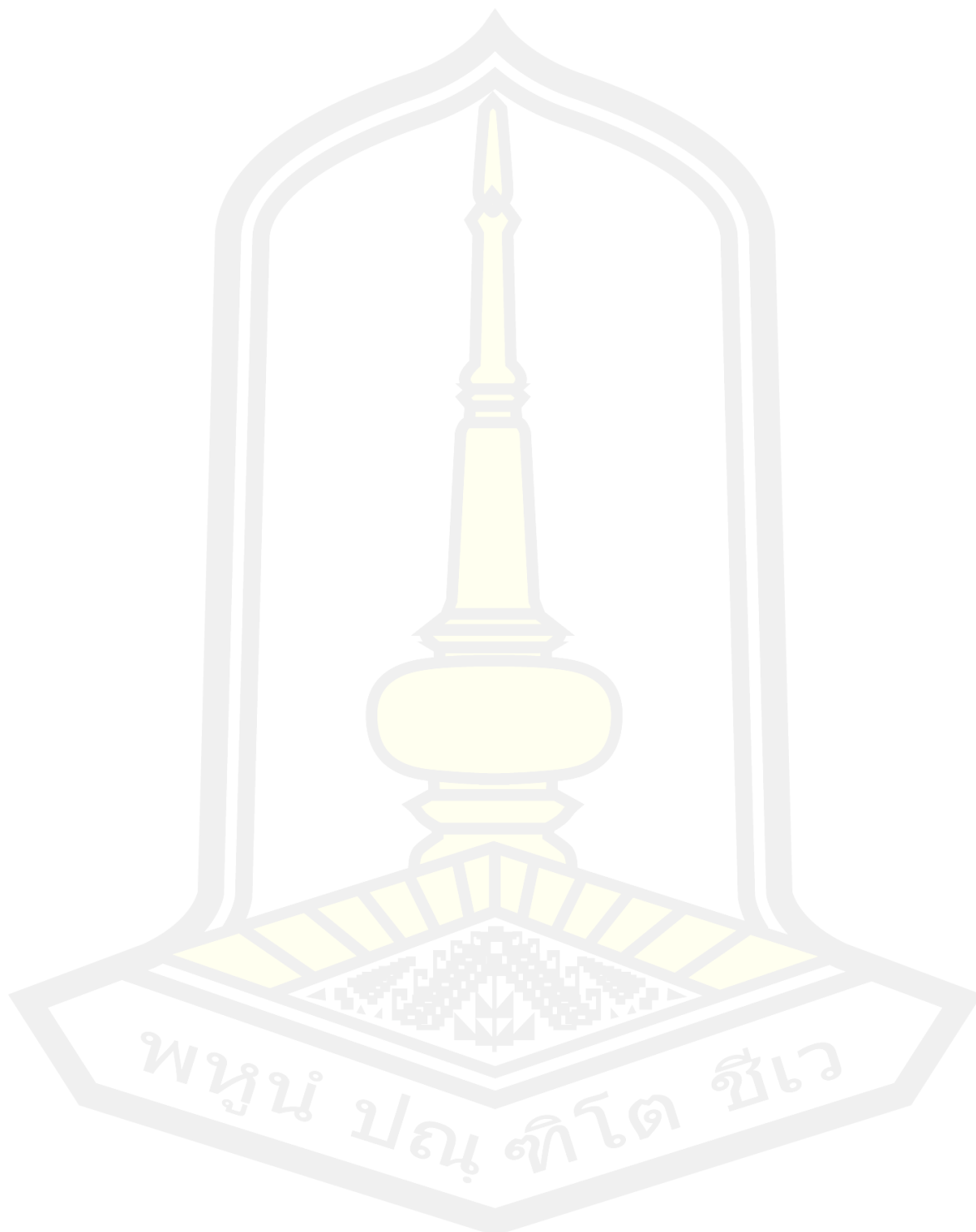
TITLE	Classification of Diabetic Retinopathy Images using Deep Learning		
AUTHOR	Pichada Saichua		
ADVISORS	Assistant Professor Olarik Surinta , Ph.D.		
DEGREE	Master of Science	MAJOR	Information Technology
UNIVERSITY	Maharakham University	YEAR	2022

ABSTRACT

In this paper, researchers have demonstrated the effectiveness of the fusion convolutional neural network (CNN) and ensemble CNN architectures for diabetic retinopathy classification. Due to the fusion and ensemble CNN architectures, authors proposed to use five CNN architectures consisting of InceptionV3, ResNet50, ResNet50V2, Xception, and DenseNet121 to find the best CNN model. The fusion and ensemble CNN architectures were then created combining two of the best CNN models and while training the CNN models, researchers also used data augmentation techniques. The data augmentation technique, researchers discovered, can increase the accuracy of CNNs. However, the data augmentation technique should not distort the retinal image. For the fusion CNNs, the Xception and InceptionV3 were combined and then attached with two dense layers with the size of 1,024 units for each dense layer. Hence, researchers selected the optimal dropout value with 0.4. For the ensemble CNNs, the output probabilities, that were calculated from the Xception and InceptionV3 models, were sent to the ensemble learning method. Researchers compared the weighted and unweighted average methods using ensemble learning methods. The results showed that the weighted average method outperformed the unweighted average method in all ensemble CNNs. In the experiments, researchers discovered that the fusion architecture slightly outperformed ensemble CNN architecture.

Keyword : Convolutional neural network, Fusion CNNs, Ensemble CNNs, Diabetic

retinopathy classification, Data augmentation technique



กิตติกรรมประกาศ

ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.โอฬาริก สุรินตะ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่มีความเมตตากรุณาให้คำแนะนำพร้อมทั้งชี้แนะข้อบกพร่องตลอดจนร่วมแก้ไข และติดตามปัญหาที่เกิดขึ้นในงานวิจัย อีกทั้งฝึกฝนให้ผู้วิจัยมีทักษะทางการคิด การอ่าน การเขียน และการนำเสนอผลงานทางวิชาการ จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณคณะกรรมการการสอบวิทยานิพนธ์ทุกท่านที่ได้ให้ความอนุเคราะห์เป็นกรรมการสอบ และให้คำแนะนำเพื่อปรับปรุงงานวิจัยจนสำเร็จเป็นวิทยานิพนธ์นี้

ขอขอบคุณครอบครัว มิตรสหาย และตัวข้าพเจ้าเอง ที่ได้ให้โอกาสทางการศึกษาและส่งเสริมข้าพเจ้าตลอดมา โดยคุณค่าและประโยชน์อันพึงมาจากวิทยานิพนธ์ฉบับนี้ผู้วิจัยขอมอบแต่ผู้มีพระคุณทุกท่าน

พิชาดา สายเชื้อ



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ช
สารบัญ.....	ซ
สารบัญรูปภาพ.....	ฅ
สารบัญตาราง.....	ท
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาของการวิจัย.....	1
1.2 ความมุ่งหมายของการวิจัย.....	3
1.3 ขอบเขตของการวิจัย.....	3
1.3.1 ชุดข้อมูล.....	3
1.3.2 กระบวนการเรียนรู้เชิงลึกสำหรับการจำแนกประเภทรูปภาพจอประสาทตา.....	4
1.3.3 การประเมินประสิทธิภาพ.....	4
1.4 ผลที่คาดว่าจะได้รับจากการวิจัย.....	4
1.5 นิยามศัพท์เฉพาะ.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ภาวะเบาหวานขึ้นตา (Diabetic Retinopathy).....	7
2.2 การจำแนกรูปภาพ (Image Classification).....	7
2.3 การเรียนรู้เชิงลึก (Deep Learning).....	8
2.4 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network).....	9
2.5 การเพิ่มประสิทธิภาพแบบจำลอง.....	14

2.5.1 เทคนิค Data Augmentation	14
2.5.2 เทคนิค Fusion Convolutional Neural Network.....	15
2.5.3 เทคนิค Ensemble Convolutional Neural Network.....	16
2.6 การประเมินประสิทธิภาพแบบจำลอง	18
2.6.1 คอนฟิวชันเมทริกซ์ (Confusion Matrix).....	18
2.6.2 ค่าความถูกต้อง (Accuracy).....	18
2.7 งานวิจัยที่เกี่ยวข้อง.....	19
บทที่ 3 วิธีดำเนินการวิจัย.....	25
3.1 ชุดข้อมูลรูปภาพที่ใช้ในการวิจัย	25
3.1.1 การแปลงขนาดรูปภาพ (Resize Image).....	27
3.1.2 การเพิ่มจำนวนชุดข้อมูลด้วยเทคนิค Data Augmentation.....	28
3.2 การสร้างแบบจำลองสำหรับการจำแนกประเภทรูปภาพจอบประสาทตา.....	29
3.2.1 ปรับค่าไฮเปอร์พารามิเตอร์ (Fine-tuning Hyperparameter).....	30
3.2.2 ทดลองเพื่อค้นหาชุดรูปแบบ Data Augmentation ที่เหมาะสม	31
3.2.3 ทดลองร่วมกับเทคนิค Ensemble CNN และ เทคนิค Fusion CNN.....	31
3.3 การประเมินประสิทธิภาพของแบบจำลอง.....	32
บทที่ 4 ผลการศึกษา.....	33
4.1 เครื่องมือและข้อมูลที่ใช้ในการทดลอง.....	33
4.1.1 คอมพิวเตอร์ที่ใช้ในการทดลอง.....	33
4.1.2 การเขียนโปรแกรมที่ใช้ในการทดลอง	34
4.1.3 ข้อมูลที่ใช้ในการทดลอง	34
4.2 วิธีการทดลอง.....	35
4.2.1 สร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter	35
4.2.2 สร้างแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation.....	35

4.2.3	สร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN	36
4.2.4	สร้างแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN.....	37
4.3	ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึกด้วย แบบจำลอง CNN.....	38
4.3.1	ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter.....	38
4.3.2	ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation.....	45
4.3.3	ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN	50
4.3.4	ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN.....	53
4.4	ประเมินและเปรียบเทียบประสิทธิภาพการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะ เบาหวานขึ้นตาของแบบจำลอง CNN	69
บทที่ 5	สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ	70
5.1	สรุปผลการวิจัย	70
5.2	อภิปรายผล.....	71
5.3	ข้อเสนอแนะ และแนวทางการวิจัยในอนาคต.....	72
บรรณานุกรม.....		73
ภาคผนวก.....		80
ภาคผนวก ก	เอกสารรับรองโครงการวิจัยในมนุษย์	81
ภาคผนวก ข	ชุดข้อมูลรูปภาพจอบประสาทตาที่ใช้ในการวิจัย.....	83
ภาคผนวก ค	บทความที่ได้เผยแพร่ในวารสารวิชาการ ICIC Express Letters	88
ประวัติผู้เขียน.....		97

สารบัญรูปภาพ

หน้า

ภาพประกอบที่ 1.1 ตัวอย่างของชุดข้อมูลรูปภาพของประสาทตาปกติ	4
ภาพประกอบที่ 1.2 ตัวอย่างของชุดข้อมูลรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา	4
ภาพประกอบที่ 2.1 โครงสร้างโครงข่ายประสาทเทียม	9
ภาพประกอบที่ 2.2 โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชัน	9
ภาพประกอบที่ 2.3 ตัวอย่างการทำคอนโวลูชัน	10
ภาพประกอบที่ 2.4 ขั้นตอนการทำ Max Pooling	11
ภาพประกอบที่ 2.5 โครงสร้างของชั้นการเชื่อมโยงแบบสมบูรณ์	11
ภาพประกอบที่ 2.6 โครงสร้าง ResNet50	12
ภาพประกอบที่ 2.7 โครงสร้าง Xception	13
ภาพประกอบที่ 2.8 โครงสร้าง InceptionV3	13
ภาพประกอบที่ 2.9 โครงสร้าง DenseNet121	14
ภาพประกอบที่ 2.10 การทำ Data Augmentation รูปแบบ Rotation	14
ภาพประกอบที่ 2.11 การทำ Data Augmentation รูปแบบ Horizontal Flip และ Vertical Flip	15
ภาพประกอบที่ 2.12 การทำ Data Augmentation รูปแบบ Zoom	15
ภาพประกอบที่ 2.13 โครงสร้าง Fusion Convolution Neural Network	16
ภาพประกอบที่ 2.14 โครงสร้าง Ensemble Convolutional Neural Network	17
ภาพประกอบที่ 2.15 คอนฟิวชันเมทริกซ์การจำแนกแบบ 2 คลาส	18
ภาพประกอบที่ 3.1 ตัวอย่างรูปภาพจอประสาทตาปกติจากชุดข้อมูล APTOS 2019 Blindness Detection	25
ภาพประกอบที่ 3.2 ตัวอย่างรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจากชุดข้อมูล APTOS 2019 Blindness Detection	26

ภาพประกอบที่ 3.3 ตัวอย่างรูปภาพจอประสาทตาปกติจากชุดข้อมูล Diabetic Retinopathy Detection.....	26
ภาพประกอบที่ 3.4 ตัวอย่างรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจากชุดข้อมูล Diabetic Retinopathy Detection.....	26
ภาพประกอบที่ 3.5 ตัวอย่างรูปภาพจอประสาทตาปกติจากชุดข้อมูล Diabetic Retinopathy Dataset.....	26
ภาพประกอบที่ 3.6 ตัวอย่างรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจากชุดข้อมูล Diabetic Retinopathy Dataset.....	27
ภาพประกอบที่ 3.7 รายละเอียดของชุดข้อมูล.....	27
ภาพประกอบที่ 3.8 ตัวอย่างการเปลี่ยนแปลงขนาดรูปภาพ	28
ภาพประกอบที่ 3.9 ตัวอย่างการ Rotation จากรูปภาพจอประสาทตา	28
ภาพประกอบที่ 3.10 ตัวอย่างการ Flip จากรูปภาพจอประสาทตา.....	29
ภาพประกอบที่ 3.11 ตัวอย่างของการ Zoom จากรูปภาพจอประสาทตา.....	29
ภาพประกอบที่ 4.1 การทำ Data Augmentation จาก Training Set	36
ภาพประกอบที่ 4.2 อัตรา Accuracy ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Fine-tuning Hyperparameter	44
ภาพประกอบที่ 4.3 อัตรา Loss ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Fine-tuning Hyperparameter.....	44
ภาพประกอบที่ 4.4 Confusion Matrix ของแบบจำลอง Xception ที่สร้างร่วมกับเทคนิค Fine-tuning Hyperparameter	45
ภาพประกอบที่ 4.5 อัตรา Accuracy ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation.....	49
ภาพประกอบที่ 4.6 อัตรา Loss ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation.....	49
ภาพประกอบที่ 4.7 Confusion Matrix ของแบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation.....	50

ภาพประกอบที่ 4.8 อัตรา Accuracy ของการเรียนรู้แบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN 67

ภาพประกอบที่ 4.9 อัตรา Loss ของการเรียนรู้แบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN 67

ภาพประกอบที่ 4.10 Confusion Matrix ของแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN 68



สารบัญตาราง

หน้า

ตารางที่ 4.1 สถิติการแบ่งข้อมูลสำหรับการทดลอง.....	34
ตารางที่ 4.2 ค่าพารามิเตอร์สำหรับการทดลอง Fine-tuning Hyper-Parameter	35
ตารางที่ 4.3 รูปแบบ Data Augmentation ที่ใช้ในการทดลอง	36
ตารางที่ 4.4 คำนี้นักสำหรับการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN.....	37
ตารางที่ 4.5 ค่าพารามิเตอร์สำหรับการทดลอง Fusion CNN.....	37
ตารางที่ 4.6 ผลลัพธ์การสร้างแบบจำลอง ResNet50 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter.....	38
ตารางที่ 4.7 ผลลัพธ์การสร้างแบบจำลอง ResNet50V2 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter.....	39
ตารางที่ 4.8 ผลลัพธ์การสร้างแบบจำลอง InceptionV3 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter.....	40
ตารางที่ 4.9 ผลลัพธ์การสร้างแบบจำลอง DenseNet121 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter.....	41
ตารางที่ 4.10 ผลลัพธ์การสร้างแบบจำลอง Xception ร่วมกับเทคนิค Fine-tuning Hyper-Parameter.....	42
ตารางที่ 4.11 สรุปผลลัพธ์การทดลองการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter.....	43
ตารางที่ 4.12 ผลลัพธ์การสร้างแบบจำลอง ResNet50 ร่วมกับเทคนิค Data Augmentation	46
ตารางที่ 4.13 ผลลัพธ์การสร้างแบบจำลอง ResNet50V2 ร่วมกับเทคนิค Data Augmentation..	46
ตารางที่ 4.14 ผลลัพธ์การสร้างแบบจำลอง DenseNet121ร่วมกับเทคนิค Data Augmentation.	47
ตารางที่ 4.15 ผลลัพธ์การสร้างแบบจำลอง InceptionV3ร่วมกับเทคนิค Data Augmentation ...	47
ตารางที่ 4.16 ผลลัพธ์การสร้างแบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation	48
ตารางที่ 4.17 สรุปผลลัพธ์การทดลอง CNN ร่วมกับเทคนิค Data Augmentation	48

ตารางที่ 4.18 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Ensemble.....	51
ตารางที่ 4.19 ผลลัพธ์การสร้างแบบจำลอง ResNet50 และ DenseNet121 ร่วมกับเทคนิค Ensemble.....	51
ตารางที่ 4.20 ผลลัพธ์การสร้างแบบจำลอง ResNet50V2 และ DenseNet121 ร่วมกับเทคนิค Ensemble.....	51
ตารางที่ 4.21 ผลลัพธ์การสร้างแบบจำลอง ResNet50 และ ResNet50V2 ร่วมกับเทคนิค Ensemble.....	52
ตารางที่ 4.22 สรุปผลลัพธ์การสร้างแบบจำลองด้วยเทคนิค Ensemble - Unweighted Average	52
ตารางที่ 4.23 สรุปผลลัพธ์การสร้างแบบจำลองด้วยเทคนิค Ensemble - Weighted Average	52
ตารางที่ 4.24 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN.....	54
ตารางที่ 4.25 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1	54
ตารางที่ 4.26 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2	55
ตารางที่ 4.27 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3	55
ตารางที่ 4.28 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4	56
ตารางที่ 4.29 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN.....	56
ตารางที่ 4.30 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN.....	57
ตารางที่ 4.31 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1	57

ตารางที่ 4.32 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2 58

ตารางที่ 4.33 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3 58

ตารางที่ 4.34 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4 59

ตารางที่ 4.35 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN..... 59

ตารางที่ 4.36 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN 60

ตารางที่ 4.37 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1 60

ตารางที่ 4.38 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2 61

ตารางที่ 4.39 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3 61

ตารางที่ 4.40 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4 62

ตารางที่ 4.41 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN..... 62

ตารางที่ 4.42 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN..... 63

ตารางที่ 4.43 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1 63

ตารางที่ 4.44 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2 64

ตารางที่ 4.45 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3	64
ตารางที่ 4.46 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4	65
ตารางที่ 4.47 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN.....	65
ตารางที่ 4.48 สรุปผลลัพธ์การสร้างแบบจำลอง Xception ร่วมกับเทคนิค Fusion CNN	66
ตารางที่ 4.49 เปรียบเทียบประสิทธิภาพในการจำแนกประเภทรูปภาพจอสภาพตาที่มีภาวะเบาหวานขึ้นตาของแบบจำลอง CNN และเทคนิคต่าง ๆ	69



บทที่ 1

บทนำ

1.1 ความเป็นมาของการวิจัย

องค์การอนามัยโลก (WHO) ระบุว่าภาวะเบาหวานขึ้นตา (Diabetic Retinopathy: DR) เป็นหนึ่งในสาเหตุสำคัญที่ทำให้คนทั่วโลกตาบอด [1] ซึ่งเป็นผลมาจากการเป็นโรคเบาหวาน และมีน้ำตาลสะสมในกระแสเลือดเป็นจำนวนมากพร้อมทั้งสะสมเป็นระยะเวลานาน ซึ่งปัญหาหลักที่ทำให้เกิดภาวะเบาหวานคืออาหารที่บริโภคเป็นมื้อหลักในชีวิตประจำวันนั้นเป็นอาหารที่มีส่วนประกอบของน้ำตาลเป็นจำนวนมาก อีกทั้งอาหารที่จำหน่ายตามร้านสะดวกซื้อส่วนใหญ่มีน้ำตาลเป็นส่วนประกอบหลัก เช่น เครื่องดื่มชูกำลัง ขนมหวาน กาแฟสำเร็จรูป เป็นต้น ซึ่งการที่ร่างกายของมนุษย์ได้รับน้ำตาลในปริมาณมากเกินไปจะส่งผลให้มีระดับน้ำตาลในเลือดสูง และทำให้เป็นอันตรายต่อร่างกายโดยเฉพาะอย่างยิ่งกับบริเวณดวงตา ซึ่งหากหลอดเลือดในจอประสาทตาเริ่มมีอาการอักเสบหรือโป่งพองจะส่งผลให้น้ำเหลืองกระจายทั่วจอประสาทตา ดังนั้นหากไม่ได้รับการตรวจวินิจฉัยและเข้ารับการรักษาอย่างรวดเร็วอาจทำให้จอประสาทตาเกิดภาวะขาดเลือด พร้อมทั้งเซลล์ที่ใช้ในการมองเห็นถูกทำลายซึ่งเป็นผลให้ความสามารถในการมองเห็นลดลงตลอดจนสูญเสียการมองเห็นไปในที่สุด

เนื่องจากอัตราการเจ็บป่วยด้วยภาวะเบาหวานขึ้นตาที่เพิ่มจำนวนมากขึ้น ทำให้จักษุแพทย์ทำการรักษาผู้ป่วยได้ล่าช้าและไม่ทั่วถึง ด้วยเหตุนี้จึงทำให้นักวิจัยหลายคนได้ทำการศึกษาและพัฒนา ระบบที่ช่วยในการตรวจวินิจฉัยภาวะเบาหวานขึ้นตาดด้วยเทคนิคการเรียนรู้เชิงลึก เพื่อเป็นการแบ่งเบาภาระงานของจักษุแพทย์ พร้อมทั้งทำให้การคัดกรองโรคในเบื้องต้นมีความรวดเร็วมากยิ่งขึ้น ซึ่ง จะทำการพิจารณาและวินิจฉัยอาการจากรูปภาพจอประสาทตา โดยในปัจจุบันทางเว็บไซต์ Kaggle ได้เผยแพร่ชุดข้อมูลรูปภาพจอประสาทตาที่แบ่งออกเป็นจอประสาทตาปกติ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตา อีกทั้งภาวะเบาหวานขึ้นตาสสามารถแบ่งตามระดับความรุนแรงของโรคออกได้เป็น 4 ระดับ [2] ได้แก่

- 1) ภาวะเบาหวานขึ้นจอประสาทตาขั้นเบา (Mild Non-Proliferative Diabetic Retinopathy)
- 2) ภาวะเบาหวานขึ้นจอประสาทตาขั้นปานกลาง (Moderate Non-Proliferative Diabetic Retinopathy)
- 3) ภาวะเบาหวานขึ้นจอประสาทตาขั้นหนัก (Severe Non-Proliferative Diabetic Retinopathy)
- 4) ภาวะเบาหวานเข้าจอประสาทตาที่มีเส้นเลือดงอกใหม่ (Proliferative Diabetic Retinopathy)

งานวิจัยที่เกี่ยวข้องกับการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา ที่ประสบความสำเร็จ และสามารถใช้งานได้จริงมีหลากหลายงานด้วยกัน เช่น

Vo และ Verma [3] นำเสนอสถาปัตยกรรมสำหรับการจำแนกรูปภาพจอประสาทตาสองรูปแบบ ได้แก่ สถาปัตยกรรม Combined Kernels with Multiple Losses Network (CKML Net) และสถาปัตยกรรม VGGNet with Extra Kernel (VNXX) ซึ่งเป็นการพัฒนามาจากสถาปัตยกรรม GoogLeNet และ สถาปัตยกรรม VGG16 ตามลำดับ โดยใน CKML Net ตั้งแต่ Convolution Layer ไปจนถึง Pooling Layer จะถูกแทนที่ด้วย Three Parallel Branches ซึ่งภายในแต่ละ Branche จะประกอบไปด้วยโครงสร้างของ GoogLeNet และ VNXX ภายใน Convolution Layer ของสถาปัตยกรรม VGGNet จะถูกแทนที่ด้วย Convolution Layer ของสถาปัตยกรรม GoogLeNet จากผลการทดสอบการจำแนกประเภทจอประสาทตาบนชุดข้อมูล EyePACS พบว่า CKML Net และ VNXX ให้อัตราความถูกต้องที่ร้อยละ 82.30 และ 83.12 ตามลำดับ ซึ่งมากกว่า GoogLeNet และ VGGNet ที่ให้อัตราความถูกต้องที่ร้อยละ 79.71 และ 82.34 ตามลำดับ

Abbas และคณะ [4] นำเสนอเทคนิคการสกัดคุณลักษณะพิเศษจากรูปภาพจอประสาทตาด้วย 2 กระบวนการหลัก ได้แก่ Scale-Invariant Feature Transform (SIFT) และ Gradient Location-Orientation Histogram (GLOH) จากนั้นนำ Feature ไปทำการแปลงค่าและทำการเลือกผลลัพธ์ที่เหมาะสมที่สุดด้วยวิธี Principle Component Analysis (PCA) และส่งเข้าไปประมวลผลใน Hidden Layer จำนวน 3 ชั้นด้วยกันซึ่งผลลัพธ์ที่ได้จากการรู้จำถูกแบ่งออกเป็น 5 กลุ่มตามระดับความรุนแรงของโรค และนำไปทดสอบกับรูปภาพจอประสาทตาจำนวน 750 รูปภาพ โดยจากการทดลองพบว่ากระบวนการดังกล่าวสามารถให้อัตรา Sensitivity (SE) และ Specificity (SP) ที่ร้อยละ 92 และ 94 ตามลำดับ

Xu และคณะ [5] ออกแบบสถาปัตยกรรม CNN ที่มีจำนวน 10 ชั้นสำหรับการจำแนกจอประสาทตาที่มีภาวะเบาหวานขึ้นตา โดยทำการเรียนรู้เพื่อรู้จำจากรูปภาพจอประสาทตาจำนวน 800 รูปภาพ และประเมินประสิทธิภาพด้วยข้อมูลจำนวน 200 รูปภาพ ทั้งนี้ เนื่องจากรูปภาพที่ใช้ในการเรียนรู้มีจำนวนจำกัด จึงได้มีการนำเทคนิค Data Augmentation เข้ามาช่วยในการเพิ่มจำนวนรูปภาพในการเรียนรู้หลายรูปแบบ ซึ่งประกอบด้วย Rotation, Flipping, Shearing, Rescaling และ Translation จากการทดลองพบว่าการทำ Data Augmentation ช่วยเพิ่มประสิทธิภาพการจำแนกประเภทของ CNN ถึงร้อยละ 3 โดยการเรียนรู้ของ CNN แบบ 10 ชั้น ร่วมกับเทคนิค Data Augmentation นั้นให้อัตราความถูกต้องมากถึงร้อยละ 94

โดยงานวิจัยการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาที่ได้กล่าวถึงนั้นประสบความสำเร็จ และเป็นประโยชน์อย่างมาก ทั้งนี้ ในงานวิจัยฉบับนี้จึงได้ทำการศึกษาระบวนการในการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึก โดยมีจุดประสงค์เพื่อค้นหาแบบจำลอง CNN และกระบวนการการเรียนรู้เชิงลึกที่เหมาะสมสำหรับการจำแนกรูปภาพจอบประสาทตา โดยสามารถแบ่งรูปแบบของกระบวนการศึกษาออกเป็น 3 ส่วน ดังต่อไปนี้

1) ปรับค่าไฮเปอร์พารามิเตอร์ (Fine-tuning Hyperparameter) ให้เหมาะสมสำหรับแบบจำลอง CNN ทั้ง 5 รูปแบบ ได้แก่ ResNet50, ResNet50V2, Xception, InceptionV3 และ DenseNet121 พร้อมทั้งทำการเรียนรู้แบบจำลองร่วมกับเทคนิค Data Augmentation หลายรูปแบบเพื่อให้ได้ชุดข้อมูลที่มีมิติมุมมองที่หลากหลายมากยิ่งขึ้น ซึ่งจะส่งผลให้ได้แบบจำลองที่มีประสิทธิภาพ และความถูกต้องแม่นยำสูงในการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา

2) สร้างแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN และ Ensemble CNN ซึ่งเป็นเทคนิคที่รวมเอาแบบจำลอง CNN หลายรูปแบบรวมเข้าด้วยกัน เพื่อเพิ่มประสิทธิภาพในการการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา

3) เปรียบเทียบประสิทธิภาพการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาของแบบจำลอง CNN เพื่อให้ได้แบบจำลองที่มีอัตราความถูกต้องสูงสุดในการจำแนกรูปภาพจอบประสาทตาปกติ และจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา

1.2 ความมุ่งหมายของการวิจัย

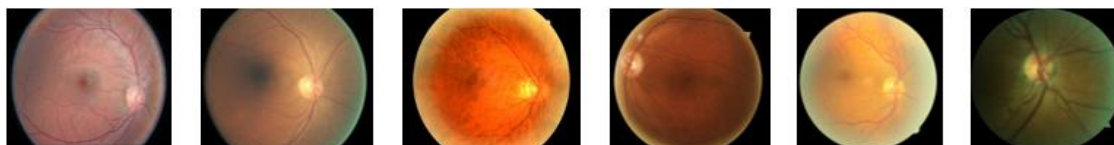
ศึกษาและเปรียบเทียบสถาปัตยกรรม CNN บนหลักการการเรียนรู้เชิงลึกสำหรับการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา

1.3 ขอบเขตของการวิจัย

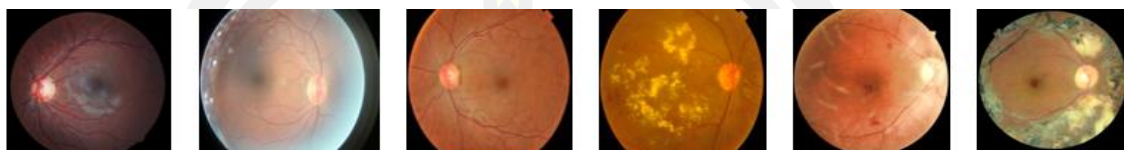
การวิจัยเรื่องการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึกได้กำหนดขอบเขตของการวิจัยออกเป็น 3 ส่วน โดยมีรายละเอียดดังต่อไปนี้

1.3.1 ชุดข้อมูล

ชุดข้อมูลรูปภาพจอบประสาทตาในการวิจัยนี้ได้ถูกรวบรวมจากข้อมูลที่เผยแพร่บนเว็บไซต์ Kaggle โดยชุดข้อมูลรูปภาพจอบประสาทตาแบ่งออกเป็นรูปภาพจอบประสาทตาปกติ และรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา ซึ่งตัวอย่างของชุดข้อมูลรูปภาพของประสาทตาปกติจะแสดงดังภาพประกอบที่ 1.1 และตัวอย่างของชุดข้อมูลรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาจะแสดงดังภาพประกอบที่ 1.2



ภาพประกอบที่ 1.1 ตัวอย่างของชุดข้อมูลรูปภาพของประสาทตาปกติ



ภาพประกอบที่ 1.2 ตัวอย่างของชุดข้อมูลรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา

1.3.2 กระบวนการเรียนรู้เชิงลึกสำหรับการจำแนกประเภทรูปภาพจอประสาทตา

1.3.2.1 การจำแนกประเภทจอประสาทตาในการวิจัยนี้ได้เลือกใช้สถาปัตยกรรม CNN จำนวน 5 รูปแบบ ได้แก่ ResNet50, ResNet50V2, Xception, InceptionV3 และ DenseNet121

1.3.2.2 เพิ่มประสิทธิภาพการจำแนกประเภทรูปภาพจอประสาทตาด้วย CNN โดยการใช้เทคนิค Data Augmentation เทคนิค Fusion CNN และเทคนิค Ensemble CNN

1.3.3 การประเมินประสิทธิภาพ

1.3.3.1 วัดประสิทธิภาพของแบบจำลองโดยการพิจารณาค่าความถูกต้อง และค่าการสูญเสียของการจำแนกรูปภาพจอประสาทตา

1.3.3.2 เปรียบเทียบประสิทธิภาพของแบบจำลอง โดยการใช้การเปรียบเทียบผลลัพธ์ค่าความถูกต้องจากข้อมูลชุดเดียวกัน

1.4 ผลที่คาดว่าจะได้รับจากการวิจัย

แบบจำลอง CNN ที่มีความถูกต้องแม่นยำสูงในการจำแนกประเภทจอประสาทตาที่มีภาวะเบาหวานขึ้นตา

1.5 นิยามศัพท์เฉพาะ

1.5.1 จอประสาทตา (Retina)

จอประสาทตา [6] เป็นอวัยวะที่อยู่บริเวณด้านหลังของผนังลูกตา และมีลักษณะเป็นเยื่อบางซึ่งทำหน้าที่เสมือนจอร์รับภาพโดยการแปลงแสงที่ผ่านเข้ามาในตาไปเป็นสัญญาณไฟฟ้า และส่งไปยังสมองผ่านทางเส้นประสาทตาเพื่อทำการรับรู้และแปลผลภาพ

1.5.2 ภาวะเบาหวานขึ้นตา (Diabetic Retinopathy)

ภาวะเบาหวานขึ้นตา [6] เป็นภาวะแทรกซ้อนจากโรคเบาหวานซึ่งเกิดจากการมีระดับน้ำตาลในเลือดสูงเป็นเวลานาน และไปอุดตันเส้นเลือดจนส่งผลให้การไหลเวียนของเลือดภายในจอประสาทตาไม่สามารถไหลเวียนได้ตามปกติ จึงเป็นผลให้เส้นเลือดที่จอประสาทตาได้รับความเสียหาย และสูญเสียการมองเห็นไปในที่สุด

