การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับการรู้จำพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ

วิทยานิพนธ์

ของ

จักรินทร์ สนุกแสน

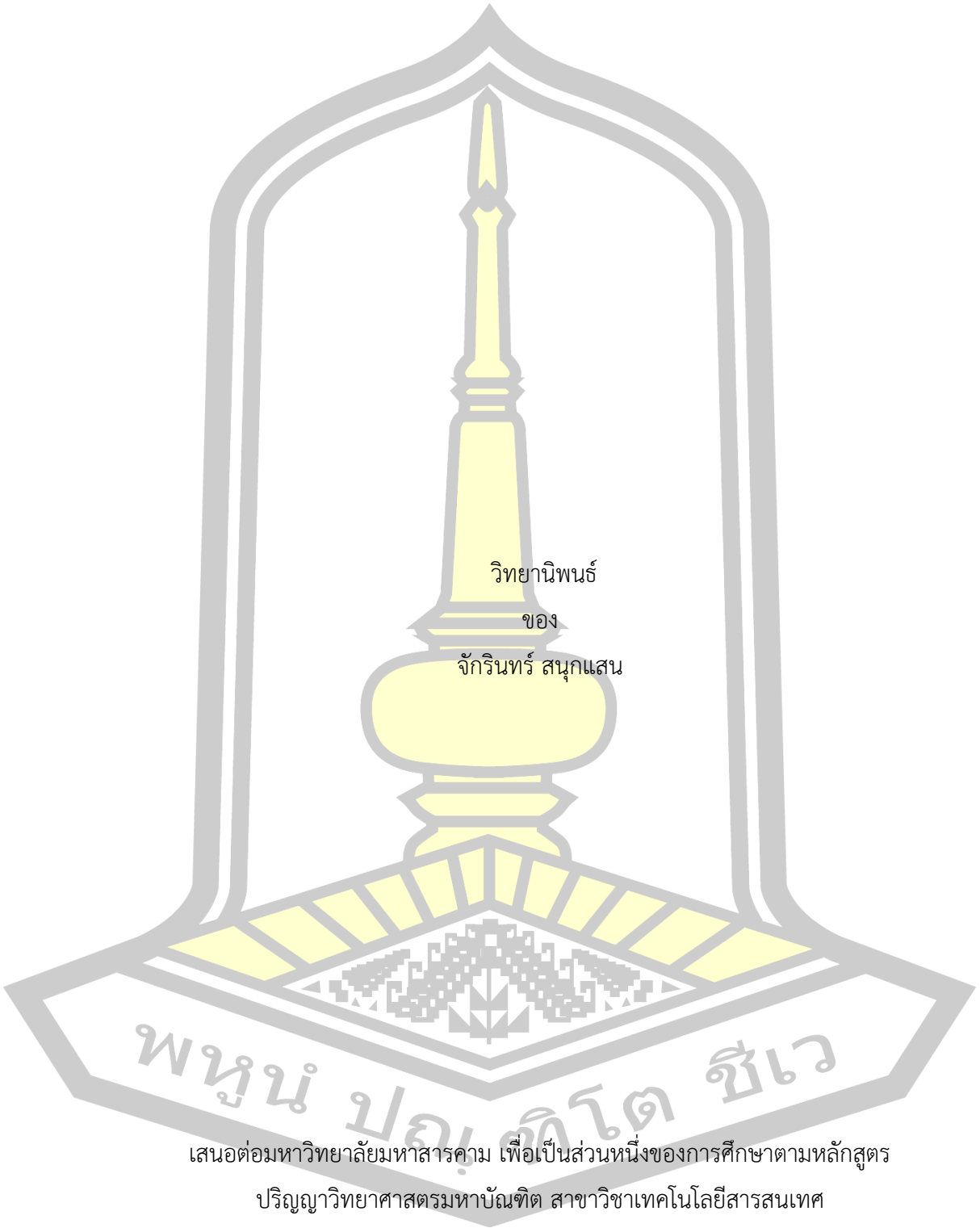
เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

สิงหาคม 2562

สงวนลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับการรู้จำพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ



วิทยานิพนธ์

ของ

จักรินทร์ สนุกแสน

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

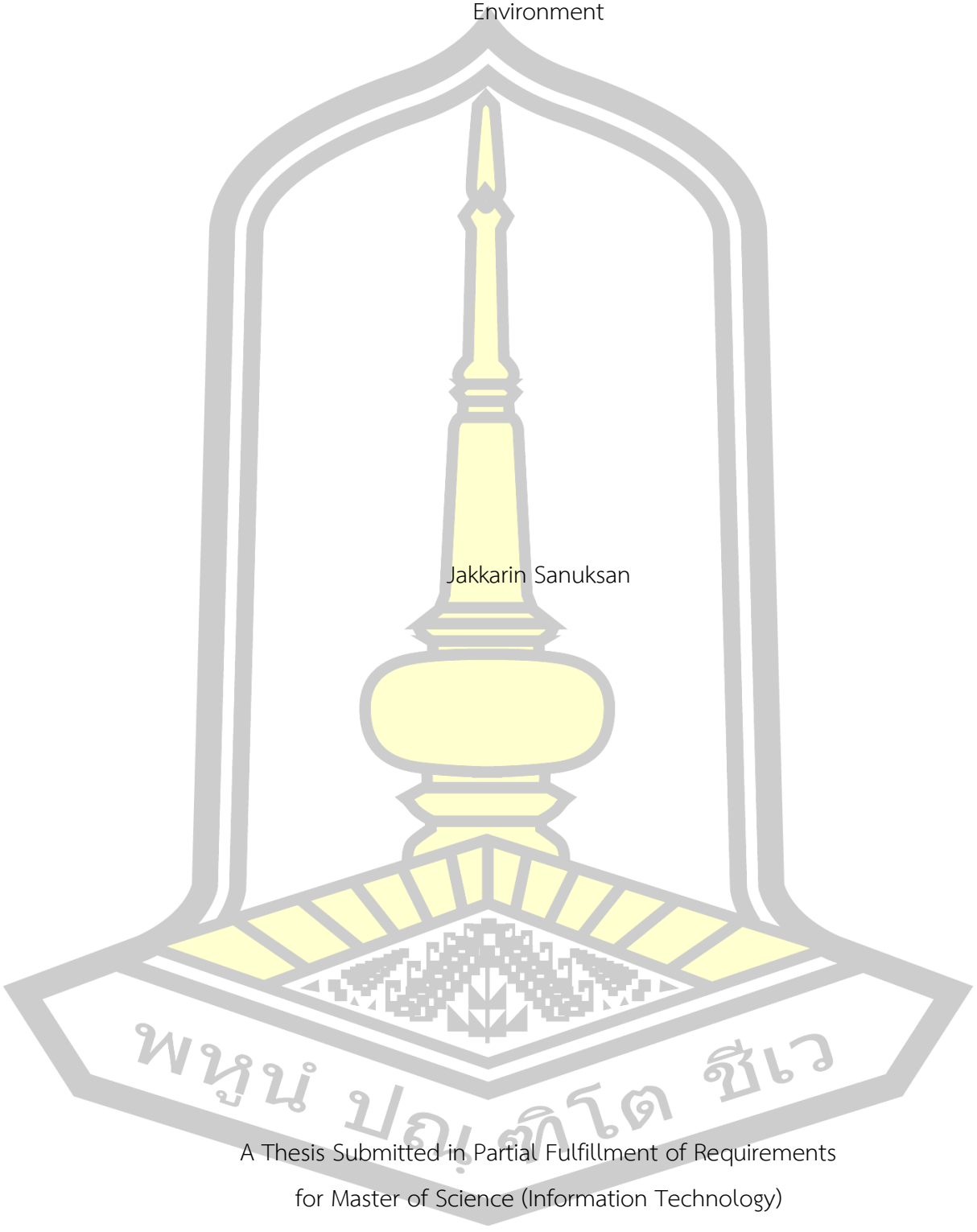
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

สิงหาคม 2562

สงวนลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Applying of Deep Learning Techniques for Plant Image Recognition in Natural Environment

Jakkarin Sanuksan



A Thesis Submitted in Partial Fulfillment of Requirements

for Master of Science (Information Technology)

August 2019

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายจักรินทร์ สนุกแสน แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(ผศ. ดร. ฉัตรตระกูล สมบัติธีระ)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผศ. ดร. โอฬาริก สุรินต๊ะ)

.....กรรมการ
(ดร. สาธิต แสงประดิษฐ์)

.....กรรมการผู้ทรงคุณวุฒิภายนอก
(รศ. ดร. สิทธิชัย บุษหมั่น)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

.....
(ผศ. ศศิธร แก้วมั่น)
คณบดีคณะวิทยาการสารสนเทศ

.....
(ผศ. ดร. กริสน์ ชัยมูล)
คณบดีบัณฑิตวิทยาลัย

พหุน ปณฺ ฑิโต ชีเว

ชื่อเรื่อง การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกสำหรับการรู้จำพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ

ผู้วิจัย จักรินทร์ สนุกแสน

อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. โอฬาริก สุรินต๊ะ

ปริญญา วิทยาศาสตรมหาบัณฑิต **สาขาวิชา** เทคโนโลยีสารสนเทศ

มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2562

บทคัดย่อ

งานวิจัยฉบับนี้ได้ศึกษาโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกเพื่อใช้สำหรับการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ โดยมีจุดประสงค์เพื่อเปรียบเทียบโครงข่ายประสาทเทียมแบบคอนโวลูชันจำนวน 4 โครงสร้าง ประกอบด้วย LeNet-5, AlexNet, GoogLeNet และ VGGNet ข้อมูลพรรณไม้ที่นำมาใช้ในการทดสอบมีจำนวนทั้งสิ้น 3 ชุดข้อมูล คือ PNE, 102 Flower และ Folio ทั้งนี้ชุดข้อมูล PNE และ 102 Flower เป็นรูปภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติทำให้มีพื้นที่พื้นหลังซับซ้อน สำหรับข้อมูลชุด Folio เป็นรูปภาพใบไม้ที่ถ่ายในห้องทดลองโดยกำหนดให้พื้นหลังของภาพเป็นสีขาว จากผลการเปรียบเทียบระหว่างโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกโดยใช้โครงสร้าง GoogLeNet และ VGGNet พบว่าโครงสร้างแบบ GoogLeNet มีประสิทธิภาพสูงสุดในชุดข้อมูล PNE และ 102 Flower และยังใช้เวลาในการเรียนรู้ที่เร็วกว่าเมื่อเทียบกับโครงสร้างแบบ VGGNet โดยใช้จำนวนรอบในการเรียนรู้ 10,000 รอบ แต่ทั้งนี้โครงสร้างแบบ VGGNet มีอัตราความถูกต้องสูงที่สุดในชุดข้อมูล Folio ซึ่งเป็นรูปภาพที่ถ่ายในห้องทดลองที่มีพื้นหลังเป็นสีขาว และใช้จำนวนรอบในการเรียนรู้เพียง 1,000 รอบ จึงสรุปได้ว่าหากต้องการที่จะสร้างโมเดลเพื่อนำไปใช้ในการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติควรใช้โครงสร้างแบบ GoogLeNet

คำสำคัญ : การรู้จำพรรณไม้, การเรียนรู้เชิงลึก, โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก, โครงสร้างแบบ AlexNet, โครงสร้างแบบ GoogLeNet, โครงสร้างแบบ VGGNet

TITLE Applying of Deep Learning Techniques for Plant Image Recognition in Natural Environment

AUTHOR Jakkarin Sanuksan

ADVISORS Assistant Professor Olarik Surinta , Ph.D.

DEGREE Master of Science **MAJOR** Information Technology

UNIVERSITY Mahasarakham University **YEAR** 2019

ABSTRACT

This paper examines a deep convolutional neural network (Deep CNN) for plant recognition in a natural environment. The primary objective is comparing 4 CNN architectures includes LeNet-5, AlexNet, GoogLeNet, and VGGNet on three plant datasets; PNE, 102 Flower, and Folio. The images in the PNE and 102 Flower dataset can be with a complicated background because it takes from a natural environment. On the other hand, the images in the Folid dataset are only leaf images that take from the laboratory environment using the white background. The comparison of deep CNN using GoogLeNet and VGGNet Architecture show that the GoogLeNet outperform while performing on the PNE and 102 Flower dataset when training time with iterations of 10,000 epochs. Consequently, the GoogLeNet also faster than the VGGNet architecture. However, the experiment showed that the VGGNet architecture outperforms the other CNN architecture on the Folio dataset and use only 1,000 epochs for training. In our experiment, we can create a model from the deep CNN using GoogleNet architecture, and this is because it showed the better result with the plant images that took from the natural environment.

Keyword : Plant Recognition, Deep Learning, Deep Convolutional Neural Network, AlexNet Architecture, GoogLeNet Architecture, VGGNet Architecture

กิตติกรรมประกาศ

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.โอฬาริก สุรินต๊ะ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่เสียสละเวลาช่วยเหลือ ให้คำปรึกษา คำแนะนำ แนวคิด ชี้แนะข้อบกพร่องและร่วมแก้ไขปัญหา ติดตามความก้าวหน้าของงานวิจัย รวมทั้งฝึกฝนให้ผู้วิจัยมีทักษะทางด้านการคิด การอ่าน การเขียนและการนำเสนอผลงานทางวิชาการ ซึ่งเป็นประโยชน์อย่างมากในการพัฒนาตนเอง อีกทั้งให้ความเอาใจใส่อย่างสูง ทำให้ผู้วิจัยสามารถดำเนินงานวิจัยจนประสบผลสำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์พัชราภรณ์ แสงโยจารย์ ผู้เชี่ยวชาญด้านพฤกศาสตร์ประจำโครงการอนุรักษ์พันธุกรรมพืชอันเนื่องมาจากพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี (อพ.สธ.) มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ ที่ให้ความช่วยเหลือในการตรวจสอบ และระบุชนิดของพรรณไม้ที่ใช้ในการทำวิจัย

ขอขอบคุณทุนสนับสนุนการศึกษาบุคลากรสายสนับสนุนจากมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์

สุดท้ายนี้ผู้วิจัยขอกราบขอบพระคุณบิดามารดาและครอบครัวของข้าพเจ้าที่ได้ให้ชีวิตและโอกาสทางการศึกษา คอยเป็นกำลังใจและให้ความห่วงใยเสมอมา ตลอดจนคุณครูและอาจารย์ทุกท่านที่กรุณาประสิทธิ์ประสาทวิชาความรู้อันเป็นประโยชน์แก่ผู้วิจัย คุณค่าและประโยชน์อันพึงมาจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่ผู้มีพระคุณทุกท่าน

จักรินทร์ สนุกแสน

พหูน ปณฺฑิโต ชีเว

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง
บทคัดย่อภาษาอังกฤษ ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฌ
สารบัญภาพ ..... ญ
บทที่ 1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของการวิจัย 1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการวิจัย 2
1.3 ขอบเขตในการวิจัย 2
1.4 ผลที่คาดว่าจะได้รับจากการวิจัย 2
1.5 นิยามศัพท์เฉพาะ 3
1.6 ข้อจำกัดของการวิจัย 3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4
2.1 ทฤษฎีทางด้านพฤกษศาสตร์ (Botany theory) 4
2.2 คอมพิวเตอร์วิทัศน์ (Computer Vision) 6
2.3 การเรียนรู้เชิงลึก (Deep Learning) 8
2.4 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) 9
2.5 การวัดประสิทธิภาพของโมเดล (Performance Measurement) 19
2.6 งานวิจัยที่เกี่ยวข้อง (Literature Review) 20
บทที่ 3 วิธีดำเนินการวิจัย 24

3.1 ชุดข้อมูลรูปภาพที่ใช้การวิจัย 24
3.2 การทดลองจำแนกพรรณไม้โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน 27
3.3 การทดสอบประสิทธิภาพ 27
บทที่ 4 ผลการศึกษา 29
4.1 เครื่องมือและชุดข้อมูลที่ใช้ในการทดลอง 29
4.2 วิธีการทดลอง 30
4.3 ผลการจำแนกพรรณไม้ด้วยวิธีการเรียนรู้เชิงลึก 31
4.4 ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม 54
บทที่ 5 สรุปผลและข้อเสนอแนะ 57
5.1 สรุปผลการวิจัย 57
5.2 การอภิปรายผล 58
5.3 ข้อเสนอแนะและงานวิจัยในอนาคต 58
บรรณานุกรม 59
ภาคผนวก 64
ประวัติผู้เขียน 98

สารบัญตาราง

ตารางที่ 2.1 แสดงตาราง Confusion Matrix ที่มีสองคลาส 19
ตารางที่ 3.1 แสดงพรรณไม้ที่รวบรวมไว้ในสวนสมุนไพรและสวนไม้หอม PNE Dataset 25
ตารางที่ 4.1 ชุดข้อมูลที่ใช้ในการทดลอง 30
ตารางที่ 4.2 พารามิเตอร์ที่ใช้ในการเตรียมชุดข้อมูล 31
ตารางที่ 4.3 พารามิเตอร์ที่ใช้ในการกำหนดค่าในการทดสอบการเรียนรู้เชิงลึก 31
ตารางที่ 4.4 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง LeNet ที่จำนวน 1,000 รอบ ... 33
ตารางที่ 4.5 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง LeNet ที่จำนวน 10,000 รอบ . 34
ตารางที่ 4.6 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ 35
ตารางที่ 4.7 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง AlexNet
ที่จำนวน 10,000 รอบ 37
ตารางที่ 4.8 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง GoogLeNet
ที่จำนวน 1,000 รอบ 38
ตารางที่ 4.9 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง GoogLeNet
ที่จำนวน 10,000 รอบ 39
ตารางที่ 4.10 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง VGGNet
ที่จำนวน 1,000 รอบ 41
ตารางที่ 4.11 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง VGGNet
ที่จำนวน 10,000 รอบ 42
ตารางที่ 4.12 การเปรียบเทียบประสิทธิภาพของโครงสร้าง CNN กับชุดข้อมูล PNE Dataset 55
ตารางที่ 4.13 การเปรียบเทียบประสิทธิภาพของโครงสร้าง CNN กับชุดข้อมูล
102 Flower Dataset 55
ตารางที่ 4.14 การเปรียบเทียบประสิทธิภาพของโครงสร้าง CNN กับชุดข้อมูล Folio Dataset 56

พหุน ปณ ฑิโต ชีเว

สารบัญภาพ

ภาพประกอบที่ 2.1 แสดงการจำแนกพรรณไม้ด้วยวิธี CNN 9
ภาพประกอบที่ 2.2 แสดงโครงสร้างการทำงานของ LeNet-5 11
ภาพประกอบที่ 2.3 แสดงการทำงานของ Sigmoid Function 12
ภาพประกอบที่ 2.4 แสดงการทำงานของ Tanh Function 12
ภาพประกอบที่ 2.5 แสดงการทำงานของ ReLU function 13
ภาพประกอบที่ 2.6 ตัวอย่างการทำงานของคอนโวลูชัน 13
ภาพประกอบที่ 2.7 การทำพูลลิ่งด้วยค่าสูงสุด 14
ภาพประกอบที่ 2.8 แสดงโครงสร้างการทำงานของ AlexNet 16
ภาพประกอบที่ 2.9 แสดงโครงสร้างการทำงานของ VGGNet 17
ภาพประกอบที่ 2.10 แสดงโครงสร้างแบบ Inception 18
ภาพประกอบที่ 2.11 แสดงโครงสร้างการทำงานของ GoogLeNet 18
ภาพประกอบที่ 2.12 แสดงการทำงานของ Cross-Validation 19
ภาพประกอบที่ 3.1 แสดงตัวอย่างภาพจากฐานข้อมูล PNE Dataset 25
ภาพประกอบที่ 3.2 แสดงตัวอย่างภาพจากฐานข้อมูล 102 Flower Dataset 26
ภาพประกอบที่ 3.3 แสดงตัวอย่างภาพจากชุดข้อมูล Folio Dataset 26
ภาพประกอบที่ 4.1 PNE Dataset กับโครงสร้าง LeNet-5 ที่จำนวน 1,000 รอบ 32
ภาพประกอบที่ 4.2 PNE Dataset กับโครงสร้าง LeNet-5 ที่จำนวน 10,000 รอบ 33
ภาพประกอบที่ 4.3 PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ 35
ภาพประกอบที่ 4.4 PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ 36
ภาพประกอบที่ 4.5 PNE Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ 37
ภาพประกอบที่ 4.6 PNE Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ 39
ภาพประกอบที่ 4.7 PNE Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ 40
ภาพประกอบที่ 4.8 PNE Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ 41
ภาพประกอบที่ 4.9 102 Flower Dataset กับโครงสร้าง LeNet ที่จำนวน 1,000 รอบ 43
ภาพประกอบที่ 4.10 102 Flower Dataset กับโครงสร้าง LeNet ที่จำนวน 10,000 รอบ 44
ภาพประกอบที่ 4.11 102 Flower Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ 44
ภาพประกอบที่ 4.12 102 Flower Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ 45
ภาพประกอบที่ 4.13 102 Flower Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ 46
ภาพประกอบที่ 4.14 102 Flower Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 10,000 รอบ 46

สารบัญภาพ (ต่อ)

ภาพประกอบที่ 4.15 102 Flower Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ 47
ภาพประกอบที่ 4.16 102 Flower Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ 48
ภาพประกอบที่ 4.17 Folio Dataset กับโครงสร้าง LeNet ที่จำนวน 1,000 รอบ 49
ภาพประกอบที่ 4.18 Folio Dataset กับโครงสร้าง LeNet ที่จำนวน 10,000 รอบ 49
ภาพประกอบที่ 4.19 Folio Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ 50
ภาพประกอบที่ 4.20 Folio Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ 51
ภาพประกอบที่ 4.21 Folio Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ 51
ภาพประกอบที่ 4.22 Folio Dataset กับโครงสร้าง GooLeNet ที่จำนวน 10,000 รอบ 52
ภาพประกอบที่ 4.23 Folio Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ 53
ภาพประกอบที่ 4.24 Folio Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ 54

พหูน ปณฺฑิโต ชีเว

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของการวิจัย

ทรัพยากรป่าไม้เป็นทรัพยากรที่สำคัญของประเทศไทย แต่ในปัจจุบันทรัพยากรป่าไม้ลดลงอย่างรวดเร็ว เนื่องจากการบุกรุก ตัดไม้ทำลายป่าทำให้พืชกลุ่มหนึ่งที่มีคุณค่าทางเศรษฐกิจได้แก่ พืชสมุนไพรถูกทำลายไปด้วย รัฐบาลตระหนักถึงความสำคัญของสมุนไพรไทยซึ่งเป็นภูมิปัญญาดั้งเดิมและเป็นทรัพยากรที่สำคัญของประเทศ จึงจัดทำแผนแม่บทแห่งชาติว่าด้วยการพัฒนาสมุนไพรไทย ฉบับที่ 1 พ.ศ.2560-2564 [1] โดยเป้าหมายหลักของแผนแม่บทฉบับนี้คือส่งเสริมให้มีการใช้ประโยชน์จากสมุนไพรอย่างมีประสิทธิภาพ เต็มประสิทธิภาพ ครบวงจร และรักษาภูมิปัญญาดั้งเดิมของไทยให้คงไว้ เพื่อเป็นแหล่งเรียนรู้ด้านภูมิปัญญาและสมุนไพรไทย

ในประเทศไทยก็มีความต้องการใช้สมุนไพรเพิ่มขึ้นเป็นจำนวนมาก ทั้งนี้เนื่องมาจากคนไทยที่ใส่ใจกับสุขภาพหันมาสนใจการดูแลสุขภาพด้วยผลิตภัณฑ์จากธรรมชาติ และตัวสมุนไพรเองก็สามารถนำมาเพิ่มมูลค่าและแปรรูปเป็นผลิตภัณฑ์ต่าง ๆ ได้หลากหลาย โดยรัฐบาลวางแผนไว้ว่าอุตสาหกรรมสมุนไพรไทยจะเป็นอุตสาหกรรมที่มีศักยภาพและมีความยั่งยืน เป็นอุตสาหกรรมที่มีส่วนช่วยขับเคลื่อนเศรษฐกิจ โดยบทสรุปของผู้บริหารแผนแม่บทแห่งชาติว่าด้วยการพัฒนาสมุนไพรไทย [1] ในยุทธศาตร์ที่ 1 มาตรการที่ 3 ให้คำสำคัญกับการอนุรักษ์และใช้พืชสมุนไพรจากป่าอย่างยั่งยืน ซึ่งยุทธศาสตร์ดังกล่าวมีความสอดคล้องกับโครงการอนุรักษ์พันธุกรรมพืชอันเนื่องมาจากพระราชดำริสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี (อพ.สธ.) ซึ่งมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ ได้เข้าร่วมสนองพระราชดำริฯ โครงการดังกล่าว ซึ่งมีวัตถุประสงค์คือดูแลรักษาพรรณไม้ และสมุนไพร รวมไปถึงใช้ประโยชน์จากพรรณไม้ต่าง ๆ มหาวิทยาลัยฯ ยังได้จัดทำสวนสมุนไพรและสวนไม้หอมในลักษณะของสวนพฤกษศาสตร์โรงเรียน เพื่อเป็นแหล่งเรียนรู้และรวบรวมพรรณไม้ที่หายาก [2-3]

ทั้งนี้ การที่บุคคลหนึ่งจะมีความรู้ความสามารถที่จะนำพรรณไม้สมุนไพรไปใช้ให้เกิดประโยชน์นั้น จำเป็นอย่างยิ่งที่จะต้องมีความรู้ทางด้านพฤกษศาสตร์ด้านการจำแนกหรืออนุกรมวิธานพืช ซึ่งเป็นเรื่องยากสำหรับบุคคลที่ไม่คุ้นเคยกับพฤกษศาสตร์ จะต้องศึกษาเรียนรู้และทำความเข้าใจถึงลักษณะพิเศษของพรรณไม้แต่ละชนิด เพื่อให้สามารถจำแนกลักษณะของพรรณไม้ได้อย่างถูกต้องแม่นยำ ไม่ให้เกิดอันตรายเมื่อนำสมุนไพรไปใช้ เนื่องจากพืชที่มีพิษบางชนิดมีลักษณะที่คล้ายกับสมุนไพรเป็นอย่างมาก [4]

จากปัญหาดังกล่าว จึงทำให้ผู้วิจัยมีแนวคิดที่จะใช้ความรู้ทางด้านการเรียนรู้เชิงลึก (Deep Learning) [5] มาเพื่อช่วยแก้ปัญหาในการจำแนกพรรณไม้ (Plant Classification) ซึ่งเทคนิคการ

เรียนรู้เชิงลึกมีจุดเด่นตรงที่สามารถสกัดลักษณะเด่น (Feature Extraction) และการจำแนกรูป (Image Classification) ได้ในขั้นตอนเดียว เหมาะกับการนำมาประยุกต์ใช้ในการวิจัยครั้งนี้ โดยกลุ่มตัวอย่างพรรณไม้ที่จะนำมาใช้ในงานวิจัยนี้เป็นรูปภาพพรรณไม้หอมที่ถ่ายในสิ่งแวดล้อมทางธรรมชาติ (Natural Environment) ซึ่งพรรณไม้หอมที่ได้รวบรวมไว้ในสวนของมหาวิทยาลัยนั้น เป็นพืชมีสรรพคุณทางยาและจัดเป็นสมุนไพรที่มีความสำคัญที่ควรค่าแก่การอนุรักษ์และนำมาใช้ประโยชน์

1.2 ความมุ่งหมายและวัตถุประสงค์ของการวิจัย

เพื่อนำเสนอกระบวนการการเรียนรู้เชิงลึกในการจำแนกพรรณไม้ และเปรียบเทียบโครงสร้างการเรียนรู้เชิงลึกแบบคอนโวลูชันนอลนิวรอลเน็ตเวิร์กโดยกลุ่มตัวอย่างพรรณไม้ที่จะนำมาใช้ในงานวิจัยนี้เป็นรูปภาพพรรณไม้หอม

1.3 ขอบเขตในการวิจัย

1.3.1 รูปภาพพรรณไม้ที่ใช้ในงานวิจัย (Plants in Natural Environment Dataset: PNE Dataset) เป็นชุดข้อมูลภาพถ่ายที่รวบรวมจากเว็บไซต์ Google โดยเป็นภาพพรรณไม้ชนิดเดียวกันกับพรรณไม้ในสวนไม้หอมของมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ [2-3] ซึ่งเป็นภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติที่มีลักษณะเด่นในภาพอย่างน้อย 1 อย่าง มีจำนวนทั้งสิ้น 10 ชนิด ชนิดละ 300 ภาพ รวมเป็น 3,000 ภาพ

1.3.2 ใช้เทคนิคการเรียนรู้เชิงลึก ด้วยวิธีคอนโวลูชันนอลนิวรอลเน็ตเวิร์ก (Convolutional Neural Networks: CNN) [5] สำหรับการรู้จำพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ

1.3.3 นำข้อมูลชุดข้อมูลรูปภาพทั้ง 3 ชุด ได้แก่ ชุดข้อมูล PNE Dataset ชุดข้อมูล 102 Flower Dataset และชุดข้อมูล Folio Dataset เปรียบเทียบประสิทธิภาพของเทคนิค Deep Learning ด้วยวิธี CNN ที่มีลักษณะของโครงสร้าง (Architecture) ที่ต่างกัน 4 แบบ ได้แก่ LeNet [6], AlexNet [7], VGGNet [8] และGoogLeNet [9]

1.3.4 ทำการตรวจสอบความถูกต้องโดยใช้ตัววัดประสิทธิภาพ คือ Precision เป็นการวัดความแม่นยำของโมเดล และ Top1-Accuracy เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกกลุ่ม (Class) [10]

1.4 ผลที่คาดว่าจะได้รับจากการวิจัย

ได้วิธีการเรียนรู้ของเชิงลึกที่เหมาะสำหรับการจำแนกพรรณไม้

1.5 นิยามศัพท์เฉพาะ

1.5.1 การจำแนกพรรณไม้ (Plant Classification) [4-5, 11] คือ การนำรูปภาพพรรณไม้ไปผ่านกระบวนการ เช่น การประมวลผลภาพ (Image Processing) การสกัดคุณลักษณะพิเศษ (Feature Extraction) และนำไปเรียนรู้โดยวิธีการเรียนรู้ของเครื่อง (Machine Learning) หรือใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ในการจำแนกพรรณไม้

1.5.2 สิ่งแวดล้อมทางธรรมชาติ (Natural Environment) คือ ภาพถ่ายพรรณไม้ในสภาวะแวดล้อมที่อยู่ทางธรรมชาติโดยไม่มีการจัดฉากหรือกระบวนการจัดแสงควบคุมแสงเหมือนในห้องทดลองหรือสตูดิโอ

1.5.3 เทคนิคการเรียนรู้เชิงลึก (Deep Learning) [5,12] คือ โครงข่ายประสาทเทียม (Neural Network: NN) เป็นแบบจำลองคณิตศาสตร์ที่พยามเลียนแบบการทำงานของโครงข่ายประสาทจริงของมนุษย์ มีสามารถในการประมาณค่าฟังก์ชันใด ๆ ก็ได้หากมีโครงข่ายที่ซับซ้อนพอ ซึ่งได้มีการพัฒนาอัลกอลิทึมสำหรับการเรียนรู้ให้ NN ที่ลึกมาก (มีหลายชั้น) ทำให้เป็นการจุดประกายความสนใจในโครงข่ายประสาทเทียมขึ้นมาใหม่ในหัวข้อการเรียนรู้เชิงลึก

1.5.4 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNN) [6] คือโครงข่ายประสาทเทียมรูปแบบหนึ่งที่มีจุดเด่นคือสามารถดึงลักษณะเด่นและจำแนกข้อมูลได้โดยตรง เหมาะสำหรับการรู้จำภาพหรือการจำแนกภาพ

1.6 ข้อจำกัดของการวิจัย

1.6.1 ข้อจำกัดด้านภาพที่ใช้เรียนรู้และทดสอบนั้น ต้องเป็นภาพพรรณไม้จำนวน 10 ชนิด (PNE Dataset) เท่านั้น หากใช้สายพันธุ์อื่น ผลลัพธ์จะเป็นเพียงพรรณไม้ 10 ชนิด เท่านั้น

1.6.2 หากต้องการให้ผลการจำแนกพรรณไม้มีอัตราความถูกต้องสูง ภาพที่นำมาทำการเรียนรู้ต้องเป็นภาพที่มีดอกไม้เป็นจุดเด่น

1.6.3 ข้อจำกัดด้านการประมวลผล เนื่องจากกระบวนการที่จะทำให้เครื่องจักรเรียนรู้ด้วยเทคนิคการเรียนรู้เชิงลึก (Deep Learning) จะใช้เวลาในการเรียนรู้ (Training) เป็นเวลานาน หากทำการทดลองโดยใช้หน่วยประมวลผลกลาง (Central Processing Unit: CPU) จะทำให้ใช้เวลาในการเรียนรู้เป็นเวลานาน หากใช้หน่วยประมวลผลด้านกราฟิกส์ (Graphics Processing unit: GPU) ได้แก่ GPU รุ่น Geforce 1080 เป็นต้นไปจะช่วยลดระยะเวลาในการเรียนรู้ได้หลายเท่า

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้ได้กล่าวนี้ได้กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการเรียนรู้เชิงลึกสำหรับการจำนวนพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ ดังนี้

2.1 ทฤษฎีทางด้านพฤกษศาสตร์ (Botany theory)

2.2 คอมพิวเตอร์วิทัศน์ (Computer Vision)

2.3 การเรียนรู้เชิงลึก (Deep Learning)

2.4 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN)

2.5 การวัดประสิทธิภาพของโมเดล (Performance Measurement)

2.6 งานวิจัยที่เกี่ยวข้อง (Literature Reviews)

2.1 ทฤษฎีทางด้านพฤกษศาสตร์ (Botany theory)

2.1.1 หลักการจำแนกพรรณไม้ตามหลักอนุกรมวิธานพืช (Plant taxonomy)

ก่องกานดา ชยามฤต และวรดลต์ แจ่มจำรูญ [4] กล่าวไว้ว่า การจำแนกพรรณไม้นั้นต้องอาศัยความรู้พื้นฐานในวิชาพฤกษศาสตร์ ซึ่งอยู่ในสาขาอนุกรมวิธานพืช (Plant Taxonomy) เป็นหลัก วิชาพฤกษศาสตร์จึงเป็นความรู้พื้นฐานที่จำเป็นอย่างยิ่งที่จะนำไปใช้จำแนกพรรณไม้ได้โดยทั่วไป ไม่ว่าพรรณไม้นั้น ๆ จะเป็นพรรณไม้ในถิ่นใด ถึงแม้ว่าจะไม่คุ้นเคยหรือไม่รู้จักพรรณไม้นั้น ๆ มาก่อนเลยก็ตาม ถ้านำเอาวิชาการด้านนี้เข้าไปช่วยแล้วก็จะจำแนกพรรณไม้ได้อย่างถูกต้องแน่นอน จึงนับได้ว่า อนุกรมธานพืชนี้เป็นหัวใจในการจำแนกพรรณไม้ ดังนั้นผู้ที่เกี่ยวข้องกับงานด้านนี้ จึงจำเป็นต้องให้ความสนใจกับวิชาการด้านนี้พอสมควร อนุกรมวิธาน ตรงกับรากศัพท์ภาษาอังกฤษว่า Taxonomy หรือ Systematics เป็นศาสตร์ที่มีขอบเขตกว้างขวางในการศึกษาวิจัยเกี่ยวกับรูปพรรณสัณฐานของสิ่งใดสิ่งหนึ่ง โดยอาศัยข้อมูลหลาย ๆ ด้านของสิ่งใดสิ่งหนึ่งเหล่านั้น เมื่อนำวิชาอนุกรมวิธานมาใช้ในวิชาพฤกษศาสตร์ จึงหมายถึงวิชา อนุกรมวิธานพืช (Plant Taxonomy หรือ Plant Systematics) ถ้าจะพูดให้เข้าใจง่ายคือ การจำแนกพรรณไม้นั่นเอง วิชาอนุกรมวิธานพืชนี้นับได้ว่าเป็นแม่บทของวิชาพฤกษศาสตร์ เพราะก่อนที่จะเรียนรู้เรื่องของพืชในด้านต่าง ๆ ไม่ว่าจะเป็นด้านสัณฐานวิทยา (Morphology) สรีรวิทยา (Physiology) กายวิภาควิทยา (Anatomy) ฯลฯ จำเป็นต้องเรียนรู้ชื่อและลักษณะเด่น ๆ ของพรรณไม้นั้น ๆ เสียก่อน จึงอาจกล่าวได้ว่าผู้ที่จะศึกษาเกี่ยวกับพรรณไม้ไม่ว่าจะเป็นทางด้านนวกรรม เกษตรกรรม เภสัชกรรม ตลอดจนการนำพรรณไม้ไปใช้ประโยชน์ในอุตสาหกรรมด้านต่าง ๆ ก็ต้องอาศัยผลงานของ นักอนุกรมวิธานพืช (Plant

Taxonomist) ด้วยกันทั้งนั้น เพื่อที่จะรู้จักชื่อพรรณไม้ต่าง ๆ อย่างถูกต้องแน่นอน ผลงานวิจัยใด ๆ ที่เกี่ยวกับพรรณไม้ ถึงแม้ว่าจะมีหลักการและการวางแผนปฏิบัติการดีเพียงใดก็ตามถ้าหากเริ่มต้นด้วยชื่อของพรรณไม้ที่ผิดพลาดหรือไม่ถูกต้องตรงตามชนิดแล้ว ผลงานวิจัยนั้นย่อมไร้คุณค่าโดยสิ้นเชิง

2.1.2 ขอบเขตของวิชาอนุกรมวิธานพืชมีดังนี้

การระบุพรรณไม้ (Plant Identification) [4] พรรณไม้มีอยู่มากมายทั้งที่รู้จักแล้วและไม่รู้จัก ซึ่งต้องมีวิธีการตรวจสอบหาชื่อให้ได้ Plant Identification คือการตรวจพิจารณาว่าพรรณไม้ที่ต้องการตรวจสอบมีลักษณะคล้ายคลึงหรือเหมือนกับพรรณไม้ที่รู้จักแล้วหรือไม่ ตรวจสอบแล้วอาจพบว่าเป็นพรรณไม้ชนิดใหม่ (New Species) หรือถ้าพรรณไม้มีอยู่แล้วแต่ไม่เคยมีรายงานการพบมาก่อนก็เรียกว่าเป็น New Record ซึ่งจะต้องมีการเขียนรายงานออกมา การตรวจสอบมีบทบาทสำคัญในการที่จะนำเสนอข้อมูลและการสื่อความหมายว่าพรรณไม้นั้น ๆ คืออะไร มีข้อมูลอะไรที่เกี่ยวข้องบ้าง การที่จะตรวจสอบได้ผู้ตรวจสอบจำเป็นต้องรู้จักส่วนประกอบต่าง ๆ ของพรรณไม้ การตรวจสอบสามารถทำได้หลายวิธี ตัวอย่างเช่น

1) ระบุพรรณไม้โดยใช้รูปวิธาน (Key) ซึ่งมีลักษณะเป็นข้อความที่บรรยายเกี่ยวกับลักษณะของพรรณไม้ที่นำมาตรวจสอบ โดยทั่วไปจะมีลักษณะเป็น Dichotomous Key คือจะมีคู่ของข้อความที่แยกเป็นสองหัวข้อ โดยมีรายละเอียดของลักษณะพรรณไม้ที่แตกต่างกัน เพื่อให้ผู้ที่ทำการระบุเลือกว่าข้อความใดตรงกับลักษณะพรรณไม้ที่นำมาตรวจสอบ คู่ของข้อความที่กล่าวถึงสิ่งเดียวกันแต่มีลักษณะต่างกันใน Key นี้เรียกว่า Couplet แต่ละข้อความเรียก Lead รูปแบบของ Dichotomous Key มีสองแบบ คือ Indented หรือ Yoked Key และ Bracketed Key ทั้งสองแบบมีข้อดีและข้อเสียแตกต่างกันไป แบบแรกได้รับความนิยมมากกว่า