1.5.3 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก [9] เป็นการจำลองรูปแบบการทำงานของระบบเซลล์ประสาทในสมองของมนุษย์ที่สามารถตัดสินใจพร้อมกับคาดการณ์ผลลัพธ์ได้ด้วยตนเองหรือที่เรียกว่าโครงข่ายประสาทเทียม ซึ่งเป็นการเรียนรู้ข้อมูลภายในโครงข่ายประสาทเทียมที่ถูกนำมาซ้อนทับกันหลายชั้น และประมวลผลออกมาเป็นคำตอบหรือผลลัพธ์การพยากรณ์

1.5.4 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network)

โครงข่ายประสาทเทียมคอนโวลูชัน [11] เป็นโครงข่ายประสาทเทียมรูปแบบหนึ่งที่เหมาะสมสำหรับการวิเคราะห์และประมวลผลเพื่อจำแนกประเภทข้อมูลที่เป็นรูปภาพ ซึ่งเป็นข้อมูลรูปแบบ 2 หรือ 3 มิติ โดยโครงสร้างพื้นฐานของโครงข่ายประสาทเทียมคอนโวลูชันนั้นประกอบไปด้วย Convolutional Layer, Pooling Layer และ Fully Connected Layer

1.5.5 การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา (Diabetic Retinopathy Classification)

การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาเป็นการนำข้อมูลรูปภาพจอประสาทตาไปประมวลผลด้วยกระบวนการเรียนรู้เชิงลึกเพื่อดึงเอาคุณลักษณะเฉพาะของวัตถุในรูปภาพ และทำการจำแนกออกเป็นประเภทจอประสาทตาดปกติ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตา

1.5.6 เทคนิค Ensemble Convolutional Neural Network

เทคนิค Ensemble Convolutional Neural Network [26] เป็นการนำแบบจำลองที่มีการเรียนรู้ด้วยเทคนิคหรือวิธีการที่แตกต่างกันตั้งแต่ 2 รูปแบบขึ้นไปมาทำจำแนกประเภทรูปภาพจากชุดข้อมูลเดียวกัน และนำคำตอบที่ได้จากแต่ละแบบจำลองไป Vote เพื่อหาค่าเฉลี่ยของความน่าจะเป็น และเลือกค่าสูงสุดเป็นคำตอบของ Ensemble

1.5.7 เทคนิค Fusion Convolutional Neural Network

เทคนิค Fusion Convolutional Neural Network [25] เป็นการนำคุณสมบัติในการสกัดคุณลักษณะเฉพาะของสถาปัตยกรรม Convolutional Neural Network ตั้งแต่ 2 รูปแบบขึ้นไปมารวมเข้าด้วยกัน และนำไปสร้างแบบจำลองรูปแบบใหม่ที่มีคุณสมบัติการสกัดคุณลักษณะเฉพาะของทุกสถาปัตยกรรมเพื่อเพิ่มประสิทธิภาพในการจำแนกประเภทของข้อมูล



บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้ได้กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึก โดยมีรายละเอียดดังต่อไปนี้

- 2.1 ภาวะเบาหวานขึ้นตา
- 2.2 การจำแนกรูปภาพ
- 2.3 การเรียนรู้เชิงลึก
- 2.4 โครงข่ายประสาทเทียมคอนโวลูชัน
- 2.5 การเพิ่มประสิทธิภาพแบบจำลอง
- 2.6 การประเมินประสิทธิภาพแบบจำลอง
- 2.7 งานวิจัยที่เกี่ยวข้อง

2.1 ภาวะเบาหวานขึ้นตา (Diabetic Retinopathy)

ภาวะเบาหวานขึ้นตา (Diabetic Retinopathy: DR) เป็นภาวะแทรกซ้อนทางตาซึ่งเกิดจากโรคเบาหวาน มักจะเกิดในผู้ป่วยควบคุมระดับน้ำตาลในเลือดได้ไม่ดี โดย Wilkinson และคณะ [6] กล่าวว่า การที่มีระดับน้ำตาลในเลือดสูงเป็นระยะเวลาานานทำให้เส้นเลือดฝอยที่หล่อเลี้ยงจอประสาทตาเกิดความผิดปกติจนนำไปสู่ความเสียหาย จนส่งผลให้มีเลือดออกในจอประสาทตาและวันลูกตาซึ่งทำให้จอประสาทตาเกิดการลอกไปในที่สุด ทั้งนี้ ภาวะเบาหวานขึ้นตาสามารถแบ่งตามระดับความรุนแรงของโรคได้ 5 ระดับ ประกอบด้วย จอประสาทตาปกติ, ภาวะเบาหวานขึ้นจอประสาทตาขั้นเบา, ภาวะเบาหวานขึ้นจอประสาทตาขั้นปานกลาง, ภาวะเบาหวานขึ้นจอประสาทตาขั้นหนัก และภาวะเบาหวานเข้าจอประสาทตาในระยะที่มีเส้นเลือดงอกใหม่ ซึ่งความผิดปกติของจอประสาทตาดังกล่าวทำให้จอประสาทตาถูกทำลายจนนำไปสู่การสูญเสียการมองเห็น โดยแนวทางการรักษาภาวะเบาหวานขึ้นตาในปัจจุบันนั้นคือการวินิจฉัยและรักษาอาการโดยจักษุแพทย์ผู้เชี่ยวชาญเท่านั้น

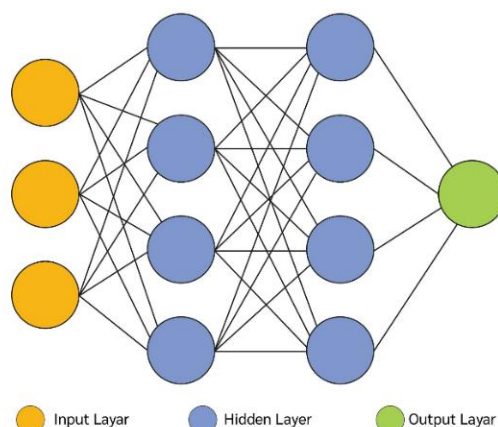
2.2 การจำแนกรูปภาพ (Image Classification)

การจำแนกรูปภาพ (Image Classification) เป็นการจัดกลุ่มหรือประเภทให้กับรูปภาพด้วยกระบวนการเรียนรู้เชิงลึกซึ่งเป็นส่วนหนึ่งในสาขาคอมพิวเตอร์วิทัศน์ (Computer Vision) [7] โดยเป็นการนำข้อมูลรูปภาพมาทำการสกัดคุณลักษณะพิเศษ (Feature) ของวัตถุภายในรูปภาพ เช่น แสง สี เส้นขอบ และความโค้ง โดยคุณลักษณะพิเศษดังกล่าวจะถูกใช้ในการเรียนรู้ของแบบจำลองเพื่อนำไปทำการจำแนกรูปภาพ โดยเป็นการจำแนกให้อยู่ในประเภท (Class) ที่มีคุณลักษณะพิเศษใกล้เคียงกัน อีกทั้งมีนักวิจัยหลายคนนำเทคนิคการจำแนกรูปภาพมาประยุกต์ใช้ร่วมกับการตรวจหาโรค และวินิจฉัยภาวะเบาหวานขึ้นตาจากรูปภาพจอประสาทตา เช่น Wu และคณะ [8] เสนอวิธีการ

จำแนกภาวะเบาหวานขึ้นตาด้วยรูปภาพจอประสาทตาซึ่งเป็นการจำแนกตามระดับความรุนแรงของโรคด้วยสถาปัตยกรรม VGG19, Resnet50 และ InceptionV3 โดยทำการเรียนรู้ร่วมกับเทคนิค Migration Learning รูปแบบ Feature-based Transfer Learning ซึ่งเป็นการนำข้อมูลเดิมที่มีการเรียนรู้แล้ว รวมเข้ากับข้อมูลที่เตรียมทำการเรียนรู้ใหม่ เพื่อช่วยให้ประหยัดเวลาในการเรียนรู้ของแบบจำลองพร้อมทั้งใช้ Keras' Built-in Pre-training Model ในการปรับแต่งพารามิเตอร์ของแบบจำลองเพื่อสามารถถึงประสิทธิภาพการทำงานสูงสุดของแบบจำลองออกมาได้ และได้มีการใช้เทคนิค Data Augmentation ได้แก่ Flipping, Cropping, Rotation, Contrast, Chromatic Aberration และ Brightness เพื่อให้ได้จำนวน และมิติของรูปภาพเพิ่มมากขึ้นซึ่งจะป้องกันการเกิดปัญหา Overfitting ซึ่งจากผลการทดลองพบว่าแบบจำลอง InceptionV3 ที่มีการเรียนรู้ 300 Epoch ได้รับอัตราความถูกต้องสูงสุดที่ร้อยละ 61

2.3 การเรียนรู้เชิงลึก (Deep Learning)

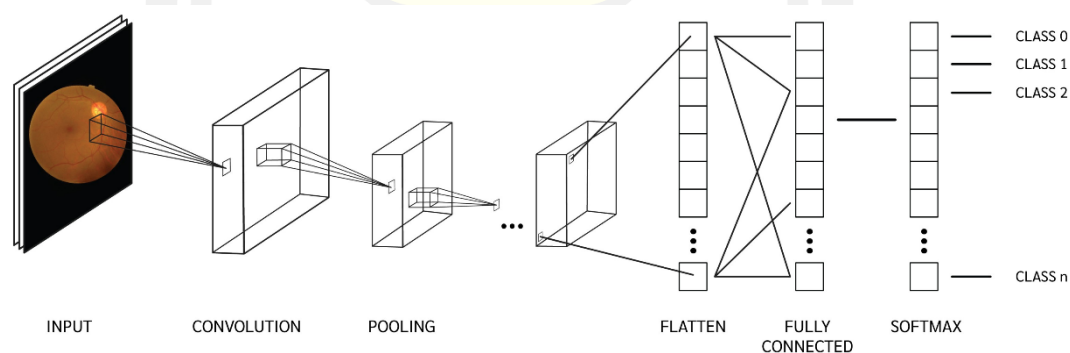
การเรียนรู้เชิงลึก (Deep Learning) เป็นสาขาหนึ่งของการเรียนรู้ของเครื่องจักร (Machine Learning) และปัญญาประดิษฐ์ (Artificial Intelligence) ซึ่งถูกกล่าวถึงโดย Smola และ Vishwanathan [9] การเรียนรู้เชิงลึกเป็นความสามารถในการใช้หลักการคิดพร้อมทั้งหลักการประมวลผลแบบสมองของมนุษย์ซึ่งเรียกว่าโครงข่ายประสาทเทียม (Artificial Neural Network) ซึ่งได้ถูกสร้างขึ้นโดยการนำระบบโครงข่ายประสาทเทียมมาทำการซ้อนกันจำนวนหลายชั้นพร้อมทั้งเรียนรู้ข้อมูลโดยการนำชุดข้อมูลสำหรับการเรียนรู้ของแบบจำลองเพิ่มเข้าไปในส่วนของชั้นรับข้อมูล (Input Layer) เพื่อส่งต่อข้อมูลไปประมวลผลที่ชั้นซ่อน (Hidden Layer) และนำเสนอผลลัพธ์ของการประมวลผลที่ชั้นแสดงผล (Output Layer) ซึ่งในโครงสร้างของโครงข่ายประสาทเทียมระหว่าง Input Layer และ Output Layer ประกอบไปด้วยชั้นจำนวนมากเนื่องจากในแต่ละชั้นนั้นจะช่วยเพิ่มประสิทธิภาพของแบบจำลองให้สามารถตรวจจับหรือจำแนกประเภทข้อมูลได้ถูกต้องแม่นยำมากยิ่งขึ้น และการเรียนรู้ข้อมูลของแบบจำลอง [10] สามารถแบ่งออกได้เป็น 3 ประเภท ได้แก่ การเรียนรู้แบบมีผู้ฝึกสอน (Supervised Learning) การเรียนรู้แบบไม่มีผู้ฝึกสอน (Unsupervised Learning) และการเรียนรู้แบบกึ่งมีผู้สอน (Semi Supervised Learning) โดยโครงข่ายประสาทเทียมถูกนำเสนอไว้หลากหลายรูปแบบ เช่น โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN) [11] โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network: RNN) [12] และโครงข่ายประสาทเทียมแบบคอนโวลูชันย้อนกลับ (Recurrent Convolutional Neural Network: RCNN) [13] เป็นต้น ซึ่งโครงสร้างของโครงข่ายประสาทเทียมจะแสดงดังภาพประกอบที่ 2.1



ภาพประกอบที่ 2.1 โครงสร้างโครงข่ายประสาทเทียม [14]

2.4 โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network)

โครงข่ายประสาทเทียมคอนโวลูชัน (Convolutional Neural Network: CNN) เป็นหนึ่งในการเรียนรู้เชิงลึกแบบมีผู้ฝึกสอน (Supervised Learning) ซึ่งถูกนำไปประยุกต์ใช้กันอย่างแพร่หลาย รวมถึงการประยุกต์ใช้กับการประมวลผลรูปภาพ และ CNN นั้นมีจุดเริ่มต้นมาจากการงานวิจัยด้านการจำแนกรูปภาพตัวอักษรของ LeCun และคณะ [11] เป็นการประมวลผลข้อมูลที่รับเข้ามาแบบ 3 มิติที่มีความกว้าง ความยาว และความลึก โดยใน CNN นั้นมีคุณสมบัติในการสกัดหรือคัดแยกคุณลักษณะพิเศษของข้อมูลหรือรูปภาพ พร้อมกับจำแนกประเภทของข้อมูลได้ และโครงสร้างทั่วไปของ CNN จะแสดงดังภาพประกอบที่ 2.2



ภาพประกอบที่ 2.2 โครงสร้างโครงข่ายประสาทเทียมคอนโวลูชัน

จากภาพประกอบที่ 2.2 แสดงให้เห็นว่าโครงสร้างของ CNN ประกอบไปด้วย 3 ชั้นหลักโดยมีรายละเอียดดังต่อไปนี้

2.4.1 ชั้นคอนโวลูชัน (Convolutional Layer) ทำหน้าที่ในการหาคุณลักษณะพิเศษของรูปภาพหรือกลุ่มข้อมูลที่มีลักษณะใกล้เคียงกัน โดยใช้วิธีการดอทเมทริกซ์ (Dot Matrix) ในการคำนวณส่วนย่อยของรูปภาพ และใช้ตัวกรอง (Filter) หรือเคอร์เนล (Kernel) เพื่อสกัดคุณลักษณะพิเศษ

(Feature Extraction) ออกมาจากรูปภาพ ซึ่งผลลัพธ์ที่ได้จะถูกเรียกว่า Feature Map ตัวอย่างการทำคอนโวลูชันจะแสดงดังภาพประกอบที่ 2.3 และสามารถคำนวณได้จากสมการ (1)

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} * I_{x+i-1,y+j-1} \quad (1)$$

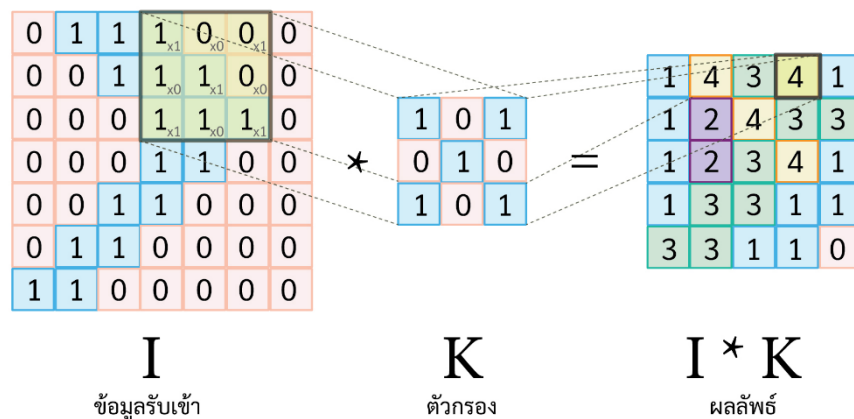
โดย

I คือกลุ่มข้อมูลที่รับเข้ามา

K คือตัวกรองหรือเคอร์เนล

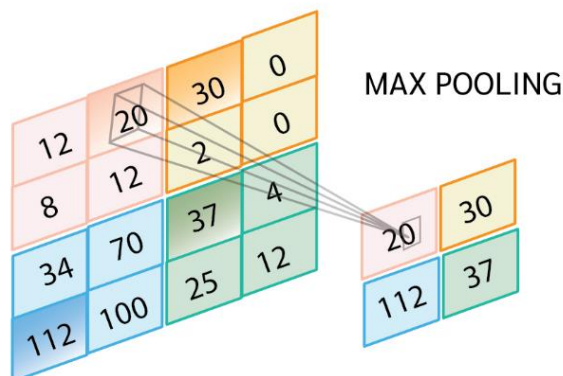
h คือความสูง

w คือความกว้าง



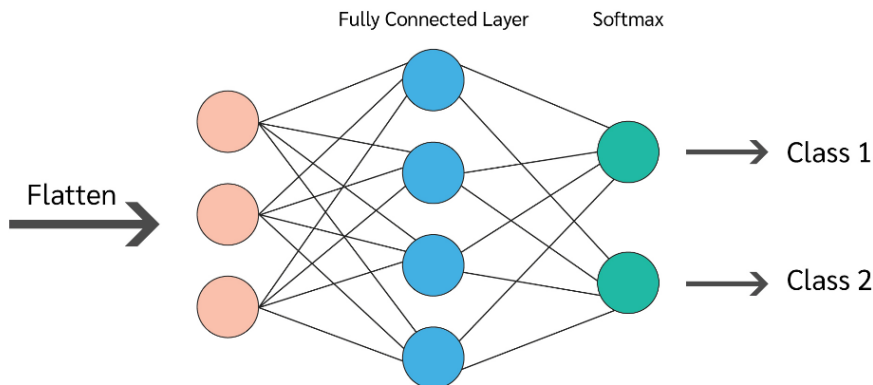
ภาพประกอบที่ 2.3 ตัวอย่างการทำคอนโวลูชัน [15]

2.4.2 ชั้นพูลลิง (Pooling Layer) ทำหน้าที่ในการลดขนาด (Down Sampling) ของ Feature Map ซึ่งนิยมนำมาต่อจาก Convolutional Layer เนื่องจากขนาดของ Feature Map ที่ทำการสกัดออกมาจาก Convolutional Layer มีขนาดของข้อมูลใหญ่ ส่งผลให้การประมวลผลล่าช้าจึงได้มีการแบ่งย่อยรูปภาพโดยใช้หลักการ Sliding Windows ซึ่งจะเป็นการนำมิติของข้อมูลเชิงพื้นที่ของรูปภาพมาทำการแบ่งออกเป็นพื้นที่ย่อย โดยส่วนใหญ่แล้วมักเลือกใช้ฟังก์ชันค่าสูงสุด (Max Pooling) ในการคำนวณเพื่อหาค่าที่มากที่สุดของแต่ละพื้นที่ย่อยเพื่อสร้างเป็น Feature Map ใหม่ที่มีขนาดเล็กลง จากภาพประกอบที่ 2.4 แสดงให้เห็นว่าภายในพื้นที่ย่อยส่วนแรกประกอบด้วยตัวเลข 12, 20, 8, 12 ดังนั้นจึงเลือกตัวเลขที่มีค่าสูงสุดคือ 20 มาพิจารณาเพื่อสร้างเป็น Feature Map ใหม่ ซึ่งตัวอย่างขั้นตอนการทำ Max Pooling จะแสดงดังภาพประกอบที่ 2.4



ภาพประกอบที่ 2.4 ขั้นตอนการทำ Max Pooling [15]

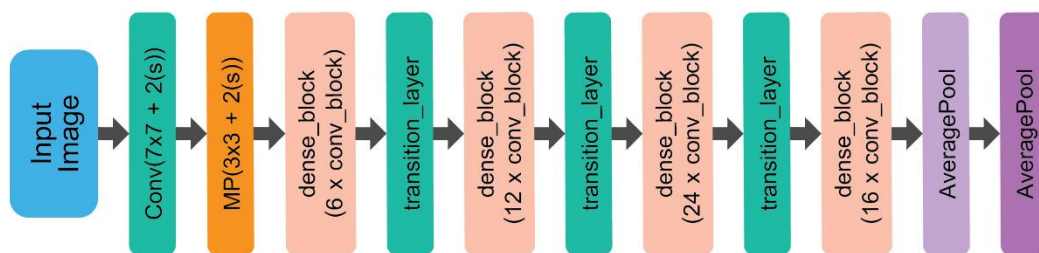
2.4.3 ชั้นการเชื่อมโยงแบบสมบูรณ์ (Fully Connected Layer) จะอยู่ในขั้นสุดท้ายของสถาปัตยกรรม CNN โดยทำหน้าที่ในการแปลงข้อมูล Feature Map 2 มิติ จาก Pooling Layer ให้เป็น Feature Map 1 มิติ ภายในขั้นนี้โดยเรียกกระบวนการนี้ว่า Flatten ก่อนจะนำไปทำการเชื่อมโยงกันระหว่าง Feature Map และชั้น Output แบบสมบูรณ์ จากนั้นจะถูกส่งต่อไปเพื่อคำนวณหรือพยากรณ์หาคำตอบในฟังก์ชัน Softmax ซึ่งตัวอย่างโครงสร้างของชั้นการเชื่อมโยงแบบสมบูรณ์จะแสดงดังภาพประกอบที่ 2.5



ภาพประกอบที่ 2.5 โครงสร้างของชั้นการเชื่อมโยงแบบสมบูรณ์ [15]

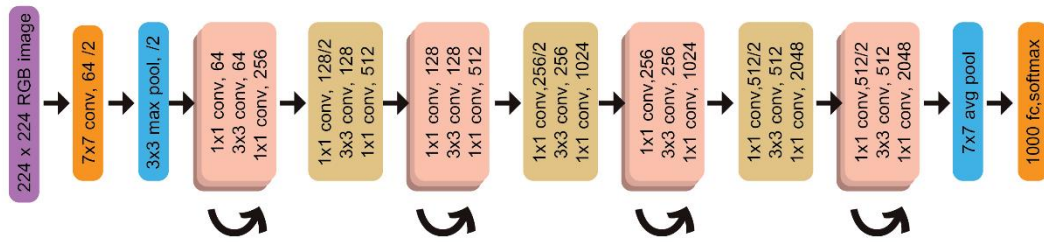
โดยในการวิจัยนี้มีการนำสถาปัตยกรรม CNN เข้ามาช่วยในการรู้จำเพื่อจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจำนวน 5 รูปแบบ ได้แก่ ResNet50, ResNet50V2, Xception, InceptionV3 และ DenseNet121 โดยมีรายละเอียดดังต่อไปนี้

2.4.4 ResNet50 และ ResNet50V2 ถูกนำเสนอโดย He และคณะ [16] [17] ในปี 2015 และ 2016 ตามลำดับ ซึ่งในสถาปัตยกรรม ResNet นั้นมีการนำเสนอการใช้ Residual Block โดยเป็นการนำเอาผลลัพธ์ของชั้นก่อนหน้ามารวมกับผลลัพธ์ของชั้นที่ทำการประมวลผลอยู่ เพื่อลดการเกิดปัญหา Vanishing Gradient เนื่องจากภายใน Residual Block มีการใช้ Skip Connection เพื่อช่วยดึงค่า Gradient จากชั้นก่อนหน้ามาใช้เพื่อช่วยลดเวลาในการเชื่อมต่อกันของแต่ละชั้น ซึ่งจะสามารถลดเวลาในการคำนวณลงโดยที่ประสิทธิภาพของในการทำงานของสถาปัตยกรรมนั้นยังคงเดิม อีกทั้งใน Fully Connected Layer มีการนำเทคนิค Global Average เข้ามาใช้ร่วมด้วย จึงทำให้สถาปัตยกรรม ResNet มีความลึกมากถึง 50 ชั้น และใน ResNet50V2 ได้มีการแก้ไขในส่วนของ Residual Block เพื่อแก้ไขข้อผิดพลาดในการเรียนรู้ของสถาปัตยกรรมซึ่งในการแก้ไวนั้นสามารถลดข้อผิดพลาดในการเรียนรู้ของแบบจำลองได้ร้อยละ 2 และโครงสร้างพื้นฐานของสถาปัตยกรรม ResNet50 จะแสดงดังภาพประกอบที่ 2.6



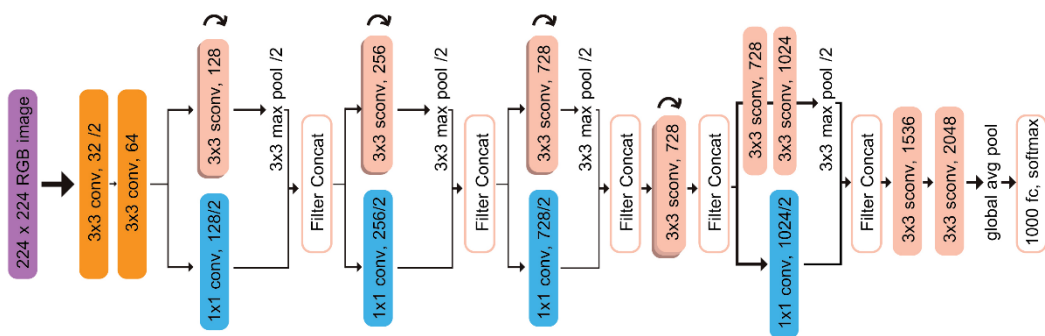
ภาพประกอบที่ 2.6 โครงสร้าง ResNet50 [18]

2.4.5 Xception ถูกคิดค้นโดย Francois Chollet [19] ในปี 2017 โดยเป็นการพัฒนามาจากสถาปัตยกรรม Inception ซึ่งในส่วนประกอบของสถาปัตยกรรม Xception มีการใช้ Depthwise Separable Convolutional Layer แทนที่ Convolutional Layer อีกทั้งมีการใช้ Skip Connection เช่นเดียวกับสถาปัตยกรรม ResNet เพื่อลดจำนวนของพารามิเตอร์ และระยะเวลาในการคำนวณพร้อมทั้งระยะเวลาในการเชื่อมต่อกันของแต่ละชั้นลดน้อยลง ด้วยเหตุนี้ผลการทดลองประสิทธิภาพการทำงานของสถาปัตยกรรม Xception บนชุดข้อมูล ImageNet พบว่ามีอัตราความถูกต้องและความเร็วมากกว่าสถาปัตยกรรม Inception โดยโครงสร้างพื้นฐานของสถาปัตยกรรม Xception จะแสดงดังภาพประกอบที่ 2.7



ภาพประกอบที่ 2.7 โครงสร้าง Xception [20]

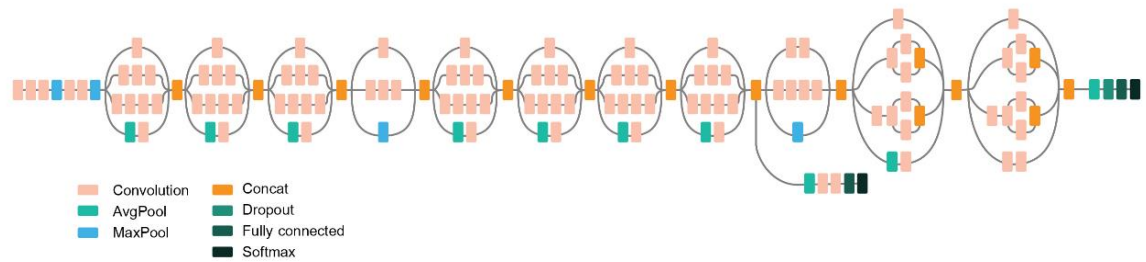
2.4.6 InceptionV3 ถูกนำเสนอโดย Szegedy และคณะ [21] ในปี 2016 ซึ่งเป็นการพัฒนามาจากสถาปัตยกรรม InceptionV1 และ InceptionV2 ทั้งนี้ แนวคิดหลักของ InceptionV3 มุ่งเน้นไปที่การทำให้แบบจำลองมีจำนวนพารามิเตอร์ลดน้อยลงแต่ประสิทธิภาพในการทำงานยังคงเดิม ซึ่งใน InceptionV3 ได้มีการลดขนาดของคอร์เนลสำหรับการสกัดคุณลักษณะพิเศษจากชุดข้อมูลเพื่อหลีกเลี่ยงปัญหาอคติที่เกิดขึ้นใน Inception รุ่นเดิม อีกทั้งยังใช้วิธีแยกตัวประกอบ (Factorization) ของ Convolutional Layer เพื่อให้มีขนาดเล็กลง ตัวอย่างเช่น จากเดิมขนาดของคอร์เนลภายใน Convolutional Layer มีขนาด 5×5 Pixels ถูกปรับให้ลดลงเหลือเพียง 3×3 Pixels และใน Pooling Layer จากเดิมคอร์เนลมีขนาด 3×3 Pixels ถูกปรับให้ลดลงเหลือ 2×2 Pixels ซึ่งการลดขนาดของคอร์เนลจะทำให้จำนวนของพารามิเตอร์ และเวลาในการคำนวณลดลง โดยโครงสร้างพื้นฐานของสถาปัตยกรรม InceptionV3 จะแสดงดังภาพประกอบที่ 2.8



ภาพประกอบที่ 2.8 โครงสร้าง InceptionV3 [22]

2.4.7 DenseNet121 ถูกนำเสนอโดย Huang และคณะ [23] ในปี 2017 เป็นสถาปัตยกรรมที่มีแนวคิดมาจากสถาปัตยกรรม ResNet แต่ในส่วนของ Residual Block นั้นได้ถูกเปลี่ยนเป็น Dense Block ซึ่งภายในมีความซับซ้อนมากกว่า เนื่องจากมีการใช้ Skip Connection ที่นำเอา Feature Map ในชั้นก่อนหน้า Concatenate กับ Feature Map ในชั้นปัจจุบันก่อนจะส่งไปเป็น Input ของ

ขั้นถัดไป จากการทดลองบนชุดข้อมูล ImageNet พบว่าสถาปัตยกรรม DenseNet มีจำนวนพารามิเตอร์น้อยกว่าสถาปัตยกรรม ResNet แต่ให้อัตราความถูกต้องแม่นยำในการจำแนกประเภทของข้อมูลสูงกว่า ซึ่งโครงสร้างพื้นฐานของสถาปัตยกรรม DenseNet121 จะแสดงดังภาพประกอบที่ 2.9



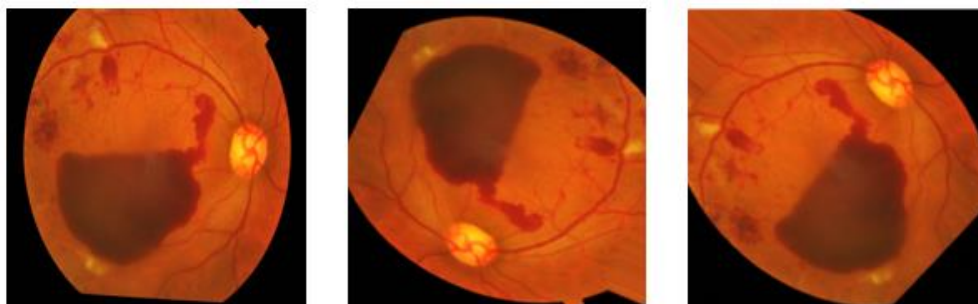
ภาพประกอบที่ 2.9 โครงสร้าง DenseNet121 [23]

2.5 การเพิ่มประสิทธิภาพแบบจำลอง

2.5.1 เทคนิค Data Augmentation

การเรียนรู้ของ CNN สำหรับการรู้จำและจำแนกรูปภาพนั้น ประสิทธิภาพของแบบจำลองจะขึ้นอยู่กับชุดข้อมูลรูปภาพที่ถูกนำไปเข้ากระบวนการรู้จำ ดังนั้นจึงจำเป็นต้องใช้ชุดข้อมูลขนาดใหญ่พร้อมทั้งมีรูปภาพที่มีมุมมองหลากหลายมิติเพื่อหลีกเลี่ยงการเกิดปัญหา Overfitting ด้วยเหตุนี้จึงมีการนำเทคนิค Data Augmentation เข้ามาช่วยในการเพิ่มจำนวนของชุดข้อมูลที่น่าไปทำการเรียนรู้แบบจำลอง โดยในการวิจัยนี้ได้ใช้ Data Augmentation สำหรับการรู้จำของแบบจำลองด้วยการเครื่องมือ Augmentor ที่นำเสนอโดย Bloice และคณะ [24] โดยมีการดัดแปลงรูปภาพให้กลายเป็นรูปภาพใหม่ให้มีความใกล้เคียงกับรูปภาพต้นฉบับ และการดัดแปลงรูปภาพสามารถทำได้หลายรูปแบบ เช่น Rotation, Flip และ Zoom โดยมีรายละเอียดดังต่อไปนี้

Rotation เป็นการนำรูปภาพมาทำการหมุนรูปภาพตามองศาที่ทำการกำหนดไว้โดยตัวอย่างการทำ Data Augmentation รูปแบบ Rotation จะแสดงดังภาพประกอบที่ 2.10

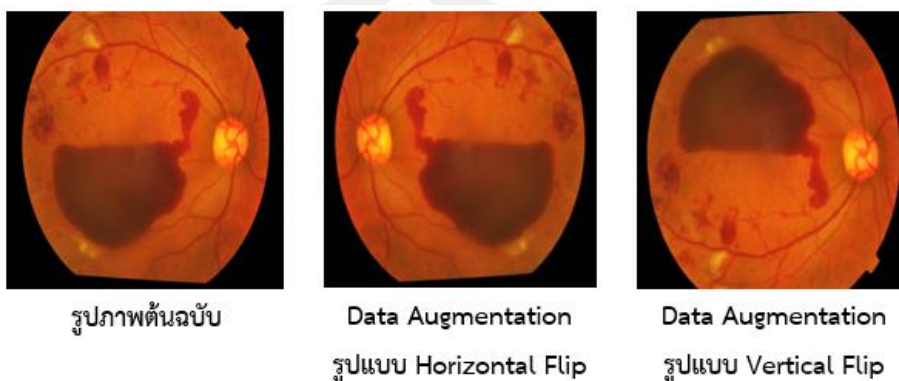


รูปภาพต้นฉบับ

Data Augmentation รูปแบบ Rotation

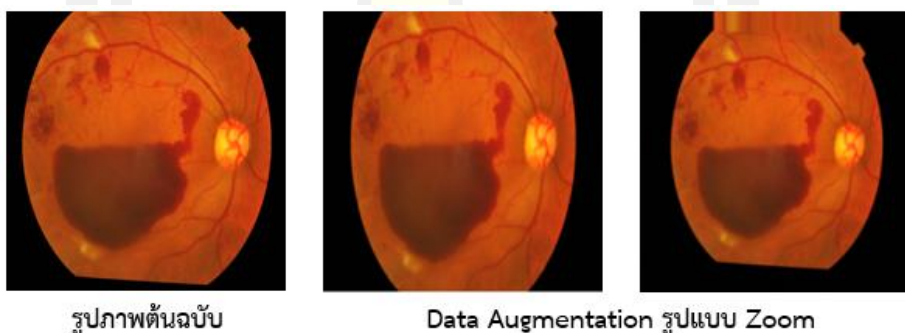
ภาพประกอบที่ 2.10 การทำ Data Augmentation รูปแบบ Rotation

Flips เป็นการพลิกรูปภาพซึ่งมีการพลิกในแนวนอน (Horizontal Flip) และแนวนอน (Vertical Flip) โดยตัวอย่างการทำ Data Augmentation รูปแบบ Horizontal Flip และ Vertical Flip จะแสดงดังภาพประกอบที่ 2.11



ภาพประกอบที่ 2.11 การทำ Data Augmentation รูปแบบ Horizontal Flip และ Vertical Flip

Zoom เป็นการขยายรูปภาพตามระยะของเขตที่กำหนดไว้ โดยตัวอย่างการทำ Data Augmentation รูปแบบ Zoom จะแสดงดังภาพประกอบที่ 2.12



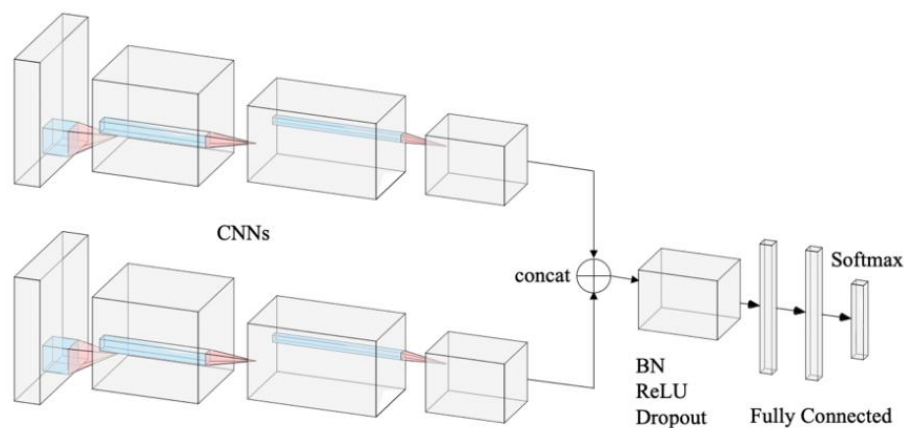
ภาพประกอบที่ 2.12 การทำ Data Augmentation รูปแบบ Zoom

2.5.2 เทคนิค Fusion Convolutional Neural Network

Fusion Convolutional Neural Network หรือ Fusion CNN [25] เป็นการนำเอาคุณสมบัติในการสกัดคุณลักษณะเฉพาะของสถาปัตยกรรม CNN ที่แตกต่างกันตั้งแต่ 2 รูปแบบขึ้นไปมารวมเข้าด้วยกัน จากนั้นนำไปเชื่อมต่อกับทุกชั้นภายใน Fully Connected Layer เพื่อสร้างเป็นแบบจำลองรูปแบบใหม่ที่มีคุณสมบัติในการสกัดคุณลักษณะเฉพาะของทุก CNN และในการทำ Fusion CNN จะประกอบด้วย 2 กระบวนการ โดยมีรายละเอียดดังต่อไปนี้

- 1) เลือกสถาปัตยกรรม CNN ที่แตกต่างกันโดยมีอัตราความถูกต้องในการจำแนกประเภทข้อมูลสูงในการทดสอบบนชุดข้อมูลเดียวกันตั้งแต่ 2 รูปแบบ ขึ้นไปเพื่อทำการคัดแยกคุณสมบัติในการสกัดคุณลักษณะเฉพาะของ CNN แต่ละรูปแบบซึ่งเรียกว่า Feature Map มาทำการ Concatenated เพื่อสร้างเป็น Feature Map รูปแบบใหม่ที่รวมคุณสมบัติของทุก CNN เข้าไว้ด้วยกันเพื่อทำการส่งไปประมวลผลในขั้นถัดไป

2) นำ Feature Map รูปแบบใหม่ส่งไปยังชั้น Batch Normalization (BN), Rectified Linear Units (ReLU) และ Dropout ตามลำดับ ซึ่งทุกชั้นในสถาปัตยกรรมจะถูกเชื่อมต่อเข้าด้วยกันแบบสมบูรณ์ใน Fully Connected Layer และใช้ฟังก์ชัน Softmax ในการประมวลผลเพื่อพยากรณ์คำตอบของการจำแนกประเภทข้อมูล โดยโครงสร้างของการทำ Fusion CNN จะแสดงดังภาพประกอบที่ 2.13



ภาพประกอบที่ 2.13 โครงสร้าง Fusion Convolution Neural Network

2.5.3 เทคนิค Ensemble Convolutional Neural Network

Ensemble Convolutional Neural Network หรือ Ensemble CNN [26] เป็นการนำแบบจำลอง CNN ที่ได้รับการเรียนรู้จากข้อมูลชุดเดียวกันจำนวน 2 CNN ไปทำการพยากรณ์ข้อมูลเพื่อให้ได้ค่าความน่าจะเป็นของแต่ละแบบจำลอง และนำค่าความน่าจะเป็นของทั้ง 2 แบบจำลองไปทำการ Vote หรือ การคำนวณหาค่าเฉลี่ยของคำตอบ เพื่อให้ได้คำตอบที่ถูกต้องมากที่สุด ซึ่งเรียกว่ากระบวนการ Average Voting [27][28] โดยจะเป็นการเพิ่มความถูกต้องแม่นยำในการจำแนกประเภทของแบบจำลอง CNN โดยโครงสร้างของการทำ Ensemble CNN จะแสดงดังภาพประกอบที่ 2.14 และการทำ Average Voting นั้นสามารถทำได้ทั้งแบบถ่วงน้ำหนัก (Weighted Average) และ ไม่ถ่วงน้ำหนัก (Unweighted Average) โดยมีรายละเอียดดังนี้

1) Weighted Average ในเทคนิคนี้จะมีการกำหนดค่าน้ำหนักของพารามิเตอร์ในแต่ละแบบจำลองเพื่อหาค่าน้ำหนักที่เหมาะสมที่สุด โดยในการกำหนดค่าน้ำหนักนั้นเมื่อนำมารวมกันจะต้องมีค่าเท่ากับ 1 ซึ่งสามารถคำนวณได้ดังสมการที่ (2)

$$\hat{y} = \frac{1}{n} \sum_{i=1}^n \alpha \vec{y}_i \quad (2)$$

โดย

α คือค่าน้ำหนักที่คูณเข้ากับเวกเตอร์ของผลลัพธ์ของความน่าจะเป็น (\vec{y}_i) และ n คือจำนวนของแบบจำลองที่นำมาทำการ Ensemble

2) Unweighted Average ในเทคนิคนี้จะทำการนำผลลัพธ์ของแต่ละแบบจำลองมาทำการคำนวณเพื่อหาค่าเฉลี่ยของความน่าจะเป็น โดยค่าเฉลี่ยที่มีค่าสูงสุดจะถูกเลือกเป็นคำตอบของ Ensemble ซึ่งสามารถคำนวณได้ดังสมการที่ (3)

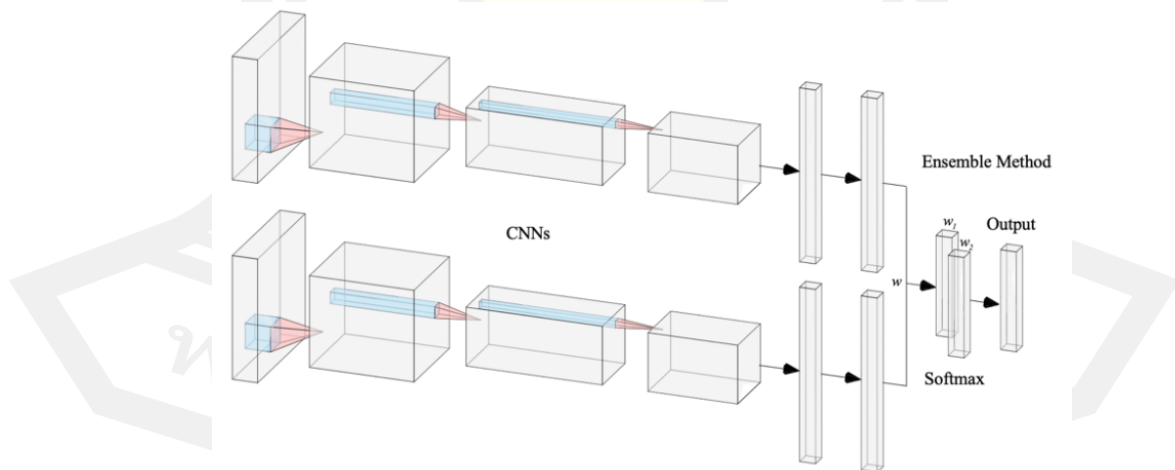
$$\hat{y} = \frac{1}{n} \sum_{i=1}^n \vec{y}_i \quad (3)$$

โดย

\vec{y}_i คือเวกเตอร์ของผลลัพธ์ของความน่าจะเป็นของแต่ละแบบจำลอง

n คือจำนวนของแบบจำลองที่นำมาทำการ Ensemble

จากนั้นใช้ฟังก์ชัน *argmax* เลือกค่าความน่าจะเป็นที่สูงที่สุดของ \vec{y}_i



ภาพประกอบที่ 2.14 โครงสร้าง Ensemble Convolutional Neural Network

2.6 การประเมินประสิทธิภาพแบบจำลอง

2.6.1 คอนฟิวชันเมทริกซ์ (Confusion Matrix)

คอนฟิวชันเมทริกซ์ [29] เป็นการประเมินผลลัพธ์การทำนายของแบบจำลอง โดยนำผลลัพธ์ที่ได้จากแบบจำลองไปทำการเปรียบเทียบกับผลลัพธ์ที่เป็นความจริง และในการวิจัยนี้เป็นการจำแนกประเภทรูปภาพออกเป็น 2 คลาส จึงสามารถเกิดเหตุการณ์ได้ดังภาพประกอบที่ 2.15

		คลาสที่ทำนาย	
		จอประสาทตาปกติ	จอประสาทตาที่มีภาวะเบาหวานขึ้นตา
คลาสจริง	จอประสาทตาปกติ	TP	FN
	จอประสาทตาที่มีภาวะเบาหวานขึ้นตา	FP	TN

ภาพประกอบที่ 2.15 คอนฟิวชันเมทริกซ์การจำแนกแบบ 2 คลาส [29]

โดย

True Positive (TP) คือค่าที่ทำนายว่าจริง และข้อมูลนั้นเป็นจริง

True Negative (TN) คือค่าที่ทำนายว่าจริง แต่ข้อมูลนั้นเป็นเท็จ

False Positive (FP) คือค่าที่ทำนายว่าเท็จ และข้อมูลนั้นเป็นเท็จ

False Negative (FN) คือค่าที่ทำนายว่าเท็จ แต่ข้อมูลนั้นเป็นจริง

2.6.2 ค่าความถูกต้อง (Accuracy)

การวัดค่าความถูกต้องเป็นการวัดค่าความถูกต้องโดยรวมของแบบจำลอง ซึ่งสามารถหาได้จากจากการพยากรณ์ทั้งหมดนั้น แบบจำลองสามารถทำนายถูกกี่ครั้ง และทำการคิดออกมาเป็นค่าร้อยละ โดยสามารถคำนวณหาค่าความถูกต้องได้จากสมการที่ (4)

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \quad (4)$$

2.7 งานวิจัยที่เกี่ยวข้อง

ในปีที่ผ่านมา มีนักวิจัยหลายคนได้นำเสนองานวิจัยที่ใช้เทคนิคการรู้จำรูปภาพเข้ามาช่วยในการคัดกรองโรคในเบื้องต้นสำหรับการแยกระดับความรุนแรงของภาวะเบาหวานขึ้นตาซึ่งในกระบวนการนี้ไม่ต้องผ่านการวินิจฉัยจากจักษุแพทย์โดยตรงส่งผลให้การคัดกรองโรคเป็นไปด้วยความรวดเร็วและอาจลดการสูญเสียการมองเห็นของผู้ป่วยรวมทั้งลดภาระงานของจักษุแพทย์อีกด้วย

จากปัญหาที่ได้กล่าวมา Hattiya และคณะ [30] ได้นำเสนอการจำแนกภาวะเบาหวานขึ้นตาจากรูปภาพจอประสาทตาด้วยสถาปัตยกรรม CNN 7 รูปแบบ ได้แก่ AlexNet, ResNet50, DenseNet201, InceptionV3, MobileNet, MnasNet และ NASNetMobile ซึ่งได้ทำการทดลองกับชุดข้อมูลที่เผยแพร่บนเว็บไซต์ Kaggle จำนวน 23,513 รูปภาพ โดยรูปภาพจอประสาทตาได้ถูกแบ่งออกเป็นรูปภาพจอประสาทตาปกติ และรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา ซึ่งในการทดลองในส่วนแรกนั้นมีการทดลองหาค่าสีที่เหมาะสมสำหรับการจำแนกรูปภาพจากค่าสี 5 รูปแบบ ได้แก่ RGB, grayscale, HSV, L^*a^*b และ YCbCr จากผลการทดลองพบว่าในชุดข้อมูลแบบค่าสี RGB มีความถูกต้องแม่นยำมากที่สุดจึงเหมาะสำหรับการเรียนรู้เพื่อสร้างแบบจำลองสำหรับการจำแนกรูปภาพจอประสาทตา ส่วนต่อมาเป็นการทดลองและเปรียบเทียบประสิทธิภาพการจำแนกของสถาปัตยกรรม CNN โดยใช้ชุดข้อมูลเดียวกัน และจากผลการทดลองพบว่าสถาปัตยกรรม AlexNet ให้อัตราความถูกต้องสูงสุดที่ร้อยละ 81.42

Burewar และคณะ [31] นำเสนอแบบจำลองสำหรับการจำแนกภาวะเบาหวานขึ้นตาตามระดับความรุนแรงของโรคจากรูปภาพจอประสาทตา ซึ่งโครงสร้างของแบบจำลองที่นำเสนอขึ้นเป็นการนำคุณสมบัติของ U-Net มาประยุกต์ใช้เพื่อนำเทคนิค Segmentation มาทำการตรวจจับขอบเขตหรือรูปร่างของหลอดเลือดภายในจอประสาทตา จากนั้นนำผลลัพธ์ที่ได้ส่งต่อไปเพื่อให้แบบจำลอง CNN ได้ทำการรู้จำ ซึ่งการนำเทคนิค Segmentation เข้ามาร่วมใช้ในการจำแนกนั้นช่วยเพิ่มคุณลักษณะพิเศษเพื่อให้แบบจำลองมีข้อมูลสำหรับการรู้จำเพิ่มมากขึ้น และส่งผลให้แบบจำลองมีประสิทธิภาพในการจำแนกประเภทได้ถูกต้องแม่นยำ อีกทั้งในการทดลองได้มีการใช้เทคนิค Data Augmentation ได้แก่ Translation, Stretching, Rotation, Flipping และ Color Augmentation เข้ามาร่วมในการเรียนรู้เพื่อเพิ่มความหลากหลายในด้านมิติของรูปภาพและเพื่อหลีกเลี่ยงการเกิดปัญหา Overfitting จากผลการทดลองพบว่าการนำเทคนิค Segmentation เข้ามาร่วมประมวลผลนั้นให้อัตราความถูกต้องสูงสุดที่ร้อยละ 93.64

Sudarmadji และคณะ [32] นำสถาปัตยกรรม VGG มาพัฒนาร่วมกับ Genetic Algorithm และ Genetic Algorithm Search Optimization อีกทั้งมีการปรับแต่งค่าพารามิเตอร์ให้เหมาะสมเพื่อให้ได้แบบจำลองที่สามารถจำแนกภาวะเบาหวานขึ้นตาพร้อมกับจำแนกได้ตามระดับความรุนแรงของโรคโดยมีความถูกต้องแม่นยำสูง และในโครงสร้างของสถาปัตยกรรมที่พัฒนาภายในแต่ละ

Convolutional Block นั้นจะประกอบไปด้วย Convolutional Layer ที่มีขนาดแตกต่างกันจึงทำให้มีประสิทธิภาพในการจำแนกได้ดีกว่า VGG แบบดั้งเดิม เมื่อทดสอบบนชุดข้อมูล Messidor และ Kaggle ซึ่งประกอบด้วยรูปภาพจอประสาทตาจำนวน 89,902 รูปภาพ พบว่าแบบจำลองที่ทำการปรับปรุงแล้วให้อัตราความถูกต้องแม่นยำที่ร้อยละ 99.66 ซึ่งมากกว่า VggNet-s ที่มีอัตราความถูกต้องที่ร้อยละ 96

Thota และคณะ [33] เสนอการจำแนกภาวะเบาหวานขึ้นตาตามระดับความรุนแรงของโรคด้วยแบบจำลอง VGG16 ที่ได้มีการปรับแต่งค่าพารามิเตอร์ให้เหมาะสมสำหรับการจำแนกรูปภาพจอประสาทตา โดยมีการกำหนดค่า Learning Rate = 0.001, Optimization = ADAM, Loss function = Categorical Cross Entropy และ Batch Size = 100 อีกทั้งมีการใช้เทคนิค Data Augmentation หลายรูปแบบ ได้แก่ Rotation ที่กำหนดค่าระยะการหมุนเป็น 0 – 360, Flipping, Changing Brightness, Contrast, Hue และ Saturation Levels จากผลการทดลองบนชุดข้อมูล EyePACS พบว่าแบบจำลอง VGG16 ที่ถูกปรับแต่งค่าพารามิเตอร์ให้เหมาะสมกับชุดข้อมูลนั้นสามารถให้อัตราความถูกต้องในการจำแนกภาวะเบาหวานขึ้นตาตามระดับความรุนแรงโรคจากรูปภาพจอประสาทตาที่ร้อยละ 74

Wan และคณะ [34] ทำการจำแนกประเภทภาวะเบาหวานขึ้นตาจากรูปภาพจอประสาทตาโดยใช้ CNN ดังนี้ AlexNet, VggNet, GoogleNet และ ResNet ซึ่งมีการใช้เทคนิค 5 Cross-Validation เพื่อแบ่งชุดข้อมูลสำหรับการทดลอง และผลลัพธ์การทดลองการจำแนกประเภทในขั้นแรกให้ผลลัพธ์ที่มีอัตราความถูกต้องต่ำจึงได้มีการใช้เทคนิค Transfer Learning เพื่อช่วยในการเพิ่มประสิทธิภาพในการจำแนกประเภทของแบบจำลอง ได้แก่ เพิ่มรอบการเรียนรู้แบบของจำลองจำนวน 30 รอบ และกำหนดให้ Optimization เป็นอัลกอริทึม Stochastic Gradient Descent (SGD) พร้อมทั้งมีการทดลองปรับค่า Learning Rate ในระหว่างช่วง 0.0001 - 0.1 เพื่อหาค่าที่เหมาะสมที่สุดของแต่ละสถาปัตยกรรม จากผลการทดลองพบว่า VggNet-s เป็นแบบจำลองที่มีความแม่นยำในการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาได้สูงสุด ซึ่งมีอัตราความแม่นยำมากถึงร้อยละ 95.68

Deepa และคณะ [35] เสนอวิธีการจำแนกระดับความรุนแรงของภาวะเบาหวานขึ้นตาจากรูปภาพจอประสาทตาแบบอัตโนมัติด้วย CNN ในการทดลองได้เลือกใช้สถาปัตยกรรม InceptionV3 และ Xception เพื่อทำการรู้จำรูปภาพจอประสาทตาทั้งหมด 2,290 รูปภาพ อีกทั้งมีการปรับค่าพารามิเตอร์ให้เหมาะสมกับแบบจำลองโดยกำหนดค่าดังนี้ Batch Size = 64, Momentum = 0.9, Optimizers = SGD และ Adagrad และ Learning Rate = 0.01 เมื่อทำการเรียนรู้และรู้จำรูปภาพแล้วนำแบบจำลองที่ได้ทั้ง 2 รูปแบบ ไปเข้ากระบวนการ Ensemble CNN และเลือกผลการ

พยากรณ์ด้วยวิธี Majority Voting และจากการทดลองพบว่าการทำ Ensemble CNN ให้อัตราความถูกต้องแม่นยำในการจำแนกประเภทรูปภาพจอประสาทตามากถึงร้อยละ 95

Li และคณะ [36] นำเสนอการจำแนกประเภทระหว่าง รูปภาพจอประสาทตาปกติ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยสถาปัตยกรรม CNN ที่แตกต่างกัน 5 รูปแบบ ซึ่งประกอบไปด้วย VGG16, ResNet18, GoogleNet, DenseNet121, และ SE-BN-Inception โดยในการทดลองได้ใช้รูปภาพจอประสาทตาในการทดลองจำนวน 13,673 รูปภาพ และใช้ร่วมกับเทคนิค Data Augmentation ได้แก่ Translation, Stretching, Rotation, Flipping, และ Color Augmentation เพื่อลดปัญหา Overfitting จากการทดลองพบว่า SE-BN-Inception มีประสิทธิภาพในการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาสองสุด โดยให้อัตราความถูกต้องที่ร้อยละ 54.18

Sharifrazi และคณะ [37] นำเสนอการใช้เทคนิค Fusion โดยในการทดลองเป็นการรวมคุณสมบัติของสถาปัตยกรรม CNN และเทคนิค Support Vector Machine (SVM) เข้าด้วยกันเพื่อสร้างเป็นสถาปัตยกรรม 2D-CNN แบบใหม่ เพื่อตรวจหาเชื้อ COVID-19 จากภาพเอ็กซเรย์ โดยภายใน Fully Connected Layer ได้ทำการเพิ่มฟังก์ชัน Sigmoid Activation และ SVM เพื่อจำแนกภาพเอ็กซเรย์ปอดที่มีเชื้อ COVID-19 อีกทั้งมีการใช้ Sobel Filter เพื่อกำหนดเส้นหรือรูปทรงของปอดก่อนนำไปประมวลผลภายในแบบจำลอง และมีการใช้เทคนิค Data Augmentation ได้แก่ Shift, Rotation และ Changing Brightness เพื่อเพิ่มจำนวนของชุดข้อมูลสำหรับการเรียนรู้ของแบบจำลองซึ่งจากเดิม 333 รูปภาพ เพิ่มเป็น 1,332 รูปภาพ และจากผลการทดลองพบว่าแบบจำลองที่นำเสนอสามารถตรวจหาผู้ป่วย COVID-19 ได้อย่างถูกต้องแม่นยำ โดยมีอัตราความถูกต้องสูงถึงร้อยละ 99.02

Yazhini และ Loganathan [38] นำเสนอแบบจำลอง FM-GLCM-VGG19 สำหรับการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาตามระดับความรุนแรงของโรค โดยใช้เทคนิค Fusion Model ซึ่งเป็นการรวมเอาคุณสมบัติการสกัดคุณลักษณะพิเศษจากรูปภาพของหลายสถาปัตยกรรมรวมเข้าด้วยกันเพื่อสร้างเป็นสถาปัตยกรรมรูปแบบใหม่ และในการทดลองได้เลือกการสกัดคุณลักษณะพิเศษของเทคนิคการวิเคราะห์เชิงพื้นที่ด้วยหลักการเมทริกซ์ระดับเทา (Gray-Level Co-occurrence Matrix: GLCM) และการสกัดคุณลักษณะพิเศษของสถาปัตยกรรม VGG19 และยังคงใช้ฟังก์ชัน Softmax ในการประมวลผลเพื่อจำแนกประเภทของรูปภาพ จากการทดลองบนชุดข้อมูลจากเว็บไซต์ Kaggle พบว่าแบบจำลอง FM-GLCM-VGG19 ให้อัตราความถูกต้องแม่นยำที่ร้อยละ 71.30

Rehman และคณะ [39] นำเสนอเครื่องมือในการตรวจหาภาวะเบาหวานขึ้นตาจากรูปภาพจอประสาทตาแบบอัตโนมัติ โดยในการทดลองได้เลือกใช้สถาปัตยกรรม CNN ที่ได้มีการเรียนรู้มาก่อน

แล้วจำนวน 3 รูปแบบ ได้แก่ AlexNet, VGG-16 และ SqueezeNet อีกทั้งได้นำเสนอสถาปัตยกรรม CNN แบบ 5 ชั้นที่ทำการปรับแต่งด้วยตนเอง ซึ่งประกอบด้วย Convolutional Layer จำนวน 2 ชั้น และ Fully Connected Layer จำนวน 3 ชั้น โดยมีรายละเอียดดังนี้ภายในชั้นแรกกำหนดให้คอร์เนลมีขนาดที่แตกต่างกันทั้งหมด 4 ขนาด สำหรับใช้ในการสกัดคุณลักษณะพิเศษของข้อมูล และยังได้ลดขนาดของ Convolutional Layer เพื่อลดระยะเวลาในการคำนวณ และในชั้นต่อมาได้ทำการรวมผลลัพธ์ของคอร์เนลทั้ง 4 รูปแบบเข้าด้วยกัน ซึ่งเทียบได้ว่าเป็น Pooling Layer และ 3 ชั้นสุดท้ายของสถาปัตยกรรม 5 ชั้น เป็นส่วนของ Fully Connected Layer โดยมีจำนวนนิวรอน 100, 50 และ 10 ตามลำดับ ซึ่งการกำหนดค่าของสถาปัตยกรรมดังกล่าวสามารถหลีกเลี่ยงการเกิดปัญหา Overfitting ได้ จากผลการทดลองบนชุดข้อมูลเดียวกัน พบว่าแบบจำลองรูปแบบ 5 ชั้น มีผลลัพธ์ที่ดีที่สุดโดยมีอัตราความไว (Sensitivity) มากถึงร้อยละ 98.94 ซึ่งเป็นการวัดความสามารถความไวในการตรวจหารูปภาพจอประสาทตา และยังให้อัตราความถูกต้องมากถึงร้อยละ 98.15

Shanthi และคณะ [40] ปรับปรุงโครงสร้างของสถาปัตยกรรม Alexnet เพื่อประยุกต์ใช้ในการจำแนกประเภทจอประสาทตาที่มีภาวะเบาหวานขึ้นตา และแยกตามระดับความรุนแรงของโรคจากรูปภาพจอประสาทตา ซึ่งในการปรับปรุงโครงสร้างของสถาปัตยกรรม Alexnet นั้นในขั้นตอนแรกได้เปลี่ยนขนาดอินพุตของรูปภาพให้เหลือเพียง 259×259 Pixels เพื่อให้สอดคล้องกับความกว้าง ความสูง และช่องสีสามช่อง ที่แสดงถึงความลึกของรูปภาพจอประสาทตา ผลลัพธ์ของชั้นก่อนหน้าจะถูกนำมาถูกคำนวณเป็นผลคูณเชิงสเกลาร์ซึ่งทำซ้ำตามความยาวและความกว้างภายใน Convolutional Layer ภายในชั้น Rectified Linear Unit (RELU) ใช้ Negative Activations ทั้งหมดเป็น 0 และ Pooling Layer จะทำการรวมผลลัพธ์ในการคำนวณทั้งหมดเข้าด้วยกันโดยขนาดของผลลัพธ์จะถูกลดลงตามพิกัดเชิงพื้นที่ และใน Fully Connected Layer จะเป็นชั้นสำหรับการจำแนกประเภทของข้อมูล โดยจะทำการเลือกค่าตอบที่ถูกต้องจากผลลัพธ์ของความน่าจะเป็นสูงสุด และจากผลการทดลองจากชุดข้อมูล Messidor พบว่าในการจำแนกจอประสาทตาปกติ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตาระยะที่ 3 มีอัตราความถูกต้องสูงสุดที่ร้อยละ 96.6

Kassani และคณะ [41] ปรับปรุงสถาปัตยกรรม Xception สำหรับการจำแนกประเภทและระดับความรุนแรงของภาวะเบาหวานขึ้นตา ซึ่งมีการปรับปรุงโครงสร้างในชั้นบนสุดโดยการเพิ่มการ Stacking High-Level ของคุณลักษณะเฉพาะจากชั้น Intermediate CNN และใน Convolutional Layer แบบเดิมจะถูกแทนที่ด้วย Depthwise Separable Convolution Layer โดยจะช่วยลดระยะเวลา และความซับซ้อนในการคำนวณ อีกทั้งใน Pooling Layer มีการเปลี่ยนไปใช้เทคนิค Min Pooling สำหรับช่วยในการสกัดคุณลักษณะพิเศษของข้อมูล และปรับปรุงประสิทธิภาพของการจำแนกหมวดหมู่ ซึ่งเป็นประโยชน์อย่างมากในการนำไปประยุกต์ใช้ในการวินิจฉัยโรคแบบเรียลไทม์

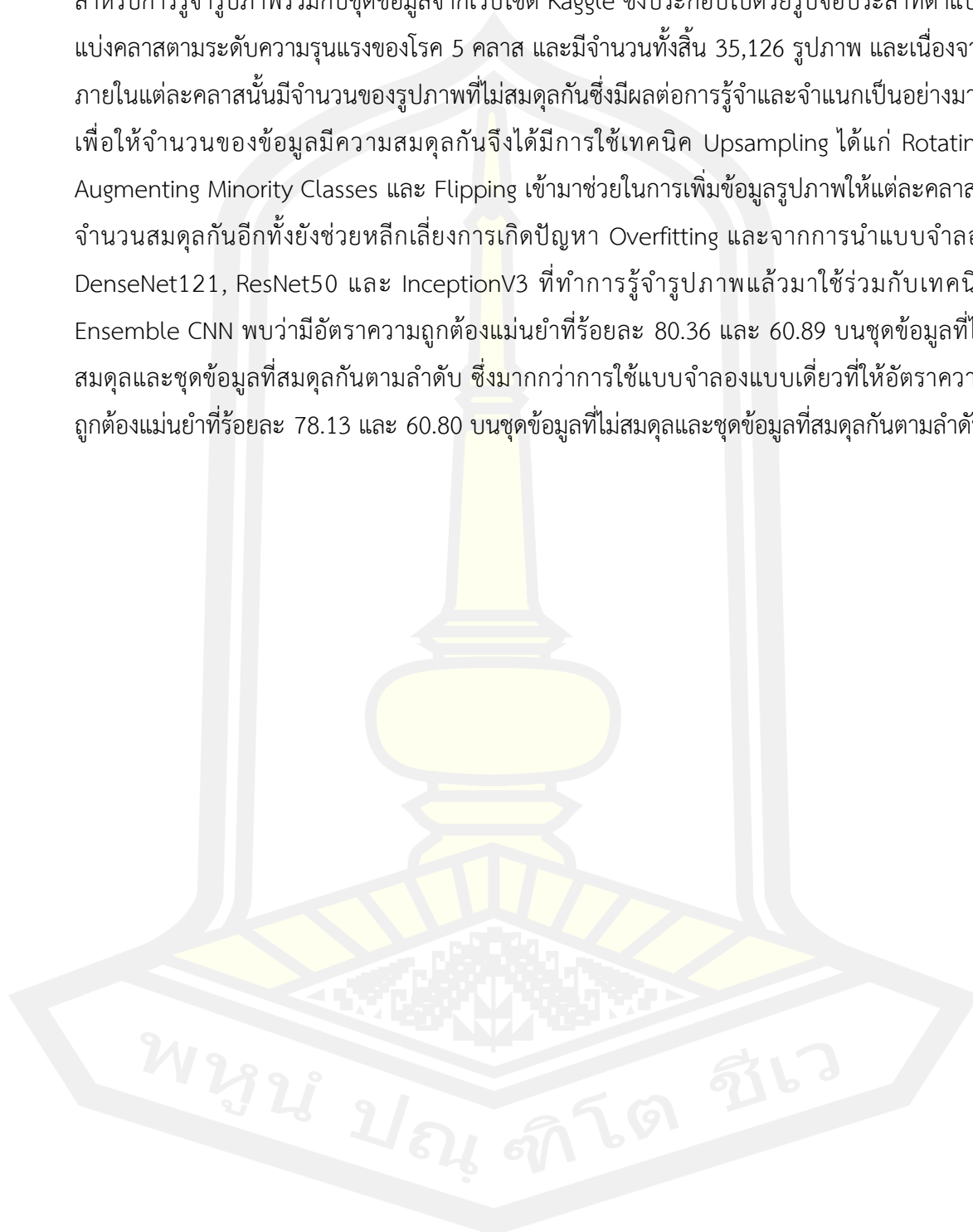
โดยผลการทดลองบนชุดข้อมูล APTOS 2019 ซึ่งประกอบด้วยรูปภาพขอประสาทตาจำนวน 3,662 รูปภาพ พบว่าสถาปัตยกรรม Xception ที่ได้รับการปรับปรุงแล้วให้อัตราความถูกต้องแม่นยำในการจำแนกประเภทรูปภาพขอประสาทตาที่ร้อยละ 83.09 และสถาปัตยกรรม Xception รุ่นดั้งเดิมให้อัตราความถูกต้องที่ร้อยละ 79.95 ซึ่งจะเห็นได้ว่า Xception ที่ได้รับการปรับปรุงแล้วมีอัตราความถูกต้องแม่นยำมากกว่าแบบดั้งเดิมมากถึงร้อยละ 3.50 ซึ่งถือว่ามีนัยสำคัญ