2) ระบุโดยการเปรียบเทียบ (Comparison) โดยนำพรรณไม้ที่ยังไม่รู้จักไปเปรียบเทียบลักษณะกับตัวอย่างแห้งของพรรณไม้ที่มีชื่อที่ถูกต้องกำกับในพิพิธภัณฑ์พืช หรือเปรียบเทียบกับภาพถ่าย ภาพวาด หรือคำบรรยายของพรรณไม้ที่รู้จักแล้ว ข้อนี้จะใช้เวลาน้อยลงถ้าได้มีการใช้ key ให้ได้ชื่อก่อน แล้วจึงอ่านคำบรรยายและเปรียบเทียบตัวอย่างเพื่อความมั่นใจ

3) ระบุโดยการถามผู้เชี่ยวชาญที่ทำการศึกษาวิจัยพรรณไม้ในกลุ่มที่กำลังตรวจสอบอยู่ กรณีนี้ต้องรู้ก่อนว่าพรรณไม้ที่จะตรวจสอบจัดอยู่ในวงศ์ใด รายชื่อผู้เชี่ยวชาญหรือผู้ที่รับผิดชอบทำการวิจัยวงศ์ต่าง ๆ สำหรับโครงการพรรณพฤกษชาติประเทศไทย สามารถตรวจสอบได้จากเว็บไซต์ของหอพรรณไม้ (http://www.dnp.go.th/botany) [13]

2.1.3 การจำแนกพรรณไม้ (Plant Classification)

การจำแนกพรรณไม้ [4] เป็นการจัดพรรณไม้ให้เป็นกลุ่มหรือหมวดหมู่โดยอาศัยลักษณะความคล้ายคลึง (Similarities) และความแตกต่าง (Differences) ของลักษณะต่าง ๆ ที่ศึกษา กลุ่มพรรณไม้ ที่ถูกจำแนกตามหลักอนุกรมวิธานนี้เรียก หน่วยอนุกรมวิธาน (Taxon) ซึ่งมีการจัดระดับเป็น

หมวดหมู่ที่มีขนาดใหญ่เล็กลดหลั่นกันไปตามลำดับ โดยหมวดหมู่ที่เป็นหลักสำคัญมี 7 ระดับได้แก่ Kingdom, Division, Class, Order, Family, Genus และSpecies ทั้งนี้รายละเอียดในการจำแนกของนักพฤกษศาสตร์ อนุกรมวิธานแต่ละคน หรือในแต่ละยุคสมัยอาจแตกต่างกัน ทำให้ระบบของการจัดจำแนก (System of Classification) อาจมีได้หลายระบบ

2.1.4 การจำแนกชนิดพรรณไม้แบบลักษณะเด่นเฉพาะ

วรดลต์ แจ่มจำรูญ [14] กล่าวไว้ว่า ในการศึกษาด้านอนุกรมวิธานพืชโดยเฉพาะในการวิเคราะห์ระดับวงศ์ สกุล และชนิด มักจะประสบปัญหาคือ ตัวอย่างที่นำมาวิเคราะห์ไม่มีส่วนของดอกหรือผลอันเป็นลักษณะสำคัญของการใช้รูปวิธาน ดังนั้นหากตัวอย่างที่นำมาวิเคราะห์ไม่มีส่วนของดอกและผลก็ยังสามารถใช้ลักษณะอื่น ๆ ของพืช เช่น ใบ น้ำยาง กลิ่น สิ่งปกคลุม เป็นต้น อันเป็นลักษณะเด่นเฉพาะตัวของพืชแต่ละวงศ์ สกุล หรือแม้แต่ชนิด โดยใช้พื้นฐานความรู้ด้านอนุกรมวิธานพืช สัณฐานวิทยา และนิเวศวิทยา เพื่อช่วยการวิเคราะห์ชื่อพืชได้ในระดับหนึ่ง และสามารถนำไปตรวจสอบหรือวิเคราะห์เพิ่มเติม เช่น การเทียบเคียงกับตัวอย่างพรรณไม้แห้งเพื่อให้การวิเคราะห์นั้นมีความถูกต้องมากที่สุด

2.2 คอมพิวเตอร์วิทัศน์ (Computer Vision)

คอมพิวเตอร์วิทัศน์ (Computer Vision) [11–13] คือเทคโนโลยีการทำให้คอมพิวเตอร์มีความสามารถในการมองเห็นภาพหรือเข้าใจภาพได้ โดยจะเปรียบเทียบการมองเห็นของมนุษย์กับการมองเห็นของคอมพิวเตอร์โดยมีหลักการง่าย ๆ คือเวลามนุษย์ใช้ตามองไปที่ฉาก ๆ หนึ่ง เช่น การมองไปที่วิวในห้อง จะสามารถรับรู้ได้เลยว่า นี่คือโต๊ะ เก้าอี้ ตู้เย็น การรับรู้ได้เพราะว่ามีตาในการมอง และมีสมองในการประมวลผลว่า นี่คือวัตถุ ต่าง ๆ แต่สำหรับคอมพิวเตอร์ เป็นเรื่องที่ยากมาก เพราะฉะนั้นจึงเกิดศาสตร์ทางด้านคอมพิวเตอร์ขึ้นมาใหม่ ที่เราเรียกว่า คอมพิวเตอร์วิทัศน์ นั่นก็คือการทำให้คอมพิวเตอร์ได้รับรู้ภาพหรือมองเห็นภาพได้เหมือนมนุษย์ โดยหลักการทำงานของคอมพิวเตอร์วิทัศน์ คือ ทำให้คอมพิวเตอร์สามารถมองเห็นได้ ซึ่งก็คือใช้กล้องถ่ายภาพนั่นเอง กล้องนี้ก็เปรียบได้เหมือนตาของมนุษย์ และสิ่งถัดมาก็คือจะต้องให้คอมพิวเตอร์เรียนรู้และรู้จักคำจำกัดความของวัตถุต่าง ๆ ซึ่งขั้นตอนนี้ก็เปรียบได้เหมือนกับสมองของมนุษย์

ในปัจจุบันก็มีการนำคอมพิวเตอร์วิทัศน์ มาใช้ในอุตสาหกรรมการผลิตที่มีความสลับซับซ้อนที่ต้องการการทำงานที่มีประสิทธิภาพสูงและต้องการความแม่นยำในการผลิตสูงด้วยการใช้คอมพิวเตอร์วิทัศน์ในการประมวลผลภาพ (Image Processing) จากกล้องถ่ายภาพดิจิตอล และทำการปรับปรุงคุณภาพ เพื่อให้คอมพิวเตอร์มองเห็นผ่านโปรแกรม เพื่อให้คอมพิวเตอร์หาความเหมือน ทำการจดจำค่าใดค่าหนึ่ง และทำการแยกแยะวัตถุ เพื่อให้คอมพิวเตอร์สามารถตัดสินใจทำงานได้ถูกต้องตามคำสั่งที่ต้องการ ระบบคอมพิวเตอร์วิทัศน์จะประกอบด้วยการทำเทรชโฮลด์ (Thresholding) คือ

กระบวนการเปลี่ยนภาพสีเทาให้เป็นภาพขาวดำ การแยกภาพออกเป็นส่วน ๆ (Segmentation) การรู้จำแบบ (Pattern Recognition) การหาขอบภาพ (Edge Detection) และการเข้าคู่รูปแบบ (Template Matching) เป็นต้น

2.2.1 ระบบคอมพิวเตอร์วิทัศน์ (Computer Vision System) [11,17] คือ ปัญญาประดิษฐ์ทางด้านวิทยาศาสตร์และเทคโนโลยีของการทำให้เครื่องจักรสามารถแยกแยะหรือดึงเอาข้อมูลต่าง ๆ จากภาพนั้น โดยเขียนโปรแกรมให้คอมพิวเตอร์เข้าใจถึงเหตุการณ์ในภาพหรือลักษณะเด่นในภาพ ข้อมูลภาพสามารถทำให้มีรูปแบบต่าง ๆ ได้หลายรูปแบบ เช่น ความต่อเนื่องของภาพ วิดีโอ หรือภาพหลายมิติจากกล้องหลาย ๆ ตัว

องค์ประกอบของระบบคอมพิวเตอร์วิทัศน์ขึ้นอยู่กับการประยุกต์ใช้งาน บางระบบอาจจะเป็นคอมพิวเตอร์เพียงเครื่องเดียวที่ไม่มีการเชื่อมต่อกับคอมพิวเตอร์เครื่องอื่น ๆ ซึ่งใช้ในการตรวจวัดงานเฉพาะหรือแก้ปัญหาการตรวจจับ ขณะที่บางระบบอาจประกอบด้วยคอมพิวเตอร์ย่อยหลาย ๆ ตัวของคอมพิวเตอร์ระบบใหญ่ เช่น คอมพิวเตอร์ระบบย่อยสำหรับการควบคุมตัวขับเคลื่อนทางกล เป็นต้น

1) การบันทึกภาพ (Image Acquisition) ภาพดิจิตอลถูกทำให้เกิดขึ้นโดยเซ็นเซอร์ภาพหนึ่งตัวหรือหลายตัว ซึ่งมีความไวต่อแสงติดตั้งอยู่ภายในกล้อง และยังประกอบด้วยเซ็นเซอร์วัดระยะ อุปกรณ์วาดภาพ ตัวขวาง เรดาร์ และ อัลตราโซนิก เป็นต้น ซึ่งขึ้นอยู่กับชนิดของเซนเซอร์ ผลลัพธ์ของข้อมูลภาพของวัตถุจะเป็นภาพ 2 มิติ หรือ 3 มิติ ซึ่งจะถูกประมวลผลภาพวัตถุก่อนวัตถุใหม่จะมาถึง ค่าพิกเซลจะสอดคล้องกับความเข้มแสง (ภาพสีเทาหรือภาพสี) แต่ก็มีความสัมพันธ์กับการวัดทางฟิสิกส์อื่น ๆ เช่น ความลึก

2) การประมวลผลภาพเบื้องต้น (Pre-Processing) หรือ การประมวลผลระดับล่าง (Low Level Processing) เราสามารถนำข้อมูลภาพมาดึงข้อมูลเพื่อหารอยต่อเฉพาะบางรอยต่อของข้อมูล ก่อนใช้วิธีการมองเห็นของคอมพิวเตอร์ ซึ่งมีความจำเป็นในการประมวลผลข้อมูล เพื่อที่จะรับประกันว่าสมมติฐานนั้นเชื่อถือและมั่นใจได้ เช่น การสุ่มซ้ำเพื่อที่จะรับประกันว่าระบบระยะพิกัดของภาพถูกต้อง ลดสัญญาณรบกวน และการแยะรูปภาพ เพื่อที่จะรับประกันว่าเซนเซอร์สัญญาณรบกวนไม่ทำให้ภาพเสียหาย

3) การสกัดคุณลักษณะพิเศษ (Feature Extraction) คือการหาคุณลักษณะพิเศษของภาพที่ระดับต่าง ๆ ของลักษณะเชิงซ้อนจะถูกดึงออกจากข้อมูลภาพตัวอย่าง เช่น เส้น ขอบ และเส้นนูนของภาพ ตำแหน่งของจุดภาพที่สนใจ เช่น มุน รอยเปื้อย หรือจุดภาพ ซึ่งอาจจะสัมพันธ์กับพื้นหลังรูปร่างหรือการเคลื่อนที่

4) การตรวจจับและการแยกข้อมูลภาพออกเป็นส่วน ๆ (Detection and Segmentation) การตัดสินใจเลือกจุดภาพในการประมวลผล หรือขอบภาพจะสัมพันธ์กับการ

ประมวลผลหาคุณลักษณะพิเศษ เช่น การเลือกปรับเฉพาะจุดภาพที่สนใจ การแยกข้อมูลของขอบเขตภาพหนึ่งจุดภาพหรือหลายจุดภาพ ซึ่งจะจำกัดจุดภาพเฉพาะของวัตถุที่สนใจ

5) การประมวลผลระดับสูง (High-level Processing) ในขั้นตอนนี้ อินพุต คือ ต้นแบบของกลุ่มข้อมูล เช่น กลุ่มของจุดภาพ หรือขอบเขตภาพซึ่งแสดงจุดจำกัดเฉพาะของวัตถุ เช่น ตำแหน่งหรือขนาดของวัตถุ การลงทะเบียนภาพเพื่อเปรียบเทียบและรวมความแตกต่างของมุมมอง 2 มุมมองของภาพเดียวกัน

2.2.2 วิธีการประมวลผลภาพ (Image Processing Method) [17] ในทางปฏิบัติของอุตสาหกรรมการผลิต โปรแกรมการมองเห็นของคอมพิวเตอร์จะมีการรวมเอาเทคนิคการประมวลผลภาพวิธีต่าง ๆ เข้าด้วยกัน เพื่อให้การทำงานมีความถูกต้อง แม่นยำ สมบูรณ์ และเหมาะสม ได้แก่

1) พิกเซล (Pixel) จุดภาพหรือพิกเซล มาจากคำว่าพิกเจอร์ (Picture) ที่แปลว่า รูปภาพ และอีเลเมนต์ (Element) ที่แปลว่า องค์ประกอบ หรือหน่วย โดยภาพหนึ่ง ๆ จะประกอบไปด้วยจุดภาพหรือพิกเซลมากมาย และแต่ละภาพที่สร้างขึ้นจะมีความหนาแน่นของจุดภาพ หรือเรียกว่า ความละเอียดหรือความคมชัดที่แตกต่างกันไป พิกเซลใช้ในการบอกสมบัติของภาพ จอภาพหรืออุปกรณ์แสดงผลภาพ ภาพที่มีจำนวนพิกเซลมากก็จะมีความละเอียดของภาพสูง การระบุจำนวนค่าพิกเซลจะบอกในลักษณะแนวนอนคูณแนวตั้ง (X x Y) เช่น 1,366 x 768 พิกเซล = 1,049,088 พิกเซล หรือ 1 ล้านพิกเซล หรือเป็นหน่วยวัดความละเอียดของภาพ เช่น 2,400 พิกเซลต่อนิ้ว 640 พิกเซลต่อบรรทัด หรือระยะห่าง 10 พิกเซล

2) การแบ่งแยกบริเวณ (segmentation) เป็นกระบวนการของการ แบ่งแยกภาพดิจิตอลออกเป็นส่วน ๆ แสดงวัตถุต่าง ๆ ในภาพทำให้ง่ายต่อการวิเคราะห์วัตถุในภาพ การแบ่งแยกภาพจะสามารถระบุตำแหน่งและขอบเขตวัตถุในภาพ เช่น เส้น เส้นโค้ง และฉากหลัง ผลลัพธ์ของการแบ่งแยกภาพคือ กลุ่มของการแบ่งแยกภาพหรือวัตถุ หรือกลุ่มของรูปทรงต่าง ๆ จะถูกครอบคลุมไว้ทั้งหมด หรือถูกดึงออกมาจากภาพการจับขอบภาพ (Edge Detection) พิกเซลแต่ละพิกเซลในพื้นที่จะบอกถึงคุณสมบัติบางส่วนหรือลักษณะของภาพ

2.3 การเรียนรู้เชิงลึก (Deep Learning)

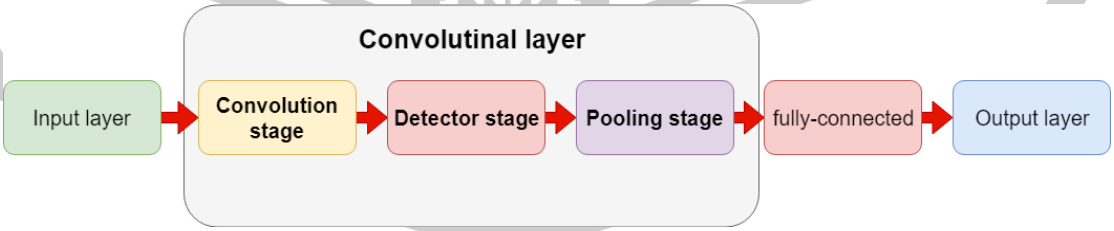
Deep Learning [5,12] คือ โครงข่ายประสาทเทียม (Artifical Neural Network) เป็นแบบจำลองคณิตศาสตร์ที่พยายามเลียนแบบการทำงานของโครงข่ายประสาทเทียมของมนุษย์ จากความรู้พื้นฐานด้านชีววิทยานำมาสู่แบบจำลองที่มีการใช้งานกันอย่างแพร่หลายในปัจจุบัน นักทฤษฎีหลายคนได้ทำการพิสูจน์ว่า NN สามารถใช้ประมาณค่าฟังก์ชันใด ๆ ก็ได้ หากเรามีโครงข่ายที่ซับซ้อนพอ จากบทพิสูจน์เหลานี้ทำให้ NN ได้รับความนิยมแพร่หลายผลการรู้จำที่ได้จากโครงสร้างนี้ก็อยู่ระดับดีมากเช่นเดียวกับ SVM แต่ทั้งนี้ในช่วงก่อนปี 2000 นั้นโครงสร้างของ NN ที่ใช้กันมักจะจำกัด

อยู่ที่โครงสร้างอย่างง่าย เหตุผลหลักคือมันเพียงพอกับงานด้านการรู้จำรู้แบบและการประมาณค่าทั่วไปแล้ว แต่อีกผลหนึ่งที่คนไม่ค่อยพูดถึงก็คือเพราะอัลกอลิทึมสำหรับสอน NN นั้นไม่มีประสิทธิภาพเพียงพอกับการสอนโครงสร้างที่มีความลึกมากกว่านั้น ในปี 2006 Goeffrey Hinton ได้นำเสนออัลกอลิทึมสำหรับการสอนโครงสร้าง NN ที่ลึกมาก (มีหลายชั้น) ได้ ซึ่งเป็นการจุดประกายความสนใจในโครงข่ายประสาทเทียมขึ้นมาใหม่ในชื่อหัวข้อ การเรียนรู้เชิงลึก (Deep Learning)

การจุดประกายเริ่มต้นนี้ทำให้เกิดโครงข่ายประสาทเทียมกลับมาเป็นที่นิยมอีกครั้งโดยเฉพาะโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) ซึ่งเป็นแบบจำลองที่ให้ผลดีกับงานเกี่ยวกับข้อมูลภาพที่มีการใช้งานมานานแล้วแต่ยังไม่เป็นที่แพร่หลายมากนัก เพราะ CNN มีความซับซ้อนในด้านการคำนวณสูงทำให้เวลาที่ต้องใช้ในการเรียนรู้และในการทำนายค่าใหม่นั้นนานกว่าโครงสร้างประเภทอื่นมาก โชคดีที่กระแสการเรียนรู้เชิงลึกนั้นมาถึงพร้อมกับพัฒนาการทางด้านฮาร์ดแวร์ โดยเฉพาะที่เกี่ยวกับหน่วยประมวลผลกราฟิกส์ (Grapics Processing Unit: GPU) ที่ในปัจจุบันได้มีการเพิ่มคำสั่งพิเศษที่ช่วยในการประมวลผลของ CNN ทำได้เร็วขึ้นมาก จนทำให้โครงข่ายประสาทเทียมแบบคอนโวลูชันเป็นแบบจำลองมาตรฐานสำหรับงานด้านการรู้จำรูปภาพในปัจจุบัน

2.4 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN)

การจำแนกรูปภาพโดยใช้ประสาทเทียมแบบคอนโวลูชัน [6,18] เป็นวิธีการที่ทำได้ทั้งการดึงลักษณะเด่นของภาพ (Feature Extraction) และการจำแนกข้อมูล (Classification) ได้ในระบบเดียวกัน ซึ่งเป็นข้อเด่นของระบบการเรียนรู้เชิงลึกโดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน ซึ่งจะแตกต่างกับการจำแนกรูปภาพแบบการเรียนรู้โดยเครื่องโดยจักรทั่วไป ดังนั้นในการวิจัยนี้ จะนำเสนอการจำแนกพรรณไม้โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันเป็นหลักในการประมวลผลข้อมูลของชุดรูปภาพทั้งหมด และเปรียบเทียบโดยใช้ลักษณะของโครงสร้าง (Architecture) ที่ต่างกัน 4 แบบ ได้แก่ LeNet [6], AlexNet [7], VGGNet [8] และ GoogLeNet [9]



ภาพประกอบที่ 2.1 แสดงการจำแนกพรรณไม้ด้วยวิธี CNN

จากภาพประกอบที่ 2.1 โครงสร้างของประสาทเทียมแบบคอนโวลูชัน จะประกอบไปด้วย 3 ส่วนใหญ่ ๆ คือ

ส่วนที่ 1 Convolution Stage เป็นขั้นตอนของการสร้าง Sliding Window (Filter) มาสแกนรูป Input เพื่อทำ Feature Map

ส่วนที่ 2 Detector Stage ชั้นนี้จะทำหน้าที่รับ Output จากชั้น Convolution Stage แปลงให้อยู่ในรูปของ Nonlinear โดยใช้ Activation อย่างเช่น Rectified Linear Units (ReLU) [7] เพื่อความง่ายในการคำนวณและประสิทธิภาพของผลลัพธ์

ส่วนที่ 3 Pooling Stage ในชั้นนี้จะทำหน้าที่ลดขนาดข้อมูลให้มีขนาดที่เล็กลงโดยที่รายละเอียดของ Input ยังครบถ้วนเหมือนเดิม หลักการทำงานขั้นตอนนี้คล้าย ๆ กับส่วนที่1. แต่ต่างที่ตรง Output ที่ได้จะมีขนาดเล็กลง Pooling มีประโยชน์ในเรื่องเพิ่มความไวในการคำนวณ และแก้ปัญหา Overfitting

Layer ถัดไปอาจจะทำซ้ำเดิมก็ได้ แต่ส่วนที่จะมาเป็น Final Layer นั้นต้องเป็น Fully-Connected NN (Feed-Forward NN) จะทำหน้าที่เป็น Model หลักของโปรแกรม และชั้นสุดท้ายจะเป็นชั้นของ Output

โครงข่ายประสาทเทียมแบบคอนโวลูชัน หรือ CNN มีจุดเด่น คือ สามารถทำการดึงลักษณะเด่นและทำการจำแนกข้อมูลได้โดยตรงในครั้งเดียว จึงเหมาะสำหรับการนำมาประยุกต์ใช้ในการจำแนกรูปภาพ โดย CNN เป็นโครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Perceptron) โดย Yann LeCun [6] เป็นคนแรกที่นำเสนอโครงสร้างการทำงานของ CNN และนำเสนอวิธีการแพร่ความผิดพลาดย้อนกลับ (Back Propagation) สำหรับการรู้จำตัวเลขของลายมือในชุดข้อมูล MNIST ซึ่งมีการทดลองในการรู้จำรูปภาพตัวเลขอยู่ในระดับดี จึงทำให้ CNN เป็นรูปแบบที่นิยมใช้จนถึงปัจจุบัน เนื่องจากโครงข่ายประสาทเทียมของระบบนี้ ไม่ได้ถูกออกแบบให้เรียนรู้แบบชั้นต่อชั้น แต่จะใช้วิธีการปรับค่าน้ำหนัก Weight เพื่อลดฟังก์ชันความคลาดเคลื่อน ด้วยวิธีการ Back Propagation นอกจากนี้จะทำการดึงลักษณะเด่นออกมาโดยอัตโนมัติ ด้วยการเรียนรู้ฟิวเจอร์คอนโวลูชันที่เกิดจากค่าน้ำหนักของแผนภาพคุณลักษณะเด่น (Feature Map) โดยโครงสร้างโดยทั่วไปของ CNN จะมีลักษณะโครงสร้างดังแสดงในภาพที่ 2.2

ภาพประกอบที่ 2.2 แสดงโครงสร้างการทำงานของ LeNet-5

2.4.1 LeNet Architectures

โครงสร้างของ LeNet โดยทั่วไปของ CNN หรือ จะประกอบได้ด้วย 5 ชั้น ได้แก่ ข้อมูลอินพุต (Input Layer) ชั้นของคอนโวลูชัน (Convolution Layer) ชั้นพูลลิ่ง (Pooling Layer) ชั้นลักษณะเด่น (Fully Connected) และชั้นเอาท์พุต (Output Layer) ซึ่งแต่ละชั้นจะมีการทำงานที่แตกต่างกันไป ซึ่งสามารถอธิบายได้ดังต่อไปนี้

1) ชั้นอินพุต (Input Layer) ชั้นนี้จะเป็นการนำเข้าภาพอินพุตเข้าสู่ระบบ

2) ชั้นคอนโวลูชัน (Convolutional Layer) ในขั้นตอนนี้จะเป็นการนำาภาพอินพุตมาคอนโวลูชันกับหน้ากาก Filter โดยที่ค่าสัมประสิทธิ์ในหน้ากากนั้นได้จากการสุ่มขึ้นมา ซึ่งการทำคอนโวลูชันสำหรับเอาท์พุตที่ j ในชั้น l มีสมการในดำเนินการแสดงได้ ดังสมการที่ 2.1

$$y_j^{(i)}(x,y) = \emptyset^1\left(\sum i=i \ \sum(u,v)\, w_{ji}^{(i)}(u,v)\, y_i^{(i-1)}(x+u, y+v) + b_j^{(i)}\right) \quad (2.1)$$

โดยที่ $K = \{(u,v) \in N^2 \mid 0 \le u \le s_x \text{ and } 0 \le v \le s_y\}$ และ (s_x,s_x) เป็นขนาดของหน้ากากคอนโวลูชัน

$w_{ji}^{(i)}(u,v)$ เป็นหน้ากากที่ใช้ในการคอนโวลูชันในชั้น l

$y_i^{(i-1)}(x+u,y+v)$ เป็นอินพุต

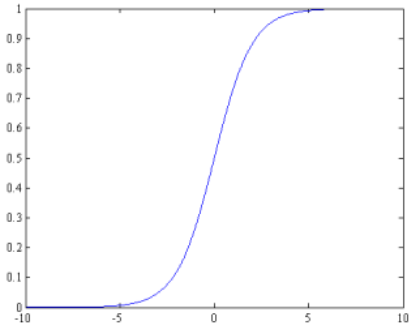
$b_j^{(i)}$ เป็นค่าไบแอส

$\emptyset^{(1)}$ เป็นฟังก์ชันกระตุ้น (Activation function)

โดยฟังก์ชันกระตุ้นที่นิยมใช้กันมี 3 ชนิดด้วยกันได้แก่

2.1) Sigmoid Function สมการในของ Sigmoid Function แสดงได้ดังสมการที่ 2.2 และแสดงการทำงานของ Sigmoid Function แสดงได้ดังภาพประกอบที่ 2.3

$$f(x) = \frac{1}{1+e^{-x}} \qquad (2.2)$$

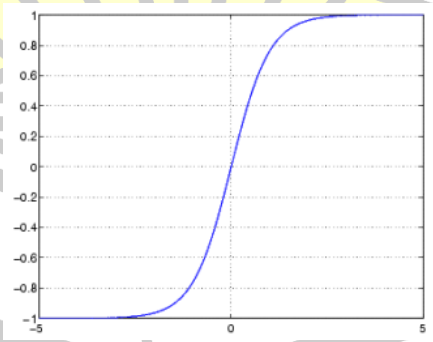


ภาพประกอบที่ 2.3 แสดงการทำงานของ Sigmoid Function

จะเห็นว่าค่าในแกน X มีค่าอยู่ในช่วง -10 ถึง 10 และแกน Y จะมีค่าตั้งแต่ 0 ถึง 1

2.2) Tan Hyperbolic Function (Tanh Function) สมการในการดำเนินการ ของ Tanh Function แสดงได้ดังสมการที่ 2.3 และผลการดำเนินการแสดงได้ดังภาพประกอบที่ 2.4

$$f(x) = \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad (2.3)$$

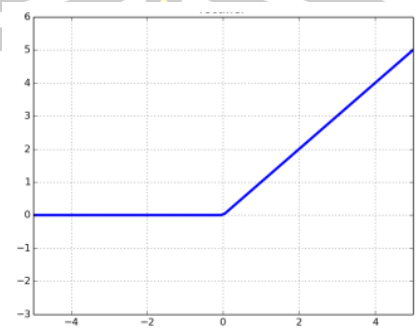


ภาพประกอบที่ 2.4 แสดงการทำงานของ Tanh Function

จะเห็นว่าค่าในแกน X มีค่าอยู่ในช่วง -5 ถึง 5 และแกน Y จะมีค่าตั้งแต่ -1 ถึง 1 ซึ่งแตกต่างจาก Sigmoid Function ที่มีค่าในแกน Y อยู่ในช่วงบวกเท่านั้น

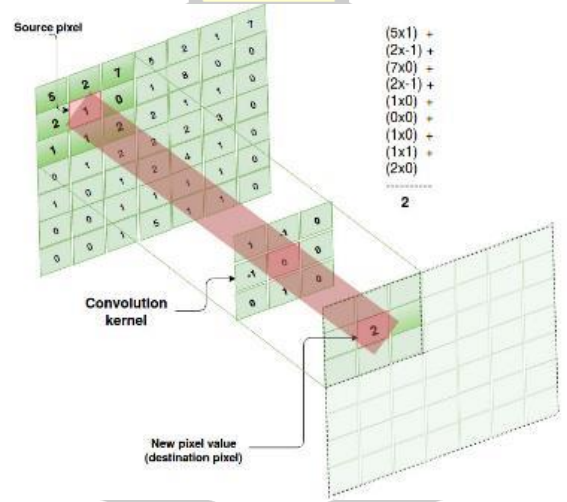
2.3) Rectified Linear Unit (ReLU Function) สมการในการดำเนินการของ ReLU Function แสดงได้ดังสมการที่ 2.4 และผลการดำเนินการแสดงได้ดังภาพประกอบ 2.5

$$f(x) = \max(0,x) \quad (2.4)$$



ภาพประกอบที่ 2.5 แสดงการทำงานของ ReLU Function

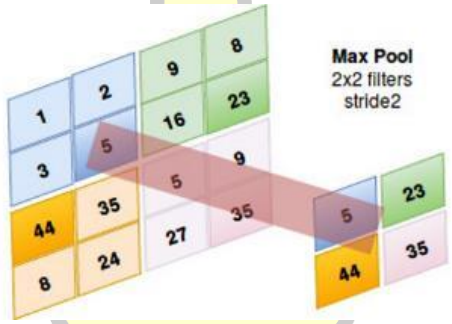
จะเห็นว่าเมื่อค่าในแกน X มีค่าติดลบ ค่าในแกน Y จะมีค่าเป็น 0 และเมื่อค่าใน แกน X เปลี่ยนแปลงเข้าสู่ค่าบวกค่าในแกน Y จะเท่ากับค่าในแกน X นั้น จากฟังก์ชันกระตุ้นทั้ง 3 แบบที่ได้กล่าวไปนั้น ReLU Function [7] ถูกนำมาใช้งานกับ CNN มากกว่าฟังก์ชันอื่น เนื่องจากลดโอกาสที่บางข้อมูลจะสูญหาย โดยที่หากข้อมูลมีค่ามากกว่า 0 จะเลือกมาเป็นเอาท์พุต แต่สำหรับกรณีข้อมูลมีค่าน้อยกว่า 0 จะให้ค่าเอาท์พุตเป็น 0 และตัวอย่างการทำคอนโวลูชัน (Convolution) ระหว่างอินพุตขนาด 7×7 พิกเซล กับหน้ากากขนาด 3×3 สามารถแสดงได้ดังรูปที่ 2.6



ภาพประกอบที่ 2.6 ตัวอย่างการทำงานของคอนโวลูชัน

เป็นตัวอย่างการคอนโวลูชันระหว่างอินพุตขนาด 7×7 พิกเซล กับหน้ากากขนาด 3×3 ทำให้ได้เอาท์พุตขนาด 3×3 พิกเซล ซึ่งเอาท์พุตเกิดจากการนำค่าสัมประสิทธิ์ของหน้ากากไปคูณกับค่าของอินพุตจากนั้นนำทุกค่าที่ได้บวกกันเพื่อเป็นเอาท์พุต จากตัวอย่างจะเห็นว่าหน้ากากมีสัมประสิทธิ์ w x y และ z ไปคอนโวลูชันกับอินพุต a b e และ f ได้เอาท์พุตเป็น aw+bx+ey+fz จากนั้นเลื่อนหน้ากากไปยังข้อมูลอินพุตถัดไป และทำกระบวนการเดิมจนครบขนาดของอินพุต โดยเอาท์พุตจะมีขนาดลดลงเนื่องจากกระบวนการนี้จะสูญเสียขอบของภาพไป

3) ชั้นพูลลิ่ง (Pooling Layer) ในขั้นตอนนี้เป็นการลดขนาดของข้อมูลลงโดยวิธีการทำพูลลิ่งที่นิยมนั้นมี 3 แบบด้วยกัน พูลลิ่งด้วยค่าสูงสุด (Max Pooling) การพูลลิ่งด้วยค่าเฉลี่ย (Average Pooling) และการพูลลิ่งด้วย L2 Norm (L2 Norm Pooling) ในการทำวิจัยนี้จะนำเสนอเพียงการทำพูลลิ่งด้วยค่าสูงสุด (Max Pooling) การทำพูลลิ่งด้วยค่าสูงสุดนั้นจะเลือกค่าสูงสุดในแต่ละบล็อคมาเป็นคำตอบ ตัวอย่างการทำงานของการทำพูลลิ่งด้วยค่าสูงสุดแสดงได้ดังรูปที่ 2.7



ภาพประกอบที่ 2.7 การทำพูลลิ่งด้วยค่าสูงสุด

การทำงานของพูลลิ่งด้วยค่าสูงสุดระหว่างอินพุตขนาด 4×4 พิกเซล กับหน้ากากพูลลิ่งขนาด 2×2 ขั้นตอนการทำงานคือ นำหน้ากากขนาด 2×2 ไปครอบที่อินพุต จากนั้นจะเลือกค่าสูงสุดออกมา ซึ่งจะได้เอาท์พุตเป็น 6, 7, 8 และ 9 ตามลำดับ โดยขนาดของเอาท์พุตจะมีขนาดลดลงครึ่งหนึ่งของขนาดอินพุต

4) ชั้นลักษณะเด่น (Fully-Connected Layer) เมื่อผ่านกระบวนการทำงานของชั้นคอนโวลูชัน ชั้นพูลลิ่งมาแล้วต่อมาจะนำเอาท์พุตจากชั้นเหล่านี้เข้าสู่ชั้นลักษณะเด่น เพื่อเตรียมตัวเข้าสู่ชั้นจำแนกต่อไป โดยชั้นนี้จะเชื่อมโยงจะนำเอาเซลล์ประสาทจากชั้นก่อนหน้า (ชั้นคอนโวลูชันและชั้นพูลลิ่ง) มาเชื่อมโยงกันกับเซลล์ประสาททุกเซลล์ที่มี โดยจะเชื่อมต่อกันเป็น 1 มิติ ดังนั้นจะไม่สามารถมีชั้นคอนโวลูชันหลังจากชั้นนี้ได้อีก โดยชั้นนี้จะทำเชื่อมต่อไปยังชั้นเอาท์พุต

5) ชั้นเอาท์พุต (Output Layer) โดยชั้นนี้จะเป็นชั้นของการจำแนกข้อมูล โดยจำนวนเซลล์ประสาทในชั้นนี้จะเท่ากับจำนวนพรรณไม้ที่ใช้ในการรู้จำภาพพรรณไม้ ซึ่งสมการที่ใช้ในการจำแนกแสดงได้สมการที่ 2.5

$$y = \frac{\exp(y_1)}{\sum_{j=1}^{n} \exp(y_1)} \qquad (2.5)$$

เมื่อ $y_1 = \sum_{i=1}^{m} x_1 \cdot w_{ij_1} + b_j$

โดยที่ m เป็นจำนวนเซลล์ประสาทในชั้นลักษณะเด่น

v_i เป็นอินพุตของเซลล์ประสาท j

y_i เป็นเอาท์พุต

n เป็นจำนวนบุคคลที่ต้องการจำแนก

หากพิจารณาภาพรวมการทำงานของ CNN พบว่าการทำงานคล้ายกับการทำงานของโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ แต่แตกต่างกันที่การคำนวณของชั้นคอนโวลูชันและชั้นพูลลิ่ง และเมื่อมีการคำนวณค่าความผิดพลาด จะนำไปปรับทั้งค่าน้ำหนัก และค่าสัมประสิทธิ์ในหน้ากากคอนโวลูชันด้วย

ในส่วนต่อมาจะกล่าวถึงเรื่องของโครงสร้างและตัวอย่างการทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชันแบบอื่น ๆ (CNN Architectures) ซึ่งเป็นโครงสร้างที่ปรับปรุงการทำงานมาจากโครงสร้างของ LeNet-5 [6] เพื่อให้การประมวลผลมีประสิทธิภาพเพิ่มมากขึ้น โดยอาศัยโครงสร้างของ CNN (LeNet-5) เป็นต้นแบบ

2.4.2 AlexNet Architectures

ปี ค.ศ. 2012 นักวิจัย Krizhevsky el al. [7] นำเสนอ Deep CNN ที่มีจำนวนทั้งสิ้น 8 Layer โดยโครงสร้างนี้เรียกว่า AlexNet ประกอบไปด้วย Convolutional Layer และ Fully-Connected Layer จำนวน 5 และ 3 Layer ตามลำดับ โดยผลลัพธ์ของ Fully-Connected Layer ถูกกำหนดให้มีจำนวน 100 โหนด เนื่องจากในงานวิจัยได้ใช้ชุดข้อมูล ImageNet LSVRC-2010 ในการทดสอบ ข้อมูลมีจำนวนทั้งสิ้น 1.2 ล้านรูปภาพ และแบ่งออกเป็น 1,000 Class ใน Layer สุดท้ายของ Fully-Connected Layer ใช้วิธี Softmax เพื่อคำนวณหาผลลัพธ์

รูปภาพที่ใช้ในการทดลองจะถูกเปลี่ยนให้มีขนาด 224x224x3 พิกเซล ดังนั้น Convolutional Layer (Conv{1}) ในชั้นแรกถูกกำหนดให้มี 96 kernel โดยที่แต่ละ Kernel มีขนาด 11x11x3 ในการคำนวณจะเลื่อน (Stride) Kernel ครั้งละ 4 พิกเซล ผลลัพธ์จากการ Convolution จะมีขนาด 55x55x96 ชั้นที่สองของ Convolutional Layer (Conv{2}) ถูกกำหนดให้มี 256 Kernel โดยที่แต่ละ Kernel มีขนาด 5x5x48 ผลลัพธ์จากการ Convolution จะมีขนาด

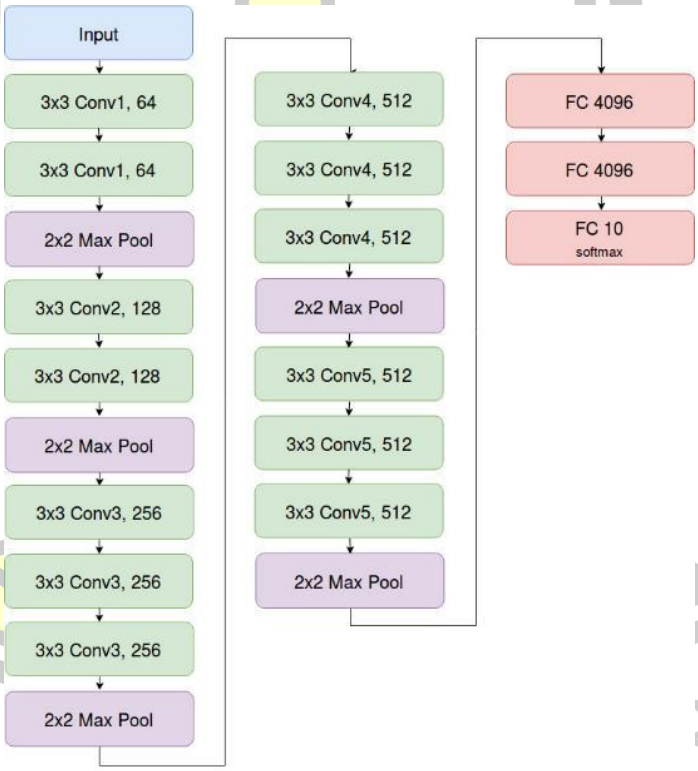
27x27x256 หลังจากการคำนวณ (Conv{1, 2}) ข้อมูลจะถูกส่งไป Pooling Layer เพื่อทำการ Normalized และใช้ Max Pooling เพื่อหาค่าสูงสุดในแต่ละ Pool (Max Pooling) โดยที่แต่ละ Pool จะถูกแบ่งเป็น 2x2 ส่วน