Gangwar และคณะ [42] นำเสนอสถาปัตยกรรม Novel Hybrid Inception-ResNet-v2 สำหรับใช้ในการตรวจจับภาวะเบาหวานขึ้นตาจากรูปภาพขอประสาทตา โดยในสถาปัตยกรรมที่นำเสนอได้มีการปรับแต่งโครงสร้างใหม่ซึ่งสามารถแบ่งออกเป็น 2 ส่วนหลัก และในส่วนแรกเป็นโครงสร้างของ InceptionResNetV2 ส่วนที่สองเป็น Custom Convolutional Block โดยโครงสร้างภายใน Custom Block จะประกอบไปด้วย Convolutional Layer จำนวน 4 ชั้น ซึ่งในแต่ละชั้นจะมีขนาดของเคอร์เนลแตกต่างกัน ในการเชื่อมต่อกันของแต่ละชั้นนั้นจะมีการใช้ Batch Normalization และ Dropout จากนั้นทำการรวม (Concatenated) ผลลัพธ์ที่ได้จากทุกชั้น และทำการแปลงให้เป็นผลลัพธ์แบบ 1 มิติใน Pooling Layer ด้วยเทคนิค Max Pooling จากนั้นนำเข้าไปประมวลผลต่อใน Fully Connected Layer พร้อมทั้งพยากรณ์คำตอบด้วยฟังก์ชัน Softmax และจากผลการทดลองบนชุดข้อมูล Messidor-1 และ APTOS 2019 Blindness Detection พบว่าให้อัตราความถูกต้องแม่นยำมากถึงร้อยละ 72.33 และ 82.18 ตามลำดับ

Mateen และคณะ [43] นำเสนอ CNN เฟรมเวิร์กสำหรับการตรวจหา และจำแนกรูปภาพขอประสาทตาที่มีภาวะเบาหวานขึ้นตา ในการทดลองได้เลือกสถาปัตยกรรม InceptionV3, Resnet50 และ VGG19 มาเพื่อทำกระบวนการ Feature Fusion หรือ Fusion CNN เนื่องจากต้องการรวมคุณสมบัติของแต่ละ CNN เข้าด้วยกันเพื่อเพิ่มประสิทธิภาพในการจำแนกรูปภาพ โดยในขั้นตอนการเตรียมข้อมูลนั้นมีการนำเทคนิค Gaussian mixture เข้ามาร่วมเพื่อใช้ในการเลือกตำแหน่งของขอประสาทตาที่มีภาวะเบาหวานขึ้นตา เช่น เส้นเลือดฝอยในจอประสาทตาเกิดการบวมหรือแตก เป็นต้น ก่อนนำไปทำการสกัดคุณลักษณะเฉพาะของรูปภาพด้วย CNN ทั้ง 3 รูปแบบ จากนั้นนำผลลัพธ์ที่ได้มาจากแต่ละ CNN รวมเข้าด้วยกันใน Fully Connected Layer และจำแนกประเภทรูปภาพขอประสาทตาที่มีภาวะเบาหวานขึ้นตาและรูปภาพขอประสาทตาปกติด้วยฟังก์ชัน Softmax จากผลการทดลองบนชุดข้อมูล DIARETDB1 พบว่าแบบจำลอง CNN ที่ผ่านกระบวนการ Feature Fusion ให้อัตราความถูกต้องแม่นยำที่ร้อยละ 98.91 ซึ่งมากกว่าแบบจำลอง InceptionV3, Resnet50 และ VGG19 ซึ่งเป็นแบบจำลองแบบเดี่ยวที่ให้อัตราความถูกต้องที่ร้อยละ 93.57, 97.90 และ 95.50 ตามลำดับ

Jinfeng และคณะ [44] นำเสนอการใช้สถาปัตยกรรม CNN ร่วมกับเทคนิค Ensemble CNN สำหรับการตรวจจับและจำแนกภาวะความรุนแรงของภาวะเบาหวานขึ้นตาจากรูปภาพขอประสาทตา

ในการทดลองได้เลือกใช้สถาปัตยกรรม DenseNet121, ResNet50 และ InceptionV3 เพื่อใช้สำหรับการรู้จำรูปภาพร่วมกับชุดข้อมูลจากเว็บไซต์ Kaggle ซึ่งประกอบไปด้วยรูปจอประสาทตาแบบแบ่งคลาสตามระดับความรุนแรงของโรค 5 คลาส และมีจำนวนทั้งสิ้น 35,126 รูปภาพ และเนื่องจากภายในแต่ละคลาสนั้นมีจำนวนของรูปภาพที่ไม่สมดุลกันซึ่งมีผลต่อการรู้จำและจำแนกเป็นอย่างมาก เพื่อให้จำนวนของข้อมูลมีความสมดุลกันจึงได้มีการใช้เทคนิค Upsampling ได้แก่ Rotating, Augmenting Minority Classes และ Flipping เข้ามาช่วยในการเพิ่มข้อมูลรูปภาพให้แต่ละคลาสมีจำนวนสมดุลกันอีกทั้งยังช่วยหลีกเลี่ยงการเกิดปัญหา Overfitting และจากการนำแบบจำลอง DenseNet121, ResNet50 และ InceptionV3 ที่ทำการรู้จำรูปภาพแล้วมาใช้ร่วมกับเทคนิค Ensemble CNN พบว่ามีอัตราความถูกต้องแม่นยำที่ร้อยละ 80.36 และ 60.89 บนชุดข้อมูลที่ไม่สมดุลและชุดข้อมูลที่สมดุลกันตามลำดับ ซึ่งมากกว่าการใช้แบบจำลองแบบเดียวที่ให้อัตราความถูกต้องแม่นยำที่ร้อยละ 78.13 และ 60.80 บนชุดข้อมูลที่ไม่สมดุลและชุดข้อมูลที่สมดุลกันตามลำดับ



บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้ได้กล่าวถึงการจำแนกรูปภาพจอประสาทตาปกติและจอประสาทตาที่มีภาวะเบาหวานขึ้นตาโดยใช้แบบจำลอง CNN ที่แตกต่างกัน 5 รูปแบบ ได้แก่ ResNet50, ResNet50V2, Xception, IncetionV3 และ DenseNet121 พร้อมทั้งเพิ่มประสิทธิภาพในการจำแนกของแต่ละแบบจำลอง และทดสอบประสิทธิภาพของแบบจำลองด้วยการหาค่าความถูกต้อง ซึ่งได้กำหนดวิธีดำเนินการวิจัยออกเป็น 3 ส่วน ดังนี้

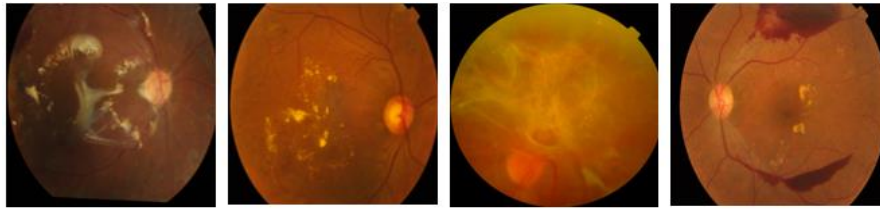
- 3.1 ชุดข้อมูลรูปภาพที่ใช้ในการวิจัย
- 3.2 การสร้างแบบจำลองสำหรับการจำแนกประเภทรูปภาพจอประสาทตา
- 3.3 การประเมินประสิทธิภาพของแบบจำลอง

3.1 ชุดข้อมูลรูปภาพที่ใช้ในการวิจัย

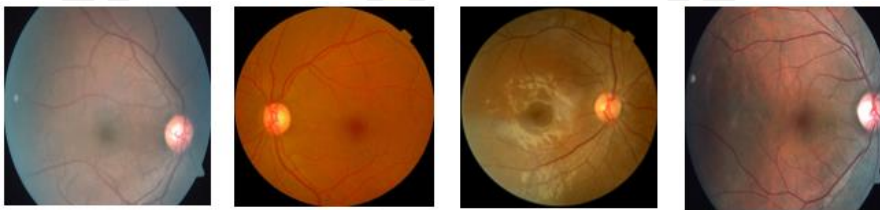
ชุดข้อมูลรูปภาพจอประสาทตาที่ใช้ในการวิจัยรวมทั้งการแบ่งจำนวนรูปภาพสำหรับการเรียนรู้ (Train) และการทดสอบแบบจำลอง (Test) ผู้วิจัยได้ทำการอ้างอิงตามงานวิจัย [30] โดย Hattiya และคณะ ซึ่งสามารถรวบรวมข้อมูลรูปภาพได้จากฐานข้อมูลสาธารณะ Kaggle โดยใช้รูปภาพจำนวน 3 ชุด ได้แก่ APTOS 2019 Blindness Detection [45], Diabetic Retinopathy Detection [46] และ Diabetic Retinopathy Dataset [47] ซึ่งเป็นข้อมูลรูปภาพแบบ RGB ที่แต่ละภาพจะถูกเลือก (Crop) เฉพาะบริเวณของจอประสาทตาเท่านั้น ประกอบไปด้วย 2 คลาส ได้แก่ จอประสาทตาปกติ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตา โดยจะแสดงตัวอย่างของชุดข้อมูลดังกล่าวประกอบที่ 3.1 – 3.6 ตามลำดับ



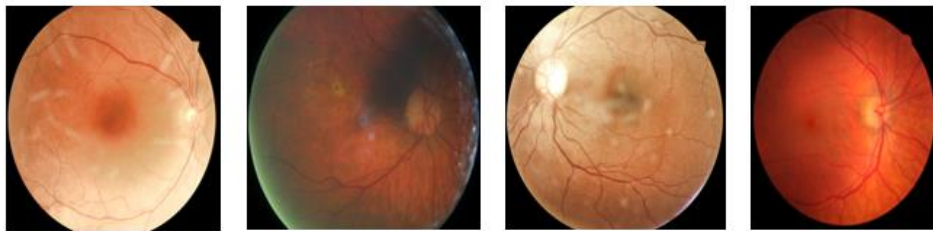
ภาพประกอบที่ 3.1 ตัวอย่างรูปภาพจอประสาทตาปกติจากชุดข้อมูล APTOS 2019 Blindness Detection



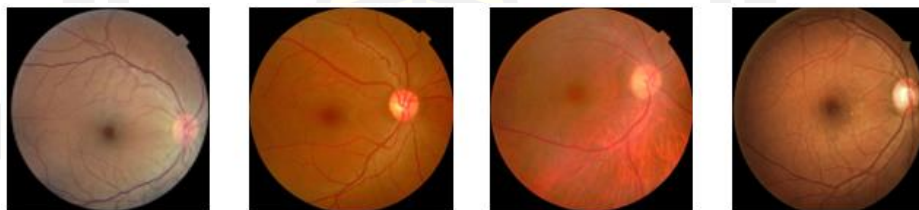
ภาพประกอบที่ 3.2 ตัวอย่างรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจากชุดข้อมูล APTOS 2019 Blindness Detection



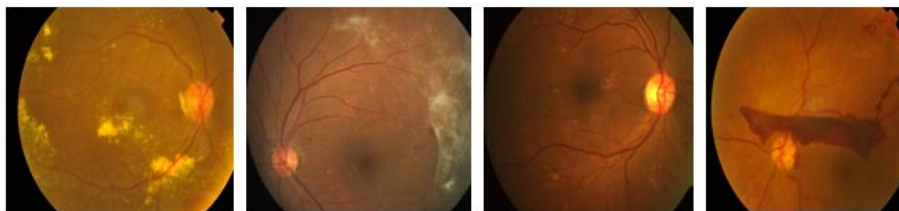
ภาพประกอบที่ 3.3 ตัวอย่างรูปภาพจอประสาทตาปกติจากชุดข้อมูล Diabetic Retinopathy Detection



ภาพประกอบที่ 3.4 ตัวอย่างรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจากชุดข้อมูล Diabetic Retinopathy Detection



ภาพประกอบที่ 3.5 ตัวอย่างรูปภาพจอประสาทตาปกติจากชุดข้อมูล Diabetic Retinopathy Dataset



ภาพประกอบที่ 3.6 ตัวอย่างรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาจากชุดข้อมูล Diabetic Retinopathy Dataset

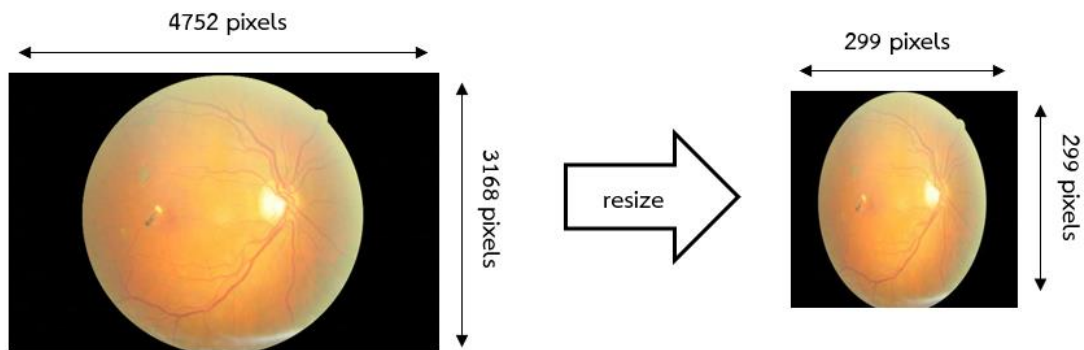
โดยการแบ่งจำนวนรูปภาพสำหรับ Train และ Test แบบจำลองในการวิจัยนี้จะแบ่งออกเป็นจอประสาทตาปกติ 11,447 รูปภาพ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตา 12,063 รูปภาพ รวมทั้งสิ้น 23,510 รูปภาพ โดยแบ่ง Train 18,808 รูปภาพ และ Test 4,702 รูปภาพ ซึ่งมีรายละเอียดดังภาพประกอบที่ 3.7

	Train	Test	รวม
จอประสาทตาปกติ	9,158	2,289	11,447
จอประสาทตาที่มีภาวะเบาหวานขึ้นตา	9,650	2,413	12,063
รวม	18,808	4,702	23,510

ภาพประกอบที่ 3.7 รายละเอียดของชุดข้อมูล

3.1.1 การแปลงขนาดรูปภาพ (Resize Image)

การแปลงขนาดรูปภาพเป็นการย่อหรือขยายรูปภาพตามขนาดที่กำหนด และชุดข้อมูลที่ถูกรวบรวมมาเพื่อทำการวิจัยนี้มีขนาดของรูปภาพที่ใหญ่ และไม่เท่ากัน ทางผู้วิจัยจึงจำเป็นต้องทำการแปลงขนาดของรูปภาพให้ตรงกับความต้องการของแบบจำลอง CNN ที่นำมาทำการทดลอง เพื่อเป็นการลดระยะเวลาในการเรียนรู้ของแบบจำลอง ลดขนาดของการใช้พื้นที่ในการเก็บข้อมูล และทำให้ขนาดของรูปภาพในชุดข้อมูลทั้งหมดอยู่ในขนาดเดียวกัน โดยในการวิจัยนี้ได้ทำการเปลี่ยนแปลงรูปภาพให้อยู่ในขนาด 299 x 299 Pixels ชนิดรูปภาพเป็นแบบ JPEG ซึ่งตัวอย่างรูปภาพจอประสาทตาที่ทำการเปลี่ยนแปลงขนาดจะแสดงดังภาพประกอบที่ 3.8

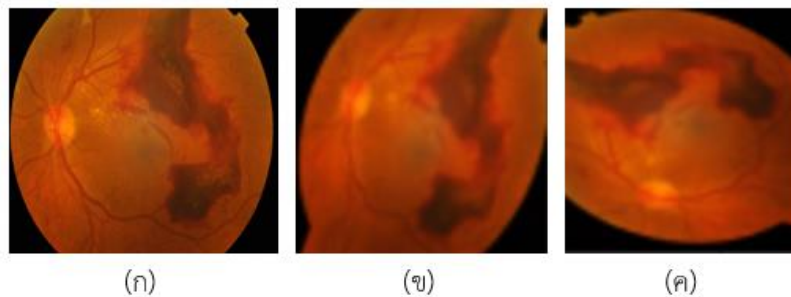


ภาพประกอบที่ 3.8 ตัวอย่างการเปลี่ยนแปลงขนาดรูปภาพ

3.1.2 การเพิ่มจำนวนชุดข้อมูลด้วยเทคนิค Data Augmentation

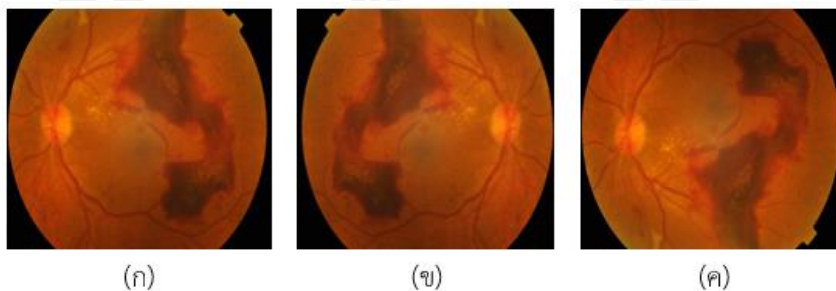
ในการวิจัยนี้ได้เลือกใช้เทคนิค Data Augmentation ในการเพิ่มจำนวนของชุดข้อมูลสำหรับการเรียนรู้ของแบบจำลอง เพื่อให้มีจำนวน มุมมอง และมิติ ของรูปภาพเพิ่มมากขึ้น อีกทั้งการที่มีจำนวนข้อมูลสำหรับการเรียนรู้จำนวนมากขึ้น ส่งผลให้ประสิทธิภาพในการจำแนกประเภทของแบบจำลองมีความถูกต้องมากยิ่งขึ้น พร้อมทั้งสามารถหลีกเลี่ยงการเกิดปัญหา Overfitting ได้อีกด้วย โดยในการวิจัยนี้ได้เลือกใช้เทคนิค Data Augmentation 3 รูปแบบ และเป็นการเรียกใช้จาก ImageDataGenerator ซึ่งเป็นไลบรารีของ Keras โดยสามารถปรับเปลี่ยนรูปภาพได้ตามการตั้งค่าที่ผู้ใช้งานกำหนด และในการวิจัยนี้ผู้วิจัยได้ทำการกำหนดค่าโดยมีรายละเอียดดังต่อไปนี้

1) รูปแบบ Rotation เป็นการหมุนรูปภาพแบบสุ่มไปตามองศาที่กำหนด โดยในการวิจัยนี้มีการกำหนดองศาในการหมุนเป็นแบบสุ่มที่มีค่าตั้งแต่ 0 - 90 องศา เพื่อให้รูปภาพที่ไม่บิดเบือนไปจากต้นฉบับมากเกินไปจนความจำเป็น ตัวอย่างของการทำ Rotation จะแสดงดังภาพประกอบที่ 3.9 โดยที่รูป (ก) เป็นรูปภาพต้นฉบับ (ข) และ (ค) เป็นรูปภาพที่มีการทำ Rotation จากรูปภาพจอประสาทตา



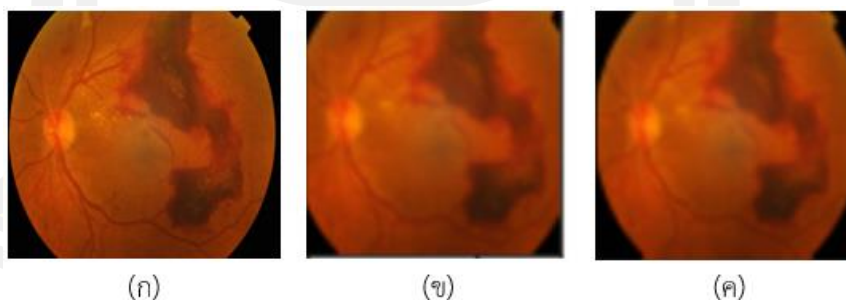
ภาพประกอบที่ 3.9 ตัวอย่างการ Rotation จากรูปภาพจอประสาทตา

2) รูปแบบ Flip เป็นการพลิกหรือบิดเบือนรูปภาพแบบสุ่ม สามารถแบ่งออกเป็น การสุ่มกลับด้านแบบแนวนอน (Horizontal Flip) และ การสุ่มกลับด้านแบบแนวตั้ง (Vertical Flip) โดยตัวอย่างของการทำ Flip จะแสดงดังภาพประกอบที่ 3.10 โดยที่รูป (ก) เป็นรูปภาพต้นฉบับ (ข) เป็นรูปภาพที่มีการทำ Horizontal Flip และ (ค) เป็นรูปภาพที่มีการทำ Vertical Flip จากรูปภาพจอประสาทตา



ภาพประกอบที่ 3.10 ตัวอย่างการ Flip จากรูปภาพจอประสาทตา

3) รูปแบบ Zoom เป็นการขยายบางส่วนในรูปภาพตามขนาดที่กำหนด โดยในการวิจัยนี้ได้ทำการกำหนดขอบเขตในการขยายรูปภาพเป็นแบบสุ่มที่มีค่าตั้งแต่ 0 - 0.2 เพื่อให้รูปภาพไม่บิดเบือนไปจากต้นฉบับมากเกินไป โดยตัวอย่างของการ Zoom รูปภาพจะแสดงดังภาพประกอบที่ 3.11 โดยที่รูป (ก) เป็นรูปภาพต้นฉบับ (ข) และ (ค) เป็นรูปภาพที่มีการ Zoom จากรูปภาพจอประสาทตา



ภาพประกอบที่ 3.11 ตัวอย่างของการ Zoom จากรูปภาพจอประสาทตา

3.2 การสร้างแบบจำลองสำหรับการจำแนกประเภทรูปภาพจอประสาทตา

ความมุ่งหมายของการวิจัยนี้คือการศึกษเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของแบบจำลอง CNN ที่แตกต่างกัน โดยในการวิจัยนี้ได้เลือกใช้สถาปัตยกรรม CNN จำนวน 5 รูปแบบ ได้แก่ ResNet50, ResNet50V2, Xception, IncetionV3 และ DenseNet121 สำหรับการเรียนรู้เพื่อรู้จำรูปภาพจอประสาทตา ซึ่งกระบวนการในการสร้างแบบจำลองจะมีรายละเอียดดังต่อไปนี้

3.2.1 ปรับค่าไฮเปอร์พารามิเตอร์ (Fine-tuning Hyperparameter)

Fine-tuning Hyperparameter สำหรับโครงข่ายประสาทเทียมคอนโวลูชันนั้นจะทำให้เกิดความถูกต้องแม่นยำสูงสุดในการจำแนกประเภท กล่าวคือการทำ Fine-tuning Hyperparameter ให้เหมาะสมสำหรับแต่ละแบบจำลองจะสามารถดึงประสิทธิภาพการทำงานในการจำแนกประเภทของแบบจำลองออกมาได้สูงสุดซึ่งในการวิจัยนี้ได้ทำการปรับ Optimization และ Learning Rate โดยมีรายละเอียดดังต่อไปนี้

1) Fine-tuning Hyperparameter Optimization

ในการวิจัยนี้ได้เลือกใช้อัลกอริทึม Optimization จำนวน 5 รูปแบบ ดังนี้

1.1) Stochastic Gradient Descent (SGD)

SGD เป็นอัลกอริทึมที่ไม่สามารถจัดการหรือจำแนกข้อมูลที่มีขนาดใหญ่และมีความแปรปรวนของข้อมูลสูง (Sparse Data) ดังนั้นอัลกอริทึมนี้จึงเหมาะสำหรับการเรียนรู้ชุดข้อมูลขนาดเล็ก

1.2) Adaptive Moment Estimation (Adam)

Adam เป็นอัลกอริทึมที่รวบรวมความสามารถหรือจุดเด่นของอัลกอริทึมรูปแบบต่าง ๆ เข้าไว้ด้วยกัน พร้อมทั้งมีความเร็วและลดปัญหาการแกว่งของค่าพารามิเตอร์ที่ส่งผลต่อ Loss Function ในการเรียนรู้ของแบบจำลองมากกว่าอัลกอริทึมรูปแบบอื่นอีกด้วย ซึ่งในปัจจุบันเป็นอัลกอริทึมที่ดีที่สุด

1.3) Adagrad

Adagrad เป็นอัลกอริทึมที่เหมาะสมสำหรับข้อมูลขนาดใหญ่ และข้อมูลที่มีความแตกต่างหรือแปรปรวนสูง เนื่องจาก Adagrad สามารถลดความเร็วพร้อมทั้งขั้นตอนในการเรียนรู้ที่ไม่จำเป็น และเพิ่มความเร็วในการเรียนรู้กับขั้นตอนที่จำเป็นหรือถูกเรียกใช้บ่อยครั้ง ส่งผลให้การเรียนรู้ของแบบจำลองมีความรวดเร็วมากยิ่งขึ้น

1.4) Adadelta

Adadelta เป็นอัลกอริทึมที่ได้พัฒนามาจาก Adagrad ซึ่งมีจุดเด่นที่การใช้งานง่ายเนื่องจากเป็นอัลกอริทึมที่ไม่ต้องมีการ Fine-tuning Hyperparameter เนื่องจากมีการคำนวณเพื่อกำหนดค่าที่เหมาะสมให้กับแบบจำลองแล้ว

1.5) Adamax

Adamax เป็นอัลกอริทึมที่ได้พัฒนามาจาก Adam โดยมีการมุ่งเน้นเพื่อแก้ปัญหาการเรียนรู้กับชุดข้อมูลขนาดใหญ่หรือชุดข้อมูลที่มีความแปรปรวนของข้อมูลสูง ซึ่งเป็นการ

ผสมความสามารถในการจัดการข้อมูลที่มีความแปรปรวนของ RMSProp เข้ากับความสามารถในการ Sparse Gradients ของ Adagrad ซึ่งสามารถเรียนรู้และจัดการชุดข้อมูลดังกล่าวได้ดีมากยิ่งขึ้น

2) Fine-tuning Hyperparameter Learning Rate

ค่า Learning Rate เป็นตัวแปรสำคัญเกี่ยวกับระยะเวลาในการเรียนรู้ของแบบจำลอง โดยหากมีการปรับค่า Learning Rate ที่ต่ำจะส่งผลให้ให้แบบจำลองสามารถเรียนรู้ชุดข้อมูลได้อย่างมีประสิทธิภาพ เนื่องจากจะเป็นการเรียนรู้ไปอย่างช้า ๆ และหากทำการปรับค่า Learning Rate สูง จะทำให้แบบจำลองเรียนรู้ได้เร็วกว่าที่มีจุดบกพร่องด้านความละเอียด ดังนั้น จึงจำเป็นต้องมีการทดลองเพื่อหาค่าที่เหมาะสมสำหรับแบบจำลองทุกรูปแบบ โดยค่าของ Learning Rate ที่ใช้ในการทดลองจะอยู่ระหว่าง 0.1 – 0.0001

3.2.2 ทดลองเพื่อค้นหารูปแบบ Data Augmentation ที่เหมาะสม

เมื่อทำการ Fine-tuning Hyperparameter และได้ค่าพารามิเตอร์ที่เหมาะสมสำหรับแต่ละแบบจำลองแล้ว ผู้วิจัยได้ทำการทดลองเพื่อค้นหารูปแบบของ Data Augmentation ที่เหมาะสมสำหรับแต่ละแบบจำลองเช่นเดียวกับขั้นตอนการปรับค่าพารามิเตอร์ เนื่องจากการกำหนดให้แบบจำลองเรียนรู้ชุดข้อมูลที่มีรูปภาพในมิติหรือมุมมองเดียวอาจทำให้ผลของการจำแนกประเภทเกิดการคลาดเคลื่อน จึงจำเป็นต้องมีการค้นหาว่า Rotation, Flip และ Zoom ซึ่งเป็นรูปแบบในเทคนิค Data Augmentation จะสามารถเพิ่มประสิทธิภาพในการจำแนกประเภทของแบบจำลองบนชุดข้อมูลรูปภาพจอสภาพตาได้มากที่สุด

3.2.3 ทดลองร่วมกับเทคนิค Ensemble CNN และ เทคนิค Fusion CNN

งานวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษาและพัฒนาแบบจำลอง CNN เพื่อให้ได้แบบจำลองที่มีประสิทธิภาพในการจำแนกประเภทจอสภาพตาที่มีภาวะเบาหวานขึ้นตา และผู้วิจัยได้ทำการทดลองโดยนำเอาคุณสมบัติการจำแนกประเภทของ CNN หลายรูปแบบรวมเข้าด้วยกัน เพื่อให้ได้แบบจำลองที่มีประสิทธิภาพเพิ่มมากขึ้นเนื่องจากเป็นแบบจำลองที่มีคุณสมบัติการจำแนกของทุก CNN ซึ่งจะสามารถพยากรณ์ได้ถูกต้องแม่นยำมากกว่า Single CNN และในการวิจัยนี้ได้มีการทดลองสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble และ Fusion โดยมีรายละเอียดดังนี้

1) เทคนิค Ensemble CNN

การทดลองสร้างแบบจำลองร่วมกับเทคนิค Ensemble CNN หรือ Ensemble Learning เป็นการนำแบบจำลองที่แตกต่างกัน 2 รูปแบบ ที่ได้รับการ Train ร่วมกับเทคนิค Data Augmentation มาทำการพยากรณ์ (Prediction) กับชุดข้อมูลสำหรับ Test เพื่อให้ได้ค่าความน่าจะเป็น (Probability) และนำไป Vote เพื่อหาค่าเฉลี่ยความน่าจะเป็นของคำตอบ ซึ่งเรียกกระบวนการ

นี้ว่า Average Voting โดยจะแบ่งการคำนวณออกเป็น Unweighted Average และ Weighted Average เพื่อให้ได้ผลลัพธ์ที่ถูกต้องแม่นยำมากที่สุด

2) เทคนิค Fusion CNN

การทดลองสร้างแบบจำลองร่วมกับเทคนิค Fusion CNN จะนำสถาปัตยกรรม CNN ที่แตกต่างกัน 2 รูปแบบ มารวมส่วนสำหรับการจำแนกประเภทเข้าด้วยกันและทำการสร้างเป็นแบบจำลองรูปแบบใหม่ซึ่งจะมีคุณสมบัติในการจำแนกคุณลักษณะเฉพาะหรือการจำแนกประเภทของทั้ง CNN ทั้ง 2 รูปแบบ จากนั้นจะนำไป Train ร่วมกับเทคนิค Data Augmentation จำนวน 100 Epoch เพื่อให้ได้แบบจำลองที่มีประสิทธิภาพและความถูกต้องแม่นยำในการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา

ทั้งนี้ จากงานวิจัยเกี่ยวข้องกับการใช้แบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN และเทคนิค Ensemble CNN พบว่าสามารถเพิ่มประสิทธิภาพในการจำแนกประเภทโดยสามารถให้อัตราความถูกต้องในการจำแนกเพิ่มมากขึ้น จึงเป็นเหตุให้ผู้วิจัยได้นำ 2 เทคนิคดังกล่าวมาทำการทดลองเพื่อเพิ่มประสิทธิภาพในการรู้จำและจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา

3.3 การประเมินประสิทธิภาพของแบบจำลอง

ในการวิจัยนี้มีความมุ่งหมายคือการเปรียบเทียบประสิทธิภาพการทำงานของแบบจำลอง CNN สำหรับการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา ซึ่งในการทดลองจะเลือกวัดประสิทธิภาพโดยการใช้ค่าความถูกต้อง (Accuracy) ที่ได้จากการทดสอบการจำแนกประเภทของแต่ละแบบจำลอง บนชุดข้อมูลที่เตรียมไว้เพื่อทำการทดสอบ (Test) และนำผลลัพธ์หรือค่า Accuracy มาทำการเปรียบเทียบกันเพื่อหาว่าแบบจำลองใดให้ค่า Accuracy ในการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาส่งสุด

บทที่ 4

ผลการศึกษา

ผลการศึกษาของการวิจัย เรื่อง การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา ด้วยการเรียนรู้เชิงลึก ผู้วิจัยได้ใช้สถาปัตยกรรม CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter, Data Augmentation, Ensemble CNN และ Fusion CNN เพื่อสร้างแบบจำลองสำหรับการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา พร้อมกับวัดประสิทธิภาพการทำงานของแบบจำลองโดยใช้ค่าความถูกต้อง (Accuracy) โดยรายละเอียดของแต่ละส่วนในงานวิจัยจะแสดงดังนี้

4.1 เครื่องมือและข้อมูลที่ใช้ในการทดลอง

4.2 วิธีการทดลอง

4.3 ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึกด้วยแบบจำลอง CNN

4.4 ประเมินและเปรียบเทียบประสิทธิภาพการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาของแบบจำลอง CNN

4.1 เครื่องมือและข้อมูลที่ใช้ในการทดลอง

4.1.1 คอมพิวเตอร์ที่ใช้ในการทดลอง

การทดลองการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึกสามารถแบ่งเครื่องมือในการศึกษาและพัฒนาแบบจำลอง CNN ออกเป็นส่วนของการรวบรวมพร้อมกับแบ่งชุดข้อมูล และส่วนของการสร้างพร้อมกับประเมินประสิทธิภาพแบบจำลอง

ในส่วนของการรวบรวมและแบ่งชุดข้อมูลสำหรับการทดลอง ผู้วิจัยทำการรวบรวมข้อมูลรูปภาพจอบประสาทตาพร้อมกับเขียนโปรแกรมเพื่อแบ่งชุดข้อมูลสำหรับฝึกฝน (Training Set) และชุดข้อมูลสำหรับการทดสอบ (Test Set) บนเครื่องคอมพิวเตอร์ที่มีหน่วยประมวลผลกลาง Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz มีหน่วยความจำขนาด 12 GB บนระบบปฏิบัติการ Windows 10 แบบ 64-bit

ส่วนของการสร้างและประเมินประสิทธิภาพจะถูกทำการทดลองบน Google Colaboratory (Google Colab) ซึ่งเป็นบริการ Software as a Service ที่ทำการติดตั้งพร้อมทั้งกำหนดค่าในระบบคลาวด์ และมีหน่วยประมวลผลด้านกราฟฟิก (Graphics Processing Unit: GPU) ที่สามารถช่วยลดระยะเวลาที่ใช้ในการสร้างและทดสอบแบบจำลอง

4.1.2 การเขียนโปรแกรมที่ใช้ในการทดลอง

การเขียนโปรแกรมที่ใช้ในการทดลองการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวาน ขึ้นตาด้วยการเรียนรู้เชิงลึก ผู้วิจัยได้เลือกใช้ภาษา Python ในการศึกษาและพัฒนาแบบจำลอง CNN โดยในส่วนแรกผู้วิจัยทำการเปลี่ยนแปลงขนาดของรูปภาพที่ใช้ในการทดลองให้อยู่ในขนาดที่กำหนด โดยใช้ไลบรารี Pillow (Python Imaging Library) ซึ่งในการวิจัยนี้ได้กำหนดขนาดของรูปภาพเป็น 299x299 Pixel และผู้วิจัยได้เลือกใช้ Keras ซึ่งเป็นไลบรารีของ Tensorflow ในการเตรียมชุดข้อมูล การกำหนดค่าพารามิเตอร์เพื่อสร้างแบบจำลอง ปรับปรุงโครงสร้างบางส่วน of แบบจำลอง รวมไปถึง การทดสอบประสิทธิภาพของแบบจำลอง

4.1.3 ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองเป็นข้อมูลรูปภาพจอประสาทตาจากฐานข้อมูลสาธารณะที่เผยแพร่บนเว็บไซต์ Kaggle จำนวน 3 ชุด ได้แก่ APTOS 2019 Blindness Detection, Diabetic Retinopathy Detection และ Diabetic Retinopathy Dataset ซึ่งเป็นรูปภาพแบบ RGB ชนิดรูปภาพเป็นแบบ JPEG ที่ประกอบไปด้วย 2 ประเภท ได้แก่ จอประสาทตาปกติ (No DR) 11,447 รูปภาพ โดยแบ่งเป็นชุดข้อมูลสำหรับ Train จำนวน 9,185 รูปภาพ และชุดข้อมูลสำหรับ Test จำนวน 2,289 รูปภาพ และจอประสาทตาที่มีภาวะเบาหวานขึ้นตา (DR) 12,063 รูปภาพ โดยแบ่งเป็นชุดข้อมูลสำหรับ Training จำนวน 9,650 รูปภาพ และชุดข้อมูลสำหรับ Test จำนวน 2,413 รูปภาพ รวมทั้งสิ้น 23,510 รูปภาพ ซึ่งการแบ่ง Training Set และ Test Set คิดเป็นร้อยละ 80 และ 20 ของจำนวนของข้อมูลทั้งหมดตามลำดับ โดยตารางสรุปสถิติการแบ่งข้อมูลสำหรับการทดลองจะแสดงดังตารางที่ 4.1

ตารางที่ 4.1 สถิติการแบ่งข้อมูลสำหรับการทดลอง

ชุดข้อมูล	ประเภทข้อมูล	จำนวนรูปภาพ	สัดส่วนการแบ่งข้อมูล (%)
Training Set	No DR	9,185	-
	DR	9,650	-
รวม		18,808	80
Test Set	No DR	2,289	-
	DR	2,413	-
รวม		4,702	20

4.2 วิธีการทดลอง

การทดลองการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาดำด้วยการเรียนรู้เชิงลึกได้มีการ Training แบบบ จำลอง ร่วมกับ เทคนิค Fine-tuning Hyperparameter, Data Augmentation, Ensemble CNN และ Fusion CNN เพื่อสร้างแบบจำลองในการจำแนกประเภทรูปภาพจอบประสาทตาซึ่งถูกประมวลผลบน GPU โดยมีรายละเอียดดังต่อไปนี้

4.2.1 สร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter

การทำ Fine-tuning Hyperparameter เป็นทดลองเพื่อค้นหาค่าพารามิเตอร์ที่เหมาะสมสำหรับแต่ละแบบจำลองซึ่งเป็นกระบวนการเตรียมตัวก่อนก่อนนำไปทำการเรียนรู้ (Train) เพื่อรู้จำข้อมูลชุด Training Set โดยใช้จำนวนรอบการเรียนรู้ (Epoch) จำนวน 100 รอบ และทำการทดสอบประสิทธิภาพของแบบจำลองด้วยการวัดค่าความถูกต้อง (Accuracy) ในการจำแนกรูปภาพจอบประสาทตาบนชุดข้อมูลสำหรับ Test ซึ่งในการทดลองนี้ผู้วิจัยได้กำหนดให้ CNN ทุกรูปแบบมีค่าพารามิเตอร์ Optimization และ Learning Rate ดังตารางที่ 4.2

ตารางที่ 4.2 ค่าพารามิเตอร์สำหรับการทดลอง Fine-tuning Hyper-Parameter

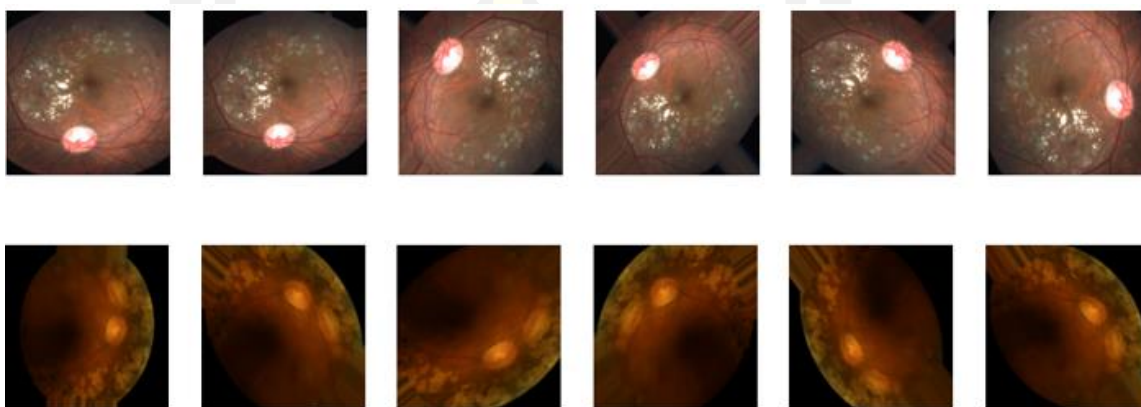
CNN	Optimization	Learning Rate			
	SGD	0.1	0.01	0.001	0.0001
Adam					
Adadelta					
Adagrad					
Adamax					

4.2.2 สร้างแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation

การสร้างแบบจำลองร่วมกับเทคนิค Data Augmentation เป็นการนำชุดข้อมูล Training Set เข้าไปทำการตัดแปลงหรือบิดเบือนรูปภาพด้วยเทคนิค Data Augmentation ที่มีรูปแบบ และการกำหนดค่าดังตารางที่ 4.3 เพื่อให้ได้ Training Set ที่มีมุมมองหลากหลายมากยิ่งขึ้น จากนั้นจะนำไปเข้ากระบวนการเรียนรู้ของแบบจำลอง ซึ่งเลือกใช้ค่าพารามิเตอร์ที่เหมาะสมจากการทดลองที่ 4.2.1 โดยใช้จำนวนรอบการเรียนรู้ (Epoch) จำนวน 100 รอบ เพื่อสร้าง และทำการทดสอบประสิทธิภาพของแบบจำลองด้วยการวัดค่าความถูกต้อง (Accuracy) ในการจำแนกรูปภาพจอบประสาทตาบนชุดข้อมูลสำหรับ Test โดยตัวอย่างของรูปภาพจอบประสาทตาที่มีการทำ Data Augmentation จะแสดงดังภาพประกอบที่ 4.1

ตารางที่ 4.3 รูปแบบ Data Augmentation ที่ใช้ในการทดลอง

	Augmentation	Range
CNN	Horizontal Flip	-
	Vertical Flip	-
	Horizontal + Vertical	-
	Rotation	90
	Zoom	0.2
	Rotation + Zoom	90, 0.2



ภาพประกอบที่ 4.1 การทำ Data Augmentation จาก Training Set

4.2.3 สร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN

การสร้างแบบจำลองร่วมกับเทคนิค Ensemble เป็นการนำแบบจำลองที่แตกต่างกัน 5 รูปแบบ โดยมีค่า Accuracy สูงสุดจากการทดลองสร้างแบบจำลองด้วยเทคนิค Data Augmentation มาทำการ Prediction กับชุดข้อมูลสำหรับ Test โดยผู้วิจัยจะนำแบบจำลอง 2 รูปแบบ ไปทำการ Prediction เพื่อให้ได้ค่า Probability ของแต่ละแบบจำลองเพื่อนำไปคำนวณหาค่าเฉลี่ยความน่าจะเป็นเพื่อให้ได้คำตอบที่ถูกต้องที่สุดซึ่งเรียกกระบวนการนี้ว่า Average Voting โดยจะแบ่งการคำนวณออกเป็น Unweighted Average และ Weighted Average ทั้งนี้ การคำนวณแบบ Weighted Average สามารถกำหนดช่วงของค่าน้ำหนักของแต่ละแบบจำลองได้ตั้งแต่ 0 – 1 แต่เมื่อรวมค่าน้ำหนักของทั้ง 2 แบบจำลองแล้วต้องมีค่าไม่เกิน 1 ดังนั้น ในการทดลองนี้จึงได้มีการกำหนดค่าน้ำหนักให้กับแบบจำลองโดยจะแสดงรายละเอียดดังตารางที่ 4.4

ตารางที่ 4.4 คำนวณน้ำหนักสำหรับการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN

CNN	Weighted Parameter	CNN	Weighted Parameter
CNN – 1	0.1	CNN – 2	0.9
	0.2		0.8
	0.3		0.7
	0.4		0.6

4.2.4 สร้างแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN

การสร้างแบบจำลองร่วมกับเทคนิค Fusion CNN เป็นการนำเอาสถาปัตยกรรม CNN ที่มีอัตราความถูกต้องสูงสุดในการทดลองสร้างแบบจำลองร่วมกับเทคนิค Data Augmentation มาทำการปรับปรุงเพิ่มเติม โดยผู้วิจัยเพิ่ม Dense Layer และ Dropout Layer เพื่อเพิ่มประสิทธิภาพของแบบจำลองและยังช่วยลดการเกิดปัญหา Overfitting ได้อีกด้วย โดยใช้จำนวนรอบการเรียนรู้ (Epoch) จำนวน 100 รอบ เพื่อสร้าง และทำการทดสอบประสิทธิภาพของแบบจำลองด้วยการวัดค่าความถูกต้อง (Accuracy) ในการจำแนกรูปภาพจอบประสาทตาบนชุดข้อมูลสำหรับ Test โดยรายละเอียดการตั้งค่า Dense Layer และ Dropout Layer จะแสดงดังตารางที่ 4.5

ตารางที่ 4.5 ค่าพารามิเตอร์สำหรับการทดลอง Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter				
1	1024	-	0.1	0.2	0.3	0.4
	2048					
	4096					
2	1024	-	0.1	0.2	0.3	0.4
	2048					
	4096					

4.3 ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยการเรียนรู้เชิงลึกด้วยแบบจำลอง CNN

เมื่อทำการสร้างแบบจำลอง CNN สำหรับการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา ร่วมกับเทคนิค Fine-tuning Hyperparameter, Data Augmentation, Ensemble CNN และ Fusion CNN ด้วยข้อมูลชุด Training Set และทดสอบประสิทธิภาพการทำงานของแบบจำลองกับชุดข้อมูล Test Set โดยผลลัพธ์ของการทดลองมีรายละเอียดดังต่อไปนี้

4.3.1 ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter

จากการทดลองสร้างแบบจำลอง ResNet50, ResNet50V2, InceptionV3, DenseNet121 และ Xception ร่วมกับเทคนิค Fine-tuning Hyperparameter เพื่อหาค่าพารามิเตอร์ที่เหมาะสมสำหรับแต่ละแบบจำลองโดยมีการประเมินประสิทธิภาพด้วยการวัดค่า Accuracy โดยจะแสดงผลการทดลองดังตารางที่ 4.6 – 4.10

ตารางที่ 4.6 ผลลัพธ์การสร้างแบบจำลอง ResNet50 ร่วมกับเทคนิค Fine-tuning Hyperparameter

CNN	Optimization	Learning Rate	Time (Min:s)	Test	
				Accuracy (%)	Loss
ResNet50	SGD	0.1	39:22	80.05	0.42
		0.01	38:50	80.22	0.44
		0.001	39:20	78.48	0.45
		0.0001	39:46	71.65	0.51
	Adam	0.1	40:43	62.76	0.63
		0.01	40:15	60.06	0.64
		0.001	38:44	79.77	0.43
		0.0001	38:57	80.73	0.41
	Adadelta	0.1	38:20	71.05	0.57
		0.01	38:06	72.50	0.56
		0.001	38:21	80.03	0.44
		0.0001	38:22	52.51	0.69

	Adagrad	0.1	38:42	68.42	0.57
		0.01	38:50	70.88	0.55
		0.001	38:25	80.56	0.42
		0.0001	38:20	73.84	0.47
	Adamax	0.1	37:51	52.51	0.69
		0.01	37:50	79.77	0.43
		0.001	50:37	80.63	0.4
		0.0001	37:32	64.48	0.65

ตารางที่ 4.7 ผลลัพธ์การสร้างแบบจำลอง ResNet50V2 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter

CNN	Optimization	Learning Rate	Time (Min:s)	Test	
				Accuracy (%)	Loss
ResNet50V2	SGD	0.1	50:08	81.56	0.4
		0.01	39:07	79.01	0.42
		0.001	44:30	79.73	0.46
		0.0001	51:16	78.37	0.45
	Adam	0.1	39:47	51.30	0.69
		0.01	40:11	51.32	0.69
		0.001	39:35	75.44	0.47
		0.0001	38:37	78.48	0.44
	Adadelta	0.1	38:17	78.52	0.46
		0.01	40:08	76.18	0.48
		0.001	38:12	79.84	0.43
		0.0001	39:44	72.56	0.56
	Adagrad	0.1	36:59	78.90	0.44
		0.01	36:15	80.63	0.41

		0.001	36:41	80.14	0.41
		0.0001	37:37	76.22	0.5
	Adamax	0.1	38:07	51.32	0.69
		0.01	38:05	64.29	0.6
		0.001	37:25	82.35	0.38
		0.0001	38:14	78.05	0.43

ตารางที่ 4.8 ผลลัพธ์การสร้างแบบจำลอง InceptionV3 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter

CNN	Optimization	Learning Rate	Time (Min:s)	Test	
				Accuracy (%)	Loss
InceptionV3	SGD	0.1	53:21	82.16	0.39
		0.01	38:35	77.56	0.44
		0.001	44:12	78.77	0.45
		0.0001	40:40	77.71	0.5
	Adam	0.1	42:01	51.32	0.69
		0.01	38:07	50.81	0.69
		0.001	38:08	83.20	0.38
		0.0001	37:44	82.03	0.39
	Adadelta	0.1	37:53	81.56	0.42
		0.01	37:50	81.86	0.41
		0.001	45:08	80.31	0.45
		0.0001	48:32	75.65	0.52
	Adagrad	0.1	43:11	82.03	0.39
		0.01	42:43	81.56	0.39
		0.001	42:43	80.63	0.41
		0.0001	43:20	79.16	0.46

	Adamax	0.1	33:05	48.98	1.45
		0.01	33:19	57.04	0.69
		0.001	32:48	81.88	0.39
		0.0001	32:42	79.77	0.56

ตารางที่ 4.9 ผลลัพธ์การสร้างแบบจำลอง DenseNet121 ร่วมกับเทคนิค Fine-tuning Hyper-Parameter

CNN	Optimization	Learning Rate	Time (Min:s)	Test	
				Accuracy (%)	Loss
DenseNet121	SGD	0.1	33:54	63.82%	0.63
		0.01	32:49	74.95%	0.47
		0.001	32:46	81.26%	0.42
		0.0001	32:47	80.50%	0.42
	Adam	0.1	36:17	57.42%	0.66
		0.01	36:32	69.08%	0.57
		0.001	35:46	80.80%	0.42
		0.0001	36:12	76.82%	0.46
	Adadelta	0.1	33:05	82.56%	0.39
		0.01	33:01	75.44%	0.51
		0.001	34:33	81.65%	0.4
		0.0001	32:38	49.62%	0.77
	Adagrad	0.1	28:21	61.78%	0.61
		0.01	35:46	79.48%	0.45
		0.001	35:41	81.79%	0.39
		0.0001	36:32	81.77%	0.39
	Adamax	0.1	34:57	64.44%	0.63
		0.01	33:28	77.52%	0.46

		0.001	33:12	82.05%	0.44
		0.0001	33:25	80.35%	0.42

ตารางที่ 4.10 ผลลัพธ์การสร้างแบบจำลอง Xception ร่วมกับเทคนิค Fine-tuning Hyper-Parameter

CNN	Optimization	Learning Rate	Time (Min:s)	Test	
				Accuracy (%)	Loss
Xception	SGD	0.1	53:37	82.62	0.38
		0.01	56:35	82.84	0.38
		0.001	56:38	82.26	0.4
		0.0001	57:55	76.07	0.57
	Adam	0.1	56:48	48.68	0.69
		0.01	55:08	80.94	0.42
		0.001	55:01	83.86	0.38
		0.0001	56:56	79.67	0.41
	Adadelta	0.1	54:52	81.52	0.43
		0.01	55:08	82.52	0.38
		0.001	56:43	80.28	0.48
		0.0001	58:55	67.63	0.63
	Adagrad	0.1	56:36	80.82	0.4
		0.01	56:48	78.09	0.43
		0.001	56:44	82.09	0.4
		0.0001	57:32	79.03	0.49
	Adamax	0.1	58:04	56.21	0.68
		0.01	57:49	79.14	0.43
		0.001	57:59	84.05	0.38
		0.0001	50:12	80.09	0.46

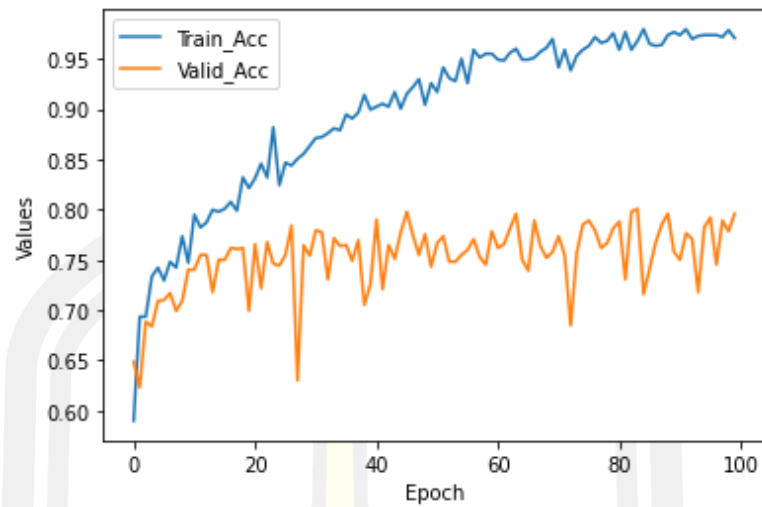
จากผลลัพธ์การทดลองการสร้างแบบจำลองเพื่อจำแนกประเภทรูปภาพจอสภาพตา ด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter บนชุดข้อมูล Test Set พบว่าแบบจำลอง Xception, InceptionV3, DenseNet121, ResNet50V2 และ ResNet50 เมื่อถูกกำหนดค่าพารามิเตอร์ให้เหมาะสมจะสามารถให้อัตราความถูกต้องสูงสุดที่ร้อยละ 84.07, 83.30, 82.77, 82.20 และ 80.73 ตามลำดับ โดยรายละเอียดของค่าที่เหมาะสมสำหรับแต่ละแบบจำลอง จะแสดงดังตารางที่ 4.11

ตารางที่ 4.11 สรุปผลลัพธ์การทดลองการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter

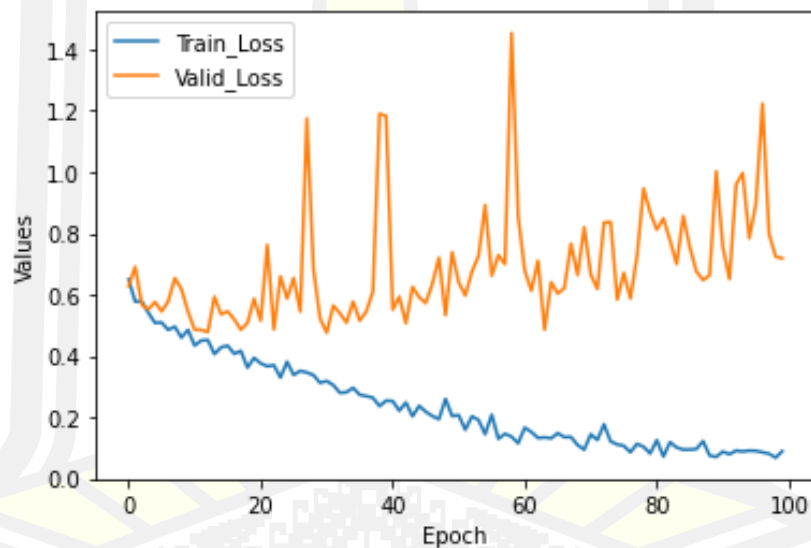
CNN	Optimization	Learning Rate	Time (Min: s)	Test	
				Accuracy (%)	Loss
Xception	Adamax	0.001	57:59	84.07	0.38
InceptionV3	Adam	0.001	38:08	83.30	0.38
DenseNet121	Adadelata	0.1	33:05	82.77	0.40
ResNet50V2	Adamax	0.001	37:25	82.20	0.38
ResNet50	Adam	0.0001	38:57	80.73	0.41

จากตารางสรุปผลลัพธ์การทดลองการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter พบว่าแบบจำลอง Xception สามารถจำแนกประเภทรูปภาพจอสภาพตาได้ดีที่สุด โดยในการ Train แบบจำลองจำนวน 100 รอบ สามารถให้อัตรา Accuracy ของการ Train สูงสุดที่ร้อยละ 97.95 และ Loss ต่ำสุดที่ 0.07

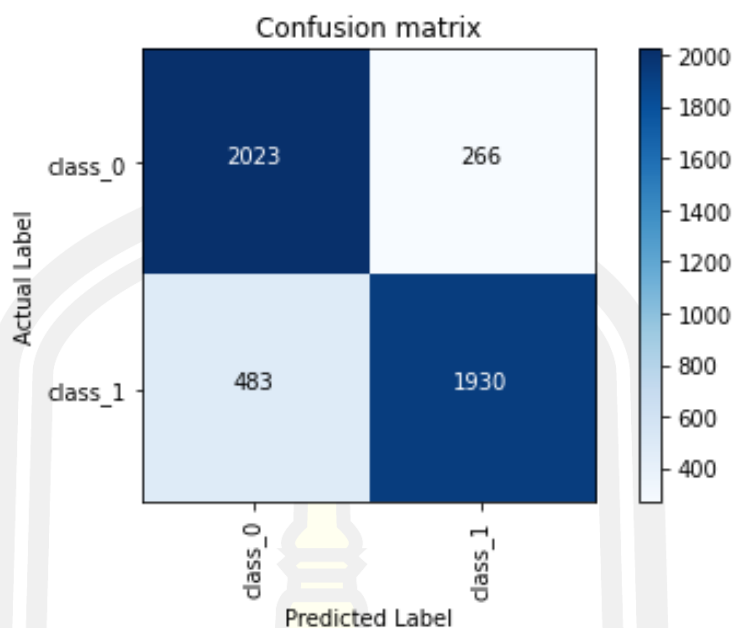
ซึ่งรายละเอียดอัตรา Accuracy ในการ Train แบบจำลอง Xception จะแสดงดังภาพประกอบที่ 4.2 และ 4.3 โดยแบ่งออกเป็นอัตรา Accuracy ในการเรียนรู้ และอัตรา Loss ในการเรียนรู้ ตามลำดับ



ภาพประกอบที่ 4.2 อัตรา Accuracy ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Fine-tuning Hyperparameter



ภาพประกอบที่ 4.3 อัตรา Loss ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Fine-tuning Hyperparameter



ภาพประกอบที่ 4.4 Confusion Matrix ของแบบจำลอง Xception ที่สร้างร่วมกับเทคนิค Fine-tuning Hyperparameter

จากภาพประกอบที่ 4.4 จะแสดงการวัดประสิทธิภาพของแบบจำลอง Xception โดยการใช้ Confusion Matrix ซึ่งเป็นการแสดงรายละเอียดการจำแนกประเภทรูปภาพระหว่างจอบประสาทตาประสาทตาปกติ (Class_0) และจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา (Class_1) บนชุดข้อมูล Test Set โดยอัตรา Accuracy ของการจำแนกประเภทรูปภาพจะคิดเป็นร้อยละ 84.07

4.3.2 ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation

จากการทดลองสร้างแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation ที่ใช้จำนวนรอบการเรียนรู้ (Epoch) 100 รอบ เพื่อหารูปแบบที่สามารถเพิ่มประสิทธิภาพการทำงานของแบบจำลองได้สูงสุด โดยมีรูปแบบดังตารางที่ 4.3 และการกำหนด Hyperparameter Optimization และ Learning Rate ตามตารางที่ 4.11 ซึ่งเป็นผลลัพธ์จากการทดลองสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter เพื่อให้ได้ค่า Hyperparameter ที่เหมาะสมที่สุด โดยผลลัพธ์และรายละเอียดของการทดลองการจำแนกรูปภาพจอบประสาทตาด้วยแบบจำลอง ResNet50, ResNet50V2, InceptionV3, DenseNet121 และ Xception ร่วมกับเทคนิค Data Augmentation จะแสดงดังตารางที่ 4.12 – 4.16 ตามลำดับ

ตารางที่ 4.12 ผลลัพธ์การสร้างแบบจำลอง ResNet50 ร่วมกับเทคนิค Data Augmentation

CNN	Augmentation	Time (Min:s)	Test	
			Accuracy (%)	Loss
ResNet50	Horizontal Flip	39:50	81.67%	0.48
	Vertical Flip	39:59	83.79%	0.46
	Rotation	32:58	82.86%	0.4
	Zoom	35:53	81.43%	0.52
	Horizontal + Vertical	24:41	83.45%	0.4
	Rotation + Zoom	26:24	83.20%	0.42

ตารางที่ 4.13 ผลลัพธ์การสร้างแบบจำลอง ResNet50V2 ร่วมกับเทคนิค Data Augmentation

CNN	Augmentation	Time (Min:s)	Test	
			Accuracy (%)	Loss
ResNet50 V2	Horizontal Flip	10:44	81.62	0.46
	Vertical Flip	10:20	81.45	0.43
	Rotation	25:26	82.18	0.39
	Zoom	25:47	80.94	0.43
	Horizontal + Vertical	21:40	83.28	0.39
	Rotation + Zoom	26:39	81.28	0.43

ตารางที่ 4.14 ผลลัพธ์การสร้างแบบจำลอง DenseNet121ร่วมกับเทคนิค Data Augmentation

CNN	Augmentation	Time (Min:s)	Test	
			Accuracy (%)	Loss
DenseNet121	Horizontal Flip	10:02	83.33	0.55
	Vertical Flip	10:13	82.92	0.42
	Rotation	26:45	83.94	0.35
	Zoom	26:31	83.33	0.46
	Horizontal + Vertical	10:53	83.86	0.45
	Rotation + Zoom	26:07	83.88	0.38

ตารางที่ 4.15 ผลลัพธ์การสร้างแบบจำลอง InceptionV3ร่วมกับเทคนิค Data Augmentation

CNN	Augmentation	Time (Min:s)	Test	
			Accuracy (%)	Loss
Inception V3	Horizontal Flip	11:10	83.88	0.37
	Vertical Flip	11:09	84.54	0.36
	Rotation	39:29	82.69	0.41
	Zoom	38:29	85.05	0.37
	Horizontal + Vertical	13:29	83.50	0.43
	Rotation + Zoom	38:34	77.58	0.46

ตารางที่ 4.16 ผลลัพธ์การสร้างแบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation

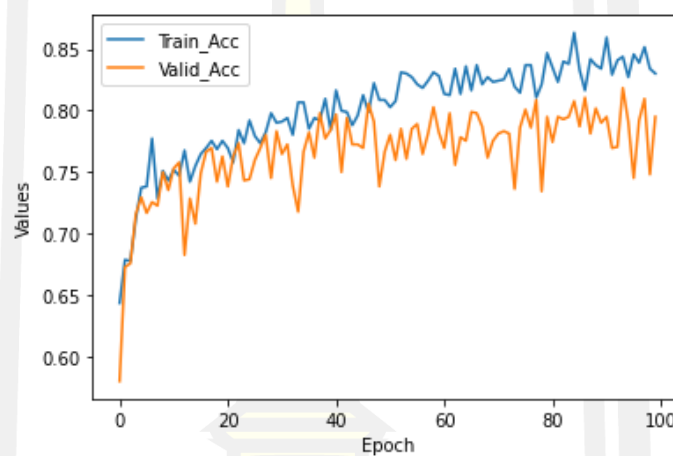
CNN	Augmentation	Time (Min:s)	Test	
			Accuracy (%)	Loss
Xception	Horizontal Flip	26:41	83.96	0.5
	Vertical Flip	25:29	81.20	0.42
	Rotation	44:14	85.37	0.34
	Zoom	40:56	83.67	0.57
	Horizontal + Vertical	11:37	84.94	0.35
	Rotation + Zoom	38:19	84.23	0.48

จากผลลัพธ์ของการทดลองการจำแนกประเภทรูปภาพจอบประสาทตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation รูปแบบ Horizontal Flip, Vertical Flip, Rotation และ Zoom บนชุดข้อมูล Test Set พบว่า เมื่อทำการรู้จำรูปภาพที่ผ่านกระบวนการ Augmentation สามารถเพิ่มประสิทธิภาพในการจำแนกประเภทรูปภาพจอบประสาทตาของทุกแบบจำลอง โดยแบบจำลอง Xception ที่ทำการเรียนรู้ร่วมกับ Rotation สามารถให้อัตราความถูกต้องในการจำแนกสูงสุด และรองลงมาเป็น InceptionV3, DenseNet121, ResNet50 และ ResNet50 V2 ตามลำดับ โดยการสรุปค่าผลลัพธ์ และรูปแบบ Data Augmentation ที่เหมาะสมสำหรับแต่ละแบบจำลอง จะแสดงดังตารางที่ 4.17

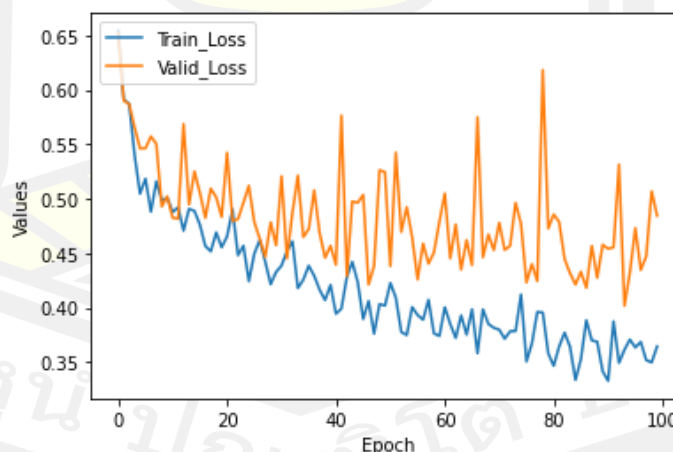
ตารางที่ 4.17 สรุปผลลัพธ์การทดลอง CNN ร่วมกับเทคนิค Data Augmentation

CNN	Augmentation	Time (Min: s)	Test	
			Accuracy (%)	Loss
Xception	Rotation	44:14	85.50	0.34
InceptionV3	Zoom	38:29	84.69	0.37
DenseNet121	Rotation	26:45	84.01	0.35
ResNet50	Horizontal + Vertical	39:59	83.26	0.41
ResNet50V2	Horizontal + Vertical	21:40	83.18	0.39