ใน Convolutional Layer ชั้นที่สาม (Conv{1}) จำนวนของ Kernel ถูกกำหนดให้มี 384 kernel โดยที่แต่ละ Kernel มีขนาด 3x3x256 ผลลัพธ์จากการ Convolution จะมีขนาด 13x13x256 โดยชั้นที่สี่ ถูกกำหนดให้มี 384 Kernel โดยที่แต่ละ Kernel มีขนาด 3x3x192 ผลลัพธ์จากการ Convolution จะมีขนาด 13x13x384 ชั้นที่ห้า มีจำนวน 256 Kernel โดยแต่ละ Kernel มีขนาด 3x3x192 สุดท้ายผลลัพธ์จากการ Convolution มีขนาด 13x13x256 โดยที่ Conv{1-5} ใช้ฟังก์ชัน Rectified Linear Units (ReLu) เป็นฟังก์ชันกระตุ้น (Activation Function) สุดท้ายแล้ว Fully-Connected Layer ถูกแบ่งออกเป็น 2 ชั้น และกำหนดให้มีจำนวนชั้นละ 4,096 Node โครงสร้าง AlexNet แสดงดังภาพประกอบที่ 2.8

ภาพประกอบที่ 2.8 แสดงโครงสร้างการทำงานของ AlexNet

2.4.3 VGGNet Architectures

ในปี ค.ศ. 2015 งานวิจัย [8] ได้นำเสนอโครงสร้างแบบ VGGNet ซึ่งเป็นโครงสร้างแบบ Deep CNN เนื่องจากมีความลึกของ Layer จำนวน 16 ชั้น โดยใช้ Kernel หรือ Mash ขนาด 3x3 ในการ Convolution โครงสร้างของ VGGNet-16 ประกอบด้วย Convolutional Layer (Conv) จำนวน 5 กลุ่ม โดย Conv กลุ่มที่ 1 หรือ Conv1 ประกอบด้วย (Conv1 {1, 2}) และมีจำนวน Kernel ชั้นละ 64 Kernel ใน Conv2 ประกอบด้วย (Conv2{1, 2}) มีจำนวน Kernel ชั้นละ 128 Kernel ใน Conv3 ประกอบด้วย (Conv3{1, 2, 3}) และมีจำนวน Kernel ชั้นละ 256 Kernel ใน Conv4 ประกอบด้วย (Conv4{1, 2, 3}) และมีจำนวน Kernel ชั้นละ 512 kernel และชั้น Conv5 ประกอบด้วย (Conv5{1, 2, 3}) ในแต่ละชั้นมีจำนวน Kernel ชั้นละ 512 Kernel โดยที่ใน (Conv{1-5}) จะตามด้วย Max Pooling Layer จากนั้นข้อมูลจะส่งไปยัง Fully-Connected Layer ที่มีจำนวนโหนด 4,096 และ 4,096 โหนด ดังภาพที่ 2.9

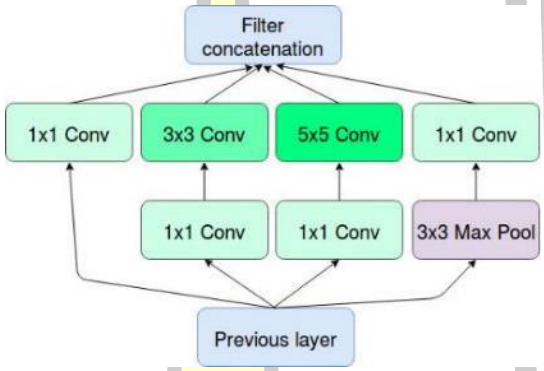


ภาพประกอบที่ 2.9 แสดงโครงสร้างการทำงานของ VGGNet

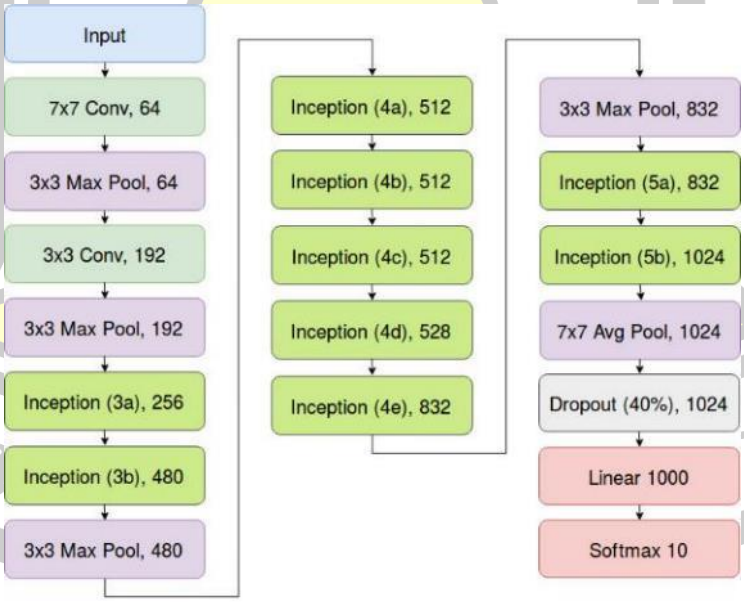
2.4.4 GoogLeNet

โครงสร้างแบบ GoogLeNet [9] ถูกนำเสนอในปี ค.ศ. 2014 โดยใช้โครงสร้าง Inception (Inception Architecture) เพื่อให้ CNN มีความลึกและกว้าง อีกทั้งยังไม่ทำให้การ

ประมวลผลช้าลง ในแต่ละ Inception ประกอบไปด้วยการคำนวณ Convolution ที่ใช้ Kernel ขนาด 1x1, 3x3 และ 5x5 ซึ่งเป็นการประมวลผลแบบขนาน (Parallel) อีกทั้งยังทำให้จำนวนมิติของข้อมูลลดลง (Dimension Reduction) (แสดงดังภาพประกอบ 2.10) โครงสร้างแบบ GoogLeNet ถูกออกแบบให้มีโครงสร้างแบบ Inception จำนวน 9 Layer จากจำนวนทั้งสิ้น 22 Layer และใช้ฟังก์ชัน ReLu เพื่อเป็นฟังก์ชันกระตุ้น และ Fully-Connected Layer กำหนดให้มีจำนวน 1024 Node ก่อนจะส่งต่อไปยัง Output Layer โครงสร้างแบบ GoogLeNet แสดงดังภาพประกอบ 2.11



ภาพประกอบที่ 2.10 แสดงโครงสร้างแบบ Inception

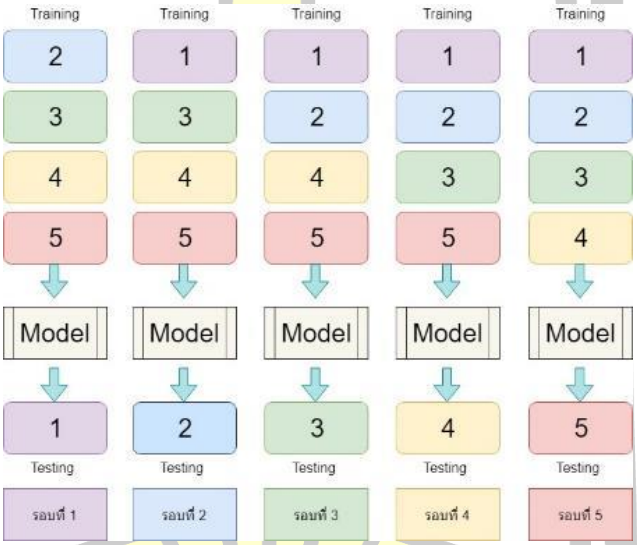


ภาพประกอบที่ 2.11 แสดงโครงสร้างการทำงานของ GoogLeNet

2.5 การวัดประสิทธิภาพของโมเดล (Performance Measurement)

Cross-Validation [10] เป็นวิธีนี้เป็นวิธีที่นิยมในการทำงานวิจัย เพื่อใช้ในการทดสอบประสิทธิภาพของโมเดลเนื่องจากผลที่ได้มีความน่าเชื่อถือ การวัดโดยแบ่งกลุ่มข้อมูลออกเป็น k กลุ่ม เช่นกำหนด k = 5-Fold Cross-Validation และแบ่งข้อมูล K1 กลุ่มสำหรับใช้ในการเรียนรู้ (Training Data) และ K1 กลุ่มสำหรับใช้ในการทดสอบ (Test Data)

Stratified Cross-Validation [19] เป็นการทดสอบไขว้ที่ปรับปรุงจากวิธี k-Fold Cross-Validation โดยในขั้นตอนการสุ่มข้อมูลจะคำนึงถึงสัดส่วนการกระจายข้อมูลในแอททริบิวต์ที่เป็นคลาสเป้าหมาย ข้อมูลที่ถูกสุ่มให้เป็น Training Data และ Test Data ในแต่ละ Fold จะมีสัดส่วนของคลาสไม่แตกต่างกันจากข้อมูลดั้งเดิม ในการวิจัยนี้จะใช้วิธี Stratified Cross-Validation ในการทดสอบประสิทธิภาพของโมเดล แสดงดังภาพประกอบ 2.12



ภาพประกอบที่ 2.12 แสดงการทำงานของ Cross-Validation

จากนั้นจึงตรวจสอบอัตราความถูกต้อง โดยใช้ตัววัดประสิทธิภาพ ดังนี้

ตารางที่ 2.1 แสดงตาราง *Confusion Matrix* ที่มีสองคลาส

Predicted/actual	**Yes**	**No**
Yes	TP	FP
No	FN	TN

โดยกำหนดให้ True Positive (TP) คือ จำนวนข้อมูลที่ทำนายถูกว่าเป็นคลาส = yes

True Negative (TN) คือ จำนวนข้อมูลที่ทำนายถูกว่าเป็นคลาส = no

False Positive (FP) คือ จำนวนข้อมูลที่ทำนายผิดมาเป็นคลาส = yes

False Negative (FN) คือ จำนวนข้อมูลที่ทำนายผิดมาเป็นคลาส = no

2.5.1 Precision เป็นการวัดความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คำนวณได้จากสมการ 2.6

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Positive}} \quad (2.6)$$

2.5.2 Accuracy เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส คือ จำนวน True Positive ของทุกคลาสรวมกัน

2.6 งานวิจัยที่เกี่ยวข้อง (Literature Review)

Liu et al. [20] ได้นำเสนอวิธีการจัดหมวดหมู่ดอกไม้ (Flower Classification) โดยใช้วิธี CNN ซึ่งได้นำไปใช้กับข้อมูลรูปภาพดอกไม้ทั้งสิ้นจำนวน 63,442 รูปภาพ โดยแบ่งออกเป็น 79 สายพันธุ์ โดยโครงสร้างของ CNN นั้นประกอบด้วย 8 ชั้น (Layer) โดยแบ่งออกเป็นชั้น Convolutional จำนวน 5 Layer และอีก 3 Layer เป็นชั้นของ Fully-Connected ซึ่ง Output สุดท้ายจากชั้น Fully-Connected จำนวน 79 Output ถูกคำนวณด้วยวิธี Softmax ซึ่งผลลัพธ์ทั้งหมดก็คือจำนวนกลุ่มของดอกไม้ โดยงานวิจัยนี้ได้ทดสอบกับข้อมูลชุด Oxford 102 Flower Species ที่มีจำนวนกลุ่มทั้งสิ้น 102 สายพันธุ์ โดยมีรูปภาพดอกไม้ทั้งสิ้น 8,189 รูปภาพ และแต่ละสายพันธุ์จะมีรูปภาพของดอกไม้ตั้งแต่ 40 ถึง 200 รูปภาพ ปรากฏว่า มีอัตราความถูกต้อง 84.02% จากนั้นได้ทดสอบกับข้อมูลชุด Large-Scale 79 Flower Species ซึ่งมีดอกไม้ทั้งสิ้น 79 สายพันธุ์ และมีจำนวนรูปภาพทั้งสิ้น 52,775 รูปภาพ โดยได้แบ่งรูปภาพจำนวน 47,500 เป็นข้อมูลชุดเรียนรู้ และ 5,275 รูปภาพเป็นข้อมูลชุดทดสอบ พบว่าวิธี CNN ที่นำเสนอในงานวิจัยนี้มีอัตราความถูกต้อง 76.54% เมื่อเทียบกับวิธี CNN แบบง่ายที่มีอัตราความถูกต้อง 70.12%

Pawara et al. [21] ได้เปรียบเทียบวิธีการหาคุณลักษณะพิเศษแบบเฉพาะพื้นที่ (Local Descriptor) และวิธี (Back of Word: BOW) กับวิธีการเรียนรู้เชิงลึกสำหรับใช้รู้จำพรรณไม้ (Plant Recognition) โดยชุดข้อมูลพรรณไม้ที่ใช้ในการทดสอบประกอบด้วย AgrilPlant (10 สายพันธุ์ 3,000 รูปภาพ), LeafSnap (184 สายพันธุ์ 7,719 รูปภาพ) และ Folio (32 สายพันธุ์ 640 รูปภาพ) ในการทดลองได้ปรับรูปภาพให้มีขนาด 256x256 พิกเซล ข้อมูลถูกแบ่งออกเป็น 80% สำหรับข้อมูลชุดเรียนรู้ และ 20% สำหรับข้อมูลชุดทดสอบ และใช้วิธี Cross-Validation โดยกำหนดให้ k มีค่า

เท่ากับ 5 เพื่อใช้สำหรับการทดสอบประสิทธิภาพของแต่ละวิธี ในงานวิจัยนี้ ได้ทดสอบวิธี HOG ซึ่งเป็น Local Descriptor ประเภทหนึ่ง และวิธี HOG-BOW ซึ่งเป็นการนำวิธี HOG และ BOW มารวมกันเพื่อสร้างคุณลักษณะพิเศษ จากนั้นจึงนำไปจัดหมวดหมู่ข้อมูลด้วยวิธี k-Nearest Neighbors (KNN), Support Vector Machine (SVM) และ Multi Layer Perceptron (MLP) ซึ่งวิธีการ HOG-BOW ร่วมกับ MLP ให้ประสิทธิภาพสูงสุดในข้อมูลชุด LeafSnap แต่ HOG-BOW ร่วมกับ SVM ให้ประสิทธิภาพสูงสุดในข้อมูลชุด AgrilPlant และ Folio และได้ทดสอบวิธีการเรียนรู้เชิงลึกโดยใช้โครงสร้างของ AlexNet และ GoogLeNet จากการทดลองพบว่า AlexNet ให้อัตราความถูกต้องสูงสุด 97.67% สำหรับข้อมูลชุด Folio สำหรับข้อมูลชุด AgrilPlant และ LeafSnap วิธี GoogLeNet มีอัตราความถูกต้องสูงสุดที่ 98.33% และ 97.66% ตามลำดับ และมีอัตรารู้จำข้อมูลชุด Folio ที่ 97.63% ซึ่งใกล้เคียงกับวิธี AlexNet

Okafor et al. [22] ได้เปรียบเทียบวิธีการที่ใช้สำหรับรู้จำสัตว์ (Animal Recognition) โดยเปรียบเทียบระหว่างวิธีการเรียนรู้เชิงลึก (Convolutional Neural Networks: Deep CNN) และวิธี (Back of Word: BOW) โดยการทดสอบวิธีการ BOW นั้นได้ทดสอบทั้งสิ้นสองรูปแบบคือวิธี BOW และ Histograms of Oriented Gradients BOW (HOG-BOW) ซึ่งได้ทดสอบทั้งภาพสีและภาพสีเทา และใช้วิธีการ L2-SVM ในการจัดหมวดหมู่ของรูปภาพสัตว์ สำหรับการทดสอบด้วยวิธี CNN นั้นได้ทดสอบกับโครงสร้าง (Architect) แบบ AlexNet และ GoogLeNet โดยการลดจำนวนของนิวรอล (Neurons) ในแต่ละชั้นของโครงสร้างลง โดยได้ทดสอบกับข้อมูลชุด Wild-Anim ซึ่งมีจำนวนทั้งสิ้น 5 กลุ่ม (Class) และมีรูปภาพทั้งสิ้น 5,000 รูปภาพ ซึ่งรูปภาพทุกรูปเป็นภาพสีแบบ RGB และปรับให้มีขนาด 250x250 พิกเซล สำหรับการทดสอบ ได้แบ่งข้อมูลออกเป็น 5 ชุด ด้วยวิธีการของ Cross-Validation เพื่อทดสอบความถูกต้องและความเบี่ยงเบนมาตรฐาน จากงานวิจัย แสดงให้เห็นว่าวิธี CNN โดยใช้โครงสร้างของ GoogLeNet ให้ประสิทธิภาพที่ดีที่สุดโดยมีความถูกต้อง 99.93% เมื่อเทียบกับวิธี BOW มีอัตราความถูกต้องที่ 84% แต่เมื่อทำการลดจำนวนของนิวรอลลงปรากฏว่าอัตราการรู้จำด้วย GoogLeNet กลับลดลงเป็น 99.38%

Sladojevic et al. [23] นำเสนอวิธีการโครงข่ายประสาทเทียมเชิงลึก (Convolutional Neural Networks: Deep CNN) เพื่อการรู้จำพันธุ์พืชที่เป็นโรค (Plant Disease Recognition) โดยการจัดหมวดหมู่จากใบไม้ (Leaf Image Classification) ซึ่งทำการรู้จำพืชที่เป็นโรคจำนวนทั้งสิ้น 13 ประเภทและรวมกับสภาพใบที่เป็นปกติอีก 2 ประเภท รวมเป็น 15 กลุ่ม โดยตรวจสอบจากใบไม้เป็นหลักโดยใช้วิธี CNN ซึ่งรูปภาพใบไม้ที่นำมาใช้ในการทดสอบนั้นถูกนำไปประมวลผลเพื่อเพิ่มเติมข้อมูลในการเรียนรู้ (Augmented) โดยใช้วิธีการประมวลผลภาพ ได้แก่ การนำภาพไปบิดให้ผิดรูป (Distortion) และการหมุนภาพ (Rotation) เป็นต้น ซึ่งข้อมูลที่ใช้ในการทดสอบนั้นประกอบด้วย

4,483 รูปภาพ และเมื่อนำไปผ่าน Data Augmented แล้วได้รูปภาพเพิ่มขึ้นเป็น 30,880 รูปภาพ จากการทดลองพบว่ามีอัตราความถูกต้อง 96.3%

Reyes et al. [24] นำเสนอวิธีการโครงข่ายประสาทเทียมเชิงลึก (Convolutional Neural Networks: CNN) เพื่อการรู้จำชนิดและประเภทของพืชในเวลาเดียวกัน โดยทำการทดสอบกับชุดข้อมูล ILSVRC 2012 ซึ่งมีชุดข้อมูลขนาดใหญ่มีภาพจำนวน 1,200,000 ภาพ และมีทั้งหมด 1,000 หมวดหมู่ โดยการปรับแต่งการตั้งค่าเพื่อให้การรู้จำพืชมีความแม่นยำสูงสุดเท่าที่จะเป็นได้ งานวิจัยนี้ได้ใช้เทคนิคการปรับแต่งการตั้งค่าจากชุดข้อมูล LifeCLEF 2015 ได้ภาพในการนำไปทำการรู้จำพืช จำนวน 91,759 ซึ่งกระจ่ายอยู่ในพืช 13,887 โดยตัวอย่างนี้มี 1000 สปีชี่ species ตัวอย่างเช่น สมุนไพร และเฟิร์น และได้ภาพสำหรับการนำไปทดสอบ 21,446 ภาพ โดยภาพที่จะทำการรู้จำนั้นประกอบด้วยการถ่ายภาพ 7 มุมมองที่ต่างกัน เช่น กิ่ง ดอก ก้านดอก ผล ใบ กลุ่มของใบ และภาพรวมของต้นไม้ จากผลการทดลองซึ่งแสดงการวัดประสิทธิภาพความถูกต้องใน 7 มุมมองนั้น ภาพของดอกไม้และกลุ่มของใบไม้มีค่าสูงสุดตามลำดับ ซึ่งมากกว่ามุมมองอื่น ๆ โดยมีค่าความแม่นยำเฉลี่ยเท่ากับ 97.20%

Mohanty et al. [25] ได้นำเสนอวิธีการวิเคราะห์และการวินิจฉัยโรคโรคพืชจากภาพถ่ายโดยใช้เทคนิคการเรียนรู้เชิงลึก (Convolutional Neural Networks: CNN) โดยในยุคของสมาร์ทโฟนเป็นที่แพร่หลายกล้องในสมาร์ทโฟนมีความชัดระดับ HD และหน่วยประมวลผลที่มีประสิทธิภาพสูงบนอุปกรณ์เคลื่อนที่ จึงนำไปสู่การวินิจฉัยโรคโรคพืชในการเรียนรู้เชิงลึกโดยอัตโนมัติ ในงานวิจัยนี้ได้ทำการทดสอบบนชุดข้อมูล PlantVillage ซึงมีจำนวน 54,306 ภาพจากพืช 14 ชนิด ที่ระบุชนิดของโรคไว้ 26 โรค เพื่อที่จะพัฒนาเทคนิคที่สามารถจำแนกประเภทโรคพืชจากภาพถ่ายที่ถูกต้องจากภาพของพืชที่เป็นโรคและพืชที่มีสุขภาพดี โดยปรับขนาดภาพเป็น 256 × 256 พิกเซลและได้ทดสอบวิธีการ CNN โดยใช้โครงสร้างของ AlexNet และ GoogLeNet ใช้กลไกในการเรียนรู้ (training) ของ Transfer Learning และ Training from Scratch เลือกใช้เทคนิคในการหาลักษณะเด่นที่สนใจ Future ได้แก่ การหาจากภาพสี (Color) ภาพโทนสี่เทา (Grayscale) และการแบ่งส่วนของภาพ (Segmented) ในส่วนของการทดสอบประสิทธิภาพ ใช้วิธี Cross-Validation โดยแบ่งกลุ่มข้อมูลออกเป็น 5 กลุ่ม จากการทดลองพบว่าการใช้โครงสร้าง GoogLeNet ใช้กลไกในการเรียนรู้ (training) ของ Transfer Learning และการแบ่งกลุ่มทดสอบ (k1)Train: 80%, Test: 20% พบว่ามีอัตราความถูกต้องสูงที่สุดเท่ากับ 99.34%

Sun et al. [26] ได้นำเสนอวิธีการจำแนกพรรณพืช (Plant Identification) โดยใช้เทคนิคการเรียนรู้เชิงลึกแบบ (Convolution Neural Network: CNN) ได้ทำการทบทวนวรรณกรรมในเรื่องของการจำแนกพรรณพืช การจำแนกพืชโดยใช้ใบจะมีอัตราการรู้จำมากที่สุด นักวิจัยบางกลุ่มจะเน้นไปที่ดอกไม้และหาลักษณะเด่นโดยใช้รูปร่าง สี และลักษณะพื้นผิวอื่น ๆ โดยในหลายปีต่อมาก็ได้มี

การนำเทคโนโลยีมาใช้ในการจดจำพืชโดยเฉพาะแอพพลิเคชั่นบนโทรศัพท์มือถือ เช่น LeafSnap [27], PlantNet [28] และแอพพลิเคชั่นสำหรับจดจำดอกไม้ของ Microsoft Garage Recognition app [29] สามารถจำแนกพืชได้โดยอัตโนมัติ แต่แอพพลิเคชั่นดังกล่าวก็ยังมีข้อจำกัดในระบบนิเวศน์วิทยาอยู่คือใช้ภาพถ่ายในห้องทดลองที่มีการควบควบปัจจัยภายนอก งานวิจัยนำเสนอการจำแนกพืชโดยใช้ภาพถ่ายพืชที่อยู่ในสภาวะแวดล้อมของธรรมชาติ ซึ่งเป็นการแก้ไขปัญหาข้อจำกัดต่าง ๆ ในการจำแนกเช่น กล้องถ่ายภาพ พื้นที่ ช่วงเวลาการเติมโตของพืชแต่ละชนิด ทำการทำสอบบนชุดข้อมูล BJFU100 เป็นภาพในสภาพแวดล้อมทางธรรมชาติ 10,000 ภาพ ซึ่งเป็นไม้ประดับจำนวน 100 ชนิดในมหาวิทยาลัยปักกิ่ง จำแนกโดยใช้เทคนิคการเรียนรู้เชิงลึก ประกอบด้วย 26 ชั้น (Layer) ทดสอบประสิทธิภาพ ใช้วิธี Cross-Validation Train: 80%, Test: 20% โดยผลการทดสอบพบว่ามีอัตราความถูกต้อง 91.78% และเมื่อนำเทคนิคที่ได้จากการวิจัยนี้ไปทดสอบกับชุดข้อมูล Flavia Dataset โดยกำหนดเลเยอร์ เป็น 18, 26, 34 และ 50 ชั้น โดยใช้โครงสร้าง ResNet, ResNet26 model มีอัตราความถูกต้อง 91.78%

มนัสกานต์ เสน่หา และคณะ [30] ได้นำเสนอวิธีการทำนายอายุการใช้งานคงเหลือของเครื่องจักรโดยใช้เทคนิคการเรียนรู้เชิงลึกแบบ (Convolution Neural Network: CNN) ที่เพิ่มประสิทธิภาพ ผู้วิจัยได้แบ่งวิธีการทดลองเป็น 2 วิธี โดยวิธีแรก ใช้โครงข่ายประสาทเทียมแบบพื้นฐานประกอบด้วยโครงสร้างแบบ (Multi-Layer Perceptron: MLP), LeNet-5 Lite, AlexNet Lite, VGG-16 Lite วิธีที่สอง ใช้โครงสร้างแบบ CNN ประกอบด้วยโครงสร้างแบบ Enhanced LeNet-5, Enhanced AlexNet, Enhanced VGG-16 มาเพิ่มเทคนิคต่างๆ ได้แก่ เพื่อเพิ่มประสิทธิภาพในการทำนาย จากการทดลองเพื่อเปรียบเทียบประสิทธิภาพการทำนายอายุการใช้งานคงเหลือของเครื่องจักร วิธี Enhanced VGG-16 มีประสิทธิภาพมากที่สุด มีอัตราความถูกต้องที่ 85 %

พหุน ปณฺฑิโต ชีเว

บทที่ 3
วิธีดำเนินการวิจัย

ในบทนี้กล่าวถึงการออกแบบของระบบจำแนกพรรณไม้โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) ที่มีโครงสร้างต่างกัน (Architecture) จำนวน 4 รูปแบบ ได้แก่ LeNet [5], AlexNet [6], VGGNet [7] และ GoogLeNet [8] โดยจะทำการทดสอบเพื่อเปรียบเทียบประสิทธิภาพ โดยใช้ชุดข้อมูลรูปภาพที่สร้างขึ้นเองจำนวน 1 ชุด คือ ชุดข้อมูล PNE Dataset ที่มีจำนวนพรรณไม้ 10 ชนิด และจะใช้ชุดข้อมูลมาตรฐาน (Benchmark) จำนวน 2 ชุด ได้แก่ ข้อมูลชุด 102 Flower Dataset ที่มีจำนวนพรรณไม้ 102 ชนิด และชุดข้อมูล Folio Dataset ที่มีจำนวนพรรณไม้ 32 ชนิด และทดสอบประสิทธิภาพของวิธี CNN ทั้ง 4 รูปแบบด้วยวิธี Stratified Cross-Validation [19,31] โดยในบทนี้จะกล่าวถึงรายละเอียดวิธีการดำเนินการวิจัย โดยแบ่งการดำเนินการออกเป็น 3 ส่วนดังนี้

3.1 ชุดข้อมูลรูปภาพที่ใช้การวิจัย

3.2 การทดลองจำแนกพรรณไม้โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน

3.3 การทดสอบประสิทธิภาพ

3.1 ชุดข้อมูลรูปภาพที่ใช้การวิจัย

ข้อมูลพรรณไม้ที่ใช้ในงานวิจัยนี้มีทั้งสิ้น 3 ชุดข้อมูล ประกอบด้วยชุดข้อมูล Plants in Natural Environment (PNE), Folio และ 102 Flower

3.1.1 ชุดข้อมูลพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ (Plants in Natural Environment Dataset: PNE Dataset)

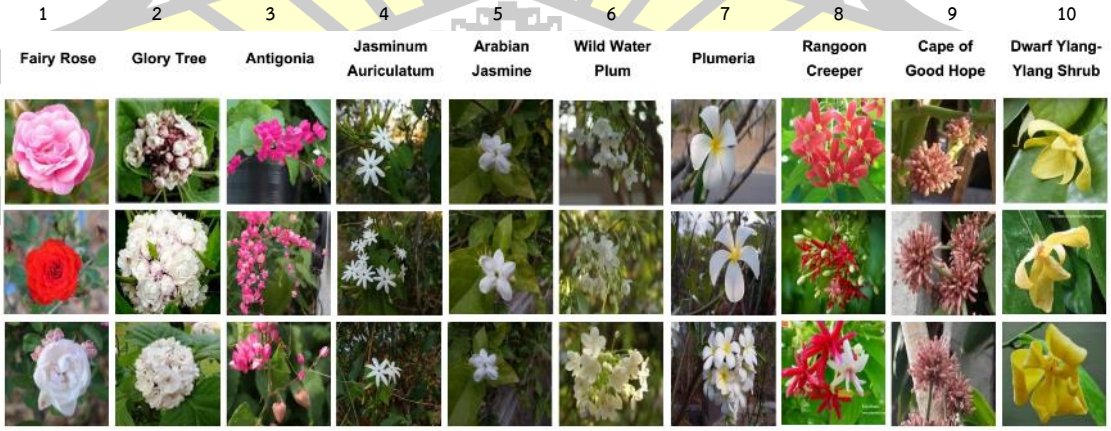
ชุดข้อมูล PNE ที่นำเสนอในงานวิจัยฉบับนี้ประกอบด้วยพรรณไม้จำนวน 10 ชนิด ชนิดละ 300 รูปภาพ ชนิดของพรรณไม้ที่อยู่ในชุดข้อมูล PNE แสดงดังในตารางที่ 3.1 และตัวอย่างของพรรณไม้แสดงดังภาพประกอบที่ 3.1 รูปภาพพรรณไม้เป็นรูปภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติที่เก็บรวบรวมจากสวนไม้หอมและบริเวณรอบมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ [2] โดยรูปภาพบางส่วนเก็บรวบรวมมาจากอินเทอร์เน็ต สุดท้ายแล้ว รูปภาพพรรณไม้ทุกรูปถูกนำมาคัดกรองและยืนยันความถูกต้องของสายพันธุ์ (Specie) โดยผู้ช่วยศาสตราจารย์พัชราพรณ์ แสงโยจารย์ นักพฤกษศาสตร์

ความท้าทายในการจำแนกประเภทรูปภาพพรรณไม้ในชุดข้อมูล PNE ได้แก่ ความแตกต่างของพรรณไม้ชนิดเดียวกัน แต่มีสีสันที่แตกต่างกัน เช่น ดอกกุหลาบหนู มีทั้งดอกสีชมพูอ่อน

ชมพู และแดง (ภาพประกอบที่ 3.1 แถวที่ 1) ทั้งนี้ มีดอกไม้จำนวน 5 ประเภทที่เป็นดอกไม้สีขาว ประกอบด้วย ดอกนางแย้ม พุทธชาด มะลิลา โมกพวง ลีลาวดี (ภาพประกอบที่ 3.1 แถวที่ 2, 4-7) ดอกพุทธชาดและมะลิลา (ภาพประกอบที่ 3.1 แถวที่ 4, 5) ที่รูปของดอกไม้ไม่ได้เป็นจุดเด่นของภาพ และดอกส่าเหล้าปัตตานี ที่แสดงใน (ภาพประกอบที่ 3.1 แถวที่ 10) แสดงถึงรูปร่าง (Shape) ที่มีความแตกต่างกัน โดยมีทั้งดอกตูม และดอกที่กำลังบานพรรณไม้ดังตารางที่ 3.1

ตารางที่ 3.1 แสดงพรรณไม้ที่รวบรวมไว้ในสวนสมุนไพรและสวนไม้หอม PNE Dataset

Class No.	Plant Name	Thai Plant Name
1	Fairy Rose	กุหลาบหนู
2	Glory Tree	นางแย้ม
3	Antigonia	พวงชมพู
4	Jasminum Auriculatum	พุทธชาด
5	Arabian Jasmine	มะลิลา
6	Wild Water Plum	โมกพวง
7	Plumeria	ลีลาวดี
8	Rangoon Creeper	เล็บมือนาง
9	Cape of Good Hope	วาสนา
10	Dwarf Ylang-Ylang Shrub	ส่าเหล้าปัตตานี

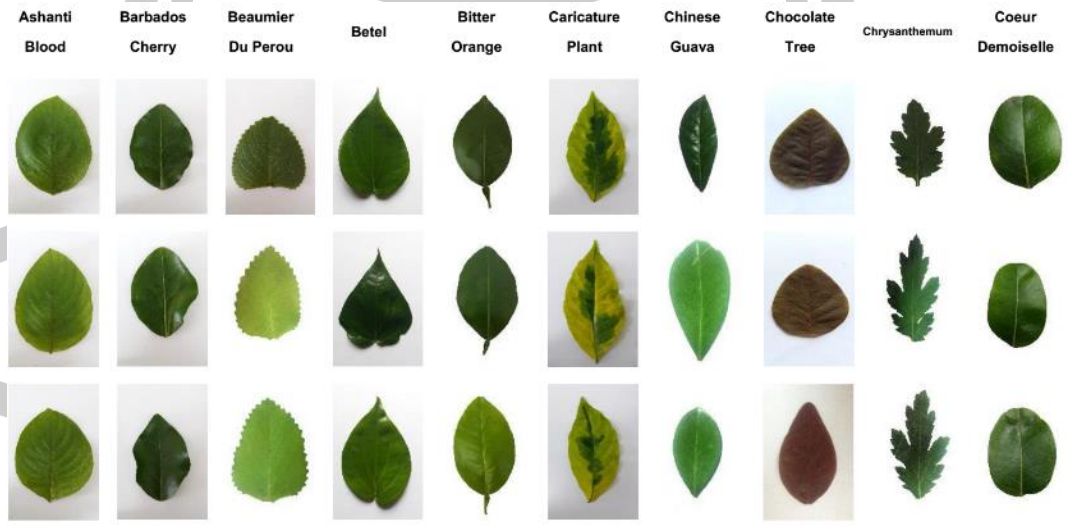


ภาพประกอบที่ 3.1 แสดงตัวอย่างภาพจากฐานข้อมูล PNE Dataset

3.1.2 ชุดข้อมูล 102 Flower (102 Flower Dataset) [33] เป็นรูปภาพของดอกไม้จำนวน 102 สายพันธุ์ โดยแต่ละสายพันธุ์ถูกเก็บรวบรวมตั้งแต่ 40-250 รูปภาพ รวมทั้งสิ้น 8,189 รูปภาพ โดย 102 Flower Dataset เป็นรูปภาพดอกไม้ที่พบทั่วไปในประเทศอังกฤษ (แสดงดังภาพประกอบที่ 3.2)

ภาพประกอบที่ 3.2 แสดงตัวอย่างภาพจากฐานข้อมูล 102 Flower Dataset

3.1.3 ชุดข้อมูล Folio (Folio Dataset) [34] ประกอบด้วยใบไม้ทั้งสิ้น 32 ชนิด ซึ่งเก็บรวบรวมชนิดละ 20 ใบ รวมแล้วมีรูปภาพใบไม้ทั้งสิ้น 640 รูปภาพ ที่เก็บรวบรวมจากบริเวณ University of Mauritius รูปภาพใบไม้ถูกถ่ายโดยกำหนดให้มีพื้นหลังเป็นฉากสีขาว ดังแสดงในภาพประกอบที่ 3.3



ภาพประกอบที่ 3.3 แสดงตัวอย่างภาพจากชุดข้อมูล Folio Dataset

3.2 การทดลองจำแนกพรรณไม้โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน

การทำงานของระบบจำแนกพรรณไม้โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน โดยใช้คอมพิวเตอร์ที่มีหน่วยประมวลผลภาพ (Graphic Processing Unit: GPU) ช่วยในการประมวลผลชุดข้อมูลรูปภาพ โดย GPU นี้จะต้องมีแพลตฟอร์มสำหรับการประมวลผลแบบคู่ขนาน (Compute Unified Device Architecture: CUDA) ซึ่ง CUDA พัฒนาโดยบริษัท Nvidia เพื่อให้นักพัฒนาสามารถดึงศักยภาพในการประมวลผลแบบขนานของ GPU สำหรับการประมวลผลในงานต่าง ๆ ที่ช่วยให้การประมวลผลข้อมูลทางด้านรูปภาพมีความรวดเร็วเพิ่มขึ้นจากการที่ใช้หน่วยประมวลผลกลาง (Central Processing Unit: CPU) ในการประมวลผลเพียงอย่างเดียว ซึ่ง GPU ที่สามารถประมวลผลแบบขนาน Parallel Computing ทำให้สามารถประมวลผลข้อมูลพร้อม ๆ กันได้ในเวลาเดียวกัน โดยเครื่องมือที่ใช้ในการทดลองในครั้งนี้ คือ NVIDIA DIGITS [35] ซึ่งเป็นซอฟต์แวร์เพื่อการสร้างโมเดลโครงข่ายประสาทเทียมแบบลึก (Deep Convolutional Neural Networks: Deep CNNs)

ในการวิจัยครั้งนี้ได้ศึกษาการจำแนกข้อมูลพรรณไม้ด้วยชุดข้อมูลขนาดเล็ก โดยมีขั้นตอนการทดลองตามลำดับดังต่อไปนี้

1) จัดการกับชุดข้อมูลรูปภาพโดยทำการแบ่งชุดข้อมูลด้วยวิธี Stratified Cross-Validation [12] ซึ่งเป็นการทดสอบไขว้ที่ปรับปรุงจากวิธี K-Fold Cross-Validation โดยกำหนดให้ k=5

2) นำชุดข้อมูลรูปภาพดังกล่าวทั้ง 3 ชุดข้อมูล ได้แก่ ชุดข้อมูล PNE Dataset, ชุดข้อมูล 102 Flower Dataset และ ชุดข้อมูล Folio Dataset มาทดสอบประสิทธิภาพของเทคนิค Deep Learning ด้วยวิธี CNNs ที่มีลักษณะของโครงสร้าง (Architecture) ที่ต่างกัน 4 แบบ ได้แก่ LeNet [5], AlexNet [6], VGGNet [7] และ GoogLeNet [8]

3) ทำการทดลองกับชุดข้อมูลรูปภาพทั้ง 3 ชุด จากนั้นก็นำไปหาประสิทธิภาพของระบบ โดยจะทำการวัดเพื่อหาค่า Precision ratio และ Accuracy

4) นำโมเดลที่ได้ผลดีที่สุดไปใช้งานในการจำแนกพรรณไม้

3.3 การทดสอบประสิทธิภาพ

ในการทดสอบประสิทธิภาพของวิธีการที่นำเสนอนั้น ใช้วิธีการทดสอบแบบ Stratified Cross-Validation [19] เป็นการทดสอบแบบไขว้ที่ปรับปรุงมาจากวิธี K-Fold Cross-Validation โดยในขั้นตอนการสุ่มนั้นข้อมูลจะคำนึงถึงสัดส่วนกระจ่ายข้อมูลในแอททริบิวต์ที่เป็นคลาสเป้าหมาย ข้อมูลที่ถูกสุ่มให้เป็นชุดข้อมูลในการเรียนรู้ Training Set และ Test Set ในแต่ละ Fold จะมีสัดส่วนของคลาสไม่แตกต่างกันจากข้อมูลดั้งเดิม ซึ่งจะทำการแบ่งข้อมูลออกเป็น 5 กลุ่ม ๆ ละเท่า ๆ กัน จากนั้นจะทำการนำข้อมูล 4 กลุ่ม มาใช้เป็นชุดข้อมูลในการเรียนรู้ (Training Data) ส่วนข้อมูลที่