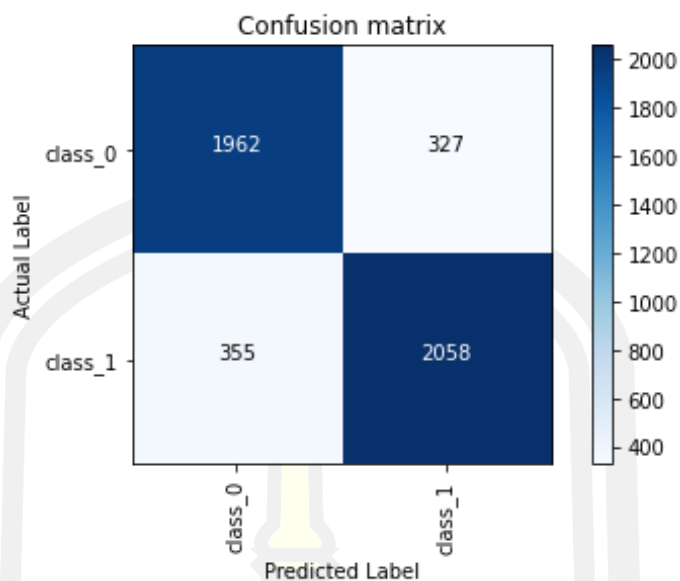
จากตารางสรุปผลลัพธ์การทดลองการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation พบว่าแบบจำลอง Xception ที่ทำการเรียนรู้เพื่อรู้จำรูปภาพจอประสาทตาบนชุดข้อมูลที่ผ่านกระบวนการ Augmentation รูปแบบ Rotation สามารถจำแนกประเภทรูปภาพจอประสาทตาได้ดีที่สุด โดยในการ Train แบบจำลองจำนวน 100 รอบ สามารถให้อัตรา Accuracy ของการ Train สูงสุดที่ร้อยละ 86.33 และ Loss ต่ำสุดที่ 0.33 ซึ่งรายละเอียดอัตรา Accuracy ในการ Train แบบจำลอง Xception จะแสดงดังภาพประกอบที่ 4.5 และ 4.6 โดยแบ่งออกเป็นอัตรา Accuracy ในการเรียนรู้ และอัตรา Loss ในการเรียนรู้ ตามลำดับ



ภาพประกอบที่ 4.5 อัตรา Accuracy ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation



ภาพประกอบที่ 4.6 อัตรา Loss ของการเรียนรู้แบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation



ภาพประกอบที่ 4.7 Confusion Matrix ของแบบจำลอง Xception ร่วมกับเทคนิค Data Augmentation

จากภาพประกอบที่ 4.7 จะแสดงการวัดประสิทธิภาพของแบบจำลอง Xception ที่มีการเรียนรู้ร่วมกับเทคนิค Data Augmentation โดยการใช้ Confusion Matrix ซึ่งเป็นการแสดงรายละเอียดการจำแนกประเภทรูปภาพระหว่างจอบประสาทตาประสาทตาปกติ (Class_0) และจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา (Class_1) บนชุดข้อมูล Test Set โดยอัตรา Accuracy ของการจำแนกประเภทรูปภาพจะคิดเป็นร้อยละ 85.50

4.3.3 ผลลัพธ์การจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN

จากการทดลองสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN เป็นการนำแบบจำลองที่ได้จากการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Data Augmentation มาทำการวิเคราะห์ด้วยเทคนิค Average Voting เพื่อหาค่าเฉลี่ยความน่าจะเป็นของคำตอบร่วมกัน โดยในการทดลองนี้แบ่งออกเป็น Unweighted Average และ Weighted Average ซึ่งมีรายละเอียดดังตารางที่ 4.18 – 4.21

ตารางที่ 4.18 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค

Ensemble

Unweight Accuracy (%)	Weight Parameter	Accuracy (%)	Weight Parameter	Accuracy (%)
85.92	0.1, 0.9	85.01	0.9, 0.1	85.69
	0.2, 0.8	85.45	0.8, 0.2	85.73
	0.3, 0.7	85.75	0.7, 0.3	86.07
	0.4, 0.6	85.92	0.6, 0.4	86.11

ตารางที่ 4.19 ผลลัพธ์การสร้างแบบจำลอง ResNet50 และ DenseNet121 ร่วมกับเทคนิค

Ensemble

Unweight Accuracy (%)	Weight Parameter	Accuracy (%)	Weight Parameter	Accuracy (%)
84.30	0.1, 0.9	84.20	0.9, 0.1	82.05
	0.2, 0.8	83.73	0.8, 0.2	82.11
	0.3, 0.7	84.81	0.7, 0.3	82.54
	0.4, 0.6	84.32	0.6, 0.4	82.92

ตารางที่ 4.20 ผลลัพธ์การสร้างแบบจำลอง ResNet50V2 และ DenseNet121 ร่วมกับเทคนิค

Ensemble

Unweight Accuracy (%)	Weight Parameter	Accuracy (%)	Weight Parameter	Accuracy (%)
84.22	0.1, 0.9	84.16	0.9, 0.1	83.48
	0.2, 0.8	84.41	0.8, 0.2	83.77
	0.3, 0.7	84.39	0.7, 0.3	84.09
	0.4, 0.6	84.28	0.6, 0.4	84.01

ตารางที่ 4.21 ผลลัพธ์การสร้างแบบจำลอง ResNet50 และ ResNet50V2 ร่วมกับเทคนิค

Ensemble

Unweight Accuracy (%)	Weight Parameter	Accuracy (%)	Weight Parameter	Accuracy (%)
83.65	0.1, 0.9	83.35	0.9, 0.1	80.90
	0.2, 0.8	83.33	0.8, 0.2	81.05
	0.3, 0.7	83.73	0.7, 0.3	81.48
	0.4, 0.6	82.69	0.6, 0.4	81.82

จากผลลัพธ์การสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN สำหรับการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาโดยการนำคำตอบหรือค่า Probability ที่ได้จากการ Prediction ของแต่ละ CNN มาทำการคำนวณร่วมกันด้วยกระบวนการ Average Voting เพื่อหาค่าเฉลี่ยของคำตอบเพื่อให้ได้คำตอบที่ถูกต้องมากที่สุด และพบว่าเมื่อ CNN ได้มีการวิเคราะห์หาคำตอบร่วมกันนั้นสามารถให้อัตราความถูกต้องได้สูงกว่า Single CNN ทุกรูปแบบ โดยเฉพาะเมื่อรวมคำตอบของแบบจำลอง Xception และ InceptionV3 เข้าด้วยกันสามารถให้อัตราความถูกต้องได้สูงสุดถึงร้อยละ 86.11 โดยตารางสรุปผลลัพธ์การสร้างแบบจำลองด้วยเทคนิค Ensemble จะแสดงดังตารางที่ 4.22 และ 4.23

ตารางที่ 4.22 สรุปผลลัพธ์การสร้างแบบจำลองด้วยเทคนิค Ensemble - Unweighted Average

CNNs		Accuracy (%)
Xception	InceptionV3	85.92
ResNet50	DenseNet121	84.30
ResNet50V2	DenseNet121	84.22
ResNet50	ResNet50V2	83.65

ตารางที่ 4.23 สรุปผลลัพธ์การสร้างแบบจำลองด้วยเทคนิค Ensemble - Weighted Average

CNNs		Weighted Parameters		Accuracy (%)
Xception	InceptionV3	0.6	0.4	86.11
ResNet50	DenseNet121	0.3	0.7	84.81
ResNet50V2	DenseNet121	0.3	0.7	84.39
ResNet50	ResNet50V2	0.3	0.7	83.73

โดยในการคำนวณวิเคราะห์เพื่อหาค่า Probability ของการจำแนกประเภทรูปภาพจอบรรยากาศที่มีภาวะเบาหวานขึ้นตาด้วยการใช้เทคนิค Average Voting รูปแบบ Unweighted Average สามารถให้อัตราความถูกต้องสูงสุดที่ร้อยละ 85.92 โดยเป็นการคำนวณค่าตอบร่วมกันระหว่าง แบบจำลอง Xception และ InceptionV3 และการจำแนกประเภทรูปภาพจอบรรยากาศที่มีภาวะเบาหวานขึ้นตาโดยการใช้รูปแบบ Weighted Average ซึ่งมีการกำหนดค่าน้ำหนักในการคำนวณเพื่อหาค่าเฉลี่ยของ Probability ของคำตอบที่ถูกต้องที่สุดโดยในการกำหนดค่าน้ำหนักนั้นจะต้องมีค่ารวมกันไม่เกิน 1 และในรูปแบบ Weighted Average ผลลัพธ์ค่าเฉลี่ยความน่าจะเป็นของคำตอบที่ได้จากการคำนวณร่วมกันของ แบบจำลอง Xception และ InceptionV3 ที่กำหนดค่าน้ำหนักไว้ที่ 0.6 และ 0.4 โดยสามารถให้อัตราความถูกต้องในการจำแนกประเภทสูงสุดที่ร้อยละ 86.11 ซึ่งเป็นค่าที่สูงสุดในการจำแนกประเภทรูปภาพจอบรรยากาศที่มีภาวะเบาหวานขึ้นตาโดยใช้เทคนิค Ensemble CNN

4.3.4 ผลลัพธ์การจำแนกรูปภาพจอบรรยากาศที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN

จากการทดลองสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN ซึ่งเป็นการรวมคุณสมบัติในการจำแนกประเภทของสถาปัตยกรรม Xception และสถาปัตยกรรม CNN รูปแบบต่าง ๆ เนื่องจาก ในการทดลองร่วมกับเทคนิค Fine-tuning Hyperparameter และการทดลองร่วมกับเทคนิค Data Augmentation สถาปัตยกรรม Xception ให้อัตรา Accuracy ในการจำแนกประเภทรูปภาพจอบรรยากาศสูงสุด

อีกทั้งมีการเพิ่ม Dense Layer และ Dropout Layer พร้อมกับปรับค่าตามที่ได้กำหนดไว้ดังตารางที่ 4.5 เพื่อให้ได้แบบจำลองที่สามารถจำแนกประเภทรูปภาพจอบรรยากาศได้ถูกต้องแม่นยำมากที่สุด โดยใช้จำนวนรอบการเรียนรู้ (Epoch) 100 รอบ ซึ่งผลลัพธ์ และรายละเอียดของการทดลองการจำแนกประเภทรูปภาพจอบรรยากาศด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN จะแสดงดังตารางที่ 4.24 – 4.47

พหุบัณฑิต

ตารางที่ 4.24 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	-	53:06	85.77	0.35
	2048		53:59	85.24	0.36
	4096		53:52	84.69	0.38
2	1024		53:47	85.09	0.34
	2048		54:16	85.24	0.35
	4096		54:50	85.30	0.39

ตารางที่ 4.25 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.1	53:21	85.60	0.49
	2048		53:32	84.11	0.54
	4096		53:09	85.58	0.36
2	1024		52:57	85.09	0.36
	2048		53:55	80.41	0.42
	4096		53:43	79.12	0.40

ตารางที่ 4.26 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.2	53:39	81.01	0.87
	2048		52:58	85.11	0.34
	4096		53:05	85.28	0.37
2	1024		53:35	85.92	0.34
	2048		54:18	84.96	0.4
	4096		54:10	85.11	0.34

ตารางที่ 4.27 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.3	53:32	85.69	0.34
	2048		53:13	83.13	0.4
	4096		53:06	84.39	0.37
2	1024		53:31	85.69	0.4
	2048		53:27	84.73	0.35
	4096		53:44	83.79	0.37

ตารางที่ 4.28 ผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.4	53:36	84.86	0.35
	2048		53:37	82.54	0.37
	4096		53:54	85.16	0.41
2	1024		53:13	86.30	0.33
	2048		53:46	84.96	0.38
	4096		53:25	83.82	0.44

ตารางที่ 4.29 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Test	
			Accuracy (%)	Loss
1024	2	0.4	86.30	0.33
1024	2	0.1	85.92	0.34
1024	2	0.2	85.92	0.34
1024	1	-	85.77	0.35
1024	1	0.3	85.69	0.34

ตารางที่ 4.30 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	-	26:20	83.48	0.37
	2048		26:55	83.71	0.36
	4096		26:38	83.24	0.38
2	1024		26:28	84.13	0.37
	2048		27:09	83.67	0.38
	4096		27:28	83.01	0.44

ตารางที่ 4.31 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.1	26:57	81.86	0.40
	2048		26:47	83.86	0.40
	4096		27:10	84.22	0.40
2	1024		28:10	84.79	0.35
	2048		27:13	83.92	0.36
	4096		27:58	83.52	0.43

ตารางที่ 4.32 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.2	33:19	82.01	0.36
	2048		33:27	81.97	0.43
	4096		33:24	84.22	0.38
2	1024		33:18	84.01	0.39
	2048		33:58	85.45	0.34
	4096		34:25	83.33	0.45

ตารางที่ 4.33 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.3	33:17	84.16	0.37
	2048		33:02	83.67	0.46
	4096		32:35	84.24	0.36
2	1024		34:30	84.47	0.36
	2048		33:54	84.64	0.36
	4096		33:47	84.09	0.37

ตารางที่ 4.34 ผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.4	33:33	84.05	0.37
	2048		33:37	83.22	0.36
	4096		33:02	84.43	0.47
2	1024		34:47	85.37	0.35
	2048		33:46	83.56	0.38
	4096		34:08	83.77	0.37

ตารางที่ 4.35 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ DenseNet121 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Test	
			Accuracy (%)	Loss
2048	2	0.2	85.45	0.34
2048	2	0.4	85.37	0.35
1024	2	0.1	84.79	0.35
2048	2	0.3	84.64	0.36
1024	2	-	84.13	0.37

ตารางที่ 4.36 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	-	31:52	78.80	0.58
	2048		33:25	83.60	0.38
	4096		32:42	82.45	0.40
2	1024		32:55	80.71	0.42
	2048		33:26	83.50	0.38
	4096		32:18	84.11	0.37

ตารางที่ 4.37 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.1	32:14	85.07	0.4
	2048		32:42	84.33	0.41
	4096		32:23	84.05	0.44
2	1024		32:47	81.20	0.42
	2048		32:48	83.35	0.42
	4096		32:14	79.99	1.64

ตารางที่ 4.38 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.2	32:49	84.41	0.37
	2048		33:05	81.33	0.42
	4096		31:47	83.48	0.39
2	1024		32:50	80.69	0.55
	2048		33:58	84.60	0.39
	4096		33:09	83.88	0.38

ตารางที่ 4.39 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.3	32:20	83.45	0.42
	2048		33:49	83.84	0.38
	4096		32:00	84.33	0.37
2	1024		33:32	81.77	0.49
	2048		33:28	80.46	2.17
	4096		31:41	85.01	0.37

ตารางที่ 4.40 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.4	32:02	83.11	0.38
	2048		31:54	84.24	0.38
	4096		31:40	84.41	0.44
2	1024		31:53	84.73	0.39
	2048		31:53	84.71	0.36
	4096		32:52	83.22	0.44

ตารางที่ 4.41 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Test	
			Accuracy (%)	Loss
1024	1	0.1	85.07	0.40
4096	2	0.3	85.01	0.37
1024	2	0.4	84.73	0.39
2048	2	0.2	84.60	0.39
4096	2	-	84.11	0.37

ตารางที่ 4.42 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	-	30:21	85.03	0.35
	2048		29:14	82.73	0.39
	4096		30:18	85.11	0.36
2	1024		30:34	84.79	0.36
	2048		29:59	79.73	0.48
	4096		30:22	79.86	0.44

ตารางที่ 4.43 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.1

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.1	30:56	83.24	0.57
	2048		30:17	84.73	0.36
	4096		30:44	79.90	1.17
2	1024		31:04	83.77	0.37
	2048		30:29	84.73	0.40
	4096		29:51	83.99	0.45

ตารางที่ 4.44 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.2

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.2	30:16	83.45	0.45
	2048		30:31	84.13	0.94
	4096		30:32	83.92	0.47
2	1024		30:58	84.01	0.5
	2048		29:10	82.48	1.87
	4096		29:55	83.99	0.45

ตารางที่ 4.45 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.3

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.3	31:41	85.09	0.36
	2048		31:59	84.41	0.36
	4096		31:52	83.79	0.38
2	1024		32:25	83.94	0.39
	2048		31:02	84.45	0.38
	4096		31:56	81.62	1.66

ตารางที่ 4.46 ผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN เพิ่ม Dropout 0.4

Dense Layer	Dense Parameter	Dropout Parameter	Time (Min:s)	Test	
				Accuracy (%)	Loss
1	1024	0.4	31:36	81.39	1.46
	2048		32:23	84.77	0.36
	4096		32:10	78.56	1.04
2	1024		32:03	83.50	0.36
	2048		31:59	84.16	0.42
	4096		32:52	84.75	0.36

ตารางที่ 4.47 สรุปผลลัพธ์การสร้างแบบจำลอง Xception และ ResNet50V2 ร่วมกับเทคนิค Fusion CNN

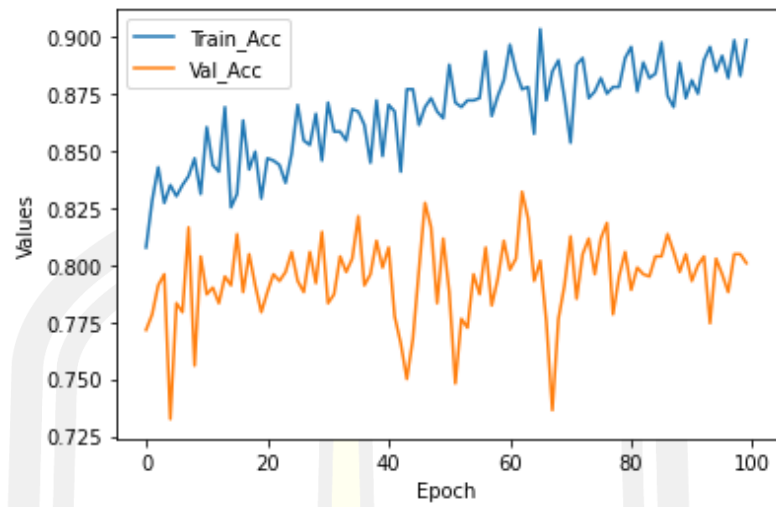
Dense Layer	Dense Parameter	Dropout Parameter	Test	
			Accuracy (%)	Loss
4096	1	-	85.11	0.36
1024	1	0.3	85.09	0.36
2048	1	0.1	84.73	0.36
2048	1	0.4	84.77	0.36
2048	1	0.2	84.13	0.94

จากผลลัพธ์ของการทดลองการจำแนกประเภทรูปภาพจอสภาพตาด้วยแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN ซึ่งในการทดลองนี้ผู้วิจัยได้นำแบบจำลอง Xception รวมเข้ากับแบบจำลอง InceptionV3, DenseNet121, ResNet50V2 และ ResNet50 เพื่อสร้างเป็นแบบจำลองรูปแบบใหม่ ที่มีคุณสมบัติของทั้ง 2 แบบจำลอง และเมื่อทำการทดสอบประสิทธิภาพการทำงานของแบบจำลองบนชุดข้อมูล Test Set พบว่าการการรวมแบบจำลองเข้าด้วยกันสามารถเพิ่มประสิทธิภาพในการจำแนกรูปภาพจอสภาพตาได้มากยิ่งขึ้น โดยแบบจำลอง Xception และ InceptionV3 สามารถให้อัตราความถูกต้องสูงสุดอยู่ที่ร้อยละ 86.30 โดยการสรุปค่าผลลัพธ์ และรายละเอียดการตั้งค่าที่เหมาะสมสำหรับแต่ละแบบจำลองจะแสดงดังตารางที่ 4.48

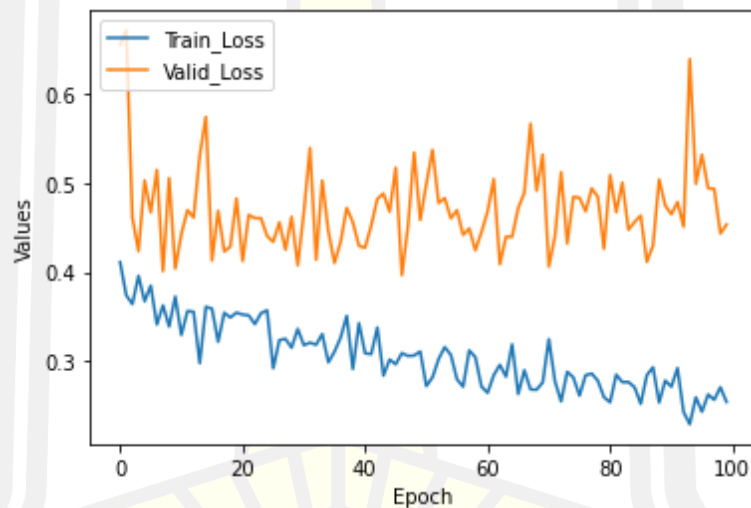
ตารางที่ 4.48 สรุปผลลัพธ์การสร้างแบบจำลอง Xception ร่วมกับเทคนิค Fusion CNN

CNN		Dense Layer	Dense Parameter	Dropout Parameter	Test	
					Accuracy (%)	Loss
Xception	InceptionV3	2	1024	0.4	86.30	0.33
	DenseNet121	2	2048	0.2	85.45	0.34
	ResNet50V2	1	4096	-	85.11	0.36
	ResNet50	1	1024	0.1	85.07	0.4

จากตารางสรุปผลลัพธ์การทดลองการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fusion CNN พบว่าเมื่อรวมคุณสมบัติของแบบจำลอง Xception และ InceptionV3 เข้าด้วยกันโดยมีการตั้งค่า Dense Layer ที่ 1024 จำนวน 2 Layer และตั้งค่า Dropout Parameter ที่ 0.4 สามารถจำแนกประเภทรูปภาพจอสภาพตาได้ดีที่สุด โดยในการ Train แบบจำลองจำนวน 100 รอบ นั้นสามารถให้อัตรา Accuracy สูงสุดที่ร้อยละ 90.33 และ Loss ต่ำสุดที่ 0.23 ซึ่งรายละเอียดการเรียนรู้ของแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN จะแสดงดังภาพประกอบที่ 4.8 และ 4.9 โดยแบ่งออกเป็นอัตรา Accuracy ในการเรียนรู้ และอัตรา Loss ในการเรียนรู้ ตามลำดับ

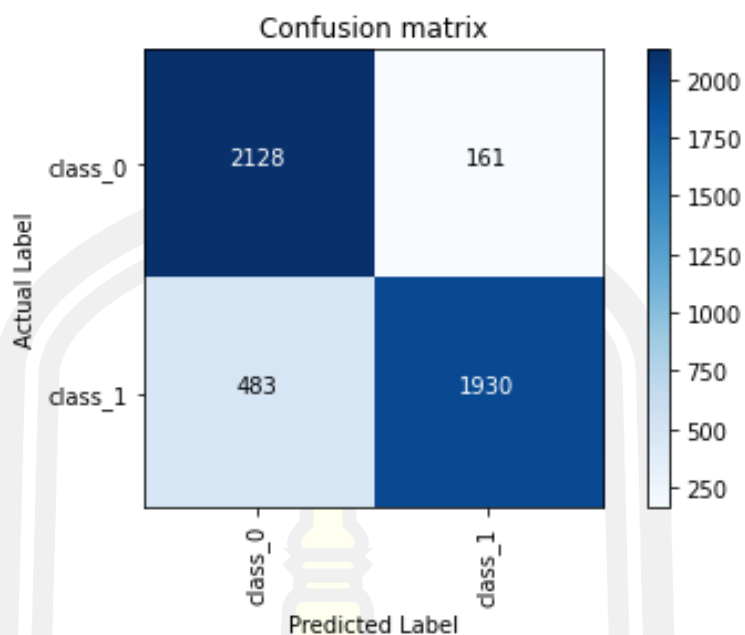


ภาพประกอบที่ 4.8 อัตรา Accuracy ของการเรียนรู้แบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN



ภาพประกอบที่ 4.9 อัตรา Loss ของการเรียนรู้แบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN

พหุบัณฑิต ชีวะ



ภาพประกอบที่ 4.10 Confusion Matrix ของแบบจำลอง Xception และ InceptionV3 ร่วมกับเทคนิค Fusion CNN

จากภาพประกอบที่ 4.10 จะแสดงการวัดประสิทธิภาพของแบบจำลอง Xception และ InceptionV3 ที่มีการเรียนรู้ร่วมกับเทคนิค Fusion CNN โดยการใช้ Confusion Matrix ซึ่งเป็นการแสดงรายละเอียดการจำแนกประเภทรูปภาพระหว่างจอบประสาทตาประสาทตาศปกติ (Class_0) และจอบประสาทตาที่มีภาวะเบาหวานขึ้นตา (Class_1) บนชุดข้อมูล Test Set โดยอัตรา Accuracy ของการจำแนกประเภทรูปภาพจะคิดเป็นร้อยละ 86.30

4.4 ประเมินและเปรียบเทียบประสิทธิภาพการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาของแบบจำลอง CNN

จากผลลัพธ์การทดสอบประสิทธิภาพการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาบนชุดข้อมูล Test Set ของแบบจำลอง CNN ที่มีการสร้างร่วมกับเทคนิค Fine-tuning Hyperparameter, Data Augmentation, Ensemble CNN และ Fusion CNN พบว่าแบบจำลอง Xception และ InceptionV3 ที่สร้างร่วมกับเทคนิค Fusion CNN มีอัตราความถูกต้องที่ร้อยละ 86.30 ซึ่งเป็นค่าความถูกต้องสูงสุด รองลงมาเป็นแบบจำลอง Xception และ InceptionV3 ที่สร้างร่วมกับเทคนิค Ensemble CNN ซึ่งมีอัตราความถูกต้องที่ร้อยละ 86.11 และแบบจำลอง Xception ที่สร้างร่วมกับเทคนิค Data Augmentation และ Fine-tuning Hyperparameter ให้อัตราความถูกต้องที่ร้อยละ 85.50, 84.07 ตามลำดับ โดยรายละเอียดการเปรียบเทียบประสิทธิภาพในการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาของแบบจำลองและเทคนิคต่าง ๆ จะแสดงดังตารางที่ 4.49

ตารางที่ 4.49 เปรียบเทียบประสิทธิภาพในการจำแนกประเภทรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาของแบบจำลอง CNN และเทคนิคต่าง ๆ

CNNs		Techniques	Accuracy (%)
Xception	InceptionV3	Fusion CNN	86.30
Xception	InceptionV3	Ensemble CNN [0.6, 0.4]	86.11
Xception		Data Augmentation	85.50
Xception		Fine-tuning Hyperparameter	84.07

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

งานวิจัยฉบับนี้มีความมุ่งหมายเพื่อศึกษาและเปรียบเทียบสถาปัตยกรรม CNN บนหลักการเรียนรู้เชิงลึกสำหรับการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาซึ่งสามารถสรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะเพิ่มเติมในการวิจัยได้ดังต่อไปนี้

5.1 สรุปผลการวิจัย

5.2 อภิปรายผล

5.3 ข้อเสนอแนะ และแนวทางการวิจัยในอนาคต

5.1 สรุปผลการวิจัย

งานวิจัยฉบับนี้ทำการศึกษาและเปรียบเทียบประสิทธิภาพการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วยแบบจำลอง CNN ที่ประกอบไปด้วย ResNet50, ResNet50V2, DenseNet121, InceptionV3 และ Xception โดยในการทดลองนั้นมีการสร้างแบบจำลองร่วมกับเทคนิคต่าง ๆ ได้แก่ Fine-tuning Hyperparameter, Data Augmentation, Ensemble CNN และ Fusion CNN เพื่อเพิ่มประสิทธิภาพในการจำแนกประเภทของแบบจำลอง ซึ่งในงานวิจัยนี้ได้เลือกใช้ Google Colaboratory (Google Colab) ทำให้สามารถประมวลผลได้บนหน่วยประมวลผลด้านกราฟฟิก (Graphics Processing Unit: GPU) เป็นเครื่องมือในการเรียนรู้ และทดสอบแบบจำลอง โดยแบบจำลองทั้งหมดถูกสร้างขึ้นจากการเรียนรู้จำนวน 100 Epoch ซึ่งเป็นการรู้จำชุดข้อมูลรูปภาพจอประสาทตาสำหรับการเรียนรู้ (Training Set) จำนวน 18,808 รูปภาพ และถูกทดสอบประสิทธิภาพบนชุดข้อมูลสำหรับการทดสอบ (Test Set) จำนวน 4,702 รูปภาพ

จากผลการทดลองพบว่าเมื่อสร้างและเรียนรู้แบบจำลองร่วมกับเทคนิค Data Augmentation สามารถเพิ่มอัตราความถูกต้องในการจำแนกประเภทจอประสาทตาให้กับทุกแบบจำลอง และสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN และ Fusion CNN นั้นสามารถให้อัตราความถูกต้องได้มากกว่า Single CNN เนื่องจากเป็นการนำเอาคุณสมบัติในการจำแนกประเภทของหลายแบบจำลองรวมเข้าด้วยกันจึงทำให้มีประสิทธิภาพสูงกว่า ซึ่งในการทดลองนั้นพบว่าเมื่อนำคุณสมบัติของ Xception และ InceptionV3 รวมเข้าด้วยกันจะให้ผลลัพธ์อัตราความถูกต้องในการจำแนกประเภทจอประสาทตาสูงสุด และจากผลการทดลองการวิจัยสามารถสรุปได้ว่าแบบจำลองที่มีการรวมคุณสมบัติของ CNN รูปแบบ Xception และ InceptionV3 ด้วยเทคนิค Fusion CNN เหมาะสมสำหรับนำไปใช้เพื่อจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา เนื่องจากมีอัตราความถูกต้องในการจำแนกประเภทที่สูงที่สุด

5.2 อภิปรายผล

งานวิจัยฉบับนี้ได้ทำการปรับรูปภาพในชุดข้อมูลให้เหลือขนาด 299x299 Pixel เนื่องจากขนาดของภาพจากต้นฉบับของข้อมูลนั้นมีขนาดที่ไม่สมดุลกัน และรูปภาพถูกแบ่งออกเป็นสัดส่วน 80:20 เพื่อให้เป็นชุดข้อมูลสำหรับการเรียนรู้ของแบบจำลอง และเป็นชุดข้อมูลสำหรับการทดสอบประสิทธิภาพของแบบจำลองตามลำดับ โดยกำหนดให้ใช้ค่าความถูกต้อง (Accuracy) ในการประเมินประสิทธิภาพของแบบจำลอง

จากผลการทดลองการสร้างแบบจำลองสำหรับการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาด้วย CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter, Data Augmentation, Ensemble CNN และ Fusion CNN แสดงให้เห็นว่าเมื่อแบบจำลองแต่ละรูปแบบได้รับการเรียนรู้ร่วมกับกระบวนการที่เหมาะสมจะสามารถถึงประสิทธิภาพในการจำแนกประเภทรูปภาพออกมาได้สูงสุด โดยจากผลลัพธ์การทดลองสร้างแบบจำลอง CNN ร่วมกับเทคนิค Fine-tuning Hyperparameter พบว่าเมื่อแต่ละแบบจำลองได้ถูกปรับแต่ง Hyperparameter Optimization และ Learning Rate ให้เหมาะสม จะสามารถให้อัตราความถูกต้องได้สูงกว่าค่าอื่น ๆ ซึ่งจะเห็นได้ว่าแบบจำลอง Xception เมื่อถูกปรับให้ใช้ Optimization เป็น Adamax ที่ Learning Rate 0.001 สามารถให้อัตราความถูกต้องในการจำแนกรูปภาพจอบประสาทตาได้สูงสุดที่ร้อยละ 84.07 แต่เมื่อมีการเรียนรู้แบบจำลอง CNN ทุกรูปแบบร่วมกับเทคนิค Data Augmentation พบว่าสามารถเพิ่มอัตราความถูกต้องในการจำแนกรูปภาพจอบประสาทตาที่มีภาวะเบาหวานขึ้นตาได้สูงขึ้น เนื่องจากชุดข้อมูลมีความหลากหลายมากในด้าน จำนวน มิติ และมุมมอง เพิ่มมากยิ่งขึ้น โดยแบบจำลอง Xception ที่เรียนรู้ร่วมกับรูปแบบ Rotation สามารถให้อัตราความถูกต้องสูงสุดร้อยละ 85.50

ทั้งนี้ การสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble CNN และ Fusion CNN พบว่าสามารถให้อัตราความถูกต้องสูงกว่าการจำแนกด้วยแบบจำลอง CNN แบบเดี่ยว (Single CNN) เนื่องจากเทคนิค Ensemble และ Fusion นั้นเป็นการนำเอาคุณสมบัติในการจำแนกคุณลักษณะพิเศษของหลายแบบจำลองรวมเข้าไว้ด้วยกัน และนำไปทำการประมวลผลเพื่อจำแนกประเภทวัตถุในรูปภาพต่อไป โดยในการสร้างแบบจำลอง CNN ร่วมกับเทคนิค Ensemble นั้นเป็นการนำเอาแบบจำลองที่ได้รับการเรียนรู้จากข้อมูลชุดเดียวกันมาทำการ Prediction เพื่อให้ได้ค่า Probability ของแต่ละแบบจำลอง และนำค่าที่ได้จากทั้ง 2 แบบจำลอง ไปทำการ Vote เพื่อหาค่าเฉลี่ยของคำตอบที่ถูกต้องที่สุด ซึ่งเรียกกระบวนการนี้ว่า Average Voting โดยจะแบ่งออกเป็น Unweighted Average และ Weighted Average โดยในการ Weighted นั้นทั้ง 2 แบบจำลอง จะมีการเฉลี่ยค่าน้ำหนักโดยที่รวมกันแล้วจะมีค่าไม่เกิน 1 และจากผลการทดลองแสดงให้เห็นว่าเมื่อนำแบบจำลอง Xception และ Inception มาทำการ Prediction และ Average Voting เพื่อหาคำตอบที่ถูกต้อง

ที่สุดด้วยรูปแบบ Weighted Average โดยมีการกำหนดค่าน้ำหนักเป็น 0.6 และ 0.4 ตามลำดับ สามารถให้อัตราความถูกต้องในการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา สูงสุดที่ร้อยละ 86.11 โดยเป็นอัตราความถูกต้องที่รองลงมาจากการสร้างแบบจำลองร่วมกับเทคนิค Fusion ที่เป็นการดึงเอาคุณสมบัติในการจำแนกคุณลักษณะพิเศษของแต่ละสถาปัตยกรรม CNN มา รวมเข้าด้วยกัน และทำการสร้างเป็นแบบจำลองรูปแบบใหม่ที่มีความสามารถในการจำแนกประเภท ของวัตถุในรูปภาพของทุก CNN โดยในการทดลองได้มีการปรับแต่งเพิ่ม Dense Layer และ Dropout Layer เพื่อลดการเกิดปัญหา Overfitting และจากการทดลองพบว่าเมื่อนำ Xception และ Inception รวมเข้าด้วยกัน พร้อมทั้งเพิ่ม Dense Layer ที่ 1024 จำนวน 2 Layer และเพิ่ม Dropout Layer ที่ 0.4 สามารถให้อัตราในการจำแนกประเภทรูปภาพจอประสาทตาที่มีภาวะ เบาหวานขึ้นตาส่งถึงร้อยละ 86.30 ซึ่งเป็นอัตราความถูกต้องสูงสุดในการวิจัยนี้ และยังเป็นผลลัพธ์ใน การจำแนกประเภทจอประสาทตาที่มีภาวะเบาหวานขึ้นตาที่มีอัตราความถูกต้องสูงกว่างานวิจัย [30] โดย Hattiya และคณะ

5.3 ข้อเสนอแนะ และแนวทางการวิจัยในอนาคต

จากงานวิจัยฉบับนี้ผู้วิจัยได้ศึกษาและเปรียบเทียบสถาปัตยกรรม CNN บนหลักการเรียนรู้เชิงลึก สำหรับการจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา โดยได้มีการใช้เทคนิคต่าง ๆ เพื่อ เพิ่มประสิทธิภาพหรืออัตราความถูกต้องในการจำแนกประเภทรูปภาพจอประสาทตาปกติ และ รูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา ในงานวิจัยครั้งถัดไปจะมีการสร้างแบบจำลอง CNN สำหรับการจำแนกประเภทรูปภาพจอประสาทตา 5 ประเภท โดยจะแบ่งตามระดับความรุนแรงของ โรค ได้แก่

- 1) จอประสาทตาปกติ
- 2) ภาวะเบาหวานขึ้นจอประสาทตาขั้นเบา (Mild Non-Proliferative Diabetic Retinopathy)
- 3) ภาวะเบาหวานขึ้นจอประสาทตาขั้นปานกลาง (Moderate Non-Proliferative Diabetic Retinopathy)
- 4) ภาวะเบาหวานขึ้นจอประสาทตาขั้นหนัก (Severe Non-Proliferative Diabetic Retinopathy)
- 5) ภาวะเบาหวานเข้าจอประสาทตาที่มีเส้นเลือดงอกใหม่ (Proliferative Diabetic Retinopathy)

บรรณานุกรม



บรรณานุกรม

- [1] Y. Shimizu, "Eye care, Vision care, Vision impairment and Blindness," *World Health Organization (WHO)*, 2021. <https://www.who.int/health-topics/blindness-and-vision-loss> (accessed Jun. 06, 2021).
- [2] D. Doshi, A. Shenoy, D. Sidhpura, and P. Gharpure, "Diabetic Retinopathy Detection using Deep Convolutional Neural Networks," in *International Conference on Computing Analytics and Security Trends (CAST)*, 2016, pp. 261–266, doi: 10.1109/CAST.2016.7914977.
- [3] H. H. Vo and A. Verma, "New Deep Neural Nets for Fine-Grained Diabetic Retinopathy Recognition on Hybrid Color Space," in *IEEE International Symposium on Multimedia (ISM)*, 2016, pp. 209–215, doi: 10.1109/ISM.2016.99.
- [4] Q. Abbas, I. Fondon, A. Sarmiento, S. Jiménez, and P. Alemany, "Automatic Recognition of Severity Level for Diagnosis of Diabetic Retinopathy using Deep Visual Features," *Med. Biol. Eng. Comput.*, vol. 55, no. 11, pp. 1959–1974, 2017, doi: 10.1007/s11517-017-1638-6.
- [5] K. Xu, D. Feng, and H. Mi, "Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy using Fundus Image," *Molecules*, vol. 22, no. 12, pp. 2054–2060, 2017, doi: 10.3390/molecules22122054.
- [6] C. P. Wilkinson *et al.*, "Proposed International Clinical Diabetic Retinopathy and Diabetic Macular Edema Disease Severity Scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003, doi: 10.1016/S0161-6420(03)00475-5.
- [7] F. David and P. Jean, *Computer vision: A modern approach*. Upper Saddle River: Prentice-Hall, 2011.
- [8] Y. Wu and Z. Hu, "Recognition of Diabetic Retinopathy Based on Transfer Learning," in *IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 2019, pp. 398–401, doi: 10.1109/ICCCBDA.2019.8725801.
- [9] A. Smola and S. Vishwanathan, *Introduction to Machine Learning*. United Kingdom: The Press Syndicate of The University of Cambridge, 2008.

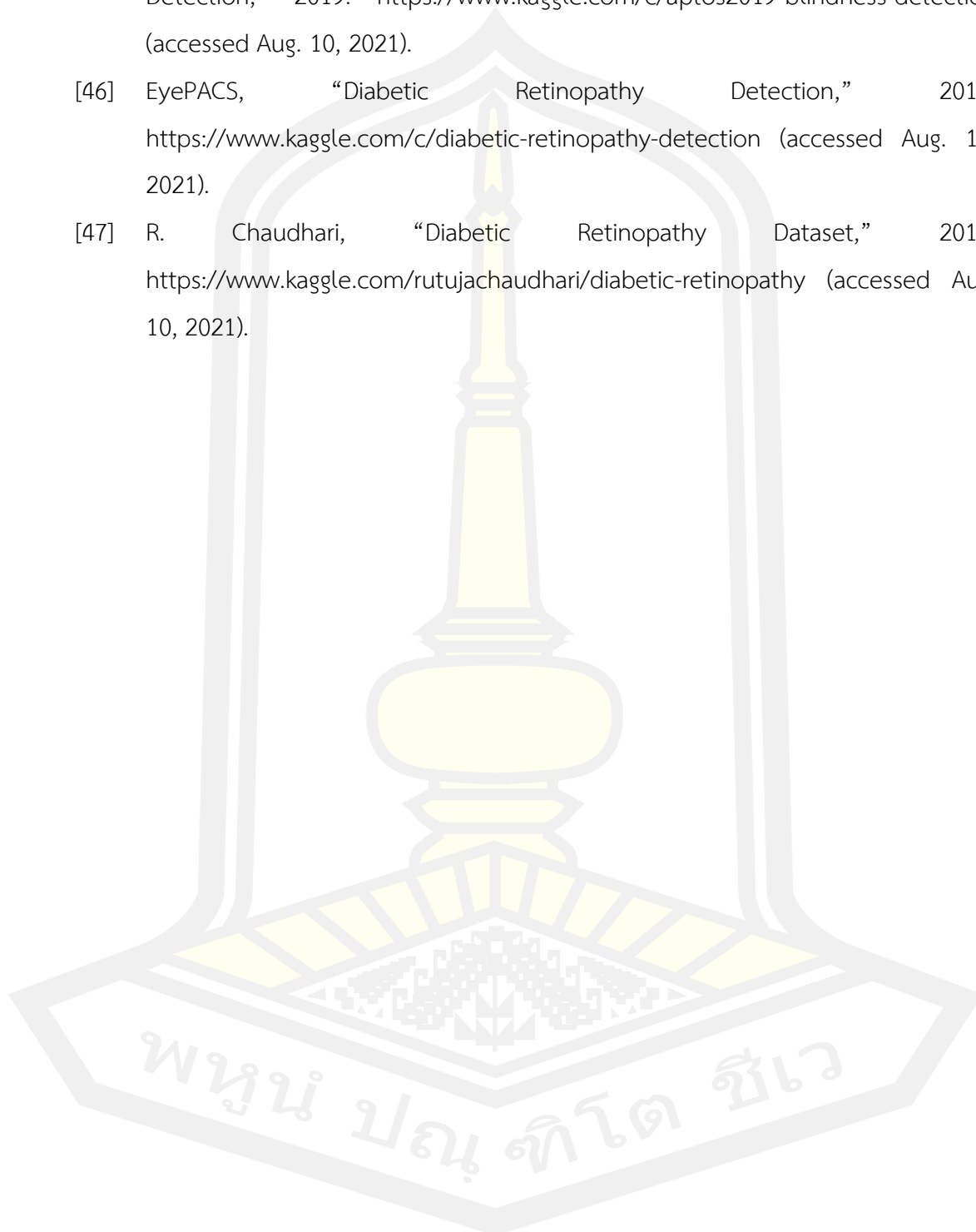
- [10] T. Oladipupo, "Types of Machine Learning Algorithms," *New Adv. Mach. Learn.*, vol. 3, no. February 2010, pp. 19–48, 2010, doi: 10.5772/9385.
- [11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.
- [12] I. Sutskever, J. Martens, and G. Hinton, "Generating text with recurrent neural networks," in *International Conference on Machine Learning (ICML)*, 2011, pp. 1017–1024.
- [13] J. Sun, R. Jin, X. Ma, J. young Park, K. ah Sohn, and T. sun Chung, "Gated Convolutional Neural Networks for Text Classification," in *Advances in Computer Science and Ubiquitous Computing*, vol. 715, 2021, pp. 309–316.
- [14] A. P. Vassilopoulos, E. F. Georgopoulos, and V. Dionysopoulos, "Artificial Neural Networks in Spectrum Fatigue Life Prediction of Composite Materials," *Int. J. Fatigue*, vol. 29, no. 1, pp. 20–29, 2007, doi: 10.1016/j.ijfatigue.2006.03.004.
- [15] P. Oncharoen, "Stock Market Movement Prediction Using Enhanced Deep Learning Model with Numerical and Textual Information," Chulalongkorn University, 2018.
- [16] K. He, Z. Xiangyu, R. Shaoqing, and S. Jian, "Deep Residual Learning for Image Recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 770–778.
- [17] K. He, Z. Xiangyu, R. Shaoqing, and S. Jian, "Identity Mappings in Deep Residual Networks," in *European Conference on Computer Vision (ECCV)*, 2016, pp. 630–645, doi: 10.1007/978-3-319-46493-0.
- [18] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. De Geus, "Malicious Software Classification using Transfer Learning of ResNet-50 Deep Neural Network," in *IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017, vol. 2017-Decem, pp. 1011–1014, doi: 10.1109/ICMLA.2017.00-19.
- [19] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1251–1258.

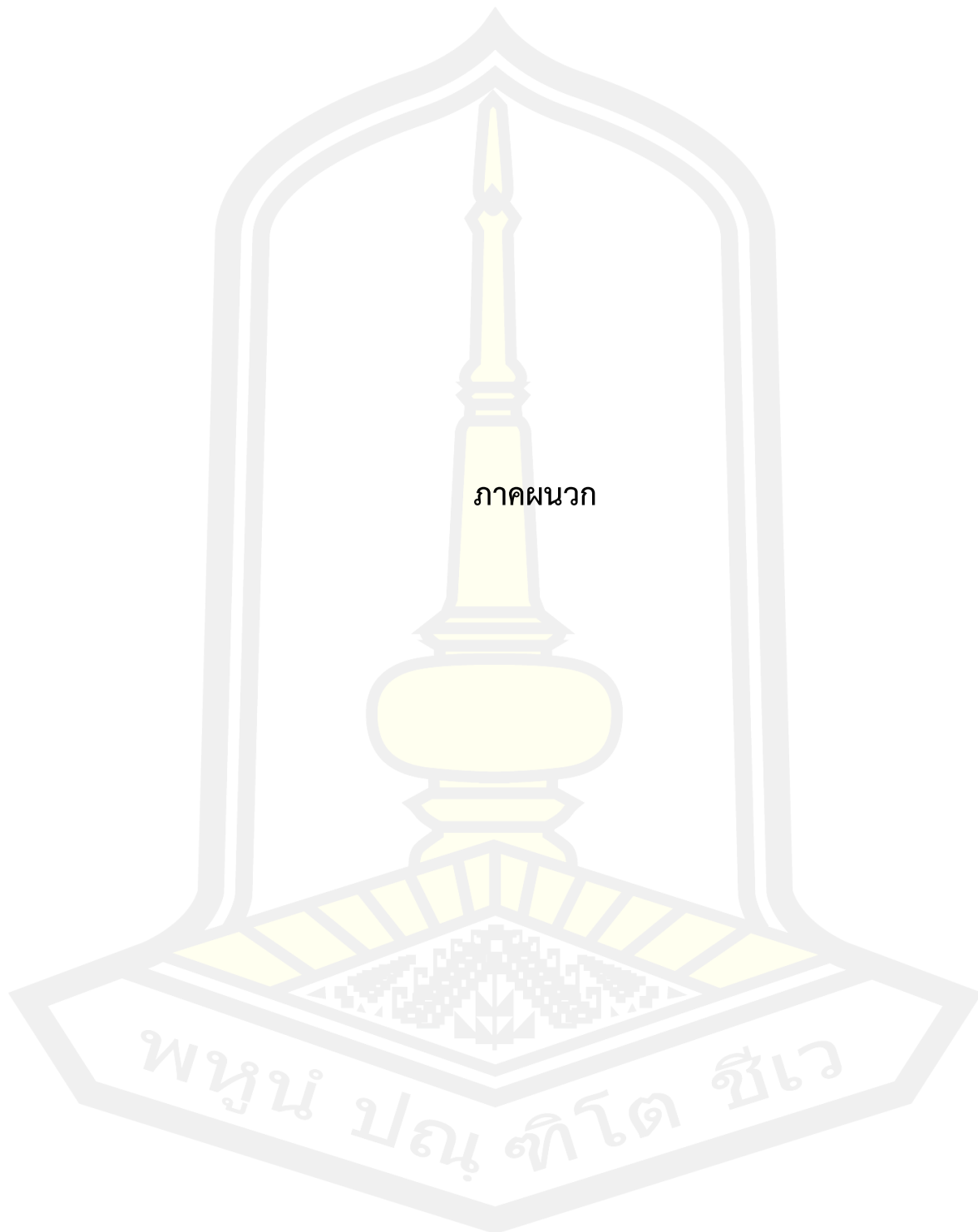
- [20] D. He *et al.*, “StNet: Local and global spatial-temporal modeling for action recognition,” in *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI)*, 2019, pp. 8401–8408, doi: 10.1609/aaai.v33i01.33018401.
- [21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, vol. 2016-Decem, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.
- [22] A. Demir, F. Yilmaz, and O. Kose, “Early Detection of Skin Cancer Using Deep Learning Architectures: Resnet-101 and Inception-v3,” in *Medical Technologies Congress (TIPTEKNO)*, 2019, vol. 2019-Janua, pp. 3–6, doi: 10.1109/TIPTEKNO47231.2019.8972045.
- [23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708, doi: 10.1109/CVPR.2017.243.
- [24] M. D Bloice, C. Stocker, and A. Holzinger, “Augmentor: An Image Augmentation Library for Machine Learning,” *J. Open Source Softw.*, vol. 2, no. 19, p. 432, 2017, doi: 10.21105/joss.00432.
- [25] A. Majid, M. A. Khan, A. Yousafzai, M. Yasmin, and U. Tariq, “Classification of Stomach Infections : A Paradigm of Convolutional Neural Network Along with Classical Features Fusion and Selection,” *Microsc. Res. Tech.*, vol. 83, no. 5, pp. 562–576, 2020, doi: 10.1002/jemt.23447.
- [26] T. Dietterich, “Ensemble Methods in Machine Learning,” in *Multiple Classifier Systems (MCS)*, 2000, vol. 1857, pp. 1–15, doi: https://doi.org/10.1007/3-540-45014-9_1.
- [27] S. Noppitak and O. Surinta, “Ensemble Convolutional Neural Network Architectures for Land use Classification in Economic Crops Aerial Images,” *ICIC Express Lett.*, vol. 15, no. 6, pp. 531–543, 2021, doi: 10.24507/icicel.15.06.531.
- [28] T. Chompookham and O. Surinta, “Ensemble Methods with Deep Convolutional Neural Networks for Plant Leaf Recognition,” *ICIC Express Lett.*, vol. 15, no. 6, pp. 553–565, 2021, doi: 10.24507/icicel.15.06.553.

- [29] J. Singh, H. Trehan, V. Kaur, and D. N. Dhillon, "Analysis techniques of Content based Image Retrieval based on Confusion Matrix," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 6, no. 3, pp. 164–169, 2017, doi: 10.17148/ijarcce.2017.6334.
- [30] T. Hattiya, K. Dittakan, and S. Musikasuwana, "Diabetic Retinopathy Detection Using Convolutional Neural Network: A Comparative Study on Different Architectures," *Maharakham Int. J. Eng. Technol.*, vol. 7, no. 1, pp. 50–60, 2021, doi: 10.1109/ACCESS.2019.2903171.
- [31] S. Burewar, A. B. Gonde, and S. K. Vipparthi, "Diabetic Retinopathy Detection by Retinal segmentation with Region merging using CNN," in *IEEE International Conference on Industrial and Information Systems (ICIIS)*, 2018, no. 978, pp. 136–142, doi: 10.1109/ICIINFS.2018.8721315.
- [32] P. W. Sudarmadji, P. Deviani Pakan, and R. Yefrenes Dillak, "Diabetic Retinopathy Stages Classification using Improved Deep Learning," in *International Conference on Informatics, Multimedia, Cyber, and Information System (ICIMCIS)*, 2020, pp. 104–109, doi: 10.1109/ICIMCIS51567.2020.9354281.
- [33] N. B. Thota and D. Umma Reddy, "Improving the Accuracy of Diabetic Retinopathy Severity Classification with Transfer Learning," in *Midwest Symposium on Circuits and Systems (MWSCAS)*, 2020, vol. 2020-Augus, pp. 1003–1006, doi: 10.1109/MWSCAS48704.2020.9184473.
- [34] S. Wan, Y. Liang, and Y. Zhang, "Deep Convolutional Neural Networks for Diabetic Retinopathy Detection by Image Classification," *Comput. Electr. Eng.*, vol. 72, pp. 274–282, 2018, doi: 10.1016/j.compeleceng.2018.07.042.
- [35] V. Deepa, C. S. Kumar, and T. Cherian, "Ensemble of Multi-stage Deep Convolutional Neural Networks for Automated Grading of Diabetic Retinopathy using Image patches," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 6, pp. 1319–1578, 2021, doi: 10.1016/j.jksuci.2021.05.009.
- [36] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, "Diagnostic Assessment of Deep Learning Algorithms for Diabetic Retinopathy Screening," *Inf. Sci. (Ny)*, vol. 501, pp. 511–522, 2019, doi: 10.1016/j.ins.2019.06.011.

- [37] D. Sharifrazi *et al.*, “Fusion of Convolution Neural Network, Support Vector Machine and Sobel Filter for Accurate Detection of COVID-19 Patients using X-ray Images,” *Biomed. Signal Process. Control*, vol. 68, no. October 2020, p. 102622, 2021, doi: 10.1016/j.bspc.2021.102622.
- [38] K. Yazhini and D. Loganathan, “An Integrated Fusion based Feature Extraction and Classification Model for Diabetic Retinopathy Diagnosis,” in *International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 1187–1193, doi: 10.1109/ICIRCA48905.2020.9183240.
- [39] Mobeen-Ur-Rehman, S. H. Khan, Z. Abbas, and S. M. Danish Rizvi, “Classification of Diabetic Retinopathy Images Based on Customised CNN Architecture,” in *Amity International Conference on Artificial Intelligence (AICAI)*, 2019, pp. 244–248, doi: 10.1109/AICAI.2019.8701231.
- [40] T. Shanthi and R. S. Sabeenian, “Modified Alexnet Architecture for Classification of Diabetic Retinopathy Images,” *Comput. Electr. Eng.*, vol. 76, pp. 56–64, 2019, doi: 10.1016/j.compeleceng.2019.03.004.
- [41] S. H. Kassani, P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, K. A. Schneider, and R. Deters, “Diabetic Retinopathy Classification Using a Modified Xception Architecture,” in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2019, pp. 1–6, doi: 10.1109/ISSPIT47144.2019.9001846.
- [42] A. K. Gangwar and V. Ravi, “Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning,” in *Frontiers of Intelligent Computing: Theory and Applications (FICTA)*, 2020, vol. 1176, pp. 679–689, doi: 10.1007/978-981-15-5788-0_64.
- [43] M. Mateen, J. Wen, N. Nasrullah, S. Sun, and S. Hayat, “Exudate Detection for Diabetic Retinopathy Using Pretrained Convolutional Neural Networks,” *Complexity*, vol. 2020, pp. 1–11, 2020, doi: 10.1155/2020/5801870.
- [44] G. Jinfeng, S. Qummar, Z. Junming, Y. Ruxian, and F. G. Khan, “Ensemble Framework of Deep CNNs for Diabetic Retinopathy Detection,” *Comput. Intell. Neurosci.*, vol. 2020, pp. 1–11, 2020, doi: 10.1155/2020/8864698.

- [45] Asia Pacific Tele-Ophthalmology Society (APTOS), “APTOS 2019 Blindness Detection,” 2019. <https://www.kaggle.com/c/aptos2019-blindness-detection> (accessed Aug. 10, 2021).
- [46] EyePACS, “Diabetic Retinopathy Detection,” 2015. <https://www.kaggle.com/c/diabetic-retinopathy-detection> (accessed Aug. 10, 2021).
- [47] R. Chaudhari, “Diabetic Retinopathy Dataset,” 2016. <https://www.kaggle.com/rutujachaudhari/diabetic-retinopathy> (accessed Aug. 10, 2021).





ภาคผนวก

พหุมนุ ทัศนิต ชีวะ



ภาคผนวก ก

เอกสารรับรองโครงการวิจัยในมนุษย์

พหุบัน ปณฺ ทิตโต สีเว



คณะกรรมการจริยธรรมการวิจัยในคน มหาวิทยาลัยมหาสารคาม

เอกสารรับรองโครงการวิจัย

เลขที่การรับรอง : 029-366/2565

ชื่อโครงการวิจัย (ภาษาไทย) การจำแนกรูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตาดำด้วยการเรียนรู้เชิงลึก
ชื่อโครงการวิจัย (ภาษาอังกฤษ) Classification of Diabetic Retinopathy Images using Deep Learning

ผู้วิจัย : นางสาวพิชิตา สายเชื้อ

หน่วยงานที่รับผิดชอบ : คณะวิทยาการสารสนเทศ

สถานที่ทำการวิจัย : คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

ประเภทการพิจารณาแบบ : แบบยกเว้น

วันที่รับรอง : 24 มกราคม 2565

วันหมดอายุ : 23 มกราคม 2566

ข้อเสนอการวิจัยนี้ ได้รับการพิจารณาและให้ความเห็นชอบจากคณะกรรมการจริยธรรมการวิจัยในคน มหาวิทยาลัยมหาสารคามแล้ว และอนุมัติในด้านจริยธรรมให้ดำเนินการศึกษาวิจัยเรื่องข้างต้นได้ บนพื้นฐานของโครงการงานวิจัยที่คณะกรรมการฯ ได้รับและพิจารณา เมื่อเสร็จสิ้นโครงการแล้วให้ผู้วิจัยส่งแบบฟอร์มการปิดโครงการและรายงานผลการดำเนินงานมายังคณะกรรมการจริยธรรมการวิจัยในคน มหาวิทยาลัยมหาสารคาม หรือ หากมีการเปลี่ยนแปลงใดๆ ในโครงการวิจัย ผู้วิจัยจักต้องยื่นขอรับการพิจารณาใหม่

..... ภาธร สว่างจิตร์

(ผู้ช่วยศาสตราจารย์ ภาธร สว่างจิตร์)

ประธานคณะกรรมการจริยธรรมการวิจัยในคน

มหาวิทยาลัยมหาสารคาม

ทั้งนี้ การรับรองนี้มีเงื่อนไขดังที่ระบุไว้ด้านหลังทุกข้อ (ดูด้านหลังของเอกสารรับรองโครงการวิจัย)



ภาคผนวก ข

ชุดข้อมูลรูปภาพจอประสาทตาที่ใช้ในการวิจัย

พหุบัณฑิตวิทยา

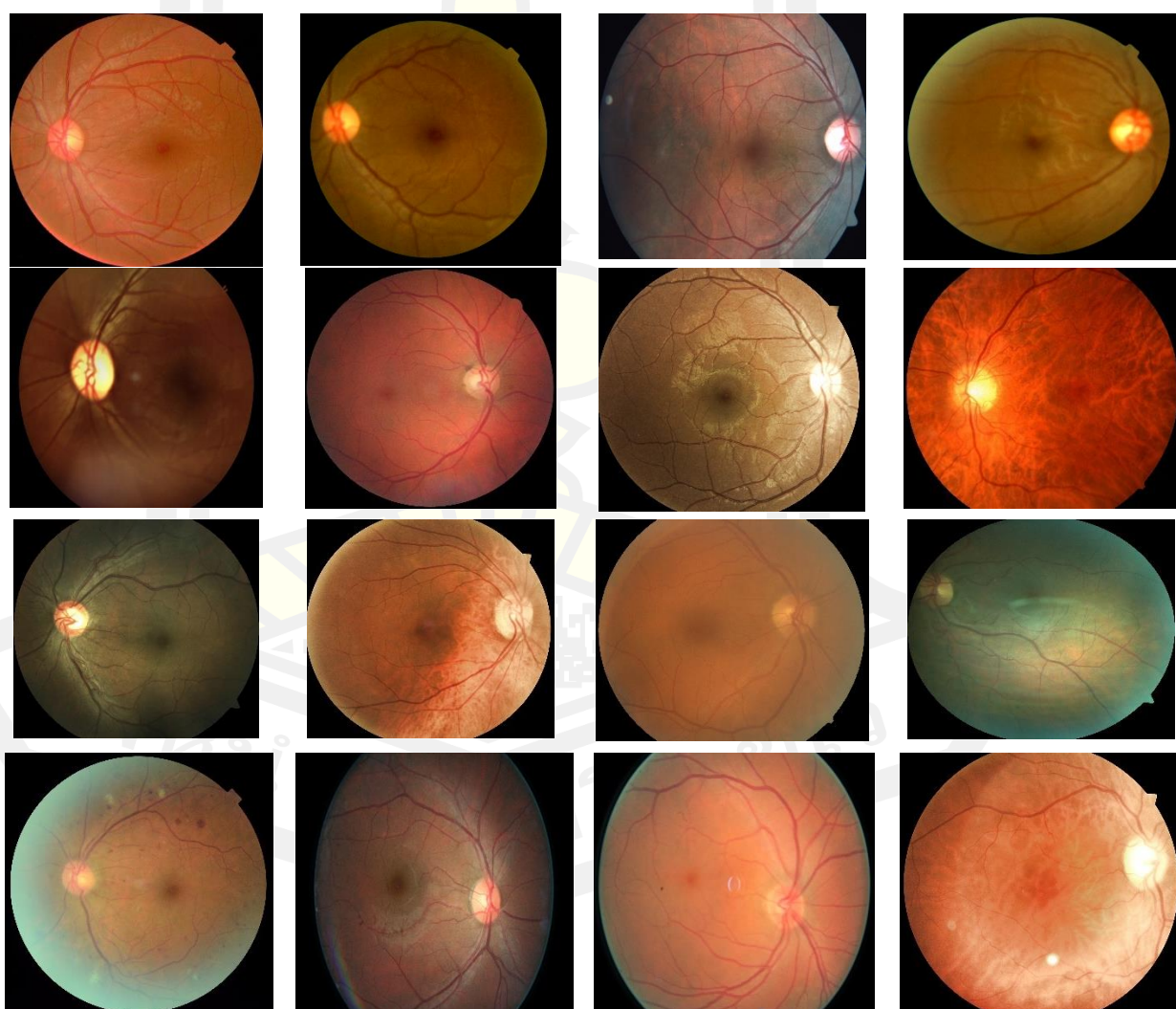
ชุดข้อมูลรูปภาพจอประสาทตาที่ใช้ในการวิจัยฉบับนี้ประกอบไปด้วยรูปภาพจอประสาทตา 2 ประเภท ได้แก่ จอประสาทตาปกติ และ จอประสาทตาที่มีภาวะเบาหวานขึ้นตา โดยผู้วิจัยได้รวบรวมมาจากฐานข้อมูลสาธารณะซึ่งไม่ต้องขออนุญาตก่อนนำมาทำการทดลองจากเว็บไซต์ Kaggle จำนวน 3 ชุด ได้แก่ APTOS 2019 Blindness Detection, Diabetic Retinopathy Detection และ Diabetic Retinopathy Dataset โดยสามารถดาวน์โหลดได้ที่

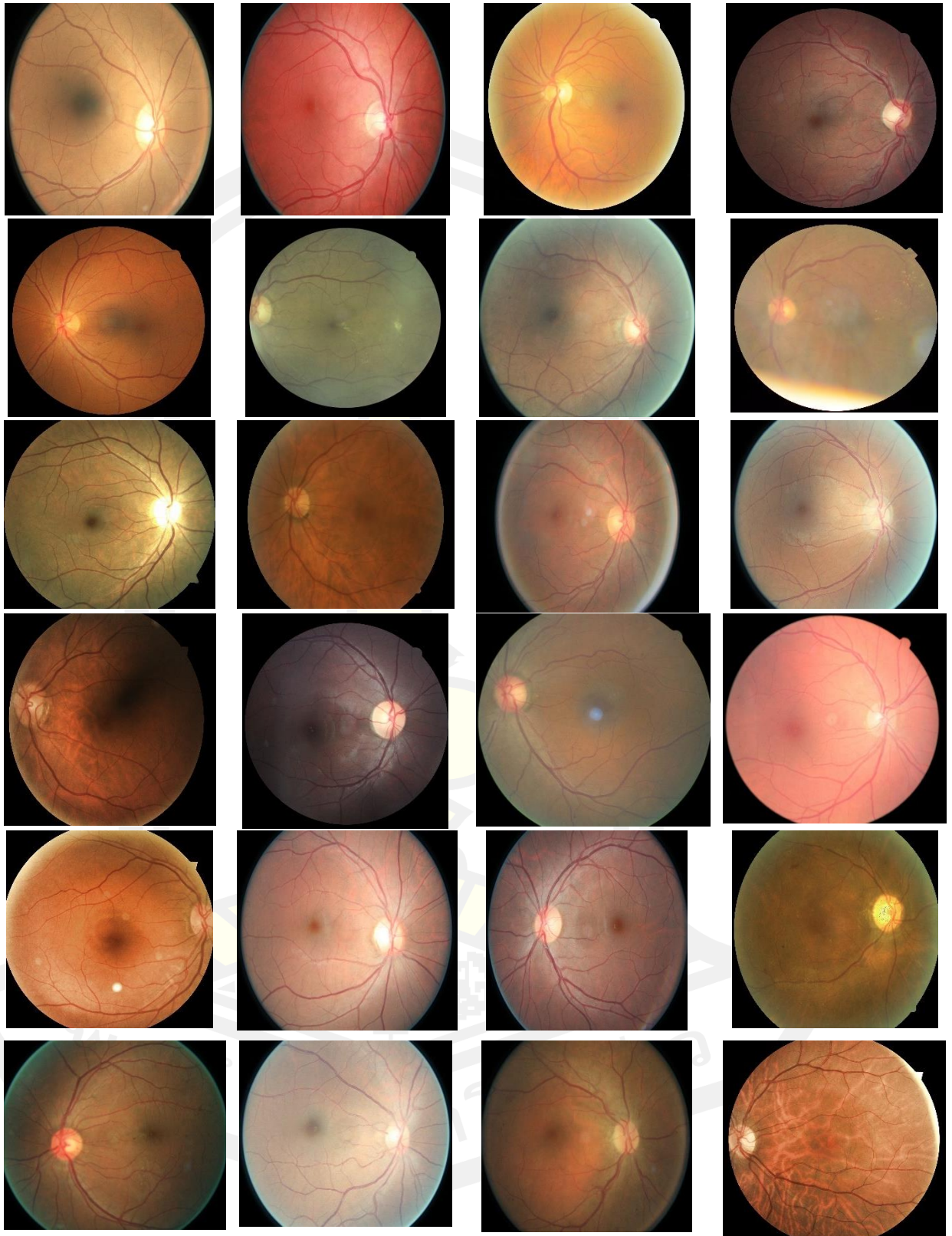
<https://www.kaggle.com/c/aptos2019-blindness-detection>