เหลือ 1 กลุ่ม จะใช้เป็นชุดข้อมูลทดสอบ (Test Data) โดยจะทำการวนสลับกันเป็นชุดข้อมูลทดสอบจนครบทุกกลุ่ม ซึ่งในที่นี้กำหนดให้ K=5 หมายถึงในข้อมูลแต่ละชุดจะมี 5 การทดลอง ดังภาพประกอบที่ 2.12

ในการวัดประสิทธิภาพของวิธีการที่นำเสนอนั้น จะนำชุดข้อมูลรูปภาพ (Image Dataset) ชุดที่ใช้เรียนรู้ มาใช้สำหรับสร้างตัวจำแนกพรรณไม้ เพื่อให้ได้ตัวจำแนกประเภทที่สามารถจำแนกประเภทชุดข้อมูลรูปภาพพรรณไม้ที่เรียนรู้ได้ถูกต้องมากที่สุด จากนั้นจะนำคำตอบที่ได้มาทดสอบกับชุดข้อมูลทดสอบซึ่งเป็นชุดข้อมูลที่ตัวจำแนกแต่ละประเภทไม่เคยรู้จักมาก่อน เพื่อวัดประสิทธิภาพของตัวจำแนกประเภทที่ได้ โดยทำการนับสะสมตัวอย่างที่จำแนกได้ถูกต้อง ของทั้ง 5 การทดลองบนชุดข้อมูลรูปภาพทดสอบ เพื่อนำมาคำนวณเป็นค่าความถูกต้องของวิธีการที่นำเสนอ

เนื่องจากวิธีการขั้นตอนวิธีทดสอบนี้เป็นวิธีการเชิงสุ่ม ทำให้ผลลัพธ์ที่ได้จากการทดสองในแต่ละรอบไม่เหมือนเดิม ในการทดลองนี้จึงได้ทำการทดลองซ้ำ โดยได้ทำซ้ำแต่ละชุดข้อมูลรูปภาพจำนวน 5 ครั้ง ในแต่ละรอบ (Epoch) ทำ 5-Folds Cross-Validation แล้วจะทำการรายงานผลด้วยค่าเฉลี่ย เพื่อให้เกิดความเชื่อมั่นตามหลักสถิติ

พหุน ปณฺฑิโต ชีเว

บทที่ 4
ผลการศึกษา

ในบทนี้จะกล่าวถึงการทดลองการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติโดยใช้โครงการประสาทเทียมแบบคอนโวลูชันเชิงลึก (Deep Convolutional Neural Network: Deep CNN) ตามขั้นตอนในการทดลองซึ่งได้กล่าวไว้แล้วในบทที่ 3 โดยจะทำการทดลองในจำนวนรอบการทดลองที่ 1,000 รอบ และ 10,000 รอบ โดยใช้โครงสร้างที่ต่างกันทั้ง 4 แบบ LeNet [6], AlexNet [7], VGGNet [8] และ GoogLeNet [9] กับฐานข้อมูลจำนวน 3 ฐานข้อมูล สุดท้ายจะเป็นการเปรียบเทียบผลการทดลองและประสิทธิภาพของอัลกอริทึม ซึ่งในบทนี้จะแสดงรายละเอียดผลการศึกษาในการดำเนินการวิจัยดังนี้

4.1 เครื่องมือและชุดข้อมูลที่ใช้ในการทดลอง

4.2 วิธีการทดลอง

4.3 ผลการจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึก

4.4 ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม

4.1 เครื่องมือและชุดข้อมูลที่ใช้ในการทดลอง

การทดลองนี้ได้ใช้เครื่องมือได้แก่ NVIDIA Deep Learning GPU Training System (DIGITS) เป็นเครื่องมือในการทดสอบการเรียนรู้เชิงลึก ซึ่งทำให้ประมวลผลได้บนหน่วยประมวลผลด้านกราฟฟิก (Graphics Processing Unit: GPU) และเลือกใช้ GPU รุ่น GeForce GTX 1060 ที่มีขนาดหน่วยความจำสำรอง 6GB และมีจำนวน NVIDIA CUDA Cores จำนวน 1152 Core ชุดข้อมูลที่ใช้ทดลองนี้จะทำการทดลองด้วยชุดข้อมูลพรรณไม้ทั้งสิ้น 3 ชุดข้อมูล ประกอบด้วยชุดข้อมูล Plants in Natural Environment (PNE), Folio และ 102 Flower ซึ่งเป็นชุดข้อมูลที่จำนวนชนิดและรูปภาพไม่เท่ากัน จึงจำเป็นต้องมีการกำหนดขนาดให้กับชุดข้อมูลที่จะทำการเรียนรู้และการทดสอบ ภาพการทดลองทุกภาพจะถูกปรับให้เป็นภาพ RGB ขนาด 256x256 Pixel โดยจะทำการแบ่งภาพ โดยจะทำการสุ่มเลือกภาพโดยร้อยละ 80 ของภาพทั้งหมดจะเป็นภาพสำหรับการเรียนรู้และที่เหลือร้อยละ 20 ของจำนวนภาพทั้งหมดจะเป็นภาพที่ใช้ในการทดสอบและหาประสิทธิภาพของโมเดล

ชุดข้อมูลที่ใช้ในการทดลองได้แก่ ชุดข้อมูล PNE Dataset มีจำนวนชนิดของพรรณไม้ทั้งหมด 10 ชนิด มีภาพทั้งหมด 3,000 ภาพ แบ่งภาพที่ใช้เรียนรู้ออกเป็น 2400 ภาพ และภาพที่ใช้ทดสอบอีก 600 ภาพ ต่อมาชุดข้อมูล Folio Dataset มีจำนวนชนิดของพรรณไม้ ทั้งหมด 32 ชนิด มี

ภาพทั้งหมด 637 ภาพ แบ่งภาพที่ใช้เรียนรู้ออกเป็น 510 ภาพ และภาพที่ใช้ทดสอบอีก 127 ภาพ ต่อมา ชุดข้อมูล 102 Flower Dataset มีจำนวนชนิดของพรรณไม้ ทั้งหมด 102 ชนิด มีภาพทั้งหมด 8,189 ภาพ แบ่งภาพที่ใช้เรียนรู้ออกเป็น 6,550 ภาพ และภาพที่ใช้ทดสอบอีก 1,639 ภาพ ดังตารางที่ 4.1

ตารางที่ 4.1 ชุดข้อมูลที่ใช้ในการทดลอง

ชุดข้อมูล	จำนวนชนิด	ภาพทั้งหมด	ภาพที่ใช้เรียนรู้	ภาพที่ใช้ทดสอบ
PNE	10	3,000	2,400	600
Folio	32	637	510	127
102 Flower	102	8,189	6,550	1,639

4.2 วิธีการทดลอง

ในขั้นตอนของการทดลองนี้จะเป็นการนำชุดข้อมูลที่ได้จัดเตรียมไว้มาทดลองกับวิธีการที่ออกแบบไว้โดยการทดลองจะแบ่งเป็น 2 ส่วน ได้แก่ 1) จะทำการทดสอบโดยจะทำการทดลองในจำนวนรอบการทดลองที่จำนวน 1,000 รอบ โดยใช้โครงสร้างที่ต่างกันทั้ง 4 แบบ LeNet, AlexNet, VGGNet และ GoogLeNet กับฐานข้อมูลจำนวน 3 ฐานข้อมูล และ 2) จะทำการทดสอบโดยจะทำการทดลองในจำนวนรอบการทดลองที่จำนวน 10,000 รอบ ซึ่งจะใช้ NVIDIA Deep Learning GPU Training System (DIGITS) [35] เป็นเครื่องมือในการทดสอบการเรียนรู้เชิงลึก โดยมีการตั้งค่าพารามิเตอร์ให้กับการทดลองดังนี้

ขั้นที่ 1 กำหนดค่าพารามิเตอร์ที่ใช้ในการเตรียมชุดข้อมูลรูปภาพที่ใช้สำหรับเรียนรู้และการทดสอบคือ ปรับภาพทั้งหมดให้เป็นภาพสี (Red, Green Blue: RGB) ทำการย่อขนาดรูปภาพให้มีขนาด 256x256 พิกเซล ตั้งค่าที่อยู่ของไฟล์ภาพ โดยตั้งไว้ที่หน้า Home Directory ทำการแบ่งข้อมูลไว้สำหรับทดสอบที่ 20% และอีก80% จะใช้เป็นภาพสำหรับการเรียนรู้และทดสอบประสิทธิภาพของโมเดลโดยใช้วิธี 5-fold cross-validation ดังตารางที่ 4.2

ขั้นที่ 2 กำหนดค่าพารามิเตอร์สำหรับให้คอมพิวเตอร์ทำการเรียนรู้โดยเริ่มจากทำการเลือกดาต้าเบสชุดข้อมูลรูปภาพที่ได้ทำการเตรียมไว้ ตามตารางที่ 4.2 ในการวิจัยเล่มนี้จะทำการทดสอบกับชุดข้อมูลทั้ง 3 แบบ PNE, 102 Flower และ Folio Dataset ใช้จำนวนในการเรียนรู้โดยแบ่งออกเป็น 2 ส่วนเพื่อทำการเปรียบเทียบ คือจำนวนรอบในการเรียนรู้ 1,000 รอบ จะทดสอบเมื่อผ่านไปทุก ๆ 50 รอบ และจำนวนรอบในการเรียนรู้ 10,000 รอบ จะทดสอบเมื่อผ่านไปทุก ๆ 100 รอบ การเคลื่อนลงตามความชัน (Gradient Descent: GD) ใช้แบบเคลื่อนลงตามความชันแบบสุ่ม (Stochastic Gradient Descent: SGD) ใช้ค่าความเร็วในการเรียนรู้ของโครงข่ายประสาทเทียม

(Base Learning Rate) กำหนดให้มีค่าเริ่มต้นที่ 0.001 การปรับเปลี่ยนรูปแบบข้อมูล (Data Transformations) ใช้การตัดส่วนของภาพกับโครงสร้าง AlexNet, VGGNet และ GoogLeNet ด้วยขนาด 227, 224 และ 224 ตามลำดับ ดังตารางที่ 4.3

ตารางที่ 4.2 พารามิเตอร์ที่ใช้ในการเตรียมชุดข้อมูล

Name	Parameter
Image Type	Color (RGB)
Image size (Width x Height)	256x256
Training Images	/Home directory
% for validation	20
% for testing	20
DB backend	LMDB
Image Encoding	JPEG

ตารางที่ 4.3 พารามิเตอร์ที่ใช้ในการกำหนดค่าในการทดสอบการเรียนรู้เชิงลึก

Name	Parameter
Dataset	PNE | 102 Flower | Folio
Solver Options	
Training epochs	1,000 | 1,0000
Snapshot interval (in epochs)	1,000 | 1,0000
Validation interval (in epochs)	50 | 100
Solver type	SGD (Stochastic gradient descent)
Base Learning Rate	0.001
Data Transformations	
Crop Size	256

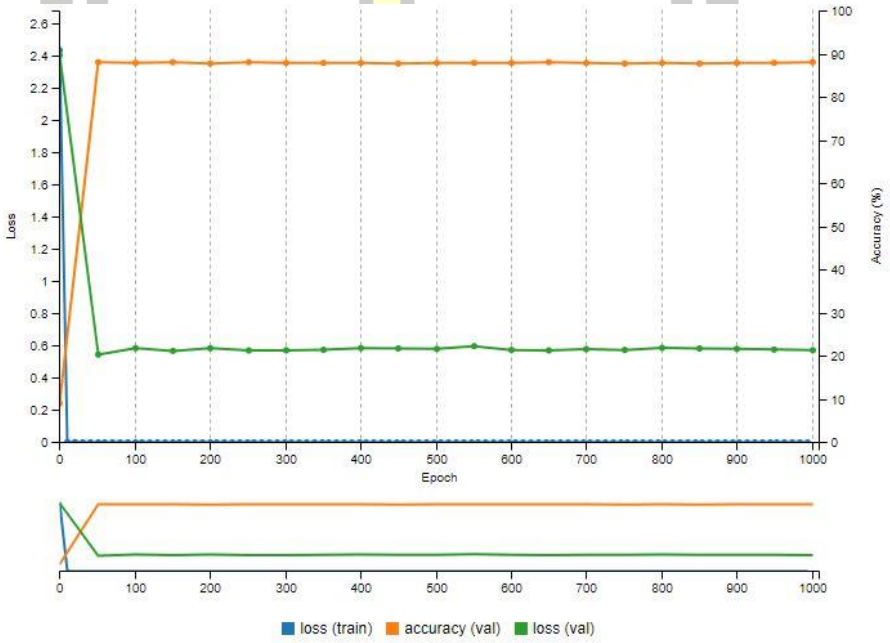
4.3 ผลการจำแนกพรรณไม้ด้วยวิธีการเรียนรู้เชิงลึก

งานวิจัยเรื่องนี้ ได้ทำการปรับรูปภาพที่ใช้ในการวิจัยให้มีขนาดเป็น 256x256 พิกเซล และชุดข้อมูลทุกชุดจะถูกแบ่งด้วยสัดส่วน 80:20 เพื่อให้ได้ชุดข้อมูลการเรียนรู้และชุดข้อมูลที่ใช้ทดสอบ ใช้วิธี Stratified Cross-Validation [19] กำหนดให้มีจำนวน 5-fold โดยจะทำการวัดความถูกต้อง

(Accuracy) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) สำหรับการประเมินประสิทธิภาพของการทดลอง โดยใช้อัตราความถูกต้องแบบ Top-1 (Top-1 Accuracy)

4.3.1 จำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset

1) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ LeNet-5 ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 88.50% ใช้เวลาในการเรียนรู้ 1 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.1



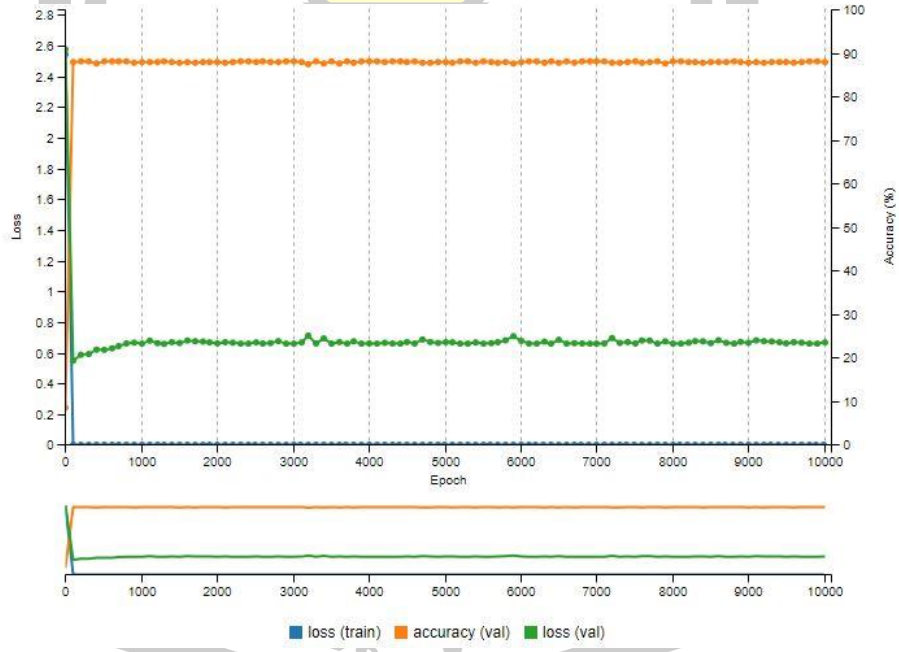
ภาพประกอบที่ 4.1 PNE Dataset กับโครงสร้าง LeNet-5 ที่จำนวน 1,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ LeNet-5 ที่จำนวนการเรียนรู้ 1,000 รอบ อัตราความถูกต้องอยู่ที่ 88.50% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 3 ซึ่งเป็นดอกพวงชมพู มีความแม่นยำน้อยที่สุด อยู่ที่ 63.33% ซึ่งทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 1,5,6,8,9 และ คลาสที่ 10 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังแสดงในตารางที่ 4.4

ตารางที่ 4.4 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง LeNet ที่จำนวน 1,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**58**					1			1		96.67
2		**55**		1	1	2	1				91.67
3	4		**38**		1	2		8	6	1	63.33
4		3		**54**	1	2					90.00
5					**59**			1			96.33
6		1	1	1	3	**53**	1				88.33
7		1				1	**58**				96.67
8	1	2	2			1		**47**	3	4	78.33
9			2					1	**56**	1	93.33
10				1		4	1			**54**	90.00

2) จำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ LeNet-5 ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ มีอัตราความถูกต้องอยู่ที่ 88.50% ใช้เวลาในการเรียนรู้ 1 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.2



ภาพประกอบที่ 4.2 PNE Dataset กับโครงสร้าง LeNet-5 ที่จำนวน 10,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ LeNet-5 ที่จำนวนการเรียนรู้ 10,000 รอบ ซึ่งการเพิ่มรอบการเรียนรู้จะทำให้อัตราความถูกต้อง

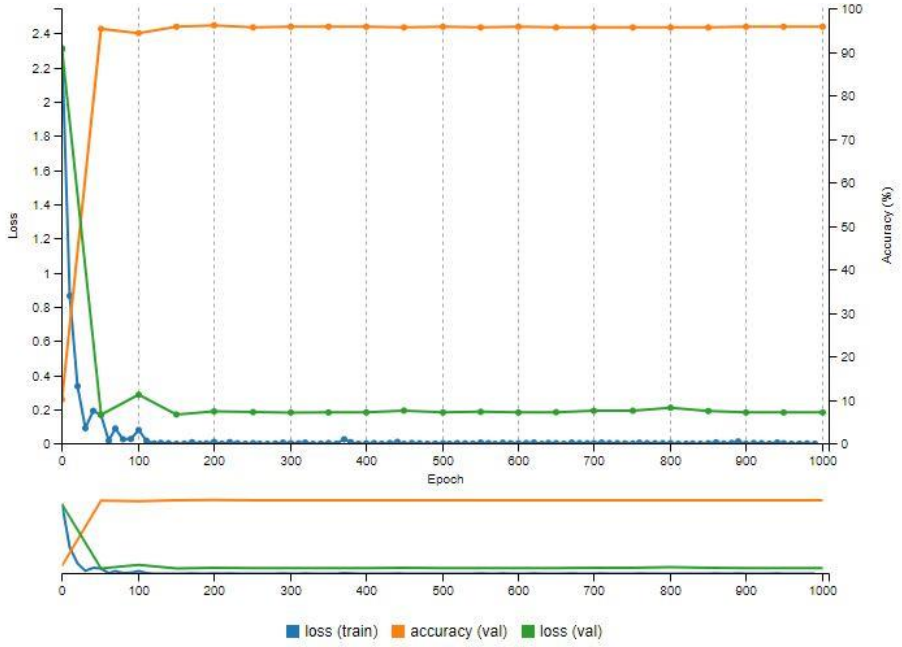
เพิ่มขึ้นด้วย อยู่ที่ 89.17% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 3 ซึ่งเป็นดอกพวงชมพู มีความแม่นยำน้อยที่สุด อยู่ที่ 66.67% ซึ่งทำนายผิดไปเป็นคลาสอื่น ได้แก่คลาสที่ 1,5,6,8,9 และ คลาสที่ 10 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังตารางที่ 4.5

ตารางที่ 4.5 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง LeNet ที่จำนวน 10,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**57**		1			1			9		95.0
2		**53**		3	1	2	1				88.33
3	4		**40**		1	2		6	6	1	66.67
4		3	1	**52**	1	2			1		86.67
5		1			**58**			1			96.67
6		1	2		3	**53**	1				88.33
7		1				1	**58**				96.67
8	1	2	2			1		**46**	3	5	76.67
9		1	3					2	**55**		91.67
10				1		4	1			**54**	90

3) จำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ AlexNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 96.67% ใช้เวลาในการเรียนรู้ 1 ชั่วโมง 50 นาที ดังภาพประกอบที่ 4.3

พหูน ปณ ฑิโต ชีเว



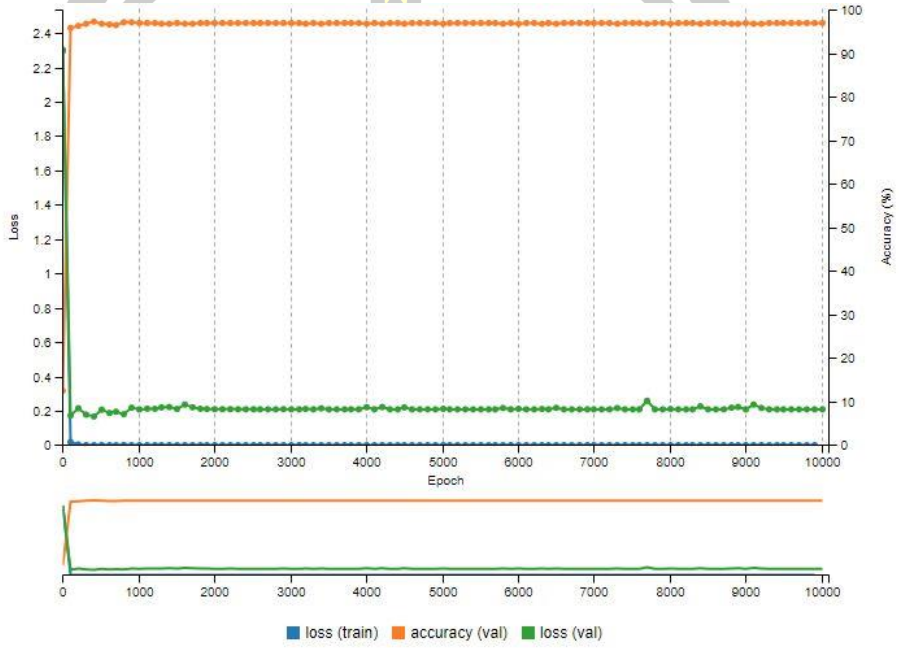
ภาพประกอบที่ 4.3 PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ AlexNet ที่จำนวนการเรียนรู้ 1,000 รอบ อัตราความถูกต้องอยู่ที่ 96.67% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 10 ซึ่งเป็นดอกส่าเหล้าปัตตานี มีความแม่นยำน้อยที่สุด อยู่ที่ 93.33% ซึ่งทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 2 และ คลาสที่ 6 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังแสดงในตารางที่ 4.6

ตารางที่ 4.6 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**59**					1					98.33
2		**58**			2						96.67
3			**60**								100
4				**60**							100
5				1	**59**						98.33
6				2		**58**					96.67
7							**60**				100
8		1	3					**54**		2	90.00
9			1					1	**58**		96.67
10		1				3				**56**	93.33

4) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ AlexNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ มีอัตราความถูกต้องอยู่ที่ 97.33% ในขณะที่อัตราความถูกต้องเพิ่มขึ้นเวลาที่ใช้ในการเรียนรู้ก็เพิ่มขึ้นด้วยอีก 5 ชั่วโมง ดังภาพประกอบที่ 4.4



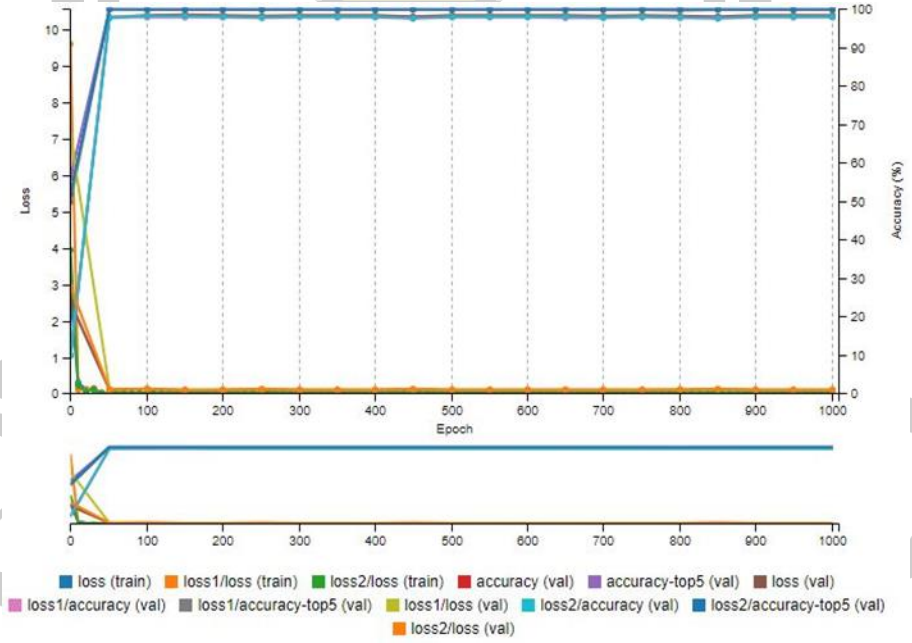
ภาพประกอบที่ 4.4 PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ AlexNet ที่จำนวนการเรียนรู้ 10,000 รอบ ซึ่งการเพิ่มรอบการเรียนรู้จะทำให้อัตราความถูกต้องเพิ่มขึ้นด้วย อยู่ที่ 97.33% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 8 ซึ่งเป็นดอกเล็บมือนาง มีความแม่นยำน้อยที่สุด อยู่ที่ 91.67% ซึ่งทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 1, 3 และคลาสที่ 10 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังตารางที่ 4.7

ตารางที่ 4.7 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**59**					1					98.33
2		**60**									100
3			**60**								100
4				**60**							100
5				1	**58**				1		96.67
6				2		**58**					96.67
7							**60**				100
8	1		2					**55**		2	91.67
9								1	**59**		98.33
10		1			1	2				**56**	93.33

5) จำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน ใช้โครงสร้างแบบ GoogLeNet ที่ใช้จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้อง อยู่ที่ 99.04% ใช้เวลาในการเรียนรู้ 4 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.5



ภาพประกอบที่ 4.5 PNE Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ AlexNet ที่จำนวนการเรียนรู้ 1,000 รอบ อัตราความถูกต้อง อยู่ที่ 96.67% แต่เมื่อพิจารณาความ

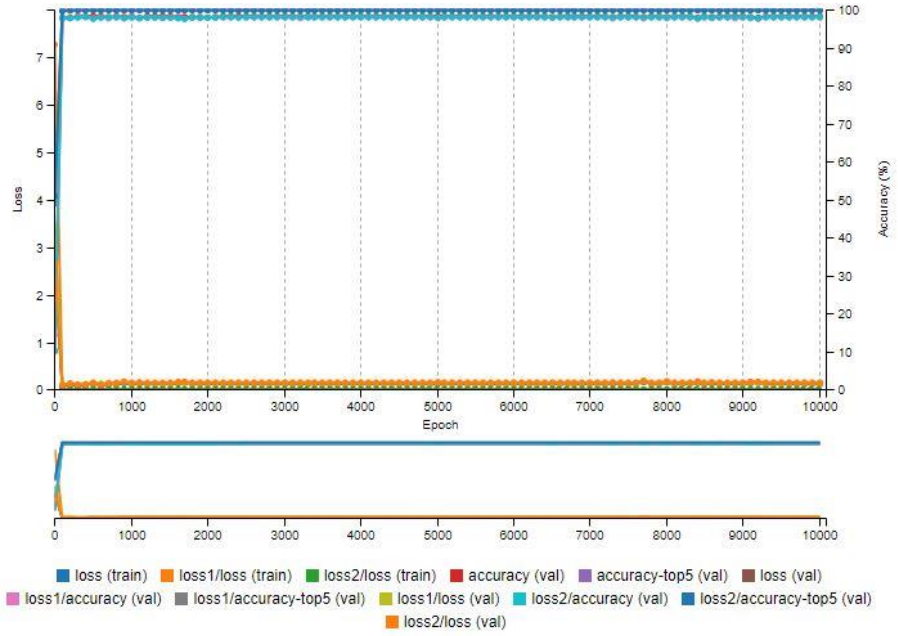
แม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 8 และ 9 ซึ่งเป็นดอกเล็บมือนาง และดอกวาสนา มีความแม่นยำน้อยที่สุด อยู่ที่ 96.67% เท่ากัน ซึ่งคลาสที่ 8 ทำนายผิดไปเป็นคลาสอื่นได้แก่ คลาสที่ 3 และ คลาสที่ 9 ทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 3 และคลาสที่ 8 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังแสดงในตารางที่ 4.8

ตารางที่ 4.8 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**59**					1					98.33
2		**55**									100
3			**38**								100
4				**54**							100
5					**59**	1					98.33
6						**59**				1	98.33
7							**60**				100
8			2					**58**			96.67
9			1					1	**58**		96.67
10	1									**59**	98.33

6) จำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน ใช้โครงสร้างแบบ GoogLeNet ที่ใช้จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ มีอัตราความถูกต้องเพิ่มขึ้นด้วย อยู่ที่ 99.17% ในขณะที่อัตราความถูกต้องเพิ่มขึ้นเวลาที่ใช้ในการเรียนรู้ก็เพิ่มขึ้นด้วยอีก 1 วัน 4 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.6

พหุน ปณ ฑิโต ชีเว



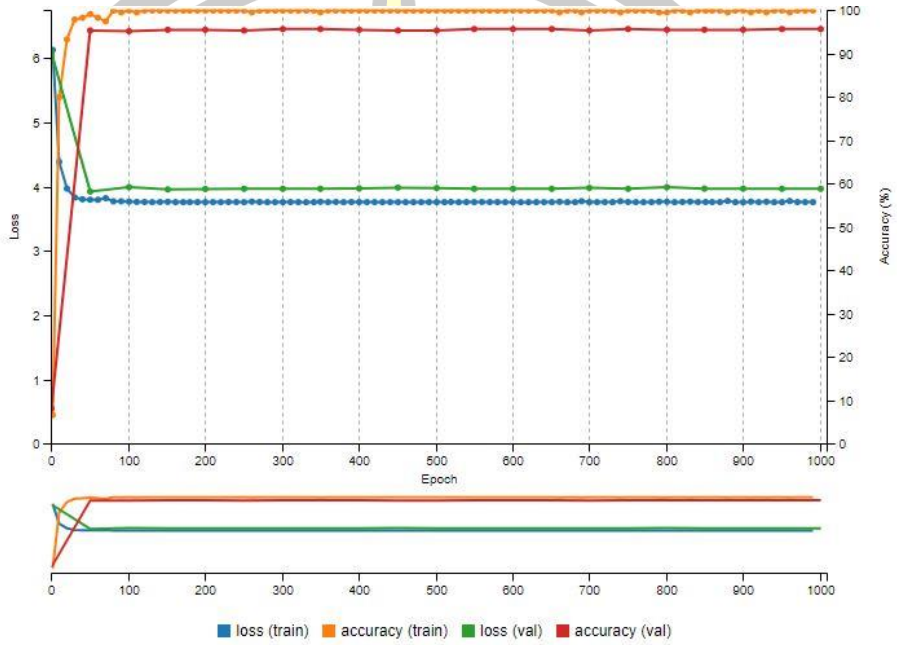
ภาพประกอบที่ 4.6 PNE Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ GoogLeNet ที่จำนวนการเรียนรู้ 10,000 รอบ ซึ่งการเพิ่มรอบการเรียนรู้จะทำให้อัตราความถูกต้องเพิ่มขึ้นด้วย อยู่ที่ 99.17% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 10 ซึ่งเป็นส่าเหล้าปัตตานี มีความแม่นยำน้อยที่สุด อยู่ที่ 96.67% ซึ่งทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 6 และคลาสที่ 8 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังตารางที่ 4.9

ตารางที่ 4.9 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 10,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**59**					1					98.33
2		**60**									100
3			**60**								100
4				**60**							100
5					**60**						100
6						**59**				1	98.33
7							**60**				100
8			1					**59**			98.33
9								1	**59**		98.33
10						1		1		**58**	96.67

7) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน ใช้โครงสร้างแบบ VGGNet ที่ใช้จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 97.17% ใช้เวลาในการเรียนรู้ 20 ชั่วโมง ดังภาพประกอบที่ 4.7



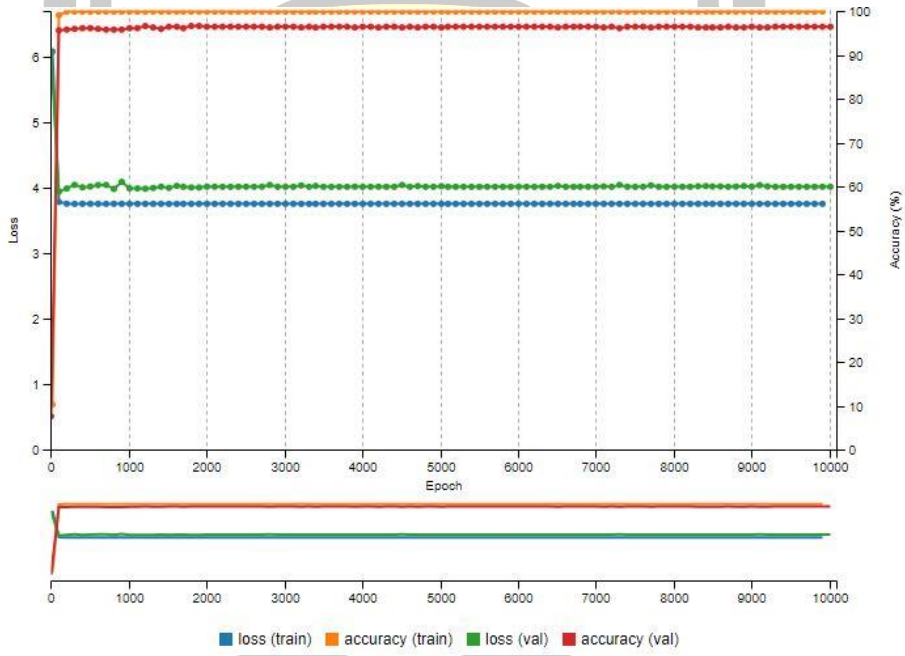
ภาพประกอบที่ 4.7 PNE Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ

จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ VGGNet ที่จำนวนการเรียนรู้ 1,000 รอบ อัตราความถูกต้องอยู่ที่ 97.17% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 8 ซึ่งเป็นดอกเล็บมือนาง มีความแม่นยำน้อยที่สุด อยู่ที่ 90.00% เท่ากัน ซึ่งคลาสที่ 8 ทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 2, 3, 6, 9 และคลาสที่ 10 พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกัน ที่สีสันและรูปร่าง ดังแสดงในตารางที่ 4.10

ตารางที่ 4.10 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**59**									1	98.33
2		**57**		1				1		1	95.00
3	2		**58**								96.67
4				**60**							100
5					**60**						100
6						**60**					100
7							**60**				100
8		1	2			1		**54**	1	1	90.00
9		1	1			1	1		**56**		93.33
10						3		1		**56**	93.33

8) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล PNE Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ VGGNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น ก็จะทำให้มีอัตราความถูกต้องเพิ่มขึ้นด้วย อยู่ที่ 97.33% ในขณะที่อัตราความถูกต้องเพิ่มขึ้นเวลาที่ใช้ในการเรียนรู้ก็เพิ่มขึ้นด้วยอีก 6 วัน 13 ชั่วโมง ดังภาพประกอบที่ 4.8



ภาพประกอบที่ 4.8 PNE Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ

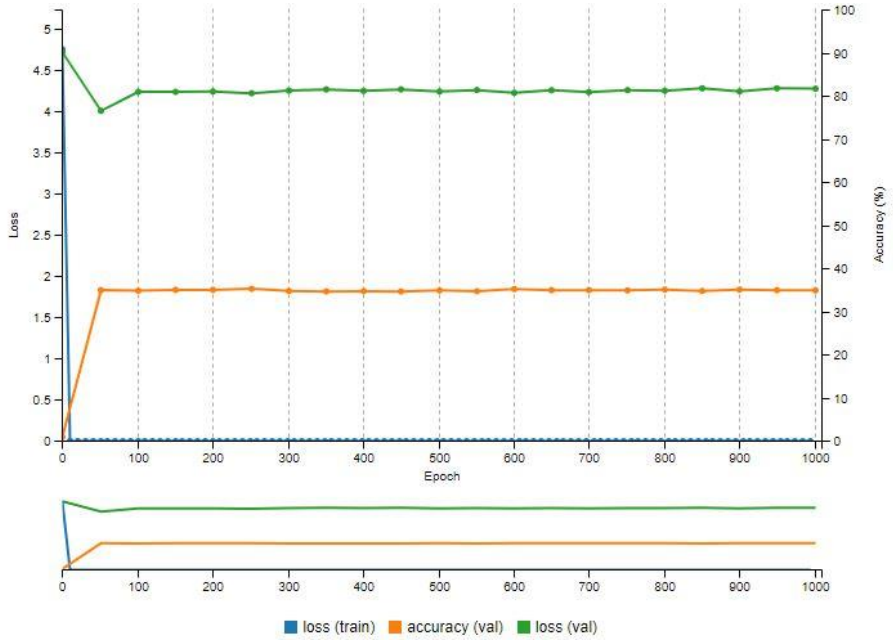
จากผลการทดสอบการจำแนกพรรณไม้ด้วยชุด PNE Dataset กับโครงสร้างแบบ VGGNet ที่จำนวนการเรียนรู้ 10,000 รอบ ซึ่งการเพิ่มรอบการเรียนรู้จะทำให้อัตราความถูกต้องเพิ่มขึ้นด้วย อยู่ที่ 97.33% แต่เมื่อพิจารณาความแม่นยำของโมเดล โดยพิจารณาแยกทีละคลาส คลาสที่ 8 ซึ่งเป็นดอกเล็บมือนาง มีความแม่นยำน้อยที่สุด อยู่ที่ 90.00% ซึ่งทำนายผิดไปเป็นคลาสอื่นได้แก่คลาสที่ 2, 3, 9 และคลาสที่ 10 ซึ่งเป็นความผิดพลาดที่เกิดจาก พรรณไม้ที่ต่างชนิดกันแต่มีความคล้ายคลึงกันที่สีสันและรูปร่าง ดังตารางที่ 4.9

ตารางที่ 4.11 Confusion Matrix ชุด PNE Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ

Class	1	2	3	4	5	6	7	8	9	10	Precision
1	**59**					1					98.33
2		**60**									100
3	1		**59**								98.33
4				**60**							100
5					**60**						100
6						**60**					100
7							**60**				100
8		1	2					**54**	1	2	90.00
9		1				1	1		**57**		95.00
10						2		1		**57**	95.00

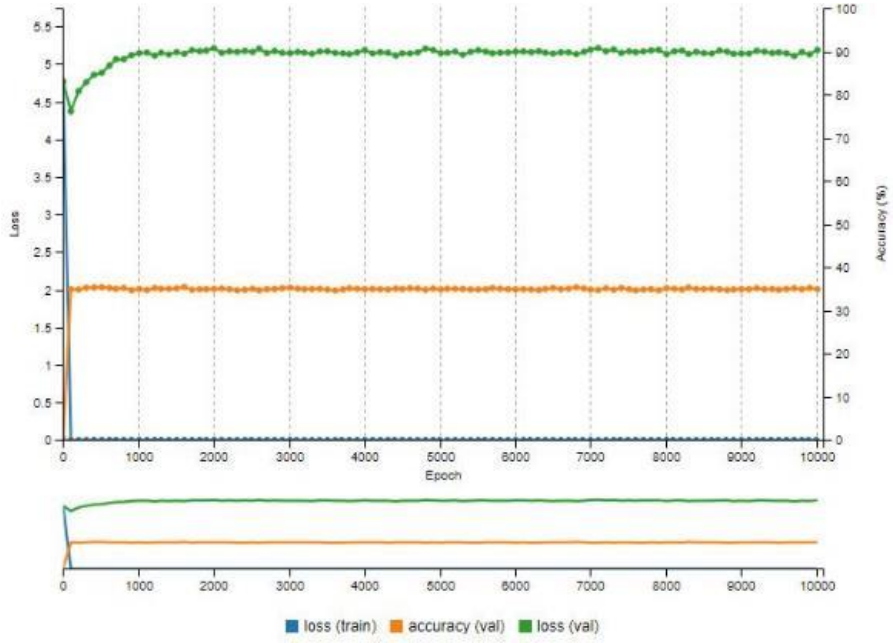
4.3.2 ผลการจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset

1) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ LeNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 33.19% ใช้เวลาในการเรียนรู้ 4 ชั่วโมง 40 นาที ดังภาพประกอบที่ 4.9



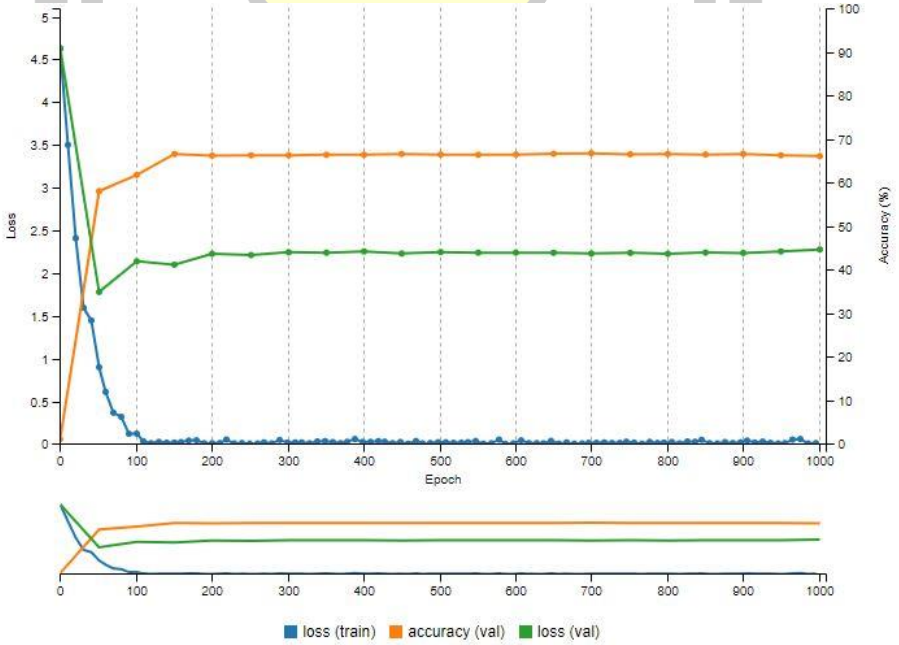
ภาพประกอบที่ 4.9 102 Flower Dataset กับโครงสร้าง LeNet ที่จำนวน 1,000 รอบ

2) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ LeNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น ทำให้มีอัตราความถูกต้องลดลงอยู่ที่ 31.42% ในขณะที่อัตราความถูกต้องลดลง เวลาที่ใช้ในการเรียนรู้นั้นกลับเพิ่มขึ้นอีก 1 วัน 12 ชั่วโมง ดังภาพประกอบที่ 4.10



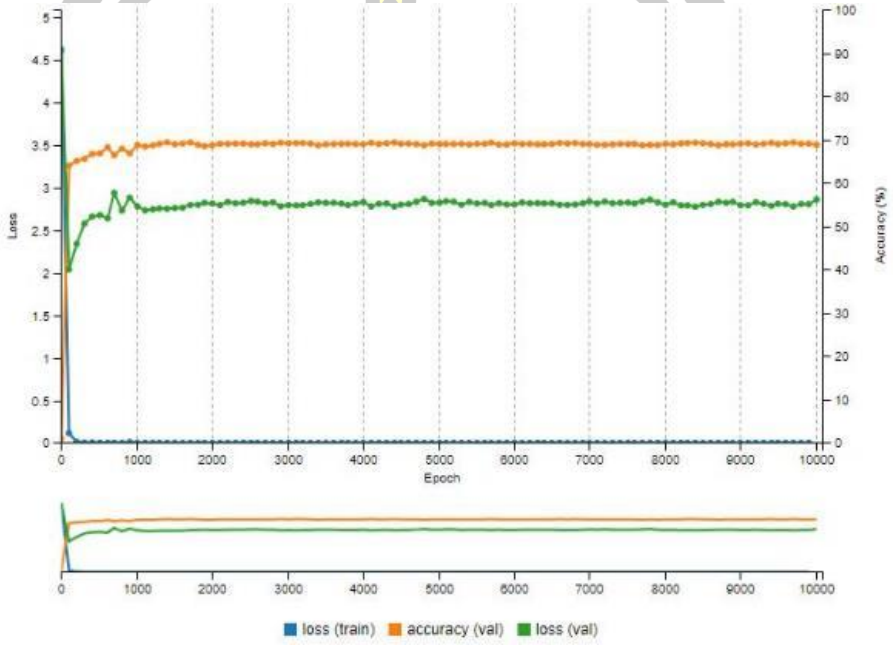
ภาพประกอบที่ 4.10 102 Flower Dataset กับโครงสร้าง LeNet ที่จำนวน 10,000 รอบ

3) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ AlexNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 66.06% ใช้เวลาในการเรียนรู้ 6 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.11



ภาพประกอบที่ 4.11 102 Flower Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ

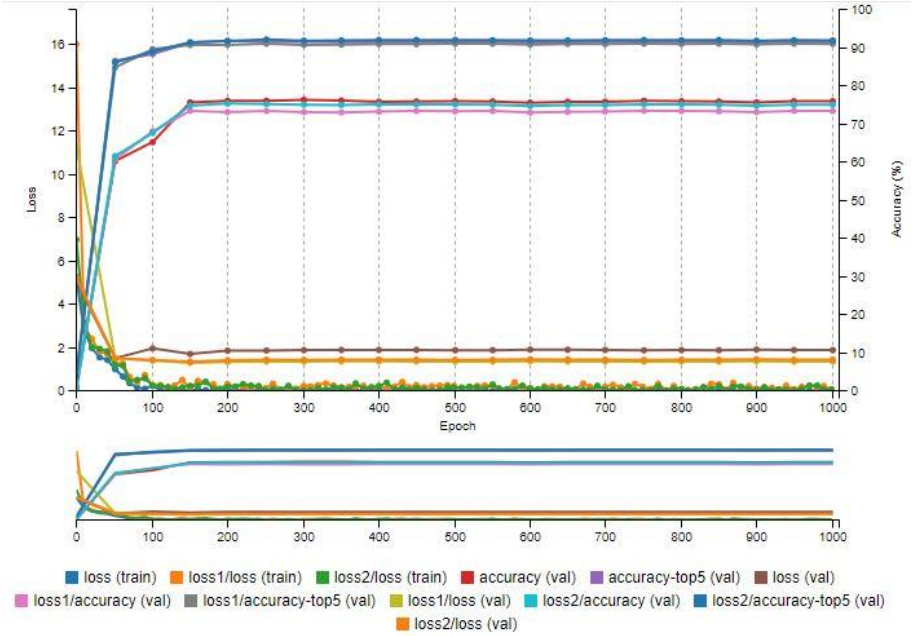
4) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ AlexNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น ทำให้มีอัตราความถูกต้องลดลงอยู่ที่ 66.38% ในขณะที่อัตราความถูกต้องลดลง เวลาที่ใช้ในการเรียนรู้นั้นกลับเพิ่มขึ้นอีก 1 วัน ดังภาพประกอบที่ 4.12



ภาพประกอบที่ 4.12 102 Flower Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ

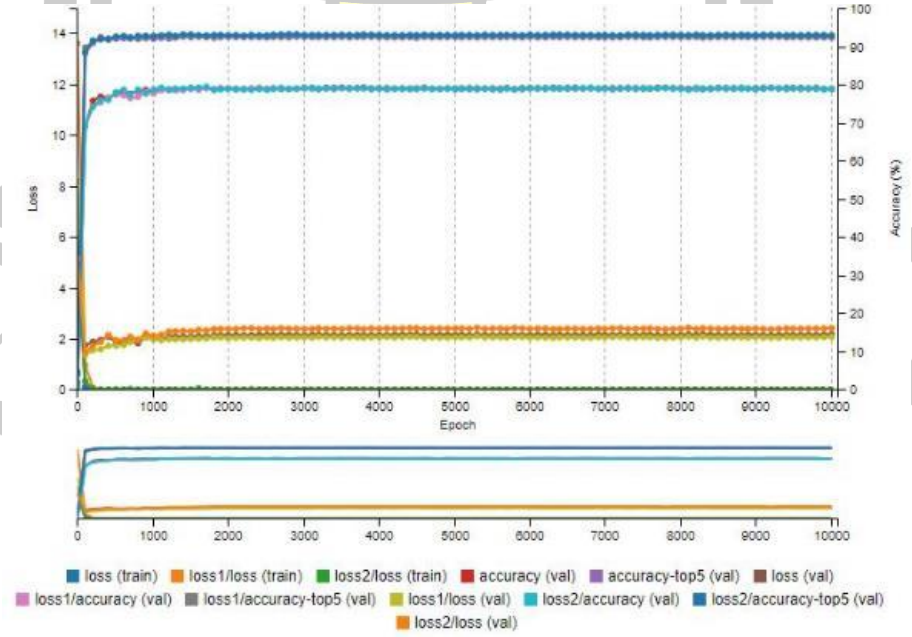
5) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ GoogLeNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 74.98% ใช้เวลาในการเรียนรู้ 1 วัน 3 ชั่วโมง ดังภาพประกอบที่ 4.13

พหุน ปณ ฑิโต ชีเว



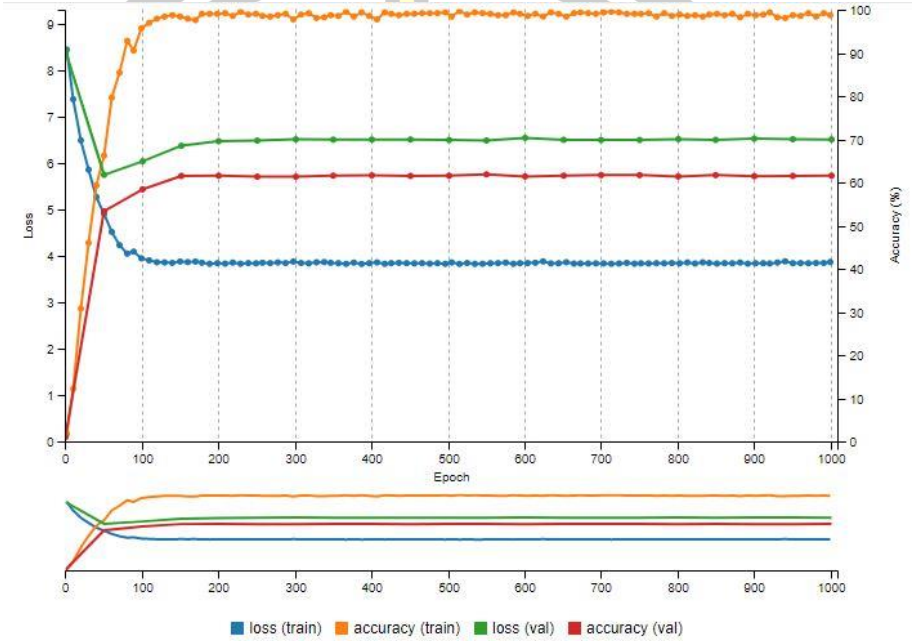
ภาพประกอบที่ 4.13 102 Flower Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ

6) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ GoogLeNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ ทำให้มีอัตราความถูกต้องเพิ่มขึ้นอยู่ที่ 78.89% ในขณะที่อัตราความถูกต้องเพิ่มขึ้น เวลาที่ใช้ในการเรียนรู้นั้นก็เพิ่มขึ้นอีก 3 วัน 1 ชั่วโมง ดังภาพประกอบที่ 4.14



ภาพประกอบที่ 4.14 102 Flower Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 10,000 รอบ

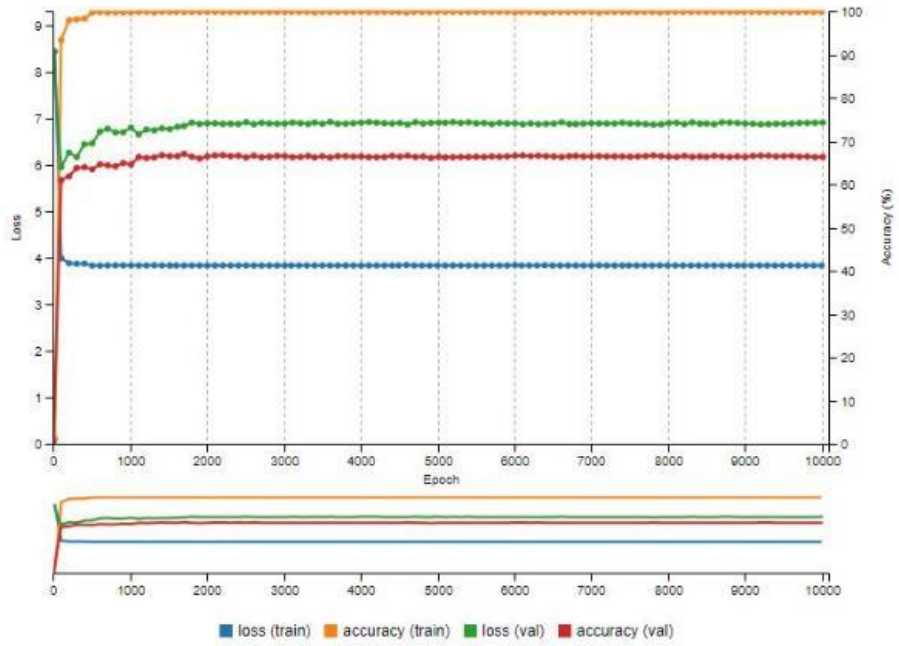
7) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ VGGNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 62.60% ใช้เวลาในการเรียนรู้ 1 วัน 19 ชั่วโมง ดังภาพประกอบที่ 4.15



ภาพประกอบที่ 4.15 102 Flower Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ

8) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล 102 Flower Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ VGGNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น ทำให้มีอัตราความถูกต้องเพิ่มขึ้นอยู่ที่ 64.12% ในขณะที่อัตราความถูกต้องเพิ่มขึ้น เวลาที่ใช้ในการเรียนรู้นั้นก็เพิ่มขึ้นอีก 3 วัน 1 ชั่วโมง ดังภาพประกอบที่ 4.16

พหุน ปณ ฑิโต ชีเว

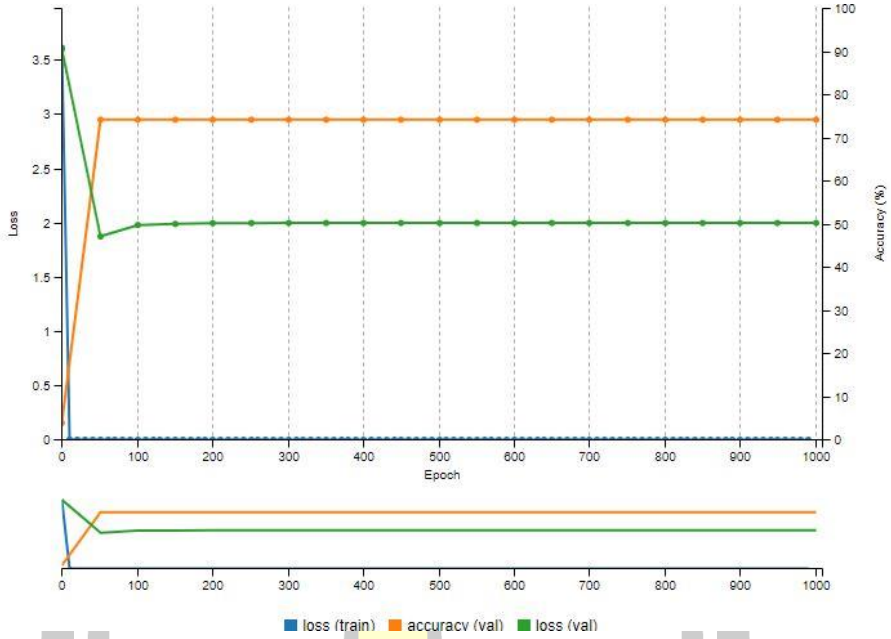


ภาพประกอบที่ 4.16 102 Flower Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ

4.3.3 ผลการจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset

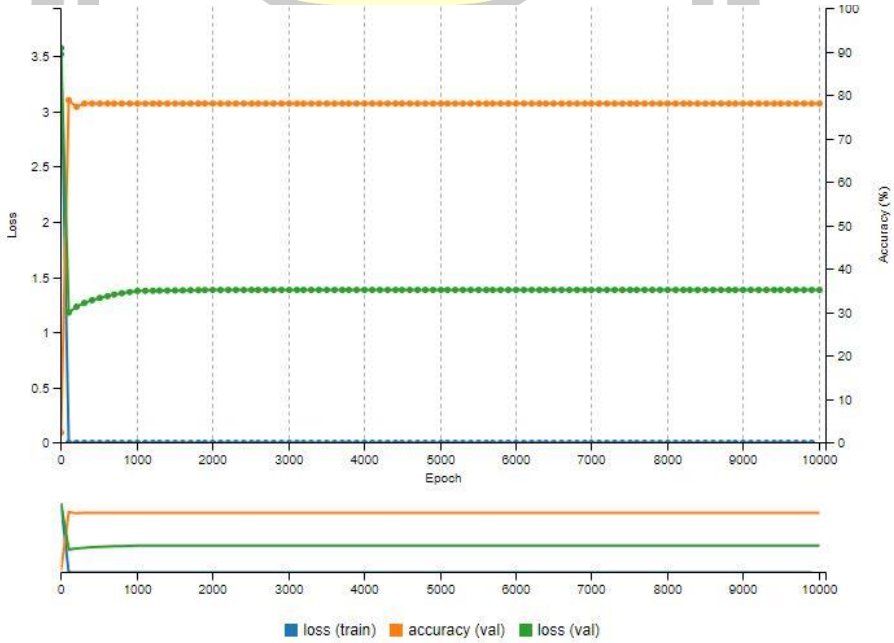
1) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ LeNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 84.25% ใช้เวลาในการเรียนรู้เพียง 20 นาที ดังภาพประกอบที่ 4.17

พหูน ปณ ฑิโต ชีเว



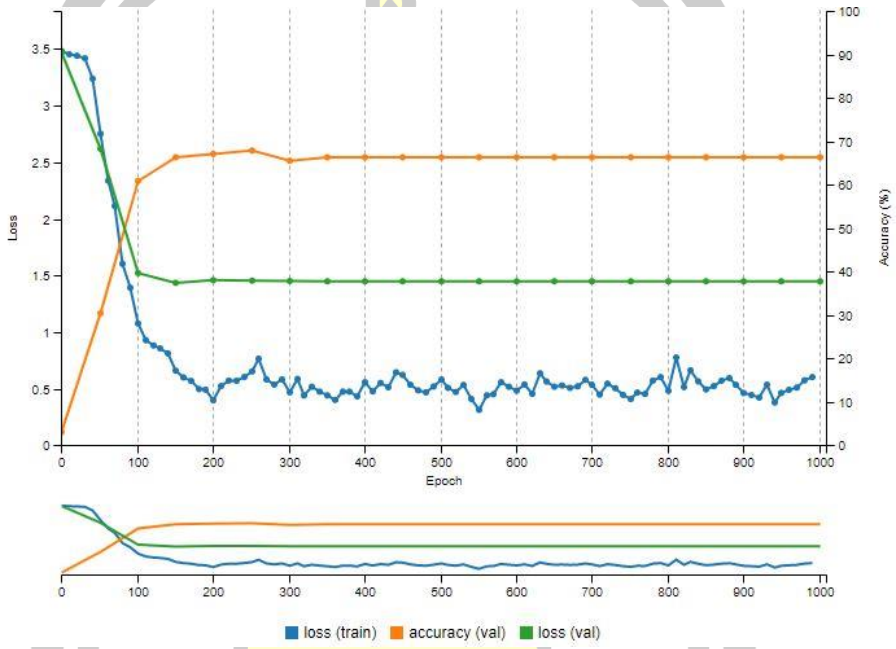
ภาพประกอบที่ 4.17 Folio Dataset กับโครงสร้าง LeNet ที่จำนวน 1,000 รอบ

2) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ LeNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น อัตราความถูกต้องกลับลดลงอยู่ที่ 74.02% ในขณะที่อัตราความถูกต้องลดลงนั้น เวลาที่ใช้ในการเรียนรู้นั้นกลับเพิ่มขึ้นอีก 2 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.18



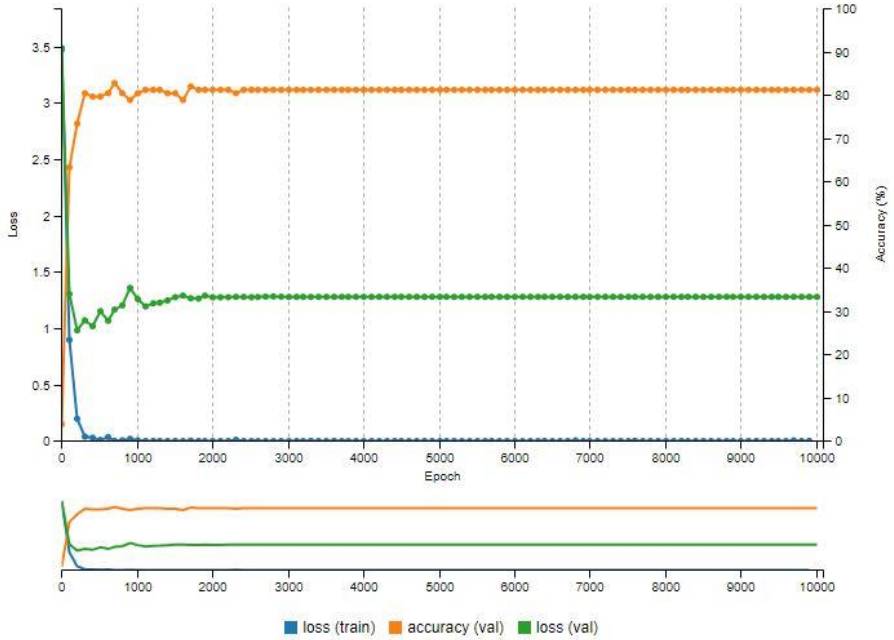
ภาพประกอบที่ 4.18 Folio Dataset กับโครงสร้าง LeNet ที่จำนวน 10,000 รอบ

3) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ AlexNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 85.04% ใช้เวลาในการเรียนรู้เพียง 30 นาที ดังภาพประกอบที่ 4.19



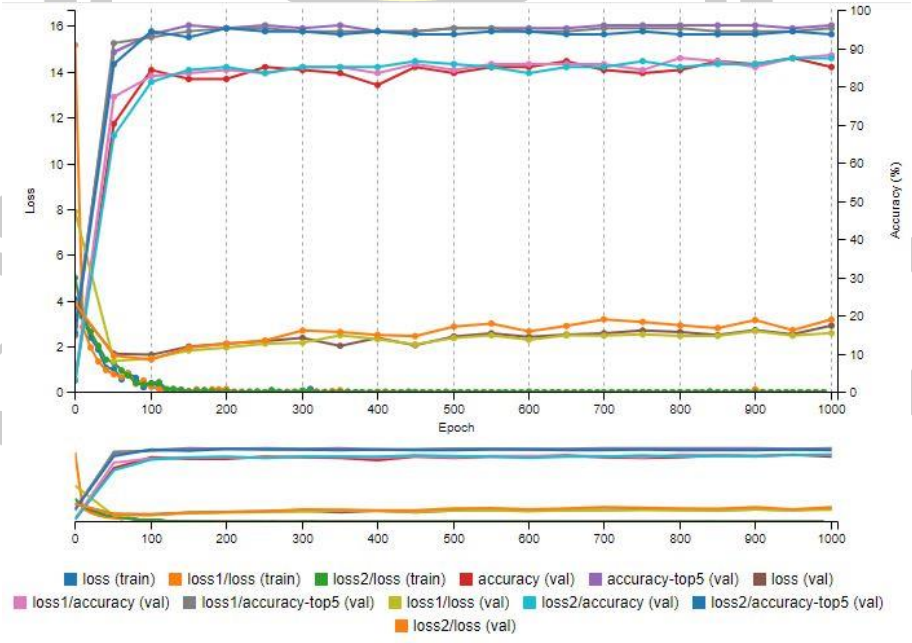
ภาพประกอบที่ 4.19 Folio Dataset กับโครงสร้าง AlexNet ที่จำนวน 1,000 รอบ

4) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ AlexNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น อัตราความถูกต้องกลับลดลงอยู่ที่ 73.23% ในขณะที่อัตราความถูกต้องลดลงนั้น เวลาที่ใช้ในการเรียนรู้นั้นกลับเพิ่มขึ้นอีก 1 ชั่วโมง 10 นาที ดังภาพประกอบที่ 4.20



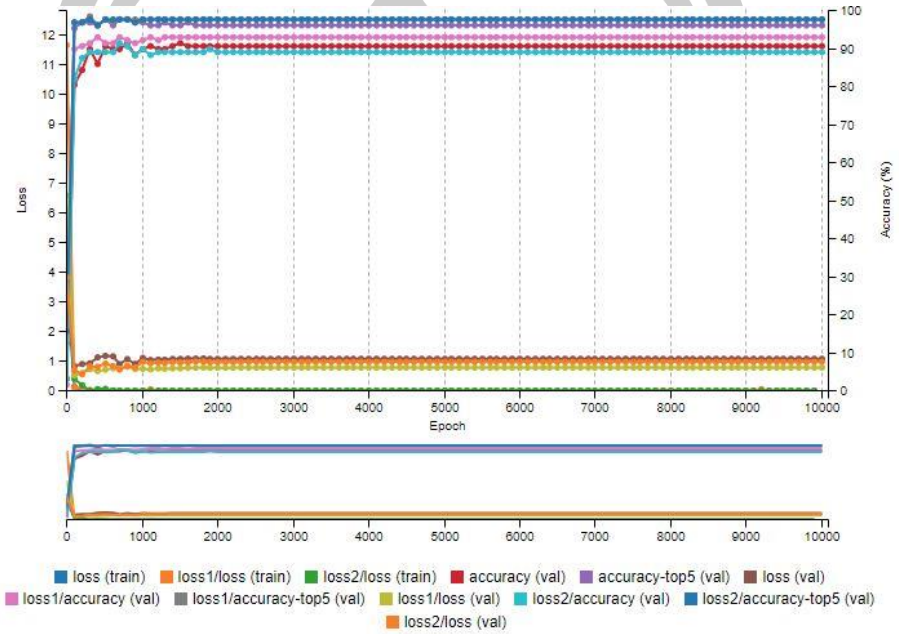
ภาพประกอบที่ 4.20 Folio Dataset กับโครงสร้าง AlexNet ที่จำนวน 10,000 รอบ

5) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ GooLeNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 87.40% ใช้เวลาในการเรียนรู้ 3 ชั่วโมง 50 นาที ดังภาพประกอบที่ 4.21



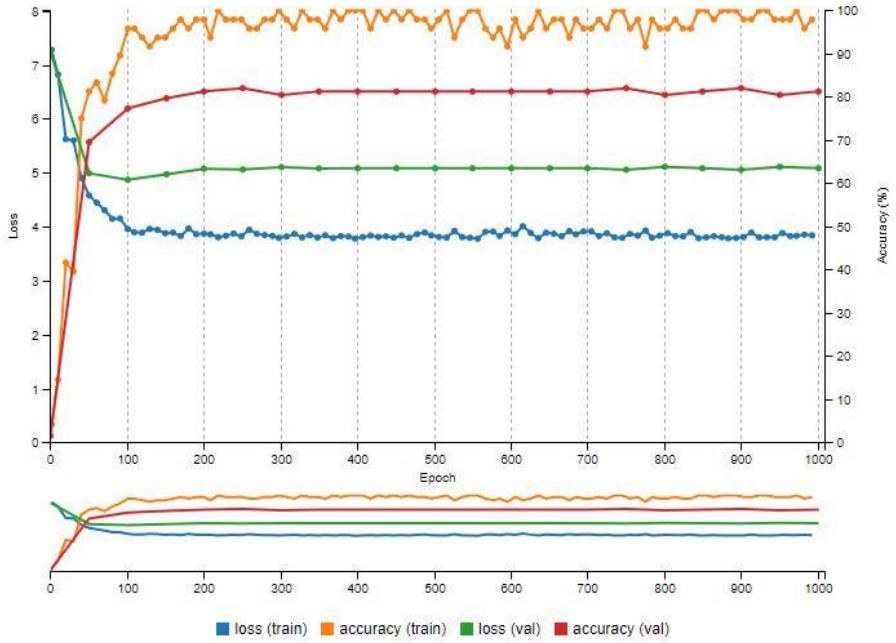
ภาพประกอบที่ 4.21 Folio Dataset กับโครงสร้าง GoogLeNet ที่จำนวน 1,000 รอบ

6) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ GooLeNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น อัตราความถูกต้องกลับลดลงอยู่ที่ 82.68% ในขณะที่อัตราความถูกต้องลดลงนั้น เวลาที่ใช้ในการเรียนรู้นั้นกลับเพิ่มขึ้นอีก 3 ชั่วโมง 30 นาที ดังภาพประกอบที่ 4.22



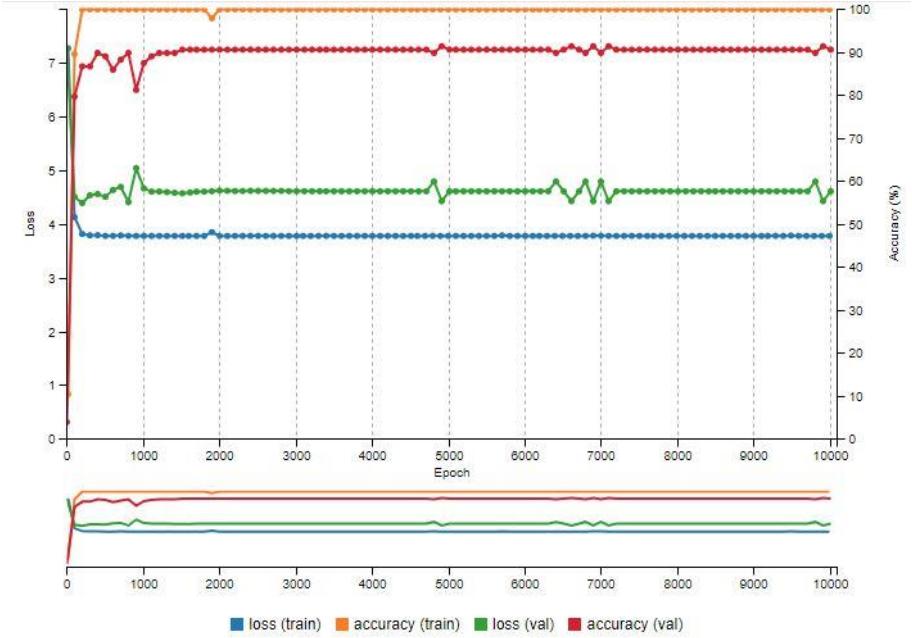
ภาพประกอบที่ 4.22 Folio Dataset กับโครงสร้าง GooLeNet ที่จำนวน 10,000 รอบ

7) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ VGGNet ที่จำนวนรอบในการเรียนรู้จำนวน 1,000 รอบ มีอัตราความถูกต้องอยู่ที่ 91.85% ใช้เวลาในการเรียนรู้ 3 ชั่วโมง 50 นาที ดังภาพประกอบที่ 4.22



ภาพประกอบที่ 4.23 Folio Dataset กับโครงสร้าง VGGNet ที่จำนวน 1,000 รอบ

8) การจำแนกพรรณไม้ด้วยวิธีเรียนรู้เชิงลึกกับชุดข้อมูล Folio Dataset เป็นการทดลองการรู้จำพรรณไม้ โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน โครงสร้างแบบ VGGNet ที่จำนวนรอบในการเรียนรู้จำนวน 10,000 รอบ เมื่อทำการเพิ่มจำนวนรอบการเรียนรู้เป็น 10,000 รอบนั้น อัตราความถูกต้องกลับลดลงอยู่ที่ 87.49% ในขณะที่อัตราความถูกต้องลดลงนั้น เวลาที่ใช้ในการเรียนรู้นั้นกลับเพิ่มขึ้นอีก 1 วัน 5 ชั่วโมง 10 นาที ดังภาพประกอบที่ 4.23



ภาพประกอบที่ 4.24 Folio Dataset กับโครงสร้าง VGGNet ที่จำนวน 10,000 รอบ

4.4 ผลการเปรียบเทียบประสิทธิภาพของอัลกอริทึม

ผลการทดสอบกับชุดข้อมูล PNE Dataset ที่ประกอบไปด้วยสมุนไพร 10 ชนิด และมีภาพสมุนไพรทั้งสิ้น 3,000 ภาพ โดยใช้วิธีการเรียนรู้เชิงลึก พบว่าวิธี CNN ที่ใช้โครงสร้างแบบ GoogLeNet ที่กำหนดให้เรียนรู้จำนวน 10,000 รอบ มีอัตราความถูกต้องที่ 99.17% ซึ่งสูงที่สุด และ GoogLeNet ที่ 1,000 รอบ มีอัตราความถูกต้องที่ 99.04% รองลงมา และตามด้วยโครงสร้างแบบ VGGNet ที่กำหนดให้เรียนรู้ จำนวน 10,000 รอบ และ 1,000 รอบ มีอัตราการเรียนรู้ 97.33% และ 97.17 ตามลำดับ แต่ทั้งนี้หากใช้การเรียนรู้จำนวน 10,000 รอบ พบว่าโครงสร้างแบบ VGGNet และ AlexNet มีอัตราความถูกต้องเท่ากันที่ 97.33% แต่ VGGNet มีค่าเบี่ยงเบนมาตรฐานที่ต่ำกว่า แสดงว่าจากการทดลองหลายรอบของ VGGNet ให้ผลความถูกต้องที่ไม่เกิดการกระจายเมื่อเทียบกับ AlexNet

หากเปรียบเทียบความเร็วในการประมวลผลของวิธีการ CNN ในการสอบด้วยการเรียนรู้จำนวน 10,000 รอบ โครงสร้างแบบ VGGNet ใช้เวลาในการประมวลผลนานที่สุดประมาณ 6 วัน 7 ชั่วโมง (~6.7d) ตามด้วยโครงสร้างแบบ GoogLeNet, Lenet-5 และAlexNet เวลาในการประมวลผล ~1.9d, ~14h และ ~6.5h ตามลำดับผลการทดลองกับชุดข้อมูล PNE Dataset ดังแสดงในตารางที่ 4.12

ตารางที่ 4.12 การเปรียบเทียบประสิทธิภาพของโครงสร้าง CNN กับชุดข้อมูล PNE Dataset

CNN Architectures	Number of Epochs / Training Time			
	1,000	Time	10,000	Time
LeNet-5	88.50±0.14	~1.3h	89.17±2.13	~14h
AlexNet	96.67±0.23	~1.5h	97.33±2.18	~6.5h
GoogLeNet	**99.04±0.11**	~4.3h	**99.17±1.75**	~1.9d
VGGNet	97.17±0.64	~20h	97.33±0.85	~6.7d

การเปรียบเทียบระหว่างวิธี CNN และวิธีที่รวมกันระหว่าง HSV+SIFT+HOG ของ Nilsback et al. [33] โดยนำมาทดสอบกับชุดข้อมูล 102 Flower พบว่าวิธี CNN ที่มีโครงสร้างแบบ GoogLeNet (เรียนรู้จำนวน 10,000 รอบ) มีอัตราความถูกต้องสูงที่สุด โดยมีความถูกต้อง 78.89% ซึ่งสูงกว่าวิธี HSV+SIFT+HOG ที่มีความถูกต้อง 72.80 %

หากเปรียบเทียบเฉพาะ CNN พบว่าจำนวนในการเรียนรู้ 10,000 รอบ ให้ผลการทดลองที่สูงกว่า จำนวน 1,000 รอบ โดยโครงสร้างที่ให้อัตราความถูกต้องสูงสุดเรียงตามลำดับคือ GoogLeNet, AlexNet, VGGNet และ LeNet-5 โดยมีอัตราความถูกต้อง 78.89, 66.38, 64.12 และ 31.42% ตามลำดับ ผลที่ได้จากทดลองแสดงดัง ตารางที่ 4.13

ตารางที่ 4.13 การเปรียบเทียบประสิทธิภาพของโครงสร้าง CNN กับชุดข้อมูล 102 Flower Dataset

CNN Architectures	Number of Epochs / Training Time			
	1,000	Time	10,000	Time
LeNet-5	33.19±0.27	~4.4h	31.42±1.12	~1.16d
AlexNet	66.06±1.74	~6.3h	66.38±2.46	~1.6d
GoogLeNet	**74.98±1.59**	~1.3d	**78.89±0.69**	~4.4d
VGGNet	62.60±0.87	~1.19d	64.12±2.74	~17d

ในงานวิจัยของ Munisami et al. [34] ใช้วิธีหาคุณลักษณะพิเศษจากรูปร่างของใบไม้ (Shape Feature) และใช้ฮีสโตแกรมของสี (Color Histogram) ร่วมกับ KNN เพื่อรู้จำใบไม้จากชุดข้อมูล Folio จากงานวิจัยระบุว่าวิธีดังกล่าวมีอัตราความถูกต้อง 87.30% ซึ่งมีอัตราความถูกต้องสูงกว่าวิธี CNN ที่ใช้โครงสร้างแบบ LeNet-5 และ AlexNet โดยโครงสร้างแบบ LeNet-5 และ

AlexNet มีอัตราความถูกต้อง 84.25 และ 85.04% เมื่อทำการเรียนรู้จำนวน 1,000 รอบ ในทางกลับกัน กรณีที่เรียนรู้ด้วยจำนวนรอบที่เพิ่มขึ้นเมื่อทดสอบที่ 10,000 รอบ อัตราความถูกต้องกลับลดลง โดยโครงสร้างแบบ LeNet-5 มีความถูกต้อง 74.02% และโครงสร้างแบบ AlexNet มีอัตราความถูกต้อง 73.23%

เมื่อนำผลการทดลองจากงานวิจัย Munisami et al. [34] มาเปรียบเทียบกับโครงสร้างแบบ GoogLeNet กลับพบว่ามีอัตราความถูกต้องใกล้เคียงกัน โดย GoogLeNet มีอัตราความถูกต้อง 87.40% ซึ่งสูงกว่าเพียง 0.1% แต่ทั้งนี้โครงสร้างแบบ VGGNet เป็นวิธีแบบ CNN ที่มีความถูกต้องสูงที่สุดจากการทดลองทั้งหมด โดยมีความถูกต้องถึง 91.85% แต่ต้องใช้เวลาในการเรียนรู้นานประมาณ 3 ชั่วโมง 50 นาที (~3.5h) จากการทดลองแสดงให้เห็นว่าโครงสร้างแบบ VGGNet ให้ผลการทดลองที่ดีกับชุดข้อมูล Folio Dataset ที่มีจำนวนรูปภาพใบไม้เพียง 637 รูป ผลการทดลองแสดงดังตารางที่ 4.14

ตารางที่ 4.14 การเปรียบเทียบประสิทธิภาพของโครงสร้าง CNN กับชุดข้อมูล Folio Dataset

CNN	Number of Epochs / Training Time			
Architectures	**1,000**	**Time**	**10,000**	**Time**
LeNet-5	84.25±0.51	~20m	74.02±0.95	~14h
AlexNet	85.04±0.48	~30m	73.23±0.81	~6.5h
GoogLeNet	87.40±1.54	~3.5h	82.68±2.41	~1.9d
VGGNet	**91.85±3.31**	~3.5h	**87.49±1.54**	~6.7d

พหุน ปณ ฑิโต ชีเว

บทที่ 5
สรุปผลและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยฉบับนี้ นำเสนอวิธีการประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks: CNN) ซึ่งประกอบด้วยโครงสร้างแบบ LeNet-5, AlexNet, GoogLeNet และ VGGNet เพื่อใช้สำหรับการจำแนกพรรณไม้ ซึ่งได้ไปทดสอบกับชุดพรรณไม้ทั้งสิ้น 3 ชุด ได้แก่ PNE, Folio และ 102 Flower Dataset งานวิจัยฉบับนี้ได้เลือกใช้ NVIDIA Deep Learning GPU Training System (DIGITS) เป็นเครื่องมือในการทดสอบการเรียนรู้เชิงลึก ทำให้สามารถประมวลผลได้บนหน่วยประมวลผลด้านกราฟฟิก (Graphics Processing Unit: GPU) และได้เลือกใช้ GPU รุ่น GeForce GTX 1060 ที่มีขนาดหน่วยความจำ 6GB และและมีจำนน NVIDIA CUDA Cores จำนวน 1152 Core

จากผลการทดลองพบว่าโครงสร้างแบบ VGGNet ให้ผลการทดลองสูงสุดกับชุดข้อมูล Folio Dataset ที่มีข้อมูลขนาดเล็กเพียง 637 ภาพ และใช้จำนวนรอบในการเรียนรู้เพียง 1,000 รอบ สำหรับโครงสร้าง แบบ GoogLeNet นั้นมีผลการทดลองสูงสุดเมื่อทดสอบกับชุดข้อมูล PNE และ 102 Flower Dataset ที่จำนวนข้อมูลในการเรียนรู้ 3,000 และ 8,189 รูปภาพ ตามลำดับ ซึ่งเป็นชุดข้อมูลที่มีจำนวนมาก แต่ต้องใช้จำนวนรอบในการเรียนรู้มากถึง 10,000 รอบ

หากเปรียบเทียบกันระหว่างโครงสร้างแบบ GoogLeNet และ VGGNet ที่มีจำนวนของชั้น (Leyer) ใกล้เคียงกันปรากฏว่าโครงสร้างแบบ VGGNet ใช้เวลาในการเรียนรู้นานกว่าโครงสร้างแบบ GoogLeNet แต่ในทางกลับกันเมื่อเปรียบเทียบอัตราความถูกต้องพบว่าโครงสร้างแบบ VGGNet มีอัตราความถูกต้องที่ต่ำกว่าแบบ GoogLeNet เมื่อทำการทดสอบกับชุดข้อมูล PNE และ 102 Flower ซึ่งข้อมูลทั้งสองชุดนั้น เป็นรูปภาพดอกไม้และพืชสมุนไพร และรูปภาพแต่ละรูปเป็นรูปภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติที่มีพื้นหลังที่ซับซ้อน แต่เมื่อทดสอบกับชุดข้อมูล Folio ที่เป็นภาพใบไม้ที่ถ่ายในห้องทดลองโดยกำหนดให้พื้นหลังเป็นสีขาวปรากฏว่าโครงสร้างแบบ VGGNet มีอัตราความถูกต้องสูงที่สุด

จากผลการทดลองการวิจัยสามารถสรุปได้ว่าโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกโดยใช้โครงสร้างแบบ GoogLeNet เหมาะสำหรับนำไปใช้เพื่อจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ เนื่องจากมีอัตราความถูกต้องสูงที่สุด และยังช่วยลดระยะเวลาในการเรียนรู้อีกด้วย

5.2 การอภิปรายผล

งานวิจัยฉบับนี้ รูปภาพที่ใช้ในการวิจัยถูกปรับให้มีขนาด 256x256 Pixel โดยชุดข้อมูลรูปภาพจะถูกแบ่งด้วยสัดส่วน 80:20 เพื่อใช้เป็นชุดข้อมูลในการเรียนรู้ (Training Set) และข้อมูลชุดทดสอบ (Test Set) ตามลำดับโดยใช้วิธี Stratified Cross-Validation กำหนดให้มีจำนวน 5-fold โดยใช้ค่าความถูกต้อง (Accuracy) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) สำหรับการประเมินประสิทธิภาพการทดลอง โดยใช้อัตราความถูกต้องแบบ Top-1 (Top-1 Accuracy) และใช้ GPU รุ่น GeForce GTX 1060 แรม 6GB ในการทดลอง ชุดข้อมูลทั้ง 3 ชุดประกอบด้วย PNE, 102 Flower และ Folio จะถูกนำมาใช้เพื่อทดสอบประสิทธิภาพของวิธีการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันโดยโครงสร้าง (Architecture) ที่นำมาเปรียบเทียบประกอบด้วย LeNet-5, AlexNet, GoogLeNet และ VGGNet [6–9] ซึ่งเป็นเทคนิคที่ได้รับความนิยมในงานรู้จำรูปแบบ (Pattern Recognition) โครงสร้างแรกคือ LeNet-5 เป็นโครงสร้างต้นแบบในงานวิจัยการรู้จำรูปแบบ ต่อมามีพัฒนาและนำเสนออัลกอริทึมอีกมากมาย โดยในปี 2012 ทีมนักวิจัยของ Krizhevsky el al. ได้นำเสนอโครงสร้างที่สองคือ AlexNet ซึ่งเป็นโครงสร้างที่ให้ผลดีที่สุดในการประกวดโครงการที่ชื่อว่า ImageNet โครงสร้างที่สามคือ VGGNet นำเสนอในปี 2014 เป็นโครงสร้างที่นำเสนอว่าเหมาะกับชุดข้อมูลขนาดใหญ่และนำเสนอโครงสร้างที่ลึกถึง 16 ชั้น และโครงสร้างที่สี่ในปีเดียวกัน GoogLeNet ถูกนำเสอนโดยใช้โครงสร้าง Inception (Inception Architecture) เพื่อให้ CNN มีความลึกและกว้างใช้การประมวลผลแบบขนานทำให้ลดเวลาในการเรียนรู้และมีประสิทธิภาพในการรูปจำแบบมาก

คุณลักษณะของชุดข้อมูล (Dataset) ที่นำมาทดสอบในการรู้จำนั้นเป็นชุดข้อมูลที่มีความแตกต่างกัน ชุดข้อมูล PNE และ 102 Flower Dataset เป็นภาพสิ่งแวดล้อมทางธรรมชาติซึ่งเป็นชุดข้อมูลที่มีจำนวนมาก และ ชุดข้อมูล Folio Dataset เป็นชุดข้อมูลที่มีขนาดเล็กเป็นภาพที่มีฉากหลังสีขาว งานวิจัยนี้จึงต้องการเปรียบเทียบให้เห็นว่าโครงสร้างแบบใด ควรเลือกใช้กับชุดข้อมูลที่มีคุณลักษณะแบบใดจึงจะเหมาะสม โดยพิจารณาจากอัตราความถูกต้องและเวลาที่ใช้ในการเรียนรู้

5.3 ข้อเสนอแนะและงานวิจัยในอนาคต

วิธีการเรียนรู้เชิงลึก (Deep Learning) จำเป็นต้องใช้ชุดข้อมูลรูปภาพจำนวนมากในการเรียนรู้ ทั้งนี้เพื่อทำให้มีอัตราในการรู้จำเพิ่มมากขึ้น งานวิจัยครั้งถัดไป จะมีการประยุกต์ใช้เทคนิคการเพิ่มจำนวนของข้อมูลชุดเรียนรู้ที่เรียกว่า Data Augmentation ซึ่งเทคนิคนี้จะทำให้มีข้อมูลสำหรับการเรียนรู้เพิ่มขึ้น และนำวิธีการถ่ายโอนความรู้ (Transfer Learning) ซึ่งจะช่วยลดระยะเวลาในการเรียนรู้และอาจส่งผลให้ประสิทธิภาพในการเรียนรู้เพิ่มขึ้นด้วย [21,36]

บรรณานุกรม

บรรณานุกรม

[1] กระทรวงสาธารณะสุขและองค์กรภาครัฐ เอกชน. “แผนแม่บทแห่งชาติว่าด้วยการพัฒนาสมุนไพร ฉบับที่ 1 พ.ศ. 2560-2564”. พิมพ์ครั้งที่ 1. บจก.ทีเอส อินเตอร์พริ้นท์; 2559.

[2] จินตนา รุ่งเรือง และวสา วงศ์สุขแสวง. “โครงการอนุรักษ์พันธุกรรมพืชอันเนื่องมาจากพระราชดำริ เรื่องการดูแลสวนสมุนไพรและสวนไม้หอม”. สุรินทร์; 2558.

[3] พัชราภรณ์ แสงโยจารย์. “ลักษณะเด่นประจำวงศ์ไม้ดอกหอมในสวนไม้หอม มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์”. การประชุมวิชาการระดับชาติ นวัตกรรมและเทคโนโลยีวิชาการ; 25-26 ธันวาคม 2560; สุรินทร์: 2560. หน้า 451–459.

[4] ก่องกานดา ชยามฤต และวรดลต์ แจ่มจำรูญ. “คู่มือจำแนกพรรณไม้”. กรุงเทพฯ: สำนักงานหอพรรณไม้ สำนักวิจัยการอนุรักษ์ป่าไม้และพันธุ์พืช. กรมอุทยานแห่งชาติ สัตว์ป่าและพันธุ์พืช; 2559.

[5] Ng A, Ngiam J, Foo CY, Mai Y, Suen C, Coates A, et al. “Deep Learning Tutorial”. University of Stanford. 2015.

[6] LeCun Y, Bottou L, Bengio Y and Haffner P. “Gradient-Based Learning Applied to Document Recognition”. in Proceedings of the IEEE; 1998. pp. 2278–2324.

[7] Krizhevsky A, Sutskever I and Hinton GE. “ImageNet Classification with Deep Convolutional Neural Networks”. Advances In Neural Information Processing Systems; 2012. pp. 1–9.

[8] Simonyan K and Zisserman A. “Very Deep Convolutional Networks for Large-Scale Image Recognition”. in International Conference on Learning Representations; 2015. pp. 1–10.

[9] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. “Going deeper with convolutions”. in IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015. pp. 1–9.

[10] เอกสิทธิ์ พัชรวงศ์ศักดา. “การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้า ไมน์นิ่ง เบื้องต้น”. ปทุมธานี: หสม.ดาต้า คิวป์; 2557.

[11] บุญธรรม ภัทราจารุกุล. “การประมวลผลภาพดิจิตอลเบื้องต้น”. กรุงเทพฯ: ซีเอ็ดยูเคชั่น; 2556.

[12] สรรพฤทธิ์ มฤคทัต. “การรู้จำรูปแบบและการเรียนรู้ของเครื่อง”. ปทุมธานี: สำงานงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ (สวทช.); 2561.

[13] ราชันย์ ภู่มา. “สารานุกรมพืชในประเทศไทย (ฉบับย่อ) เฉลิมพระเกียรติสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี ทรงเจริญพระชนมายุ 60 พรรษา”. กรุงเทพฯ: สำนักงานหอพรรณไม้ สำนักวิจัยการอนุรักษ์ป่าไม้และพันธุ์พืช กรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม; 2559.

[14] วรดลต์ แจ่มจำรูญ. “คู่มือการจำแนกพรรณไม้แบบลักษณะเด่นเฉพาะ”. กรุงเทพฯ: สำนักงานหอพรรณไม้ สำนักวิจัยการอนุรักษ์ป่าไม้และพันธุ์พืช กรมอุทยานแห่งชาติ สัตว์ป่าและพันธุ์พืช; 2558.

[15] กนกพรรณ เลิศนิพนธ์พันธุ์. “การจำแนกกริยาของมนุษย์โดยใช้ลักษณะการเคลื่อนที่และสภาพปรากฏ เพื่อการเข้าใจกิจกรรมและการตรวจหาความผิดปกติในการเฝ้าระวังจากภาพ” [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต]. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย; 2558.

[16] วิศรุต ขวัญคุ้ม และไพศาล มุณีสว่าง. “การรู้จำท่ารำวงมาตรฐานโดยการวิเคราะห์ข้อมูล 3 มิติ”. วิศวกรรมสาร มหาวิทยาลัยนเรศวร 2559; 11(1): หน้า 75–84.

[17] Szeliski R. “Computer Vision: Algorithms and Applications”. 2010.

[18] Goodfellow I, Bengio Y and Courville A. “Deep Learning”. MIT Press; 2016.

[19] ศักดิ์ เพิ่มหรรษา นิตยา เกิดประสพ และกิตติศักดิ์ กิดประสพ. “การประยุกต์เทคนิคแปลงข้อความแบบเวกเตอร์และการกระจายแบบเบย์ เพื่อทำนายการบุกรุกเครือข่ายคอมพิวเตอร์". การประชุมวิชาการระดับชาติมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 6; 23-25 กรกฎาคม 2557; พระนครศรีอยุธยา: มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ศูนย์หันตรา; 2557. หน้า 75–84.

[20] Liu Y, Zhou D, Tang F, Meng Y and Dong W. “Flower Classification via Convolutional Neural Network”. in IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA 2016); 7-11 November, 2016; Qingdao, China. 2016. pp. 110–116.

[21] Pawara P, Okafor E, Surinta O, Schomaker L and Wiering M. “Comparing Local Descriptors and Bags of Visual Words to Deep Convolutional Neural Networks for Plant Recognition”. in 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM); 24 – 26 February, 2017; Porto, Portugal. 2017. pp. 479–486.

[22] Okafor E, Pawara P, Karaaba F, Surinta O, Codreanu V, Schomaker L, et al. “Comparative Study Between Deep Learning and Bag of Visual Words for Wild-Animal Recognition”. in IEEE Symposium Series on Computational Intelligence (SSCI); Athens, Greece; 2016. pp. 1–8.

[23] Sladojevic S, Arsenovic M, Anderla A, Culibrk D and Stefanovic D. “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification”. Comput Intell Neurosci 2016; pp.1–11.

[24] Reyes AK, Caicedo JC and Camargo JE. “Fine-tuning deep convolutional networks for plant recognition”. in Conference on Conference and Labs of the Evaluation forum (CLEF); 2015. pp. 1–9.

[25] Mohanty SP, Hughes DP and Salathé M. “Using Deep Learning for Image-Based Plant Disease Detection”. Front Plant Sci 2016; 7: pp.1–10.

[26] Sun Y, Liu Y, Wang G and Zhang H. “Deep Learning for Plant Identification in Natural Environment”. Comput Intell Neurosci 2017; pp.1–7.

[27] Kumar N, Belhumeur PN, Biswas A, Jacobs DW, Kress WJ, Lopez IC, et al. “Leafsnap: A Computer Vision System for Automatic Plant Species Identification”. Computer Vision – ECCV 2012; pp.502–516.

[28] Joly A, Goëau H, Bonnet P, Bakić V, Barbe J, Selmi S, et al. “Interactive Plant Identification Based on Social Image Data”. Ecological Informatics 2014; 23: September 2014. pp.22–34.

[29] Xi Li, Xin Hua, Fang Yin, Chenxi Wang, Long Chen, Zhaoyang Zeng, et al. “Microsoft Garage Recognition app”. [Online]. 2018 [cited 23 October 2017]; Available from: https://www.microsoft.com/en-us/garage/profiles/flower-recognition/

[30] Sanayha M and Vateekul P. “Remaining Useful Life Prediction Using Enhanced Convolutional Neural Network on Multivariate Time Series Sensor Data”. Walailak Journal of Science and Technology (WJST) 2019; 16[9]: pp.1–11.

[31] Kohavi R. “Cross validation number”. 1995;5.

[32] Kohavi R. “A study of cross-validation and bootstrap for accuracy estimation and model selection”. Appears in the International Joint Conference on Articial Intelligence (IJCAI 1995) 2: pp. 1137–1143.

[33] Nilsback ME and Zisserman A. “Automated Flower Classification over a Large Number of Classes”. in Proceedings - 6th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2008); 16-19 December 2008. pp. 722–729.

[34] Munisami T, Ramsurn M, Kishnah S and Pudaruth S. “Plant Leaf Recognition Using Shape Features and Colour Histogram with K-nearest Neighbour Classifiers”. Procedia Computer Science 2015; 58: pp.740–747.

[35] Yeager L. “DIGITS: the Deep learning GPU Training System”. ICML AutoML Work. 2015;

[36] Pawara P, Okafor E, Schomaker L and Wiering MA. “Data Augmentation for Plant Classification Data Augmentation for Plant Classification”. in 18th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2017); p. 615–26

ภาคผนวก

ภาคผนวก ก

บทความวิจัย

เรื่อง “โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกสำหรับการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ”

เผยแพร่ในวารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม ปีที่ 38 ฉบับที่ 2

พหูน ปณฺฑิโต ชีเว

โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกสำหรับการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ

Deep Convolutional Neural Networks for Plant Recognition in the Natural Environment

จักรินทร์ สนุกแสน[1], โอพาริก สุรินต๊ะ[2]

Jakkarin Sanuksan[1], Olarik Surinta[2]

Received: 15 May 2018 ; Revised : 20 August 2018 ; Accepted: 5 November 2018

บทคัดย่อ

งานวิจัยฉบับนี้ได้ศึกษาโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก (Deep Convolutional Neural Network: Deep CNN) เพื่อใช้จำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ โดยมีจุดประสงค์เพื่อเปรียบเทียบ โครงข่ายประสาทเทียมแบบคอนโวลูชันจำนวน 4 โครงสร้าง ประกอบด้วย LeNet-5, AlexNet, GoogLeNet และ VGGNet ข้อมูลพรรณไม้ที่นำมาใช้ในการทดสอบมีจำนวนทั้งสิ้น 3 ชุดข้อมูลคือ PNE, 102 Flower และ Folioทั้งนี้ชุดข้อมูล PNE และ 102 Flower เป็นรูปภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติทำให้ที่มีพื้นหลังที่ซับซ้อน สำหรับข้อมูลชุด Folio เป็นรูปภาพใบไม้ที่ถ่ายในห้องทดลองโดยกำหนดให้พื้นหลังของภาพเป็นสีขาว จากผลเปรียบเทียบระหว่างโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกโดยใช้โครงสร้าง GoogLeNet และ VGGNet พบว่าโครงสร้างแบบ GoogLeNet มีประสิทธิภาพสูงที่สุดในชุดข้อมูล PNE และ 102 Flower และยังใช้เวลาในการเรียนรู้ที่เร็วกว่าเมื่อเทียบกับโครงสร้างแบบ VGGNet โดยใช้จำนวนรอบในการเรียนรู้ 10,000 รอบ แต่ทั้งนี้โครงสร้างแบบ VGGNet มีอัตราความถูกต้องสูงที่สุดในชุดข้อมูล Folio ซึ่งเป็นรูปภาพที่ถ่ายในห้องทดลองที่มีพื้นหลังเป็นสีขาว และใช้จำนวนรอบในการเรียนรู้เพียง 1,000 รอบ จึงสรุปได้ว่าหากต้องการที่จะสร้างโมเดลเพื่อนำไปใช้ในการจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติควรจะใช้โครงสร้างแบบ GoogLeNet

คำสำคัญ: การรู้จำพรรณไม้ การเรียนรู้เชิงลึก โครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก โครงสร้างแบบ AlexNet โครงสร้างแบบ GoogLeNet โครงสร้างแบบ VGGNet

Abstract

This paper examines a deep convolutional neural network (Deep CNN) for plant recognition in the natural environment. The primary objective was to compare 4 CNN architectures including LeNet-5, AlexNet, GoogLeNet, and VGGNet on three plant datasets; PNE, 102 Flower, and Folio. The images in the PNE and 102 Flower dataset include a complicated background because they were taken in a natural environment. On the other hand, the images in the Folid dataset are only leaf images that were taken in a laboratory environment using a white background. The comparison of deep CNN using GoogLeNet and VGGNet Architecture show that GoogLeNet outperformed while working on the PNE and 102 Flower dataset when using a training time with iterations of 10,000 epochs. GoogLeNet

[1] *นิสิตปริญญาโท ห้องปฏิบัติการมัลติเอเจนท์ ระบบอัจฉริยะ และการจำลองสถานการณ์ สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม มหาสารคาม 44150*

[2] *อาจารย์ ห้องปฏิบัติการมัลติเอเจนท์ ระบบอัจฉริยะ และการจำลองสถานการณ์ สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม มหาสารคาม 44150*

[1] *Master Student, Multi-agent Intelligent Simulation Laboratory (MISL), Department of Information Technology Faculty of Informatics, Mahasarakham University, Maha Sarakham 44150, Thailand. E-mail: jakkarin.san@msu.ac.th*

[2] *Lecturer, Multi-agent Intelligent Simulation Laboratory (MISL), Department of Information Technology Faculty of Informatics, Mahasarakham University, Maha Sarakham 44150, Thailand. E-mail: olarik.s@msu.ac.th*

also faster than the VGGNet architecture. However, the experiment showed that the VGGNet architecture outperforms the other CNN architectures on the Folio dataset and used only 1,000 epochs for training. In our experiment, we can create a model from the deep CNN using GoogleNet architecture, and this is because it showed better results with the plant images that were taken in the natural environment.

Keywords: Plant Recognition, Deep Learning, Deep Convolutional Neural Network, AlexNet Architecture, GoogLeNet Architecture, VGGNet Architecture

บทนำ

โครงข่ายประสาทเทียม (Neural Network) เป็นโครงข่ายที่จำลองหลักการทำงานโครงข่ายประสาทที่อยู่ในสมองของมนุษย์ โดยมีการแบ่งลักษณะของโครงสร้างการทำงานออกเป็นชั้น (Layer) การทำงานของโครงข่ายประสาทเทียมประกอบด้วย ชั้นนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นแสดงผลลัพธ์ (Output Layer) ซึ่งจำนวนหน่วย (Unit หรือ Node) ของชั้นแสดงผลลัพธ์ขึ้นอยู่กับจำนวนหมวดหมู่ (Category) ที่ทำการจำแนกประเภท (Classification) เช่น หากต้องการจำแนกประเภทข้อมูลตัวเลขอารบิก จำนวนของหมวดหมู่จะมีค่าเป็น 10 หมวด ซึ่งหมายถึงค่าความน่าจะเป็น (Probability) ที่จะเป็นคำตอบของตัวเลข 0-9 ดังนั้น หากค่าความน่าจะเป็นที่สูงที่สุดอยู่ในหน่วยใดก็ถือว่าเป็นคำตอบที่ได้จากการคาดการณ์หรือการพยากรณ์ (Predict)

นักวิจัยได้ปรับปรุงโครงข่ายให้มีจำนวนชั้นซ่อนเพิ่มขึ้น เนื่องจากจะช่วยในการคำนวณให้มีประสิทธิภาพมากขึ้น แต่ในทางกลับกัน จะส่งผลให้กระบวนการในการเรียนรู้นานขึ้น ซึ่งวิธีนี้เรียกว่า การเรียนรู้เชิงลึก (Deep Learning)[1-3] โดย LeCun et al.[1,2] ได้นำเสนอโครงข่ายที่มีลักษณะการทำงานแบบคอนโวลูชัน (Convolutional Networks) โครงข่ายลักษณะนี้จึงถูกเรียกว่า โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) โครงสร้าง (Architecture) ถูกกำหนดไว้ 5 ชั้น โครงสร้างนี้มีชื่อเรียกว่า LeNet-5 โดยโครงข่ายประสาทเทียมแบบคอนโวลูชัน จะตัดขั้นตอนในส่วนของการหาคุณลักษณะพิเศษ (Feature Extraction) ทำให้นักวิจัยไม่จำเป็นต้องใช้วิธีการหาคุณลักษณะพิเศษแบบเดิม เช่นการใช้รูปแบบ (Pattern) การใช้รูปร่าง (Shape) และการใช้สี (Colour) แต่ทั้งนี้จะใช้ขั้นตอนในการคอนโวลูชันเป็นขั้นตอนหนึ่งในการคำนวณหาคุณลักษณะพิเศษ การทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชันจะทำงานควบคู่กับชั้นพูลลิ่งเพื่อช่วยลดขนาดของเมทริก (Matrix) ที่ใช้ในการคำนวณลงดังนั้น การเรียนรู้เชิงลึกที่ใช้โครงสร้างแบบ CNN จึงเหมาะสมกับข้อมูลประเภทรูปภาพ ทำให้นักวิจัยสามารถนำการเรียนรู้เชิงลึกไปประยุกต์ใช้ในงานด้านการค้นหา (Detection)[4]we propose a Spatially Constrained Convolutional Neural Network (SC-CNN การจำแนกประเภท[5] contribute to a better understanding of the relationship between environmental factors and healthy crops, reduce the labor costs for farmers and increase the operation speed and accuracy. Implementing machine learning methods such as deep neural networks on agricultural data has gained immense attention in recent years. One of the most important problems is automatic classification of plant species based on their types. Automatic plant type identification process could offer a great help for application of pesticides, fertilization and harvesting of different species on-time in order to improve the production processes of food and drug industries. In this paper, we propose a Convolutional Neural Network (CNN และการจัดกลุ่ม (Clustering)[6] รูปภาพ

จากงานวิจัยของ Liu et al.[7] ได้นำเสนอวิธีการจำแนกประเภทดอกไม้ (Flower Classification) โดยใช้วิธี CNN ที่ปรับเปลี่ยนโครงสร้างให้มีจำนวน 8 ชั้น (Layer) ซึ่งมีชั้นคอนโวลูชัน (Convolutional Layer) จำนวน 5 ชั้น และ ชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-Connected Layer) จำนวน 3 ชั้น และใช้ฟังก์ชันซอฟต์แมกซ์ (Softmax Function) ในการคำนวณหาผลลัพธ์ โดยได้ทดสอบกับชุดข้อมูลดอกไม้ที่มีจำนวนทั้งสิ้น 63,442 รูปภาพ แบ่งออกเป็นดอกไม้ 79 สายพันธุ์ และยังได้ทดสอบกับข้อมูลชุด Oxford 103 Flower Species ที่มีจำนวนทั้งสิ้น 103 สายพันธุ์ ประกอบด้วยรูปภาพดอกไม้ทั้งสิ้น 8,189 รูปภาพ แต่ละสายพันธุ์จะมีรูปภาพของดอกไม้ตั้งแต่ 40 ถึง 200 รูปภาพ จากการทดลองพบว่าวิธีการ CNN โดยกำหนดให้มีจำนวนชั้นทั้งสิ้น 8 ชั้น มีอัตราความถูกต้อง 84.02% จากนั้นจึงได้ทดสอบกับข้อมูลชุด Large-scale 79 Flower Species ซึ่งมีดอกไม้ทั้งสิ้น 79 สายพันธุ์ มีจำนวนรูปภาพทั้งสิ้น 52,775 รูปภาพ ในการทดสอบได้แบ่งรูปภาพจำนวน 47,500 เป็นข้อมูลชุดเรียนรู้ และ 5,275 รูปภาพเป็นข้อมูลชุดทดสอบ พบว่าวิธี CNN ที่นำเสนอมีอัตราความถูกต้อง 76.54% เมื่อเทียบ

กับวิธี CNN แบบ LeNet-5 ที่มีอัตราความถูกต้อง 70.12% ซึ่งมีความแม่นยำเพิ่มขึ้นถึง 6%

Pawara *et al.*[8] ได้เปรียบเทียบวิธีการหาคุณลักษณะพิเศษแบบเฉพาะพื้นที่ (Local Descriptor) และวิธี Bag of Word (BOW) กับวิธี Deep CNNs สำหรับรู้จำพรรณไม้ (Plant Recognition) โดยชุดข้อมูลที่ใช้ในการทดสอบประกอบด้วย AgrilPlant (10 สายพันธุ์ 3,000 รูปภาพ) LeafSnap (184 สายพันธุ์ 7,719 รูปภาพ) และ Folio (32 สายพันธุ์ 640 รูปภาพ) โดยรูปภาพที่ใช้ในการทดลองถูกปรับให้มีขนาด 256x256 พิกเซล วิธี Cross Validation ถูกนำมาใช้ในการแบ่งชุดข้อมูล โดยกำหนดให้ k มีค่าเท่ากับ 5 เพื่อแบ่งข้อมูลเป็น 80% สำหรับเรียนรู้ และ 20% สำหรับทดสอบ โดยทดสอบด้วยวิธี Histogram of Oriented Gradients (HOG) และ HOG-BOW เพื่อสร้างคุณลักษณะพิเศษ จากนั้นจึงนำคุณลักษณะพิเศษที่ได้ไปทำการเรียนรู้ด้วยวิธีการคำนวณเพื่อนบ้านใกล้ที่สุด k ตำแหน่ง (k-Nearest Neighbors: KNN), ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) และเพอร์เซปตรอนแบบหลายชั้น (Multi Layer Perceptron: MLP) จากการทดลองวิธี HOG-BOW ร่วมกับ MLP ให้ประสิทธิภาพสูงสุดในข้อมูลชุด LeafSnap และวิธี HOG-BOW ร่วมกับSVM ให้ประสิทธิภาพสูงสุดในข้อมูลชุด AgrilPlant และ Folio นอกจากนั้นยังได้ทดสอบวิธี Deep CNNs โดยใช้โครงสร้างแบบ AlexNet และ GoogLeNet จากการทดลองพบว่า โครงสร้างแบบ AlexNet มีอัตราความถูกต้องสูงสุด 97.67% สำหรับข้อมูลชุด Folio และ GoogLeNet มีอัตราความถูกต้องสูงที่สุดในข้อมูลชุด AgrilPlant และ LeafSnap โดยมีความถูกต้องที่ 98.33% และ 97.66% ตามลำดับ

Reyes *et al.*[9] นำเสนอวิธี Deep CNNs เพื่อรู้จำประเภทของพืช โดยใช้ชุดข้อมูล LifeCLEF 2015 ในการทดสอบ งานวิจัยนี้ใช้วิธีการถ่ายโอนความรู้ (Transfer Learning) ซึ่งใช้ชุดข้อมูล ILSVRC2012 ซึ่งประกอบด้วยรูปภาพ 1.2 ล้านรูป และมีหมวดหมู่มากถึง 1,000 หมวดในการเรียนรู้เพื่อสร้างโมเดล จากนั้นจึงนำค่าน้ำหนัก (Weight) ที่ได้มาเรียนรู้อีกครั้งกับข้อมูล LifeCLEF 2015 วิธีนี้ทำให้ลดระยะเวลาในการเรียนรู้ ชุดข้อมูล LifeCLEF 2015 ที่ใช้ในการเรียนรู้มีจำนวน 91,759 รูปภาพ โดยเป็นพันธุ์ไม้จำนวน 1,000 สายพันธุ์ ภายในชุดข้อมูลได้แบ่งแยกรูปภาพของพืช โดยแบ่งออกเป็น 7 มุมมอง ประกอบด้วย ภาพถ่ายกิ่ง ดอก ก้านดอก ผล ใบไม้ กลุ่มของใบไม้ และภาพรวมของต้นไม้ โดยได้นำมุมมองทั้ง 7 มุมไปทดสอบ ผลการทดสอบพบว่ากลุ่มภาพถ่ายดอกมีอัตราความแม่นยำสูงที่สุดประมาณ 65% โดยรวมแล้วมีความแม่นยำเฉลี่ยเท่ากับ 48.6%

ในงานวิจัย[10,11] โครงข่ายประสาทเทียมเชิงลึก (Deep Neural Networks) ถูกนำมาใช้เพื่อรู้จำพันธุ์พืชที่เป็นโรค (Plant Disease Recognition) โดย Sladojevic *et al.*[10] ได้นำรูปภาพใบไม้ (Leaf) มาเพื่อทดลอง โดยมีรูปตัวอย่างใบไม้ที่เป็นโรคจำนวน 13 ประเภท และรวมกับใบไม้ที่อยู่ในสภาพปกติอีก 2 ประเภท รวมเป็น 15 ประเภท ในการทดลองใช้เทคนิคในการเพิ่มข้อมูลในการเรียนรู้ (Augmented Data) ด้วยวิธีในการประมวลผลภาพ (Image Processing) ประกอบด้วย การนำภาพไปบิดให้ผิดรูป (Distortion) และการหมุนภาพ (Rotation) เป็นต้น ข้อมูลที่ใช้ในการทดสอบมีจำนวนทั้งสิ้น 4,483 รูปภาพ และเมื่อนำไปผ่านวิธี Augmented จะได้รูปภาพเพิ่มขึ้นเป็น 30,880 รูปภาพ จากการทดลองพบว่ามีอัตราความถูกต้อง 96.3%

งานวิจัย[11] วิธีการเรียนรู้เชิงลึกถูกนำมาช่วยในการวินิจฉัยโรคพืชจากภาพถ่าย ซึ่งในการทดลองใช้ชุดข้อมูล PlantVillage ที่มีจำนวน 54,306 รูปภาพจากพืชทั้งสิ้น 14 ชนิด โดยระบุชนิดของโรค ไว้ทั้งสิ้น 26 โรค โดยรูปภาพได้ถูกปรับให้มีขนาด 256x256 พิกเซล และใช้วิธีการเรียนรู้เชิงลึกที่มีโครงสร้างแบบ AlexNet และ GoogLeNet โดยทำการทดสอบในสองรูปแบบประกอบด้วย เริ่มต้นเรียนรู้จากศูนย์ (Training from Scratch) และเรียนรู้แบบวิธีการถ่ายโอนความรู้ในการทดสอบประสิทธิภาพของการเรียนรู้เชิงลึกใช้วิธี Cross Validation เพื่อแบ่งข้อมูลออกเป็นข้อมูลชุดเรียนรู้ 80, 60, 50, 40 และ 20% ตามลำดับ จากการทดลองพบว่าเมื่อแบ่งข้อมูลออกเป็นชุดเรียนรู้จำนวน 80% และชุดทดสอบจำนวน 20% โครงสร้างแบบ GoogleNet ที่ใช้การเรียนรู้แบบ วิธีการถ่ายโอนความรู้มีอัตราความถูกต้องสูงที่สุดคือ 99.34%

งานวิจัยของ Sun *et al.*[12] นำเสนอวิธีการจำแนกพรรณพืช (Plant Identification) ที่เป็นรูปถ่ายที่อยู่ในสภาวะแวดล้อมทางธรรมชาติโดยเป็นพันธุ์ไม้โดยข้อมูลชุด BJFU100 ประกอบไปด้วยพันธุ์ไม้จำนวน 100 ชนิด ซึ่งเป็นรูปภาพดอกไม้ทั้งสิ้น 10,000 รูปภาพ โดยใช้วิธี CNN ที่มีโครงสร้างแบบ ResNet โดยทดลองจำนวนชั้นตั้งแต่ 18, 26, 34 และ 50 ชั้น ในการทดลองได้แบ่งข้อมูลชุดเรียนรู้จำนวน 80% และชุดทดสอบ 20% จากการทดลองปรากฏว่าโครงสร้างแบบ ResNet ที่มีจำนวน 26 ชั้น มีความถูกต้อง 91.78% และเมื่อนำไปทดสอบกับชุดข้อมูล Flavia มีอัตราความถูกต้อง 99.65%

งานวิจัยที่เกี่ยวข้องที่ได้กล่าวมาข้างต้น สามารถนำวิธีการเรียนรู้เชิงลึกไปใช้เพื่อจำแนกรูปภาพพรรณไม้ ทั้งที่จำแนกจากดอกไม้ (Flower) ใบไม้ (Leaf) และภาพถ่ายดอกไม้ที่อยู่ในสภาวะแวดล้อมทางธรรมชาติ และยังสามารถนำไป

ช่วยในการวินิจฉัยโรคที่เกิดขึ้นกับพรรณไม้ได้

งานวิจัยฉบับนี้ได้ศึกษาโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึกซึ่งเปรียบเทียบกับโครงสร้างที่มีความแตกต่างกันจำนวน 3 รูปแบบ ประกอบด้วย AlexNet, GoogLeNet และ VGGNet[2,13-15] เพื่อนำมาใช้สำหรับจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ (Plant Classification in Natural Environment) อีกทั้งยังได้นำเสนอชุดข้อมูล PNE ที่เก็บรวบรวมพรรณไม้ทั้งสิ้น 10 ชนิด โดยทุกรูปเป็นภาพถ่ายดอกไม้ที่อยู่ในสภาวะแวดล้อมทางธรรมชาติซึ่งยากต่อการจำแนก เนื่องจากในรูปภาพอาจมีทั้งรูปดอกไม้ ใบไม้ และอาจรวมถึงรูปองค์ประกอบอื่นเช่น ก้าน หรือพื้นหลังที่เป็นวัตถุอื่นเป็นต้น

โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN)

โดยทั่วไปแล้ว โครงสร้างของโครงข่ายประสาทเทียมประกอบด้วย ชั้นนำเข้า (Input Layer), ชั้นซ่อน (Hidden Layer) และชั้นแสดงผลลัพธ์ (Output Layer) หรือเรียกว่า Multi-Layer Perceptron (MLP) แต่งานวิจัยที่นำเสนอโดย LeCun *et al.*[2] ได้เพิ่มการคำนวณที่มีลักษณะแบบคอนโวลูชัน (Convolutional) เข้าไปในโครงข่าย จึงเรียกว่าโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network: CNN) โดยโครงข่าย CNN ประกอบด้วยชั้นคอนโวลูชัน ชั้นพูลลิ่ง และชั้นเชื่อมโยงแบบสมบูรณ์ ซึ่งชั้นเชื่อมโยงแบบสมบูรณ์ ก็คือชั้นซ่อน และชั้นแสดงผลลัพธ์ที่ปรากฏในโครงข่ายประสาทเทียม

ทั้งนี้ โครงข่าย CNN สามารถทำได้ทั้งการสกัดคุณลักษณะพิเศษของรูปภาพ และการจำแนกประเภท[2] ทำให้เป็นจุดเด่นของการเรียนรู้ด้วยของโครงข่าย CNN ซึ่งมีความแตกต่างกับวิธีการเรียนรู้เครื่องจักร (Machine Learning) ทั่วไปที่จะทำหน้าที่เพียงจำแนกประเภทข้อมูลหรือจัดกลุ่มข้อมูลเพียงเท่านั้น

ชั้นคอนโวลูชัน (Convolutional Layer)

คอนโวลูชันเป็นการคำนวณเพื่อหาผลลัพธ์ของ นิวรอล (Neuron) ที่เชื่อมต่อ (Connected) มาจากพื้นที่ส่วนย่อย (Local Region) ของรูปภาพ โดยรูปภาพจะถูกกำหนดให้เป็นสี่เหลี่ยมจัตุรัสขนาด $W \times W \times M$ โดยที่ W คือความสูงและความกว้างของรูปภาพ และ M คือมิติของรูปภาพ เช่น ภาพสี RGB จะถูกกำหนดให้ $M = 3$

การคำนวณในชั้นคอนโวลูชัน พื้นที่ส่วนย่อยจะถูกนำมาคำนวณแบบ dot product กับเคอร์นอล (Kernel) ที่มีขนาด $H \times H \times K$ โดยเคอร์นอลต้องมีขนาดเล็กกว่าขนาดของรูปภาพ และ K คือจำนวนของเคอร์นอลที่ต้องการคำนวณผลลัพธ์ที่ได้จากการคอนโวลูชัน เรียกว่า Feature Map การคำนวณ คอนโวลูชัน เพื่อหา Feature Map แสดงดัง Equation (1)

$$u_{ijm} = \sum_{k=0}^{K-1} \sum_{p=0}^{H-1} \sum_{q=0}^{H-1} x_{i+p,j+q,k} \, h_{pqkm} + b_{ijm} \quad (1)$$

โดยที่ x คือ รูปภาพที่ใช้ในการคำนวณ
h คือ เคอร์นอลขนาด $H \times H$
b คือ Bias

Figure 1 แสดงให้เห็นถึงวิธีการคอนโวลูชัน ระหว่าง Local Region และเคอร์นอล

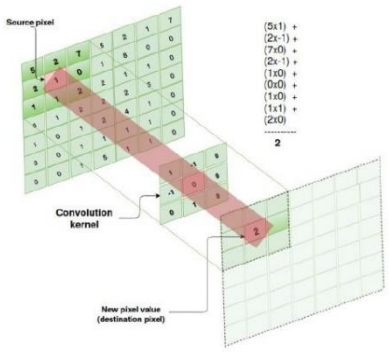


Figure 1 The illustration of the convolutional Layer. The dot production calculation between sub-region of the image and convolution kernel.