<https://www.kaggle.com/c/diabetic-retinopathy-detection>

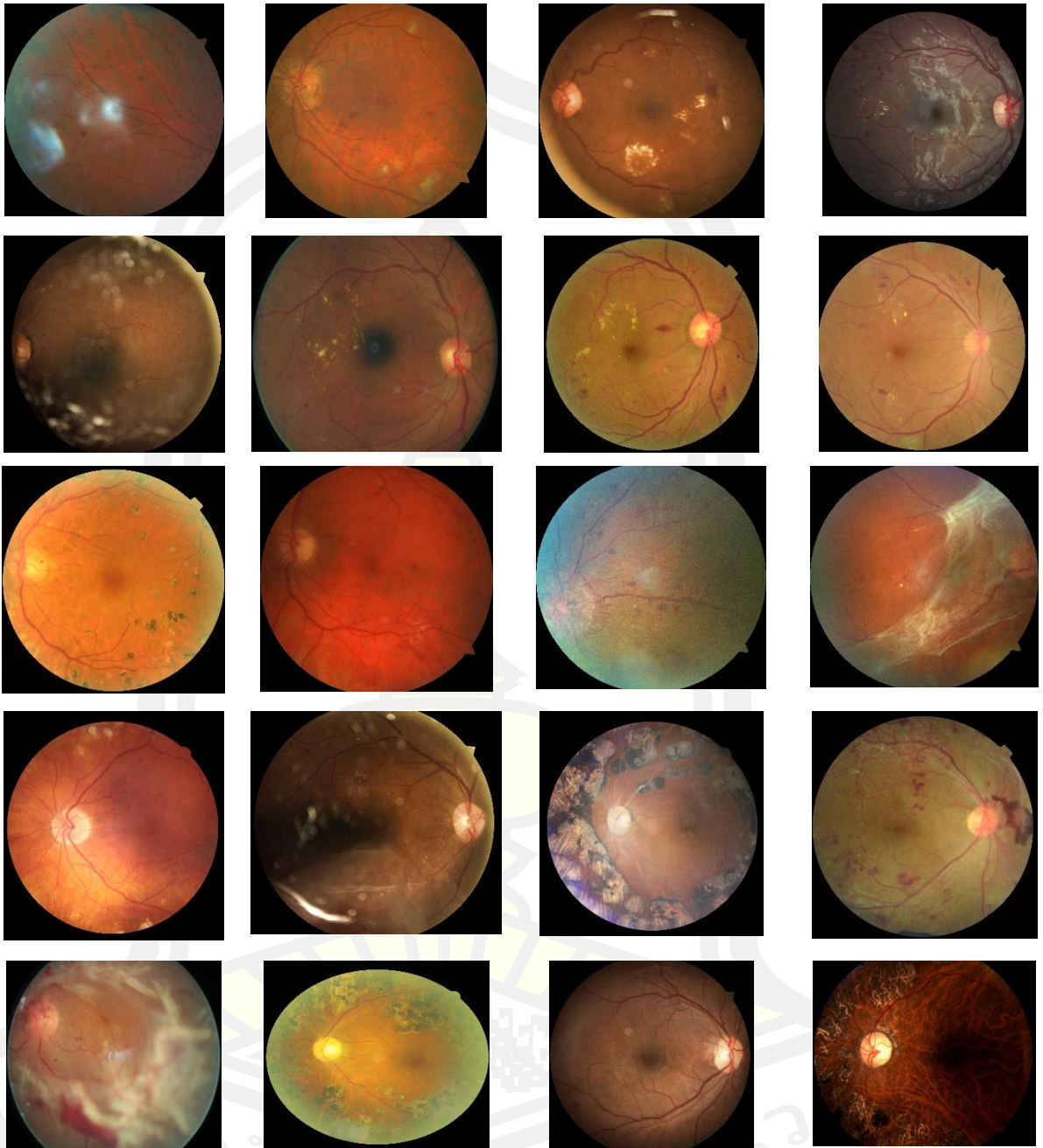
<https://www.kaggle.com/rutujachaudhari/diabetic-retinopathy>

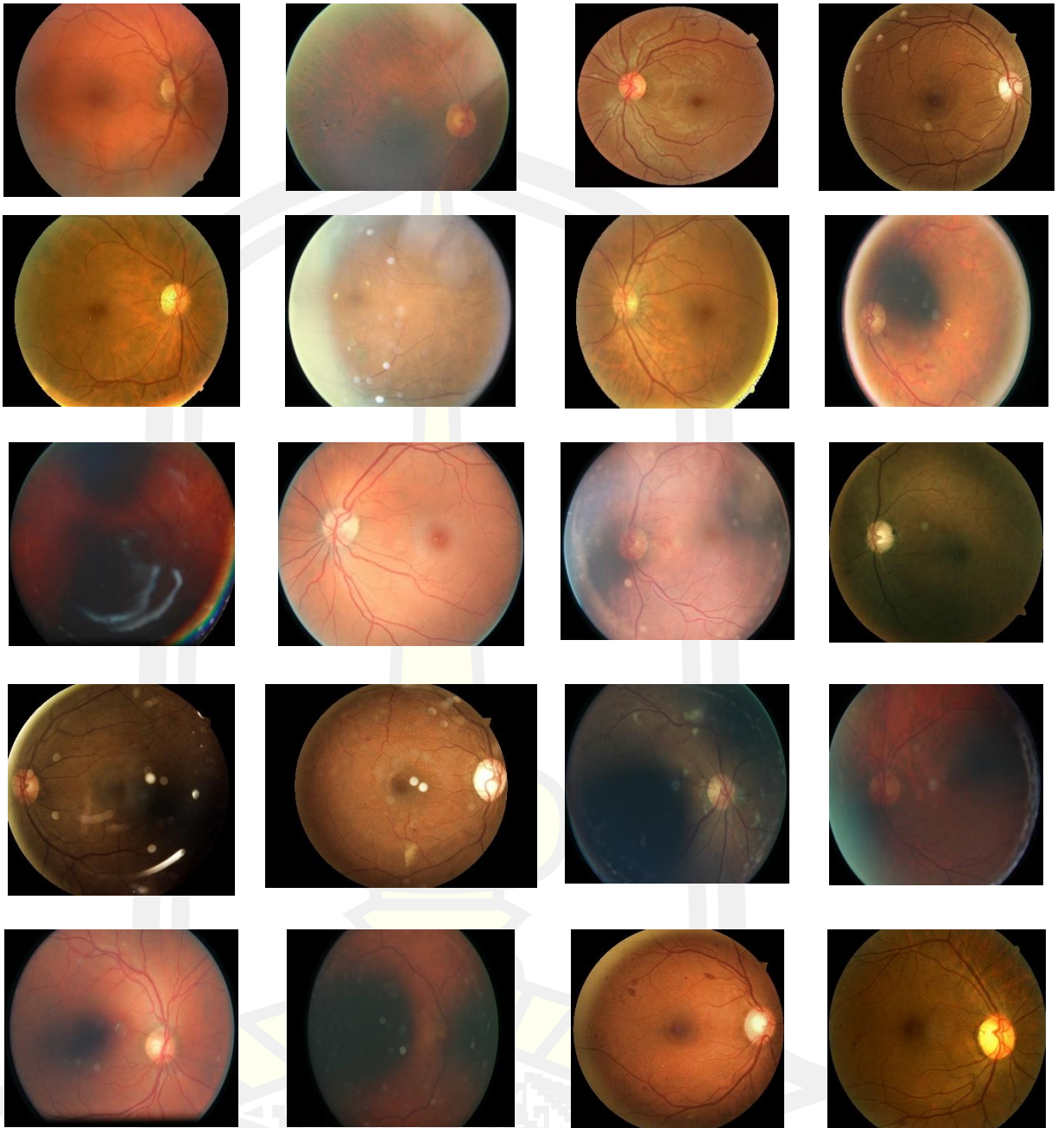
1. รูปภาพจอประสาทตาปกติ





2. รูปภาพจอประสาทตาที่มีภาวะเบาหวานขึ้นตา





พหุบัณฑิต ชีวะ



ภาคผนวก ค

บทความที่ได้เผยแพร่ในวารสารวิชาการ ICIC Express Letters

เรื่อง “Comparative Study between Ensemble and Fusion Convolutional Neural
Networks for Diabetic Retinopathy Classification”

พหุบัณฑิต โท ชีวะ

COMPARATIVE STUDY BETWEEN ENSEMBLE AND FUSION CONVOLUTIONAL NEURAL NETWORKS FOR DIABETIC RETINOPATHY CLASSIFICATION

PICHADA SAICHUA AND OLARIK SURINTA*

Multi-agent Intelligent Simulation Laboratory (MISL)
Department of Information Technology
Faculty of Informatics
Mahasarakham University

Khamriang Sub-District, Kantarawichai District, Mahasarakham 44150, Thailand
pichada.sai@msu.ac.th; *Corresponding author: olarik.s@msu.ac.th

Received August 2021; accepted October 2021

ABSTRACT. *In this paper, we have demonstrated the effectiveness of the fusion convolutional neural network (CNN) and ensemble CNN architectures for diabetic retinopathy classification. Due to the fusion and ensemble CNN architectures, we proposed to use five CNN architectures consisting of InceptionV3, ResNet50, ResNet50V2, Xception, and DenseNet121 to find the best CNN model. Two of the best CNN models were then selected for creating the fusion and ensemble CNN architectures. We also performed data augmentation techniques while training the CNN models. We found that the data augmentation technique can increase the accuracy of the CNNs. However, the data augmentation technique should not distort the retinal image. For the fusion CNNs, Xception and InceptionV3 were combined and then attached with two dense layers with the size of 1,024 units for each dense layer. Hence, we selected the optimal dropout value with 0.4. For the ensemble CNNs, the output probabilities that were calculated from the Xception and InceptionV3 models, were sent to the ensemble learning method. Using ensemble learning methods, we also compared the weighted and unweighted average methods. The results showed that the weighted average method outperformed the unweighted average method in all ensemble CNNs. From our experiments, we found that the fusion CNN architecture slightly outperformed ensemble CNN architecture.*

Keywords: Convolutional neural network, Fusion CNNs, Ensemble CNNs, Diabetic retinopathy classification, Data augmentation technique

1. Introduction. The World Health Organization (WHO) has listed diabetic retinopathy as one of the leading causes of blindness worldwide [1]. The main issue of diabetic retinopathy arises from consuming a main meal that contains much sugar. There are many foods that are commonly sold in convenience stores that have a high level of sweetness, such as beverages, sweets and coffee. If we receive more sweetness than what we need, the excess will be harmful to the human body, especially to the eyes. However, if the blood vessels in the retina begin to become inflamed and aneurysm develops, there will be lymphatic distribution throughout the retina. If untreated, this causes retinal ischemia and this leads to the cells that are used for vision being destroyed continuously. It eventually leads to decreased ability to see and maybe loss of vision.

Due to the increasing number of morbidities associated with diabetic retinopathy, there may be an insufficient number of ophthalmologists to treat patients. Here we report the development of a system that helps to detect and classify diabetic retinopathy from retinal images using deep learning methods. There are two types of diabetic retinopathy that we can classify from the retinal images: diabetic retinopathy and non-diabetic retinopathy.

DOI: 10.24507/icicel.16.04.401

Consequently, we can deeply classify diabetic retinopathy into five levels; these are normal, mild non-proliferative diabetic retinopathy (NPDR), moderate NPDR, severe NPDR, and proliferative DR [2,3].

In this research, we proposed a convolutional neural network (CNN) framework for diabetic retinopathy classification. First, we aimed to find the best CNN architectures: InceptionV3, ResNet50, ResNet50V2, Xception, and DenseNet121. Then we selected the two best CNN models to create the new CNN frameworks, called fusion CNNs and ensemble CNNs. Finally, we then compared the performance of these two CNN frameworks to classify the retinal images of diabetic retinopathy. In addition, the ensemble CNNs combine the weighted parameters from multiple CNN models that are classified using the softmax function, while the fusion CNNs combine the feature maps from multiple CNN models before sending to classify using the softmax function. The advantage of the fusion and ensemble CNN architectures reduced the generalization error and increased prediction performance [4].

2. Related Work. In this section, we survey deep learning techniques that have been proposed for recognition of diabetic retinopathy (DR) from retinal images. Yazhini and Loganathan [5] proposed a framework called integrated fusion that combined GLCM and VGG19 architecture to extract the feature vector from the retinal images. In the integrated fusion framework, the feature vector extracted by GLCM and VGG-19 was first concatenated and then sent to classify using the softmax function. The integrated fusion framework predicted the output of the DR as five levels. This integrated fusion method provided an accuracy of 71.30% and sensitivity of 50.43%.

Hattiya et al. [6] proposed to use seven CNN architectures consisting of AlexNet, ResNet50, DenseNet201, InceptionV3, MobileNet, MnasNet, and NASNetMobile for diabetic retinopathy recognition. The dataset used in the experiment was downloaded from the Kaggle website. It included 23,513 retinal images with two classes: diabetic retinopathy and non-diabetic retinopathy. They first evaluated the CNN architecture with five different color spaces: RGB, grayscale, HSV, L^*a^*b , and YCbCr. The result showed that training the CNN model with RGB color space gave the highest recognition accuracy. They then trained seven CNN architectures with the RGB color space. The AlexNet architecture achieved a high accuracy of 81.42%.

Vives-Boix and Ruiz-Fernández [7] proposed CNN architectures that updated the weighted parameters in the convolutional layer using the synaptic metaplasticity method. In their experiments, InceptionV3 with synaptic metaplasticity achieved an accuracy of 95.56% on a public small diabetic retinopathy dataset.

The CNN architectures can be combined to create a new framework, called ensemble CNNs that include the ensemble learning method in the last layer. It learns from the output probabilities of each CNN that is included in the ensemble CNNs. Chompookham and Surinta [8] proposed ensemble learning methods that were created from three and five deep CNN models. Further, the output probabilities that calculated from each CNN model were then transferred to the ensemble learning layer. The CNN model was trained using data augmentation techniques: height shift, vertical flip, and fill mode. In the experiment, the ensemble CNNs with the weighted average method were evaluated on plant leaf disease datasets. High accuracies above 99% were obtained from the ensemble CNN architecture. Also, an accuracy of 94.7% was obtained on a mulberry leaf dataset.

Noppitak and Surinta [9] proposed ensemble CNN architecture to enhance the efficiency of land use classification. First, they discovered the best CNN model from eight CNN architectures: InceptionResNetV2, MobileNetV2, DenseNet201, Xception, ResNet152V2, NASNetLarge, VGG16, and VGG19. Second, the data augmentation techniques, including rotation, width shift, and height shift were used while training the CNN models. Third, the ensemble CNN architecture was then generated with the 3 CNN models that were

found in the first step. For the ensemble learning method, the weighted average method was proposed. Hence, the grid-search method was proposed to optimize the weighted parameters. The experimental results showed that the ensemble CNN architecture achieved an accuracy of 92.80%.

Deepa et al. [10] proposed a multi-stage deep CNN to learn to distinguish the diabetic retinopathy image from the whole image and random patches as the input images. First, the input images were sent to the CNN architectures: InceptionV3 and Xception. Second, the probability vectors from the CNN architectures were then classified using artificial neural networks (ANNs). Third, in the ensemble classifier process, the outputs of the ANNs were classified using a support vector machine classifier. As a result, their proposed method showed an accuracy of 96.2%.

3. Methodology. This paper reports on the objective to improve the efficacy of CNN frameworks by applying various optimization algorithms to reducing training losses. Moreover, the comparative study methods are presented based on two CNN frameworks: 1) fusion CNNs and 2) ensemble CNNs.

3.1. Convolutional neural network architectures.

InceptionV3. InceptionV3 was proposed by Szegedy et al. [11] in 2016. InceptionV3 was modified from the previous inception architecture and focused on providing less computational cost. In the InceptionV3 architecture, first, the factorized convolutional layers and the small convolutional layers were proposed to reduce the number of parameters involved in a network and also reduce the computation cost. Second, the symmetric convolutions were replaced by asymmetric convolutions. Next, an auxiliary classifier was proposed as the regularizer. Finally, to avoid a representational bottleneck, the grid size reduction of the feature maps was proposed.

ResNet50 and ResNet50V2. ResNet and ResNetV2 were proposed by He et al. [12, 13] in 2016. In very deep networks, the numbers of stacked layers were increased. The deep networks showed the high accuracy results on the challenging ImageNet dataset. ResNet50 and ResNet50V2 include 50 parameter layers. In ResNet architecture, residual learning was proposed to allow the stacked network to jump over one or more building blocks, called shortcut connections. In ResNetV2, a new residual unit was proposed. It was shown that the new residual unit decreased the error while training around 2%.

Xception. Chollet invented an extreme version of Inception architecture, called Xception [14] in 2017. It focused on modifying the depthwise separable convolution layer, namely a depthwise convolution. In the modified depthwise convolution, the order of the depthwise convolution was pointwise convolution and then followed by a depthwise convolution. So, the number of connections is fewer and the model is lighter. As a result, the Xception architecture showed improvement in accuracy performance when compared with the InceptionV3.

DenseNet121. In 2017, Huang et al. [15] proposed densely connected convolutional networks, called DenseNet. DenseNet is like the feed-forward architecture where each dense block layer connects to all other dense block layers. For example, the first dense block layer is connected to the 2nd, 3rd, and so on until the last dense block layer. The second dense block layer is also connected to the 3rd block and so on until the last layer. If considered, the connection of the 2nd layer, the feature maps of the 1st layer and the feature maps of the 2nd layer were combined and used as inputs into the 3rd layer.

3.2. Fusion CNNs. The proposed fusion CNNs comprise two parts. The two best CNN architectures, that were selected from Section 3.1, are chosen to create temporal features, called feature maps. We then concatenated feature maps before sending them to another layer. These feature maps were first transferred to the batch normalization layer (BN). Second, the rectified linear unit (ReLU) was proposed as a nonlinear function followed

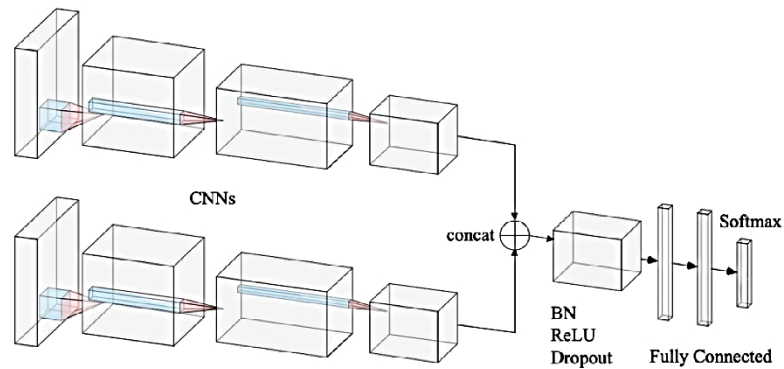


FIGURE 1. Illustration of the fusion CNN architecture

by the dropout layer to avoid overfitting. Finally, two fully connected layers and softmax activation function were attached to the network. The architecture of fusion CNNs is illustrated in Figure 1.

3.3. Ensemble CNNs. The ensemble CNN architecture consists of two parts. In the first part, we trained and evaluated five CNN architectures, including InceptionV3, ResNet50, ResNet50V2, Xception, and DenseNet121. After that, we chose only two CNN architectures to create the ensemble CNN framework. In the second part, the output probabilities (w) from each CNN were then calculated using ensemble learning methods [8,9]. Figure 2 illustrates the proposed ensemble CNN architecture. The ensemble learning methods are described as follows.

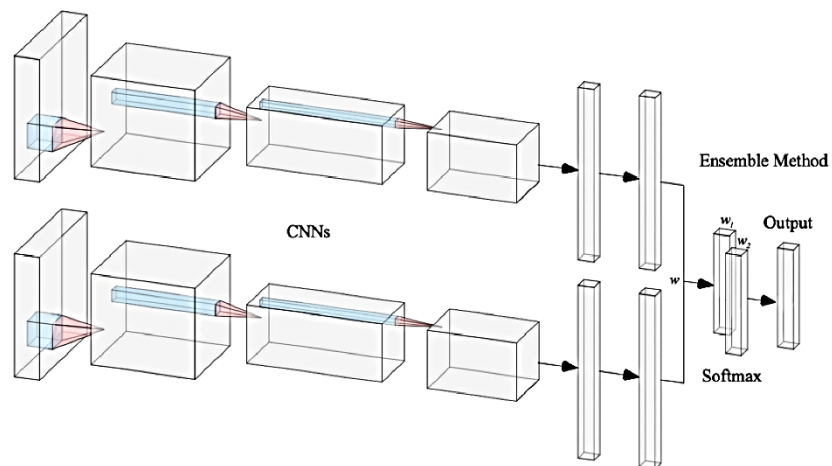


FIGURE 2. Illustration of the ensemble CNN architecture

Weighted average method. In the weighted average method, the different weight parameters are calculated with the output probabilities. We decided to calculate the higher weight with the output probabilities of the CNN model that achieved a higher classification rate. Note that the sum of all weight parameters is equal to one. The equation

of the weighted average method is given by $\hat{y} = \frac{1}{n} \sum_{i=1}^n \alpha \vec{y}_i$, where α is the weight value that multiplies with the vector of output probabilities (\vec{y}) and n is the number of ensemble CNN models.

Unweighted average method. In this learning method, the output probabilities of each CNN model are computed by average of the probability values. Then, the maximum value of the probabilities is selected as an output of the ensemble learning. The unweighted average method is calculated by $\hat{y} = \frac{1}{n} \sum_{i=1}^n \vec{y}_i$, where \vec{y} is the vector of the output probabilities of each CNN model and n is the number of ensemble CNN models. We then used the arg max function to select the highest probability value of \hat{y} .

4. Experimental Setup and Results. All experiments were evaluated in the same environment. We used the TensorFlow v2.5.0 as the deep learning framework that runs on Google Colab platform.

4.1. Diabetic retinopathy dataset. We collected retinal images from various DR datasets that were available on the Kaggle website, including APTOS 2019 blindness detection, diabetic retinopathy detection, and diabetic retinopathy. These datasets were collected by a collaboration between the Aravind Eye Hospital and Kaggle website. The hospital screened for diabetic retinopathy of the patient from the retinal images and then created a label for each image [2,3].

The retinal images used in this study were stored in RGB color space with a JPEG format. The DR dataset included 2 classes and contained 23,510 images. The number of diabetic retinopathy and non-diabetic retinopathy images was 12,063 and 11,447 images, respectively. We then divided the DR dataset into a training set (18,808 images) and test set (4,702 images). Some examples of the DR dataset are shown in Figure 3.

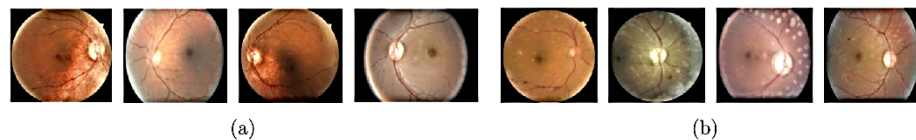


FIGURE 3. Illustration of the retinal images from the diabetic retinopathy dataset: (a) Non-diabetic retinopathy class and (b) diabetic retinopathy class

4.2. Experiments with CNNs and data augmentation techniques. In this study, data augmentation techniques [16] were proposed. The data augmentation techniques used in the experiments were flip (horizontal and vertical flips), rotation (randomly with value between 0-90°), and zoom (randomly with value between [1-0.2, 1+0.2]) techniques. We proposed these data augmentation techniques because they did not distort the retinal images. Examples of the data augmentation techniques are shown in Figure 4.

We evaluated the performance of the CNN architectures using several factors as follows: five CNN architectures (InceptionV3, ResNet50, ResNet50V2, Xception, and DenseNet121), five optimization algorithms (SGD, Adam, Adadelata, Adagrad and Adamax), and learning rates (0.1, 0.01, 0.001, and 0.0001). In addition, our experiments evaluated several optimization algorithms because the fittest optimizer assures more reliable accuracy performance [8].

We show the results obtained with the CNN architecture with the best optimal parameters on the DR dataset in Table 1. It can be seen that Xception was the best CNN architecture in our experiments. The Xception architecture achieved an accuracy of 84.07%. It slightly outperformed InceptionV3 by about 0.6%. However, it spent more computation time on the training.

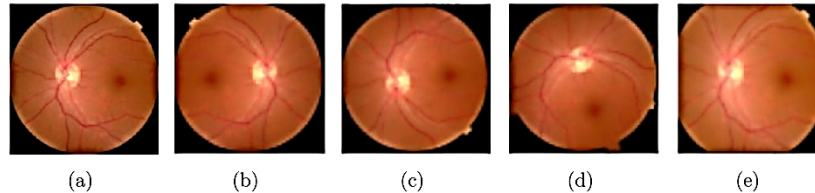


FIGURE 4. Illustration of the data augmentation techniques, including (a) original image, (b) horizontal flip, (c) vertical flip, (d) rotation, and (e) zoom

TABLE 1. The best optimizers and the accuracy (%) of each CNN model

CNNs	Image size	Optimizers	Learning rate	Training time	Test	
					Accuracy (%)	Loss
Xception	299×299	Adamax	0.001	57 min	84.07	0.38
InceptionV3	299×299	Adam	0.001	38 min	83.30	0.38
DenseNet121	224×224	Adadelata	0.1	33 min	82.77	0.4
ResNet50V2	224×224	Adamax	0.001	37 min	82.20	0.38
ResNet50	224×224	Adam	0.0001	38 min	81.28	0.41

The experimental results of the data augmentation techniques are shown in Table 2. It is important to emphasize that the data augmentation techniques can improve the performance of the classification of diabetic retinopathy. Xception combined with the rotation technique still outperformed all CNN architectures. It achieved 85.50% and was slightly better than training without data augmentation technique. It is quite surprising that it requires less computation time when compared with training without data augmentation techniques.

TABLE 2. The optimal data augmentation techniques of each CNN model

CNNs	Data augmentation techniques	Training time	Test	
			Accuracy (%)	Loss
Xception	Rotation	44 min	85.50	0.34
InceptionV3	Zoom	38 min	84.69	0.37
DenseNet121	Rotation	26 min	84.01	0.35
ResNet50	Flip	40 min	83.26	0.41
ResNet50V2	Flip	21 min	83.18	0.39

To summarize the results, the Xception architecture without data augmentation obtained the accuracy of 84.07%, outperforming all CNN architectures, except only the InceptionV3 with zoom technique that achieved the accuracy of 84.69%.

4.3. Experiments with fusion CNN architecture. In the experiment of fusion CNNs, we combined the Xception architecture with other CNNs, because we found in previous experiments that the Xception architecture provided the highest accuracy with lower loss value. The compared results of each fusion CNN are shown in Table 3.

From Table 3, our results show that the fusion CNNs between Xception+InceptionV3 obtained 86.30% accuracy with the loss value of 0.33. This fusion CNNs also required to be fully connected with two dense layers with only 1024 units of each dense layer. We can emphasize that the Xception architecture when combined with the other CNNs, still obtained an accuracy above 85%.

TABLE 3. Performance evaluation of the fusion CNNs

Fusion CNNs	Dense sizes	No. of dense layers	Dropout	Training time	Test	
					Accuracy (%)	Loss
Xception+InceptionV3	1024	2	0.4	53 min	86.30	0.33
Xception+DenseNet121	2048	2	0.2	33 min	85.45	0.34
Xception+ResNet50V2	4096	1	No	30 min	85.11	0.36
Xception+ResNet50	1024	1	0.1	32 min	85.07	0.4

4.4. Experiments with ensemble CNN architecture. For the ensemble CNNs experiments, the accuracy results of the ensemble CNNs are shown in Table 4. The performance of the ensemble CNNs was greater than the single CNN architecture. However, only ensemble CNNs between ResNet50 and ResNet50V2 (83.73%) performed with lower accuracy than the Xception architecture (84.07%). As a result, the ensemble CNNs between Xception and InceptionV3 provided an accuracy of 86.11%. When we compared the ensemble CNNs and the fusion CNNs, the fusion CNN architecture slightly outperformed the ensemble CNNs.

TABLE 4. Performance evaluation of the ensemble CNNs

Ensemble CNNs	Accuracy (%) of ensemble learning methods		
	Unweighted	Weighted	Weight
	average method	average method	parameters
Xception+InceptionV3	85.92	86.11	0.6, 0.4
ResNet50+DenseNet121	84.30	84.81	0.3, 0.7
ResNet50V2+DenseNet121	84.22	84.39	0.3, 0.7
ResNet50+ResNet50V2	83.65	83.73	0.3, 0.7

5. Conclusions. In this paper, we evaluated two convolutional neural network (CNN) frameworks: fusion CNN and ensemble CNN architectures, for diabetic retinopathy classification. We discover the best CNN architecture from five CNN architectures: Inception-V3, ResNet50, ResNet50V2, Xception, and DenseNet121. To optimize the parameters, we considered five optimization algorithms: SGD, Adam, Adadelta, Adagrad and Adamax. The various learning rates between 0.1 and 0.0001 were also evaluated. In the course of the training policy, the data augmentation techniques were also performed during the training, including flip, rotation, and zoom techniques. The result showed that the Xception architecture with the Adamax optimizer and a learning rate of 0.001 achieved the best accuracy performance. Interestingly, the data augmentation techniques can increase the accuracy of every CNN architecture. The fusion and the ensemble CNN architectures were compared. We also found that the combination between Xception and InceptionV3 architectures performed very well on both architectures. From the experimental results, we conclude that the performance of the fusion CNN architecture was slightly better than the ensemble CNN architecture.

In future work, to improve the performance of diabetic retinopathy classification, we will concentrate on experiments with the other CNN frameworks, such as snapshot ensemble CNN [17], Siamese network [18], and Hybrid CNN [19].

Acknowledgment. This research project was financially supported by Mahasarakham University.

REFERENCES

- [1] Y. Shimizu, Eye care, vision care, vision impairment and blindness, *World Health Organization (WHO)*, <https://www.who.int/health-topics/blindness-and-vision-loss>, Accessed on Jun. 06, 2021.

- [2] D. Doshi, A. Shenoy, D. Sidhpura and P. Gharpure, Diabetic retinopathy detection using deep convolutional neural networks, *International Conference on Computing, Analytics and Security Trends (CAST)*, pp.261-266, 2016.
- [3] S. Burewar, A. B. Gonde and S. K. Vipparthi, Diabetic retinopathy detection by retinal segmentation with region merging using CNN, *The 13th International Conference on Industrial and Information Systems (ICIIS)*, pp.136-142, 2018.
- [4] M. A. Ganaie, M. Hu, M. Tanveer and P. N. Suganthan, Ensemble deep learning: A review, *arXiv.org*, arXiv: 2104.02395, 2021.
- [5] K. Yazhini and D. Loganathan, An integrated fusion based feature extraction and classification model for diabetic retinopathy diagnosis, *The 2nd International Conference on Inwention Research in Computing Applications (ICIRCA)*, pp.1187-1193, 2020.
- [6] T. Hattiya, K. Dittakan and S. Musikasuwani, Diabetic retinopathy detection using convolutional neural network: A comparative study on different architectures, *Mahasarakham Int. J. Eng. Technol.*, vol.7, no.1, pp.50-60, 2021.
- [7] V. Vives-Boix and D. Ruiz-Fernández, Diabetic retinopathy detection through convolutional neural networks with synaptic metaplasticity, *Comput. Methods Programs Biomed.*, vol.206, pp.1-8, 2021.
- [8] T. Chompookham and O. Surinta, Ensemble methods with deep convolutional neural networks for plant leaf recognition, *ICIC Express Letters*, vol.15, no.6, pp.553-565, 2021.
- [9] S. Noppitak and O. Surinta, Ensemble convolutional neural network architectures for land use classification in economic crops aerial images, *ICIC Express Letters*, vol.15, no.6, pp.531-543, 2021.
- [10] V. Deepa, C. S. Kumar and T. Cherian, Ensemble of multi-stage deep convolutional neural networks for automated grading of diabetic retinopathy using image patches, *J. King Saud Univ. - Comput. Inf. Sci.*, vol.33, no.6, DOI: 10.1016/j.jksuci.2021.05.009, 2021.
- [11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, Rethinking the inception architecture for computer vision, *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2818-2826, 2016.
- [12] K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for image recognition, *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770-778, 2015.
- [13] K. He, X. Zhang, S. Ren and J. Sun, Identity mappings in deep residual networks, in *Computer Vision - ECCV 2016. ECCV 2016. Lecture Notes in Computer Science*, B. Leibe, J. Matas, N. Sebe and M. Welling (eds.), Cham, Springer, 2016.
- [14] F. Chollet, Xception: Deep learning with depthwise separable convolutions, *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1251-1258, 2017.
- [15] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, Densely connected convolutional networks, *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.4700-4708, 2017.
- [16] C. Shorten and T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J. Big Data*, vol.6, no.1, 2019.
- [17] G. Huang, Y. Li, G. Pleiss, Z. Liu, J. E. Hopcroft and K. Q. Weinberger, Snapshot ensembles: Train 1, get M for free, *The 5th International Conference on Learning Representations (ICLR)*, pp.1-14, 2017.
- [18] X. Zeng, H. Chen, Y. Luo and W. Ye, Automated diabetic retinopathy detection based on binocular Siamese-like convolutional neural network, *IEEE Access*, vol.7, pp.30744-30753, 2019.
- [19] Y. T. Hafiyah, Afiahayati, R. D. Yanuarieska, E. Anarossi, V. M. Sutanto, J. Triyanto and Y. Sakakibara, A hybrid convolutional neural network-extreme learning machine with augmented dataset for DNA damage classification using comet assay from buccal mucosa sample, *International Journal of Innovative Computing, Information and Control*, vol.17, no.4, pp.1191-1201, 2021.

ประวัติผู้เขียน

ชื่อ	นางสาวพิชาดา สายเชื้อ
วันเกิด	29 ตุลาคม 2540
สถานที่เกิด	จังหวัดร้อยเอ็ด
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 8 หมู่ 10 ตำบลสระนกแก้ว อำเภอโพนทอง จังหวัดร้อยเอ็ด 45110
ตำแหน่งหน้าที่การงาน	ลูกจ้างชั่วคราว
สถานที่ทำงานปัจจุบัน	คณะแพทยศาสตร์ มหาวิทยาลัยมหาสารคาม
ประวัติการศึกษา	พ.ศ. 2557 โรงเรียนสาธิตมหาวิทยาลัยมหาสารคาม (ฝ่ายมัธยม) พ.ศ. 2562 ปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม พ.ศ. 2565 ปริญญาวิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม
ผลงานวิจัย	1) P. Saichua, S. Khunthi, and T. Chomsiri (2019). “Design of Blockchain Lottery for Thai Government” in International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON), pp. 9–12. 2) S. Khunthi, P. Saichua, and O. Surinta (2019). “Effective Face Verification Systems Based on the Histogram of Oriented Gradients and Deep Learning Techniques” in International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP), pp. 1–6.

พหุบัณฑิต ชีวะ