ชั้นพูลลิ่ง (Pooling Layer)

โดยทั่วไปแล้วชั้นพูลลิ่ง เป็นชั้นที่คั่นกลางระหว่างชั้นคอนโวลูชัน โดยมีไว้เพื่อลดขนาด (Downsample) ของ Feature Map ให้เล็กลง โดยใช้ฟังก์ชันค่าเฉลี่ย ฟังก์ชันค่าต่ำสุด และฟังก์ชันค่าสูงสุดในการคำนวณ หากเลือกใช้ฟังก์ชันค่าสูงสุดในการคำนวณ ดังนั้น ชั้นนี้เรียกว่า Max Pooling นั่นหมายถึง Feature Map จะถูกแบ่งออกเป็น Local Region จากนั้น Local Region จะถูกแบ่งออกเป็น p x p ส่วน p โดยจะกำหนดให้อยู่ในช่วงระหว่าง 2 และ 5 ดังนั้น ค่าที่มากที่สุดในแต่ละ Pool จะถูกเลือกเพื่อนำมาเป็นตัวแทน จากนั้นระบบจะเลื่อน (Stride) ไปยัง Local Region ต่อไปโดยจะเลื่อนไปเรื่อย ๆ เช่น ครั้งละ 2 พิกเซล ไปจนกระทั่งจุดสุดท้ายของ Feature Map การคำนวณ Max Pooling แสดงดัง Figure 2

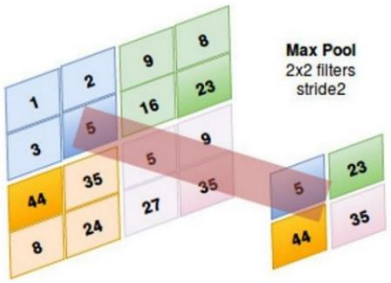


Figure 2 The illustration of max pooling with 2 x 2 kernel and stride 2.

ชั้นเชื่อมโยงแบบสมบูรณ์ (Fully-Connected Layer)

ชั้นเชื่อมโยงแบบสมบูรณ์ เป็นชั้นที่ทำหน้าที่ในการจำแนกประเภทของวัตถุ (Object) ซึ่งทุก Neuron ที่อยู่ในชั้นเชื่อมโยงแบบสมบูรณ์จะถูกเชื่อมโยงกับชั้นคอนโวลูชัน และชั้นพูลลิ่ง อย่างสมบูรณ์ ซึ่งก็คือโครงสร้างของนิวรอลเน็ตเวิร์ก (Neural Network) โดยจำนวนผลลัพธ์คือจำนวนของกลุ่มที่ต้องการจำแนก ซึ่งคำนวณด้วยการหาค่าความน่าจะเป็น (Probability) ด้วยฟังก์ชันซอฟต์แมกซ์ แสดงดังสมการ 2

$$f_i(x) = \frac{\exp(x_i)}{\sum_j^n \exp(x_j)} \qquad (2)$$

โดยที่คือผลลัพธ์ของแต่ละโหนด

โครงสร้างแบบ LeNet-5 (LeNet-5 Architecture)

ปี ค.ศ. 1989 นักวิจัย LeCun *et al.*[1] ได้นำเสนอโครงข่าย CNN โดยมีจุดประสงค์เพื่อการรู้จำตัวเลขรหัสไปรษณีย์ในการทดลอง รูปภาพของตัวเลขจะถูกปรับให้มีขนาด 16x16 พิกเซล (256 Input Unit) โครงสร้างถูกกำหนดให้มีชั้นซ่อนจำนวน 3 ชั้น (H1-H3) โดยชั้น H1 ประกอบด้วยเคอร์นอลขนาด 5x5 จำนวน 12 เคอร์นอล เมื่อผ่านการคำนวณจำนวนของ Hidden Unit จะมีขนาด 768 (12x64) และในชั้น H2 มีจำนวนทั้งสิ้น 12 เคอร์นอล แต่ละเคอร์นอล มีขนาด 5x5x8 และเมื่อผ่านการคำนวณชั้น H2 จะมีชั้นซ่อนขนาด 192 (12x16) โหนด จากนั้นโหนดในชั้นซ่อน ทั้งหมดจะถูกส่งไปยังชั้น H3 ซึ่งเป็นชั้นของการเชื่อมโยงแบบสมบูรณ์ที่กำหนดให้ชั้นซ่อนมีจำนวน 30 โหนด และสุดท้ายส่งออกไปยังชั้นแสดงผลลัพธ์ที่กำหนดให้เป็น 10 โหนด เพื่อให้สอดคล้องกับตัวเลข 0-9 โดยในชั้นเชื่อมโยงแบบสมบูรณ์เป็นการทำงานแบบโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back-propagation Neural Network)

ปี ค.ศ. 1998 ในงานวิจัย[2] ได้นำเสนอโครงข่าย CNN ที่กำหนดโครงสร้างแบบ LeNet-5 โดยโครงสร้างประกอบไปด้วยชั้นคอนโวลูชันจำนวน 2 ชั้น และชั้นพูลลิ่ง จำนวน 2 ชั้น และชั้นสุดท้ายคือชั้นเชื่อมโยงแบบสมบูรณ์ เพื่อใช้สำหรับการจำแนกวัตถุ ดังนั้น โครงสร้าง LeNet-5[2] จึงมีความแตกต่างกับโครงสร้างที่นำเสนอในงานวิจัย[1] คือการเพิ่มชั้นพูลลิ่งลงไปเพื่อเป็นการลดขนาดของ Feature Map โครงสร้างแบบ LeNet-5 แสดงดัง Figure 3

Figure 3 The architecture of LeNet-5.

โครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก (Architecture of the Deep Convolutional Neural Networks)

โครงสร้างของโครงข่ายแบบ CNN สามารถเพิ่มจำนวนของชั้นคอนโวลูชัน ได้ไม่จำกัดจึงทำให้เกิดโครงสร้างลักษณะใหม่เกิดขึ้น เช่น AlexNet, GoogLeNet และ VGGNet เป็นต้น

โครงสร้างแบบ AlexNet (AlexNet Architecture)

ปี ค.ศ. 2012 นักวิจัย Krizhevsky *el al.*[13] นำเสนอ Deep CNN ที่มีจำนวนทั้งสิ้น 8 ชั้น โดยโครงสร้างนี้เรียกว่า AlexNet ประกอบไปด้วยชั้นคอนโวลูชัน และชั้นเชื่อมโยงแบบ

สมบูรณ์ จำนวน 5 และ 3 ชั้น ตามลำดับ โดยผลลัพธ์ของชั้นเชื่อมโยงแบบสมบูรณ์ ถูกกำหนดให้มีจำนวน 100 โหนด เนื่องจากในงานวิจัยได้ใช้ชุดข้อมูล ImageNet LSVRC-2010 ในการทดสอบ ข้อมูลมีจำนวนทั้งสิ้น 1.2 ล้านรูปภาพ และแบ่งออกเป็น 1,000 กลุ่ม ในชั้นสุดท้ายของชั้นเชื่อมโยงแบบสมบูรณ์ ใช้วิธีซอร์ฟแม็กซ์เพื่อคำนวณหาผลลัพธ์

รูปภาพที่ใช้ในการทดลองจะถูกเปลี่ยนให้มีขนาด 224x224x3 พิกเซล ดังนั้น ชั้นคอนโวลูชัน (Conv{1}) ในชั้นแรกถูกกำหนดให้มี 96 เคอร์นอลโดยที่แต่ละเคอร์นอลมีขนาด 11x11x3 ในการคำนวณจะเลื่อน (Stride) เคอร์นอลครั้งละ 4 พิกเซล ผลลัพธ์จากการคอนโวลูชันจะมีขนาด 55x55x96 ชั้นที่สองของ ชั้นคอนโวลูชัน (Conv{2}) ถูกกำหนดให้มี 256 เคอร์นอลโดยที่แต่ละเคอร์นอลมีขนาด 5x5x48 ผลลัพธ์จากการคอนโวลูชันจะมีขนาด 27x27x256หลังจากการคำนวณ Conv{1, 2} ข้อมูลจะถูกส่งไปชั้นพูลลิ่งเพื่อทำการ Normalized และหาค่าสูงสุดในแต่ละ Pool (Max Pooling) โดยที่แต่ละ Pool จะถูกแบ่งเป็น 2x2 ส่วน

ในคอนโวลูชันชั้นที่สาม จำนวนของเคอร์นอลถูกกำหนดให้มี 384 เคอร์นอล โดยที่แต่ละเคอร์นอลมีขนาด 3x3x256 ผลลัพธ์จากการคอนโวลูชันจะมีขนาด 13x13x256 โดยชั้นที่สี่ ถูกกำหนดให้มี 384 เคอร์นอล โดยที่แต่ละเคอร์นอล มีขนาด 3x3x192 ผลลัพธ์จากการคอนโวลูชัน จะมีขนาด 13x13x384 ชั้นที่ห้า มีจำนวน 256 เคอร์นอล โดยแต่ละเคอร์นอลมีขนาด 3x3x192 สุดท้ายผลลัพธ์จากการคอนโวลูชัน มีขนาด 13x13x256 โดยที่ Conv{1-5} ใช้ฟังก์ชัน Rectified Linear Units (ReLu) เป็นฟังก์ชันกระตุ้น (Activation Function)

สุดท้ายแล้ว ชั้นเชื่อมโยงแบบสมบูรณ์ถูกแบ่งออกเป็น 2 ชั้น และกำหนดให้มีจำนวนชั้นละ 4,096 โหนด (Node) โครงสร้าง AlexNet แสดงดัง Figure 4

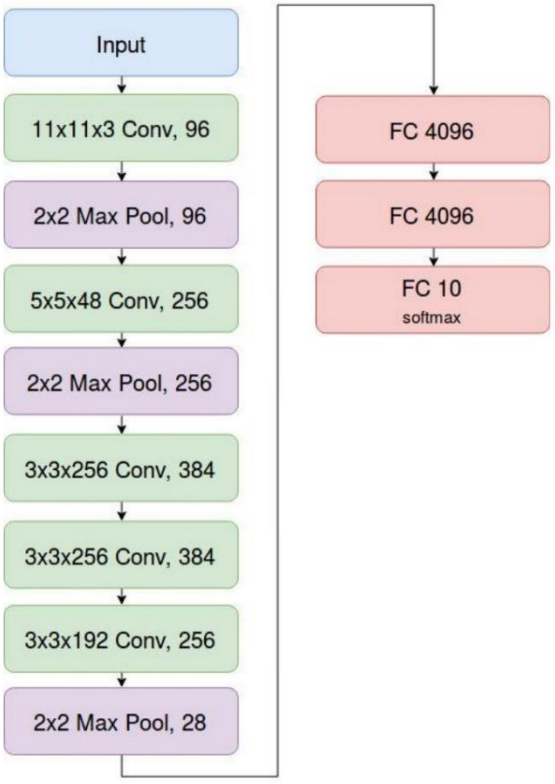


Figure 4 The AlexNet architecture.

โครงสร้างแบบ GoogLeNet (GoogLeNet Architecture)

โครงสร้างแบบ GoogLeNet[14] ถูกนำเสนอในปี ค.ศ. 2014 โดยใช้โครงสร้าง Inception (Inception Architecture) เพื่อให้ CNN มีความลึกและกว้าง อีกทั้งยังไม่ทำให้การประมวลผลช้าลง ในแต่ละ Inception ประกอบไปด้วยการคำนวณคอนโวลูชันที่ใช้เคอร์นอล ขนาด 1x1, 3x3 และ 5x5 ซึ่งเป็นการประมวลผลแบบขนาน (Parallel) อีกทั้งยังทำให้จำนวนมิติของข้อมูลลดลง (Dimension Reduction) (แสดงดัง Figure 5) โครงสร้างแบบ GoogLeNet ถูกออกแบบให้มีโครงสร้างแบบ Inception จำนวน 9 ชั้น จากจำนวนทั้งสิ้น 22 ชั้น และใช้ฟังก์ชัน ReLu เพื่อเป็นฟังก์ชันกระตุ้น และชั้นเชื่อมโยงแบบสมบูรณ์กำหนดให้มีจำนวน 1024 โหนด ก่อนจะส่งต่อไปยังชั้นแสดงผลลัพธ์ โครงสร้างแบบ GoogLeNet แสดงดัง Figure 6

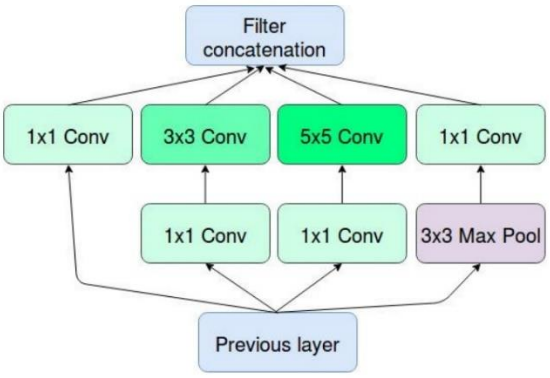


Figure 5 Inception Structure

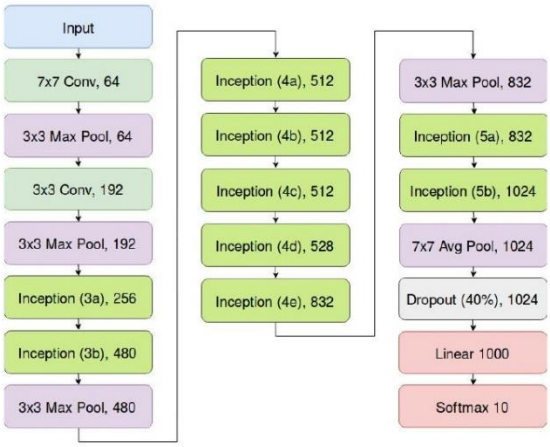


Figure 6 The googLeNet architecture.

โครงสร้างแบบ VGGNet (VGGNet Architecture)

ในปี ค.ศ. 2015 งานวิจัย[15] ได้นำเสนอโครงสร้างแบบ VGGNet ซึ่งเป็นโครงสร้างแบบ Deep CNN เนื่องจากมีความลึกของชั้นจำนวน 16 โดยใช้ Kernel ขนาด 3x3 ในการ Convolution

โครงสร้างของ VGGNet-16 ประกอบด้วย ชั้นคอนโวลูชัน (Conv) จำนวน 5 กลุ่ม โดย Conv กลุ่มที่ 1 หรือ Conv1 ประกอบด้วย Conv1 {1, 2} และมีจำนวน เคอร์นอล ชั้นละ 64 เคอร์นอล ใน Conv2 ประกอบด้วย Conv2{1, 2} มีจำนวนเคอร์นอลชั้นละ 128 เคอร์นอล ใน Conv3 ประกอบด้วย Conv3 {1, 2, 3} และมีจำนวน Kernel ชั้นละ 256 เคอร์นอล ใน Conv4 ประกอบด้วย Conv4{1, 2, 3} และมีจำนวน เคอร์นอล ชั้นละ 512 เคอร์นอล และชั้น Conv5 ประกอบด้วย Conv5{1, 2, 3} ในแต่ละชั้นมีจำนวน เคอร์นอล ชั้นละ 512 เคอร์นอล โดยที่ใน Conv1-5 จะตามด้วยชั้นแมกซ์พูลลิ่ง (Max Pooling) จากนั้นข้อมูลจะส่งไปยัง ชั้นเชื่อมโยงแบบสมบูรณ์ ที่มีจำนวนโหนด 4,096 และ 4,096โหนด และในชั้นสุดท้ายของชั้นเชื่อมโยงแบบสมบูรณ์ ถูกกำหนดให้มีจำนวน 1,000 โหนด โดยคำนวณหาผลลัพธ์ด้วยวิธีซอฟต์แมกซ์ โครงสร้างแบบ VGGNet-16 แสดงดัง Figure 7

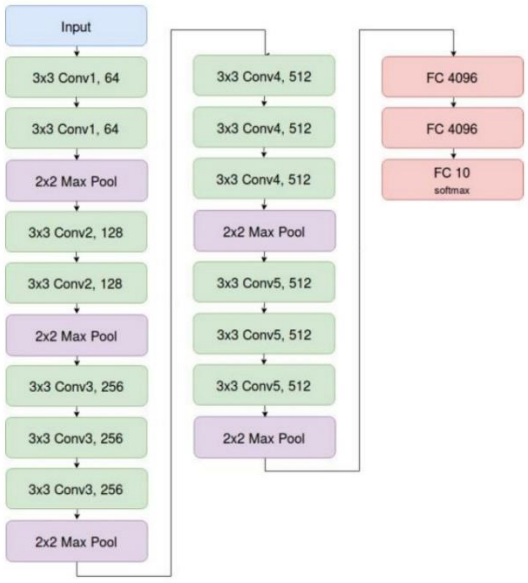


Figure 7 Architecture of VGGNet16.

ข้อมูลพรรณไม้ (Plant Datasets)

ข้อมูลพรรณไม้ที่ใช้ในงานวิจัยนี้มีทั้งสิ้น 3 ชุดข้อมูล ประกอบด้วยชุดข้อมูล Plants in Natural Environment (PNE), Folioและ 102 Flower

ชุดข้อมูลพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ (Plants in Natural Environment Dataset: PNE Dataset)

ชุดข้อมูล PNE ที่นำเสนอในงานวิจัยฉบับนี้ประกอบด้วยพรรณไม้จำนวน 10 ชนิด ชนิดละ 300 รูปภาพ ชนิดของพรรณไม้ที่อยู่ในชุดข้อมูล PNE แสดงดัง Table 1 และตัวอย่างของพรรณไม้แสดงดัง Figure 8 รูปภาพพรรณไม้เป็นรูปภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติที่เก็บรวบรวมจากสวนไม้หอมและบริเวณรอบมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์[16] โดยรูปภาพบางส่วนเก็บรวบรวมมาจากอินเทอร์เน็ต สุดท้ายแล้ว รูปภาพพรรณไม้ทุกรูปถูกนำมาคัดกรองและยืนยันความ ถูกต้องของสายพันธุ์ (Specie)

ความท้าทายในการจำแนกประเภทรูปภาพพรรณไม้ในชุดข้อมูล PNE ได้แก่ ความแตกต่างของพรรณไม้ชนิดเดียวกัน แต่มีสีสัน (Color) ที่แตกต่างกัน เช่น ดอกกุหลาบหิน มีทั้งดอกสีชมพูอ่อน ชมพู และแดง (Figure 8 แถวที่ 1) ทั้งนี้มีดอกไม้จำนวน 5 ประเภทที่เป็นดอกไม้สีขาว ประกอบด้วย ดอกนางแย้ม พุทธชาด มะลิลา โมกพวง ลีลาวดี (Figure 8 แถวที่ 2, 4-7) ดอกพุทธชาดและมะลิลา (Figure แถวที่ 4, 5)

ที่รูปของดอกไม้ไม่ได้เป็นจุดเด่นของภาพ และดอกส่าเหล้าปัตตานี ที่แสดงใน Figure 8 แถวที่ 10 แสดงถึงรูปร่าง (Shape) ที่มีความแตกต่างกัน โดยมีทั้งดอกตูม และดอกที่กำลังบาน

Table 1 List of Thai Plants in the PNE Dataset.

Class No.	Plant Name	Thai Plant Name
1	Fairy Rose	กุหลาบหนู
2	Glory Tree	นางแย้ม
3	Antigonia	พวงชมพู
4	Jasminum Auriculatum	พุทธชาด
5	Arabian Jasmine	มะลิลา
6	Wild Water Plum	โมกพวง
7	Plumeria	ลีลาวดี
8	Rangoon Creeper	เล็บมือนาง
9	Cape of Good Hope	วาสนา
10	Dwarf Ylang-Ylang Shrub	ส่าเหล้าปัตตานี

ชุดข้อมูล 102 Flower (102 Flower Dataset)

ชุดข้อมูล 102 Flower[17] เป็นรูปภาพของจำนวน 102 สายพันธุ์ โดยแต่ละสายพันธุ์ถูกเก็บรวบรวมตั้งแต่ 40-250 รูปภาพ รวมทั้งสิ้น 8,189 รูปภาพ โดย 102 Flower Dataset เป็นรูปภาพดอกไม้ที่พบทั่วไปในประเทศอังกฤษ (แสดงดัง Figure 9) และถูกนำไปใช้ในการจำแนกประเภทดอกไม้ (Flower Classification) โดยการจำแนกประเภทนั้น ใช้วิธีการ SIFT, HOG และ HSV Color Space เพื่อหาคุณลักษณะพิเศษ และใช้วิธี SVM ในการจำแนกประเภทดอกไม้

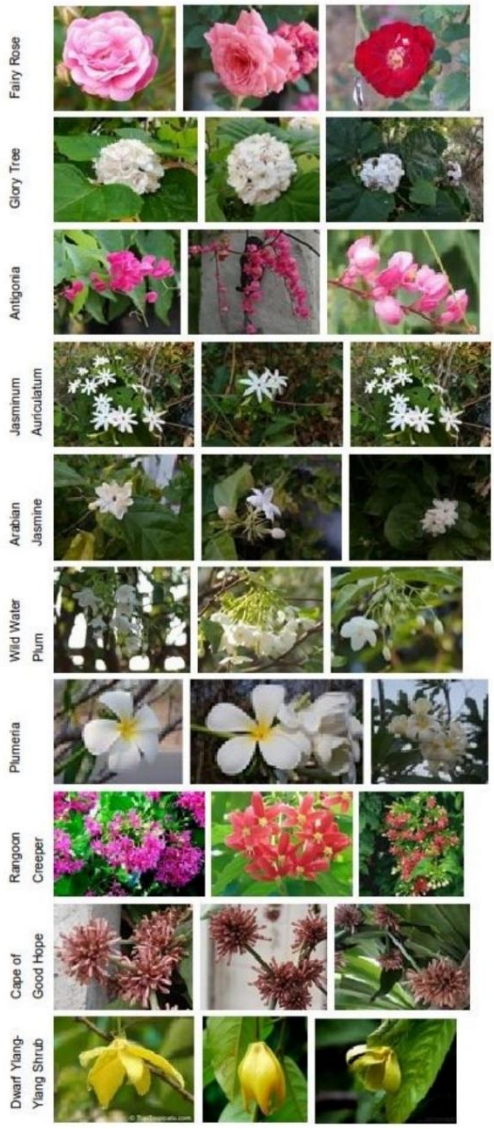


Figure 8 Images from the PNE dataset. Each row shows 3 images from the same category.

Figure 9 Sample images from 102 Flower dataset.

ชุดข้อมูลFolio (Folio Dataset)

Folio Dataset[18] ประกอบด้วยใบไม้ทั้งสิ้น 32 ชนิด ซึ่งเก็บรวบรวมชนิดละ 20 ใบ รวมแล้วมีรูปภาพใบไม้ทั้งสิ้น 640 รูปภาพ ที่เก็บรวบรวมจากบริเวณ University of Mauritius รูปภาพใบไม้ถูกถ่ายโดยกำหนดให้มีพื้นหลังเป็นฉากสีขาว โดย Folio Dataset ถูกเก็บรวบรวมเพื่อนำรูปภาพใบไม้ไปหาคุณลักษณะพิเศษ โดยใช้ลักษณะรูปร่างของใบ (Shape Feature) และ ฮิสโตแกรมสี (Color Histogram) และใช้ kNN เพื่อรู้จำประเภทใบไม้ (Plant Leaf Recognition) ตัวอย่างของ Folio Dataset แสดงใน Figure 10

Figure 10 Sample images from Folio dataset.

ผลการทดลองและอภิปรายผล (Result and Discussion)

งานวิจัยฉบับนี้ รูปภาพที่ใช้ในการวิจัยถูกปรับให้มีขนาด 256x256 พิกเซล โดยชุดข้อมูลถูกแบ่งด้วยสัดส่วน 80:20 เพื่อใช้เป็นข้อมูลชุดเรียนรู้ (Training Set) และข้อมูลชุดทดสอบ (Test Set) ตามลำดับ ใช้วิธี Stratified Cross-Validation[19] กำหนดให้มีจำนวน 5-fold โดยใช้ค่าความถูกต้อง (Accuracy) และค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) สำหรับการประเมินประสิทธิภาพการทดลอง โดยใช้อัตราความถูกต้องแบบ Top-1 (Top-1 Accuracy) และใช้ GPU รุ่น GeForce GTX 1060 แรม 6GB ในการทดลอง

ชุดข้อมูลที่ใช้ในการทดลองจำนวน 3 ชุด ประกอบด้วย ชุดข้อมูล PNE, 102 Flower และ Folio (รายละเอียดของชุดข้อมูลแสดงดัง Table 2) ถูกนำมาใช้เพื่อทดสอบประสิทธิภาพของวิธีการเรียนรู้เชิงลึก (Deep Learning) ด้วยโครงข่ายประสาทเทียมแบบ

คอนโวลูชัน (Convolutional Neural Network: CNN) โดยโครงสร้าง (Architecture) ที่นำมาเพื่อเปรียบเทียบประกอบด้วย LeNet-5, AlexNet, GoogLeNet และ VGGNet[2,13-15]

Table 2 Overview of plant datasets.

Dataset	Category	Image	Training	Test
PNE	10	3,000	2400	600
Folio	32	637	510	127
102 Flower	102	8,189	6,564	1,625

การประเมินผลการทดลองในชุดข้อมูล PNE (PNE Dataset Evaluation)

สำหรับการทดสอบกับชุดข้อมูล PNE ที่ประกอบไปด้วยสมุนไพร 10 ชนิด และมีรูปภาพสมุนไพรทั้งสิ้น 3,000 รูปภาพ โดยใช้วิธี Deep CNN พบว่าวิธี CNN ที่ใช้โครงสร้างแบบ GoogLeNet ที่กำหนดให้เรียนรู้จำนวน 10,000 รอบ (Epoch) มีอัตราความถูกต้อง 99.17% ซึ่งสูงที่สุด ตามด้วย VGGNet (เรียนรู้จำนวน 1,000 และ 10,000 รอบ) มีอัตราการเรียนรู้ 97.17 และ 97.33% ตามลำดับ แต่ทั้งนี้หากใช้การเรียนรู้จำนวน 10,000 รอบ พบว่า VGGNet และ AlexNet มีอัตราความถูกต้องเท่ากันที่ 97.33% แต่ VGGNet มีค่าความเบี่ยงเบนมาตรฐานที่ต่ำกว่า แสดงว่าจากการทดลองหลายรอบ VGGNet ให้ผลความถูกต้องที่ไม่เกิดการกระจายเมื่อเทียบกับ AlexNet

หากเปรียบเทียบความเร็วในการประมวลผลของ Deep CNN ในการทดสอบด้วยการเรียนรู้จำนวน 10,000 รอบ โครงสร้างแบบ VGGNet ใช้เวลาในการประมวลผลนานที่สุดประมาณ 6 วัน 7 ชั่วโมง (~6.7d) ตามด้วย GoogLeNet, LeNet-5 และ AlexNet เวลาในการประมวลผล ~1.9d, ~14h และ ~6.5h ตามลำดับ ผลการทดลอง PNE Dataset แสดงดัง Table 3

Table 3 Test accuracy comparison of CNN architectures on PNE dataset.

CNN Architectures	Number of Epochs / Training Time			
	1,000	Time	10,000	Time
LeNet-5	88.50±0.14	~1.3h	89.17±2.13	~14h
AlexNet	96.67±0.23	~1.5h	97.33±2.18	~6.5h
GoogLeNet	**99.04±0.11**	~4.3h	**99.17±1.75**	~1.9d
VGGNet	97.17±0.64	~20h	97.33±0.85	~6.7d

การประเมินผลการทดลองในชุดข้อมูล 102 Flower (102 Flower Dataset Evaluation)

การเปรียบเทียบระหว่างวิธี Deep CNN และวิธีที่รวมกันระหว่าง HSV+SIFT+HOG[17] โดยนำมาทดสอบกับชุดข้อมูล 102 Flower พบว่าวิธี Deep CNN ที่มีโครงสร้างแบบ GoogLeNet (เรียนรู้จำนวน 10,000 รอบ) มีอัตราความถูกต้องสูงที่สุด โดยมีความถูกต้อง 78.89% ซึ่งสูงกว่าวิธี HSV+SIFT+HOG ที่มีความถูกต้อง 72.8%

หากเปรียบเทียบเฉพาะ Deep CNN พบว่าจำนวนในการเรียนรู้ 10,000 รอบให้ผลการทดลองที่สูงกว่าจำนวน 1,000 รอบ โดยโครงสร้างที่ให้อัตราความถูกต้องสูงสุดเรียงตามลำดับคือ GoogLeNet, AlexNet, VGGNet และ LeNet-5 โดยมีอัตราความ ถูกต้อง 78.89, 66.38, 64.12 และ 31.42% ตามลำดับ ผลลัพธ์ที่ได้จากการทดลองแสดงดัง Table 4

Table 4 The accuracy of CNN architectures on 102 Flower dataset.

CNN Architectures	Number of Epochs / Training Time			
	1,000	Time	10,000	Time
LeNet-5	33.19±0.27	~4.4h	31.42±1.12	~1.16d
AlexNet	66.06±1.74	~6.3h	66.38±2.46	~1.6d
GoogLeNet	**74.98±1.59**	~1.3d	**78.89±0.69**	~4.4d
VGGNet	62.60±0.87	~1.19d	64.12±2.74	~17d

การประเมินผลการทดลองในชุดข้อมูล Folio (Folio Dataset Evaluation)

ในงานวิจัย[18] ใช้วิธีการหาคุณลักษณะพิเศษจากรูปร่างของใบไม้ (Shape Feature) และใช้ ฮีสโต แกรมของสี (Color Histogram) ร่วมกับ KNN เพื่อรู้จำใบไม้จากชุดข้อมูล Folio จากงานวิจัยระบุว่าวิธีดังกล่าวมีอัตราความถูกต้อง 87.3% ซึ่งมีอัตราความถูกต้องสูงกว่าวิธี CNN ที่ใช้โครงสร้างแบบ LeNet-5 และ AlexNet โดยโครงสร้างแบบ LeNet-5 และ AlexNet มีอัตราความถูกต้อง 84.25 และ 85.04% เมื่อทำการเรียนรู้จำนวน 1,000 รอบ ในทางกลับกัน กรณีที่เรียนรู้ด้วยจำนวนรอบที่เพิ่มขึ้น เมื่อทดสอบที่ 10,000 รอบ อัตราความถูกต้องกลับลดลง โดยโครงสร้างแบบ LeNet-5 มีความถูกต้อง 74.02% และโครงสร้างแบบ AlexNet มีความถูกต้อง 73.23%

เมื่อนำผลการทดลองจากงานวิจัย[18] มาเปรียบเทียบกับโครงสร้างแบบ GoogLeNet กลับพบว่ามีอัตราความถูกต้องที่ใกล้เคียงกัน โดย GoogLeNet มีอัตราความถูกต้อง 87.40% ซึ่งสูงกว่าเพียง 0.1% แต่ทั้งนี้โครงสร้างแบบ VGGNet เป็นวิธี Deep CNN ที่มีความถูกต้องสูงที่สุดจากการทดลองทั้งหมด โดยมีความถูกต้องถึง 91.85% แต่ต้องใช้เวลาในการเรียนรู้นานประมาณ 3 ชั่วโมง 50 นาที (~3.5h) จากการทดลองแสดงให้เห็นว่าโครงสร้างแบบ VGGNet ให้ผลการทดลองที่ดีกับชุดข้อมูลที่มีจำนวนข้อมูลไม่มาก ชุดข้อมูล Folio มีจำนวนรูปภาพใบไม้เพียง 637 รูป ผลการทดลองแสดงดัง Table 5

Table 5 Test accuracy comparison of CNN architectures on Folio dataset.

CNN Architectures	Number of Epochs / Training Time			
	1,000	Time	10,000	Time
LeNet-5	84.25±0.51	~20m	74.02±0.95	~2.5h
AlexNet	85.04±0.48	~30m	73.23±0.81	~1.4h
GoogLeNet	87.40±1.54	~3.5h	82.68±2.41	~7.2h
VGGNet	**91.85±3.31**	~3.5h	**87.49±1.54**	~1.8d

สรุปผลการทดลอง (Conclusions)

งานวิจัยฉบับนี้ นำเสนอวิธีการประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks) ซึ่งประกอบไปด้วยโครงสร้างแบบ LeNet-5, AlexNet, GoogLeNet และ VGGNet เพื่อใช้สำหรับการจำแนกพรรณไม้ ซึ่งได้นำไปทดสอบกับชุดข้อมูลพรรณไม้ทั้งสิ้น 3 ชุด ได้แก่ PNE, Folio และ 102 Flower Dataset

งานวิจัยฉบับนี้ได้เลือกใช้ NVIDIA Deep Learning GPU Training System (DIGITS) เป็นเครื่องมือในการทดสอบการเรียนรู้เชิงลึก ทำให้สามารถประมวลผลได้บนหน่วยประมวลผลด้านกราฟฟิก (Graphics Processing Unit: GPU) และได้เลือกใช้ GPU รุ่น GeForce GTX 1060 ที่มีขนาดหน่วยความจำ 6GB และมีจำนวน NVIDIA CUDA Cores จำนวน 1152 Core

จากผลการทดลองพบว่าโครงสร้างแบบ VGGNet ให้ผลการทดลองสูงที่สุดกับชุดข้อมูล Folio ที่มีข้อมูลขนาดเล็กเพียง 637 ชุดข้อมูล และใช้จำนวนรอบในการเรียนรู้เพียง 1,000 รอบ สำหรับโครงสร้างแบบ GoogLeNet นั้นมีผลการทดลองสูงที่สุดเมื่อทดสอบกับชุดข้อมูล PNE และ 102 Flower ที่มีจำนวนข้อมูลในการเรียนรู้ 3,000 และ 8,189 รูปภาพ ตามลำดับ ซึ่งเป็นข้อมูลจำนวนมาก แต่ต้องใช้จำนวนรอบในการเรียนรู้มากถึง 10,000 รอบ

หากเปรียบเทียบกันระหว่างโครงสร้างแบบ GoogLeNet และ VGGNet ที่มีจำนวนของชั้น (Layer) ที่ใกล้เคียงกันปรากฏว่าโครงสร้างแบบ VGGNet ใช้เวลาในการเรียนรู้นานกว่าโครงสร้างแบบ GoogLeNet แต่ในทางกลับกันเมื่อเปรียบเทียบอัตราความถูกต้องพบว่าโครงสร้างแบบ VGGNet มีอัตราความถูกต้องที่ต่ำกว่าโครงสร้างแบบ GoogLeNet เมื่อทดสอบกับข้อมูลชุด 102 Flower และ PNE ซึ่งข้อมูลทั้งสองชุดนั้นเป็นรูปภาพดอกไม้และพืชสมุนไพร และรูปภาพแต่ละรูปเป็นรูปภาพที่อยู่ในสิ่งแวดล้อมทางธรรมชาติที่มีพื้นหลังที่ซับซ้อน แต่เมื่อทดสอบกับชุดข้อมูล Folio ที่เป็นรูปภาพใบไม้ที่ถ่ายในห้องทดลองโดยกำหนดให้พื้นหลังเป็นสีขาวปรากฏว่าโครงสร้างแบบ VGGNet มีอัตราความถูกต้องสูงที่สุด

จากผลของการวิจัยสรุปได้ว่าโครงข่ายประสาทเทียมแบบคอนโวลูชันเชิงลึก โดยใช้โครงสร้างแบบ GoogLeNet เหมาะสำหรับนำไปใช้เพื่อจำแนกพรรณไม้ที่อยู่ในสิ่งแวดล้อมทางธรรมชาติ เนื่องจากมีอัตราความถูกต้องที่สูง และยังช่วยลดระยะเวลาในการเรียนรู้

วิธีการเรียนรู้เชิงลึกจำเป็นต้องใช้ชุดข้อมูลรูปภาพจำนวนมากในการเรียนรู้ ทั้งนี้เพื่อทำให้มีอัตราการรู้จำเพิ่มสูงขึ้น งานวิจัยครั้งถัดไป จะนำเทคนิคการเพิ่มจำนวนของข้อมูลชุดเรียนรู้ที่เรียกว่า Data Augmentation ทำให้มีข้อมูลสำหรับการเรียนรู้เพิ่มขึ้น และใช้วิธีการถ่ายโอนความรู้ (Transfer Learning) ซึ่งจะช่วยลดระยะเวลาในการเรียนรู้และอาจส่งผลให้มีประสิทธิภาพในการเรียนรู้เพิ่มขึ้น[8,20]

กิตติกรรมประกาศ

รับทุนสนับสนุนการทำวิจัย และการศึกษาจากมหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ และขอบคุณผู้ช่วยศาสตราจารย์พัชราพรณ์ แสงโยจารย์ นักพฤกษศาสตร์ ที่ช่วยคัดกรองและยืนยันความถูกต้องของสายพันธุ์จากรูปภาพพรรณไม้ทุกรูป

เอกสารอ้างอิง

1. Le Cun Y, Matan O, Boser B, Denker JS, Henderson D, Howard RE, et al. Handwritten Zip Code Recognition with Multilayer Networks. In: 10th International Conference on Pattern Recognition. 1990. pp. 35–40.
2. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-Based Learning Applied to Document Recognition. In: Proceedings of the IEEE. 1998. pp. 2278–2324.
3. Ciresan D, Meier U, Schmidhuber J. Multi-column Deep Neural Networks for Image Classification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2012. pp. 3642–9.
4. Sirinukunwattana K, Raza SEA, Tsang YW, Snead DRJ, Cree IA, Rajpoot NM. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Trans Med Imaging. 2016;35(5):1196–206.
5. Yalcin H, Razavi S. Plant Classification using Convolutional Neural Networks. In: International Conference on Agro-Geoinformatics. 2016. pp. 1–5.
6. Li Y, Hong H, Fang T. Hierarchical Segmentation of Remote Sensing Images by Unsupervised Deep Learning Features. In: International Symposium on Computational Intelligence and Design (ISCID). 2017. pp. 448–53.
7. Liu Y, Zhou D, Tang F, Meng Y, Dong W. Flower Classification via Convolutional Neural Network. In: IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications. 2017. pp. 110–6.
8. Pawara P, Okafor E, Surinta O, Schomaker L, Wiering M. Comparing Local Descriptors and Bags of Visual Words to Deep Convolutional Neural Networks for Plant Recognition. In: 6th International Conference on Pattern Recognition Applications and Methods (ICPRAM). 2017. pp. 479–86.
9. Reyes AK, Caicedo JC, Camargo JE. Fine-tuning deep convolutional networks for plant recognition. In: Conference on Conference and Labs of the Evaluation forum (CLEF). 2015. pp. 1–9.
10. Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Comput Intell Neurosci. 2016:1–11.
11. Mohanty SP, Hughes DP, Salathé M. Using Deep Learning for Image-Based Plant Disease Detection. Front Plant Sci. 2016(7):1–10.
12. Sun Y, Liu Y, Wang G, Zhang H. Deep Learning for Plant Identification in Natural Environment. Comput Intell Neurosci. 2017:1–7.
13. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Net-

works. In: Advances In Neural Information Processing Systems. 2012. pp. 1–9.

14. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. pp. 1–9.
15. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In: International Conference on Learning Representations. 2015. pp. 1–10.
16. จินตนา รุ่งเรือง และวสา วงศ์สุขแสวง. รายงานผลการดำเนินงานโครงการอนุรักษ์พันธุกรรมพืชอันเนื่องมาจากพระราชดำริ เรื่องการดูแลสวนสมุนไพรและสวนไม้หอม. สุรินทร์; 2558.
17. Nilsback ME, Zisserman A. Automated Flower Classification over a Large Number of Classes. In: Proceedings - 6th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP). 2008. p. 722–9.
18. Munisami T, Ramsurn M, Kishnah S, Pudaruth S. Plant Leaf Recognition Using Shape Features and Colour Histogram with K-nearest Neighbour Classifiers. Procedia Comput Sci. 2015(58):740–7.
19. ศักดิ์ เพิ่มหรรษา นิตยา เกิดประสพ และกิตติศักดิ์ เกิดประสพ. การประยุกต์เทคนิคแปลงข้อความแบบเวกเตอร์และการกระจายแบบเบย์ เพื่อทำนายการบุกรุกเครือข่ายคอมพิวเตอร์. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 6. รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 6; 2558. pp. 75–84.
20. Pawara P, Okafor E, Schomaker L, Wiering MA. Data Augmentation for Plant Classification Data Augmentation for Plant Classification. In: 18th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS). 2017. pp. 615–26.

ภาคผนวก ข

ชุดข้อมูลพรรณไม้ที่ใช้ในการวิจัย

ดาว์นโหลดได้จาก

www.rmuti.ac.th/user/jakharin.sa/plant

1.กุหลาบหนู (Fairy Rose)

ชื่อวิทยาศาสตร์: *Rosachinensis Jacq.var.minima voss*

ชื่อวงศ์: Rosaceae

ชื่อสามัญ: Fairy Rose, Pygmy Rose

ลักษณะทั่วไป:

กุหลาบหนู เป็นกุหลาบที่มีความสวยงามน่ารักและสดใส เป็นไม้พุ่ม สูง 20 - 50 เซนติเมตร ลำต้นมีหนาม

ใบประกอบออกสลับ ใบย่อย 5 ใบ เป็นรูปรี ขอบหยัก ปลายแหลม โคงมน หูใบติดกับก้านใบ สีเขียวสด

ดอก มีหลายสี เช่น สีแดง สีขาว สีชมพู และดอกเดียว 2 สี ออกเป็นดอกเดี่ยวๆที่ปลายยอด กลีบดอกมีทั้งชั้นเดียวและหลายชั้น มีกลีบดอกห้ากลีบ เกสรตัวผู้และตัวเมียแยกที่อยู่กัน ชนิดสองสี ปลายกลีบดอกเป็นสีแดง โคนกลีบดอกเป็นสีขาว

คุณประโยชน์:

ปลูกประดับสวน ปลูกเป็นไม้กระถาง ตัดดอกปักแจกัน เป็นดอกไม้ทานได้ มีสรรพคุณทางยาอีกด้วย อาทิ บำรุงหัวใจ คลายเครียด มีฤทธิ์เป็นยานอนหลับธรรมชาติ

2.นางแย้ม (Glory Tree)

ชื่อวิทยาศาสตร์: *Clerodendrum Philippinium*

ชื่อวงศ์: VERBENACEAE

ชื่อสามัญ: Glory Bower

ลักษณะทั่วไป:

นางแย้ม เป็นไม้เนื้ออ่อนทรงพุ่ม ตรงและเตี้ย ขนาดของลำต้นนางแย้มสูงประมาณ 1-2 เมตร ใบมีลักษณะคล้ายใบโพธิ์ ออกตรงกันข้ามเป็นคู่ๆ ขอบใบหยักปลายใบแหลม ใบของนางแย้มจะมีกลิ่นเหม็นเมื่อนำมาขยี้ ดอกของนางแย้มมีสีขาวขนาดเล็กคล้ายดอกมะลิซ้อน มักออกเป็นช่อซ้อนกันที่ปลายกิ่งสวยงามมาก และมีกลิ่นหอมเย็นชื่นใจ ในช่อหนึ่งๆ จะมีดอกติดกันจนแน่น แต่ละช่อจะมีความกว้างประมาณ 4-5 นิ้ว ส่วนดอกที่บานเต็มที่แล้วจะมีเส้นผ่า

คุณประโยชน์:

ใบ-ใช้ตำพอกเพื่อแก้โรคผิวหนัง

ราก-ใช้ต้มดื่มแก้อาการปวดข้อ แก้ริดสีดวงทวาร แก้หลอดลมอักเสบ หรือใช้ฝนกับน้ำปูนใสทาเพื่อรักษาโรคเริม และงูสวัด

ต้น-ใช้ต้มดื่มแก้อาการกระดูกสันหลังอักเสบเรื้อรัง แก้ปวดข้อ ช่วยขับปัสสาวะ แก้อาการไตพิการ

3.พวงชมพู (Antigonia)

ชื่อวิทยาศาสตร์: *Antigonon leptopus Hook. & Arn.*

ชื่อวงศ์: POLYGONACEAE

ชื่อสามัญ: Mexican Creeper

ลักษณะทั่วไป:

จัดเป็นไม้เถาเลื้อยที่มีเถาเนื้ออ่อน ขนาดเล็ก แต่ตรงส่วนโคนจะแข็งแรงมาก ลำเถาเป็นสีเขียวอ่อน มีมือสำหรับยึดเกาะพันต้นไม้หรือกิ่งอื่นเพื่อการทรงตัว และสามารถเลื้อยพันสิ่งต่าง ๆ ไปได้ไกลประมาณ 40 ฟุต ขยายพันธุ์ด้วยวิธีการเพาะเมล็ด การตอน และการปักชำกิ่ง พวงชมพูเป็นไม้กลางแจ้งที่ต้องการแสงแดดมาก ต้องการน้ำปานกลาง เจริญเติบโตได้ดีในดินแทบทุกชนิดที่มีความชื้นพอสมควร นิยมนำมาปลูกเป็นไม้ประดับทั่วไปในเขตร้อน

คุณประโยชน์:

รากและเถาใช้เป็นยากล่อมประสาท ช่วยทำให้นอนหลับ ด้วยการใช้เถาประมาณ 1 กำมือ หรือใช้รากประมาณ 1/2 กำมือ นำมาต้มกับน้ำ 4 ถ้วยแก้ว แล้วต้มให้เหลือ 2 ถ้วยแก้ว ใช้รับประทานก่อนนอนครั้งละ 3 ช้อนแกง

ยอดอ่อนและช่อดอกที่ยังไม่บานเต็มที่ อาจนำมาลวกให้สุกเพื่อใช้รับประทานเป็นผักจิ้มหรือชุบแป้งทอดรับประทานก็ได้

4.พุทธชาด (Jasminum Auriculatum)

ชื่อวิทยาศาสตร์: *Jasminum auriculatum Vahl*

ชื่อวงศ์: OLEACEAE

ชื่อสามัญ: Jusmine Vine, Star jasmine, Angel-hair jasmine

ลักษณะทั่วไป:

เป็นพันธุ์ไม้ต่างประเทศ มีถิ่นกำเนิดในประเทศอินเดียและจีน

ต้น ไม้พุ่มรอเลื้อย สูง 1-2 เมตร กิ่งอ่อนสีเขียวสด กิ่งแก่สีน้ำตาล เถารูปทรงกระบอกแคบ ผิวเกลี้ยง กิ่งเปราะหักง่าย

ใบ ใบเดี่ยว เรียงตรงข้าม รูปไข่ปลายมนหรือทู่ อาจมีติ่งแหลมเล็กน้อย แผ่นใบสีเขียวเป็นมัน

ดอก ออกดอกเป็นช่อแบบช่อกระจุกซ้อนประกอบ ตามส่วนยอดของกิ่งและซอกใบตามข้อ ดอกสีขาว กลีบเลี้ยงสีเขียวเข้ม กลีบดอกโคนเชื่อมติดกันเป็นหลอด ปลายแยกเป็น 6-8 แฉก มีกลิ่นหอม ออกดอกตลอดปี

คุณประโยชน์:

ปลูกเป็นไม้ประดับ ดอกมีกลิ่นหอม ดอกนำมาสกัดน้ำมันหอมระเหยได้

5.มะลิลา (Arabian Jasmine)

ชื่อวิทยาศาสตร์: *Jasminum sambac (L.) Aiton*

ชื่อวงศ์: OLEACEAE

ชื่อสามัญ: Arabian Jasmine, Jusmine, Kampopot

ลักษณะทั่วไป:

ใบ ใบเรียงตรงข้าม ใบประกอบแบบขนนก มีใบย่อยใบเดียว รูปไข่ รีหรือรีขอบขนาน กว้าง 3-5 ซม. ยาว 6-10 ซม. โคนใบมนหรือสอบ ปลายใบแหลม ขอบใบเรียบ ผิวใบเรียบเป็นมัน ด้านท้องใบเห็นเส้นใบชัดเจน เส้นใบขนาดใหญ่มี 4-6 คู่ ก้านใบสั้นมากและมีขน

ดอก ดอกช่อออกตามซอกใบและปลายกิ่ง มีทั้งดอกลาและดอกซ้อน โคนกลีบดอกจะเชื่อมกันเป็นหลอด ยาวประมาณ 1.5 ซม. ปลายแยกเป็น 5-8 กลีบ เมื่อดอกย่อยบานมีเส้นผ่านศูนย์กลาง 2-3 ซม. ดอกที่อยู่ตรงกลางจะบานก่อน มีกลิ่นหอมแรง

คุณประโยชน์:

ดอกสดหรือดอกแห้ง ใบ ต้มกินแก้โรคบิด ปวดท้อง ดอกสดตำพอกแก้ปวดศีรษะ ผิวหนังผื่นคัน ดอกใช้ทำพวงมาลัย ทำน้ำหอม และบูชาพระ ดอกแห้งใช้เป็นยาแต่งกลิ่น ดอกมะลาลินี้ใช้เป็นดอกไม้ของวันแม่แห่งชาติ

6.โมกพวง (Wild Water Plum)

ชื่อวิทยาศาสตร์: *Wrightia religiosa Benth.*

ชื่อวงศ์: APOCYNACEAE

ชื่อสามัญ: Moke

ลักษณะทั่วไป:

ไม้พุ่มเตี้ยขนาดเล็ก สูงประมาณ 2 - 3 ม. แตกกิ่งก้านสาขามากไม่ค่อยเป็นระเบียบ ออกดอกตลอดปี

ควบคุมการออกดอกได้ด้วยการควบคุมการให้น้ำและปุ๋ยที่เหมาะสม ออกดอกเป็นพวงพร้อมๆ กันเกือบทั้งต้น ใบไม่ค่อยร่วงและใบบาง ทำให้ย่อยสลายได้ง่ายจึงไม่มีภาระในการเก็บใบทิ้ง ใช้ประโยชน์ได้หลากหลาย เช่น ทำเป็นไม้ดัด ไม้ประดับกระถาง ทำรั้วบ้าน

ข้อแนะนำ:

ต้องการการดูแลรักษาที่ดีจึงจะออกดอกได้ตามต้องการ เป็นไม้กลางแจ้ง ต้องการแสงแดดตลอดวันจึงจะออกดอกได้ดี การทำให้แตกใบใหม่ จะทำให้การออกดอกเร็วขึ้น

คุณประโยชน์:

ยาง ใช้แก้โรคบิดที่มีอาการเลือดออก ใช้แก้พิษงูและแมลงกัดต่อย

ราก รักษาโรคเรื้อน

7.ลีลาวดี (Plumeria)

ชื่อวิทยาศาสตร์: *Plumeria spp.*

ชื่อวงศ์: APOCYNACEAE

ชื่อสามัญ: Plumeria, Frangipani, Temple tree

ลักษณะทั่วไป:

"ลีลาวดี"เป็นชื่อพระราชทานจากสมเด็จพระนางเจ้าสิริกิติ์ พระบรมราชินีนาถ ซึ่งมีความหมายว่า "ดอกไม้ที่มีท่วงท่าสวยงามและอ่อนช้อย" และในปัจจุบันนี้ต้นลีลาวดีได้รับความนิยมและปลูกกันอย่างแพร่หลายไม่ว่าจะในบ้านหรือนอกบ้านก็ตาม

คุณประโยชน์:

ดอกลีลาวดีใช้ผสมกับพลู ทำเป็นยาแก้ไข้และไข้มาลาเรีย (ดอกลีลาวดี, เปลือกต้น) ช่วยรักษาไข้หวัด (ราก) ใช้ปรุงเป็นยาแก้ไอ (เนื้อไม้) ช่วยถ่ายเสมหะและโลหิต (ยางและแก่น) ช่วยขับเหงื่อ แก้ร้อนใน (ราก) ช่วยรักษาโรคหืดหอบ ด้วยการใช้ใบลีลาวดีแห้งนำมาชงกับน้ำร้อนดื่ม (ใบแห้ง) ยางจากต้นลีลาวดีใช้ผสมกับไม้จันทร์และการบูรทำเป็นยาแก้อาการปวดฟัน (ยางจากต้น) มีการนำมาใช้ปรุงเป็นยารักษาโรคลำไส้พิการของม้า (ต้น) ใช้ปรุงเป็นยาถ่าย (เนื้อไม้, ยางจากต้น, เปลือกราก, เปลือกต้น) ฯลฯ

8.เล็บมือนาง (Rangoon Creeper)

ชื่อวิทยาศาสตร์: *Combretum indicum (L.) DeFilipps*

ชื่อวงศ์: COMBRETACEAE

ชื่อสามัญ: Rangoon Creeper, Chinese honey Suckle, Drunen sailor

ลักษณะทั่วไป:

ต้นเล็บมือนาง จัดเป็นไม้เถาเลื้อยเนื้อแข็งขนาดกลาง เลื้อยพาดพันไปกับต้นไม้อื่น แตกกิ่งก้านสาขาเป็นพุ่มหนาทึบ เถาอ่อนเป็นสีเขียว ตามลำต้นและเถาอ่อนมีขนสีเหลือหรือสีน้ำตาลอมเทาปกคลุมอยู่ แต่ต้นแก่ผิวจะเกลี้ยง โดยเถาแก่เปลือกลำต้นเป็นสีน้ำตาลปนแดง เปลือกค่อนข้างเรียบหรือมีหนามเล็กน้อย

คุณประโยชน์:

รากและใบมีรสเมาเบื่อ เป็นยาสุขุม ส่วนเมล็ดมีรสชุ่มเป็นยาร้อน สรรพคุณเป็นยาบำรุงธาตุ (ราก,ใบ, เมล็ด) ช่วยทำให้เจริญอาหาร (ราก,ใบ) ทั้งต้นมีสรรพคุณช่วยแก้ตานขโมยพุงโร (ทั้งต้น) หรือจะใช้รากผสมกับสมุนไพรชนิดอื่น ๆ จะมีสรรพคุณเป็นยาแก้ตานขโมย แก้เด็กเป็นซาง แก้ซางแห้ง แก้ธาตุวิปริต แก้อุจจาระพิการ แก้ตับทรุด และช่วยทำให้เจริญอาหาร (ราก)

9.วาสนา (Cape of Good Hope)

ชื่อวิทยาศาสตร์: *Dracaena fragrans (L.) Ker-Gawl.*

ชื่อวงศ์: AGAVACEAE

ชื่อสามัญ: Cape of Good Hope, Dracaena

ลักษณะทั่วไป:

ลักษณะวิสัย : ไม้พุ่ม เรือนยอด ทรงพุ่ม : ทรงกระบอก ลำต้น : ลำต้นเหนือดิน ตั้งตรงเองได้ เปลือกลำต้น : เรียบ สีน้ำตาลอมเทา ชนิดของใบ : ใบเดี่ยว สีเขียวแกมเหลือง ขนาดใบ กว้าง 7 ซ.ม. ยาว 50 ซ.ม. ลักษณะพิเศษของใบ ขอบใบขนานปลายใบแหลม กลีบดอก : กลีบดอกแยกกัน มี 6 กลีบ สี ขาวออกเหลืองอ่อนๆ ผลสดเป็นผลเมล็ดเดียวแข็ง

คุณประโยชน์:

วาสนานี้จัดเป็นไม้มงคลที่ขึ้นชื่อ อีกทั้งยังมีชื่อต้นที่เป็นมงคล จึงเป็นต้นไม้ยอดนิยมที่ใคร ๆ ต่างนิยมปลูกเพื่อเสริมโชคลาภตามความเชื่อนั่นเอง ใบ แก้ปวดท้อง ราก บรรเทาอาการปวดในการคลอดบุตร

10.ส่าเหล้าปัตตานี (Dwarf Ylang-Ylang Shrub)

ชื่อวิทยาศาสตร์: *Desmos cochinchinensis Lour*

ชื่อวงศ์: ANNONACEAE

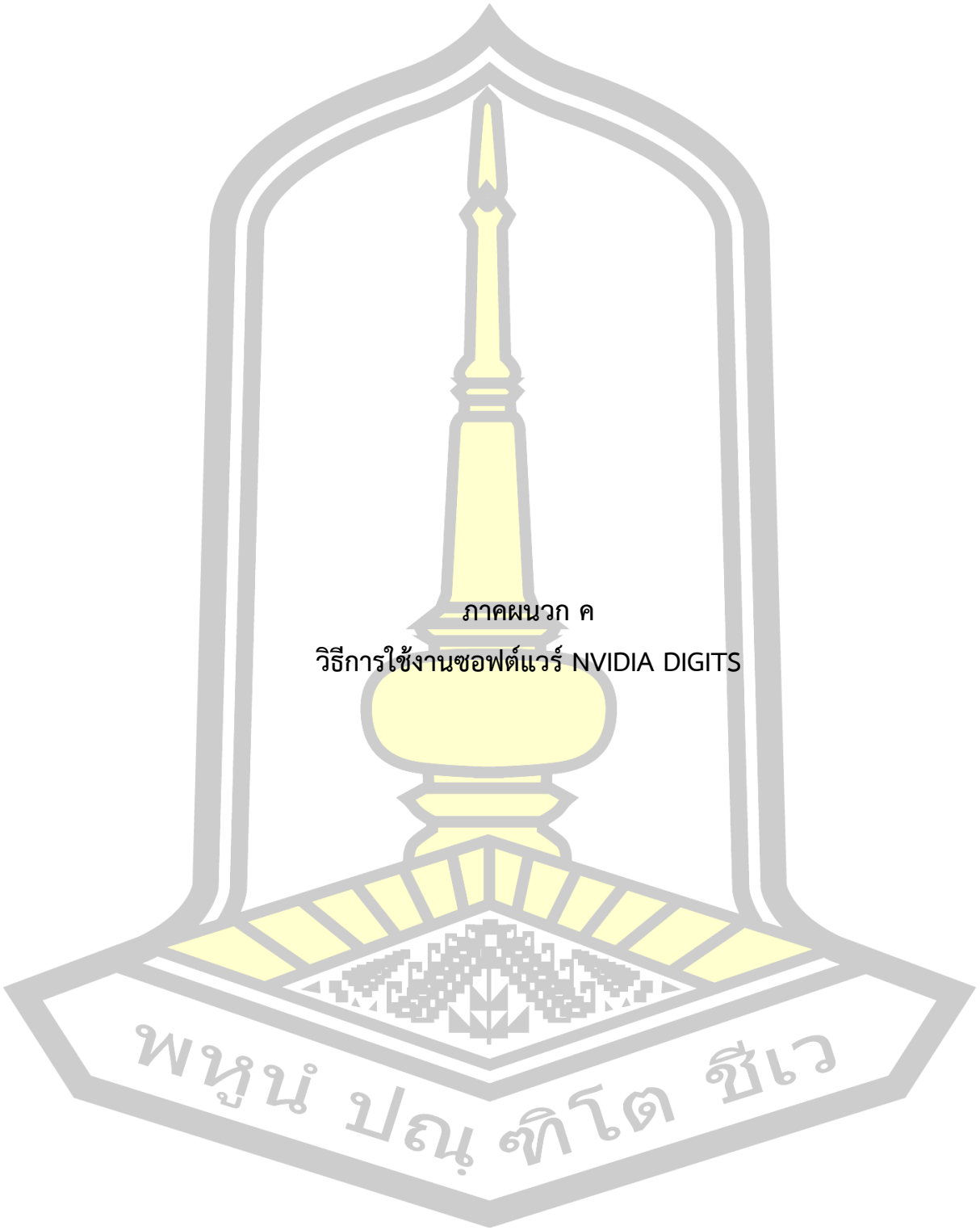
ชื่อสามัญ: นางดำ โยม ส่าเหล้าช้าง ส่าเหล้าปัตตานี

ลักษณะทั่วไป:

ไม้เถาเลื้อยเนื้อแข็ง ใบเดี่ยว เรียงสลับ รูปรีแกมรูปขอบขนาน โคนใบมน ปลายใบแหลม ขอบใบเรียบ ดอกเดี่ยว ออกนอกซอกใบ ห้อยลง มีกลิ่นหอม กลีบเลี้ยง 3 กลีบ สีเขียว รูปสามเหลี่ยม กลีบดอก 6 กลีบ เรียงเป็นสองวง สีเหลือง รูปหอก ปลายกลีบงุ้มเข้า กลีบวงในมีขนาดเล็กและสั้นกว่า ก้านดอกเรียวยาว สีม่วงแดง ผลเป็นผลกลุ่ม มีผลย่อย 6-15 ผล รูปทรงกระบอก สีดำ ผิวผลเรียบเป็นมันและมีรอยคอดในแต่ละช่วงเมล็ด แต่ละผลย่อยมี 2-6 เมล็ด

คุณประโยชน์:

ส่าเหล้าปัตตานี หรือ ดอกบุหรงน่ารัก (ดอกสาละ) ตามประสาชาวบ้านเรียกขานเป็นพืชตระกูลเดียวกับกระดังงา หรือ สายหยุด เป็นดอกไม้ที่มีกลิ่นหอมอ่อนๆ บานอยู่ประมาณ 2 – 3 วัน เกสรของส่าเหล้าปัตตานี มีสรรพคุณทางยา คือ ยาบำรุงหัวใจ บำรุงกำลังโลหิต และสารสกัดในการต้านมาลาเรียด้วย



ภาคผนวก ค

วิธีการใช้งานซอฟต์แวร์ NVIDIA DIGITS

ซอฟต์แวร์ที่ใช้ในการทำวิทยานิพนธ์ คือ NVIDIA DIGITS

NVIDIA เปิดตัวซอฟต์แวร์สำหรับฝึกโครงข่าย Deep Learning ของตัวเองชื่อว่า Deep Learning GPU Training System (DIGITS) เป็นซอฟต์แวร์เพื่อการสร้างโมเดลโครงข่ายประสาทเทียมแบบลึก (deep neural network - DNN)

DIGITS ทำให้นักวิทยาศาสตร์สามารถออกแบบโครงข่ายประสาทเทียมด้วยตัวเองด้วยการคอนฟิกเครือข่ายว่าต้องการรู้จำรูปร่างแบบใด แล้วใส่ชุดข้อมูลเพื่อฝึกโครงข่ายประสาทเพื่อฝึกให้ได้โมเดลโครงข่ายประสาทเทียมไปใช้งาน ตัวซอฟต์แวร์ DIGITS เป็นโอเพนซอร์สสามารถดาวน์โหลดแบบคอมไพล์แล้วได้จาก NVIDIA หรือเอาโค้ดจาก GitHub มาแก้ไขเองก็ได้

ความต้องการของระบบ DIGITS (Prerequisites)

- ระบบปฏิบัติการ Ubuntu 14.04 ขึ้นไป (16.04)
- ติดตั้ง NVIDIA driver
  https://github.com/NVIDIA/DIGITS/blob/master/docs/InstallCuda.md#driver
- ติดตั้งโปรแกรม NVIDIA CUDA

```
# For Ubuntu 16.04
CUDA_REPO_PKG=http://developer.download.nvidia.com/compute/cuda/repos/ubuntu1604/x86_64/cuda-repo-ubuntu1604_8.0.61-1_amd64.deb
ML_REPO_PKG= http: / / developer. download. nvidia. com/ compute/ machine-learning/repos/ubuntu1604/x86_64/nvidia-machine-learning-repo-ubuntu1604_1.0.0-1_amd64.deb

# Install repo packages
wget "$CUDA_REPO_PKG" -O / tmp/ cuda-repo. deb && sudo dpkg -i / tmp/ cuda-repo.deb && rm -f /tmp/cuda-repo.deb
wget "$ML_REPO_PKG" -O /tmp/ml-repo.deb && sudo dpkg -i /tmp/ml-repo.deb && rm -f /tmp/ml-repo.deb

# Download new list of packages
sudo apt-get update
```

- ติดตั้งแพคเกจเสริมที่จำเป็น

```
python-flask python-flaskext.wtf python-gevent python-h5py python-numpy python-pil python-pip python-scipy python-tk
```

- ดาว์นโหลดซอร์ดโค้ดโปรแกรม

```
# example location - can be customized
DIGITS_ROOT=~/digits
git clone https://github.com/NVIDIA/DIGITS.git $DIGITS_ROOT
```

ติดตั้ง Python packages

```
sudo pip install -r $DIGITS_ROOT/requirements.txt
```

เปิดใช้งาน plug-ins

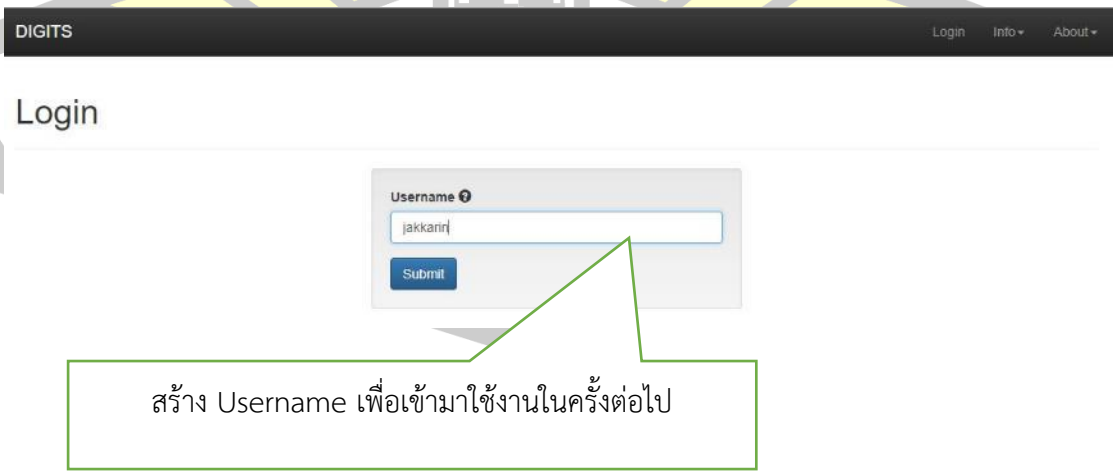
```
sudo pip install -e $DIGITS_ROOT
```

เปิดใช้งาน server

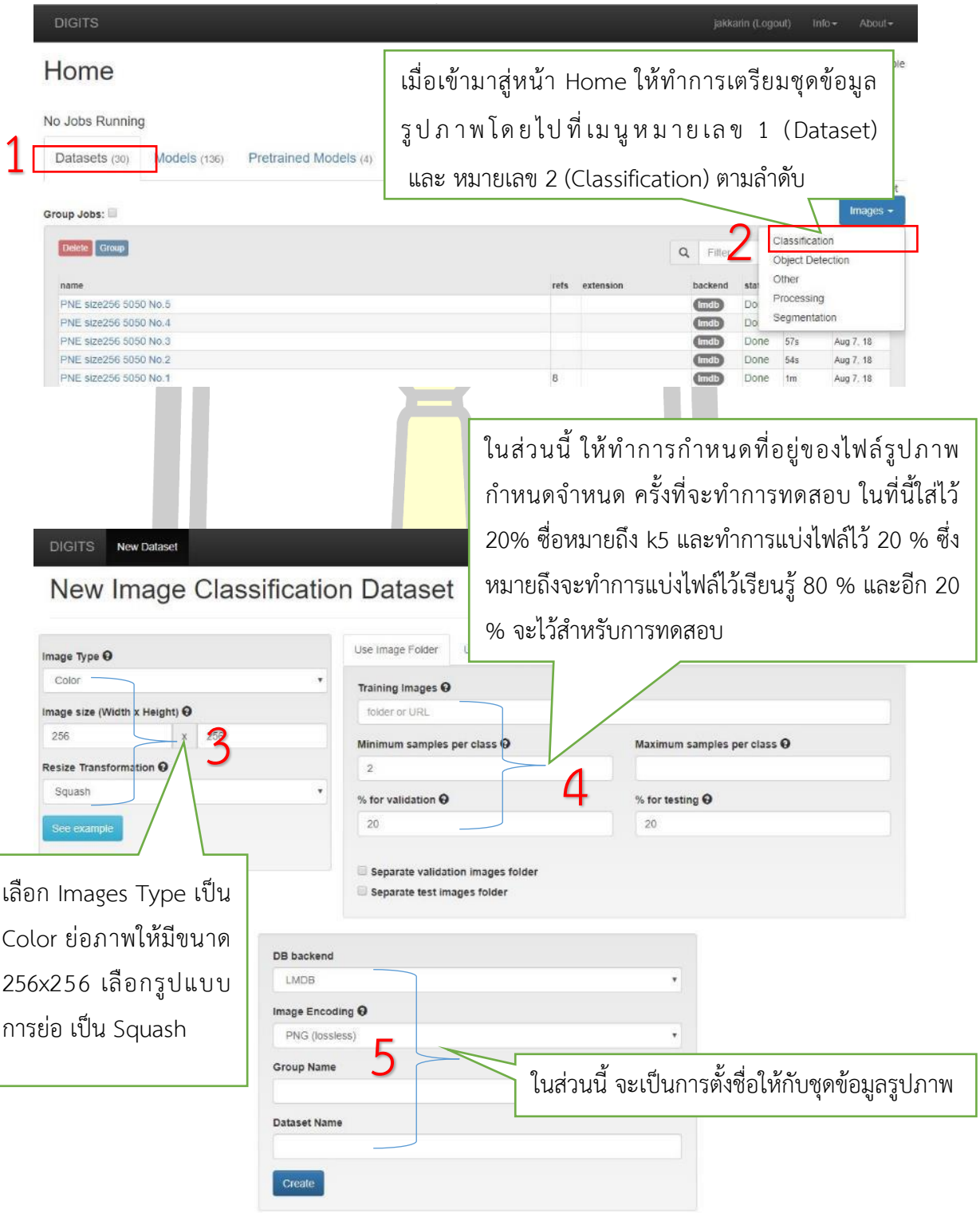
```
./digits-devserver
```

ขั้นตอนการใช้งานซอฟต์แวร์ NVIDIA DIGITS

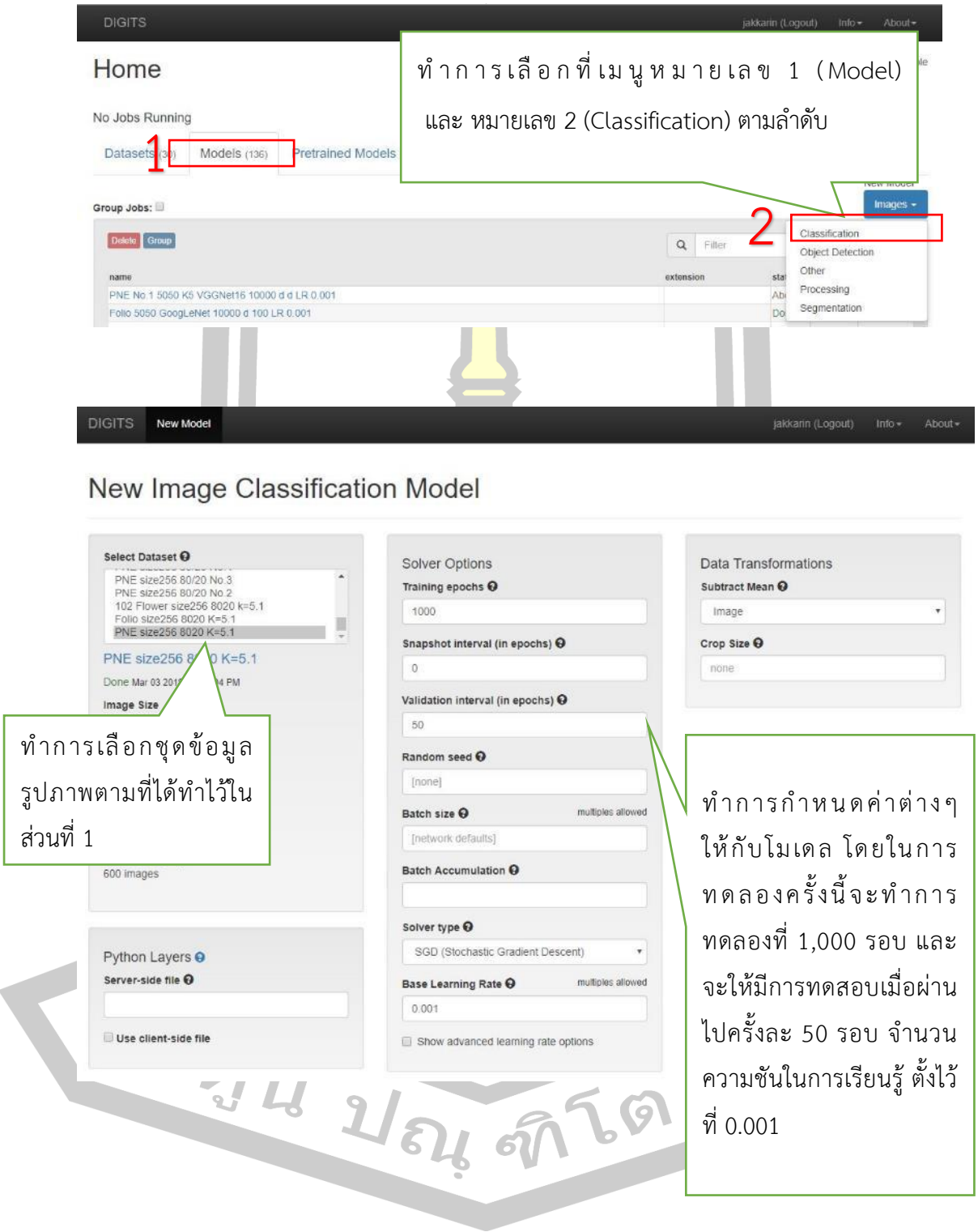
เรียกใช้งานโปรแกรม ผ่านเว็บบราวเซอร์ http://localhost:5000/



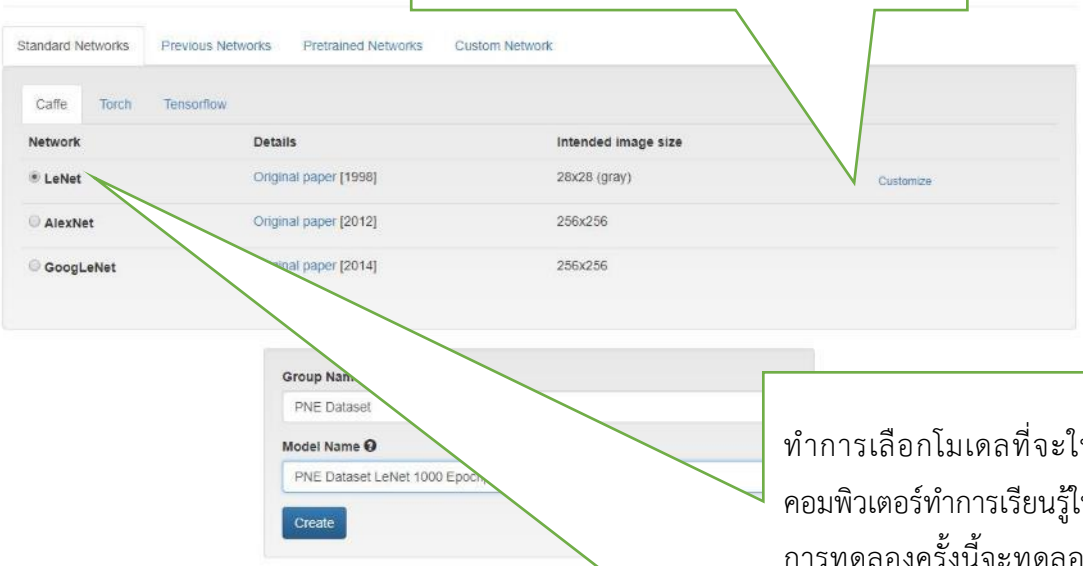
ส่วนที่ 1 จะเป็นส่วนของการเตรียมชุดข้อมูลรูปภาพ Images Dataset



ส่วนที่ 2 จะเป็นส่วนของการตั้งค่าพารามิเตอร์ให้คอมพิวเตอร์ทำการเรียนรู้จำพรรณไม้



ทั้งนี้ผู้ใช้งานสามารถเข้าไปจัดการแก้ไขโครงสร้างของโครงข่ายประสาทเทียมของแต่ละโมเดลได้ โดยคลิกที่ Customize

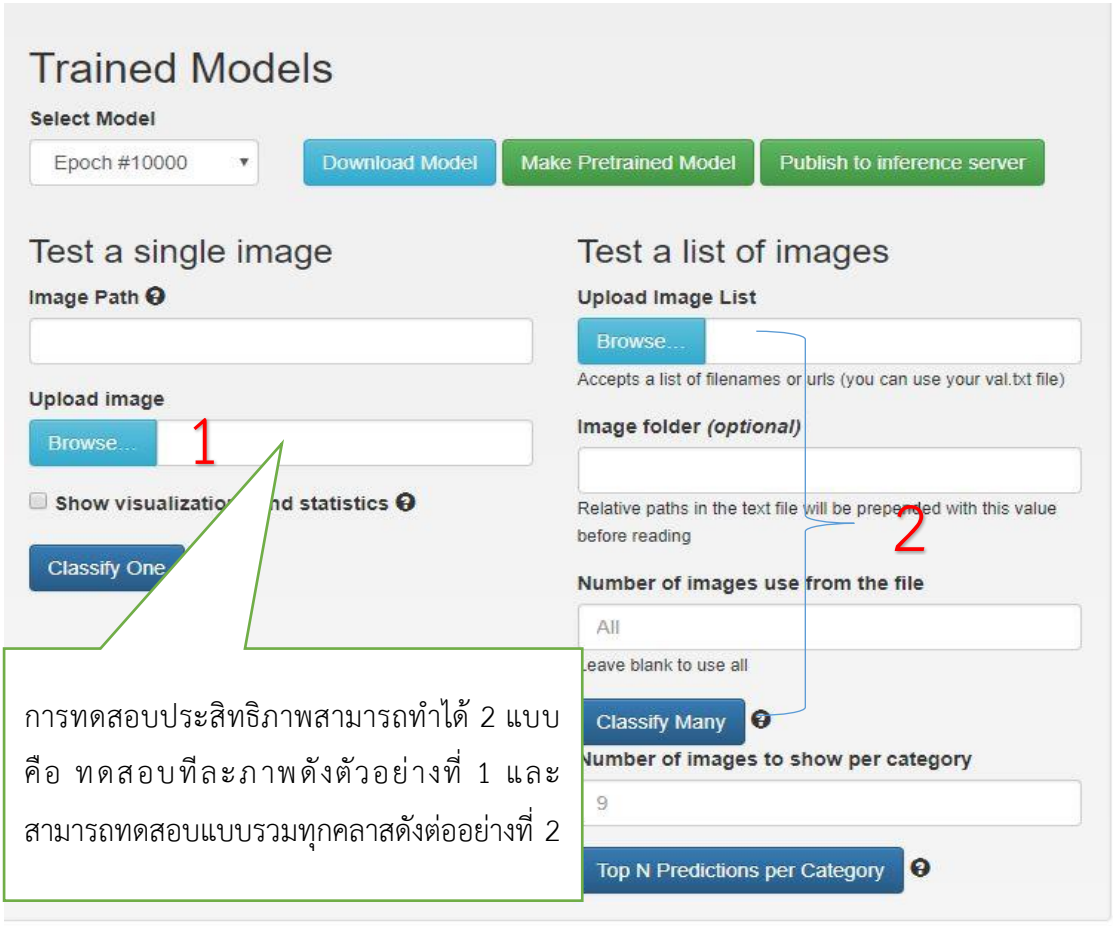


ทำการเลือกโมเดลที่จะให้คอมพิวเตอร์ทำการเรียนรู้ในการทดลองครั้งนี้จะทดลองในโมเดล LeNet

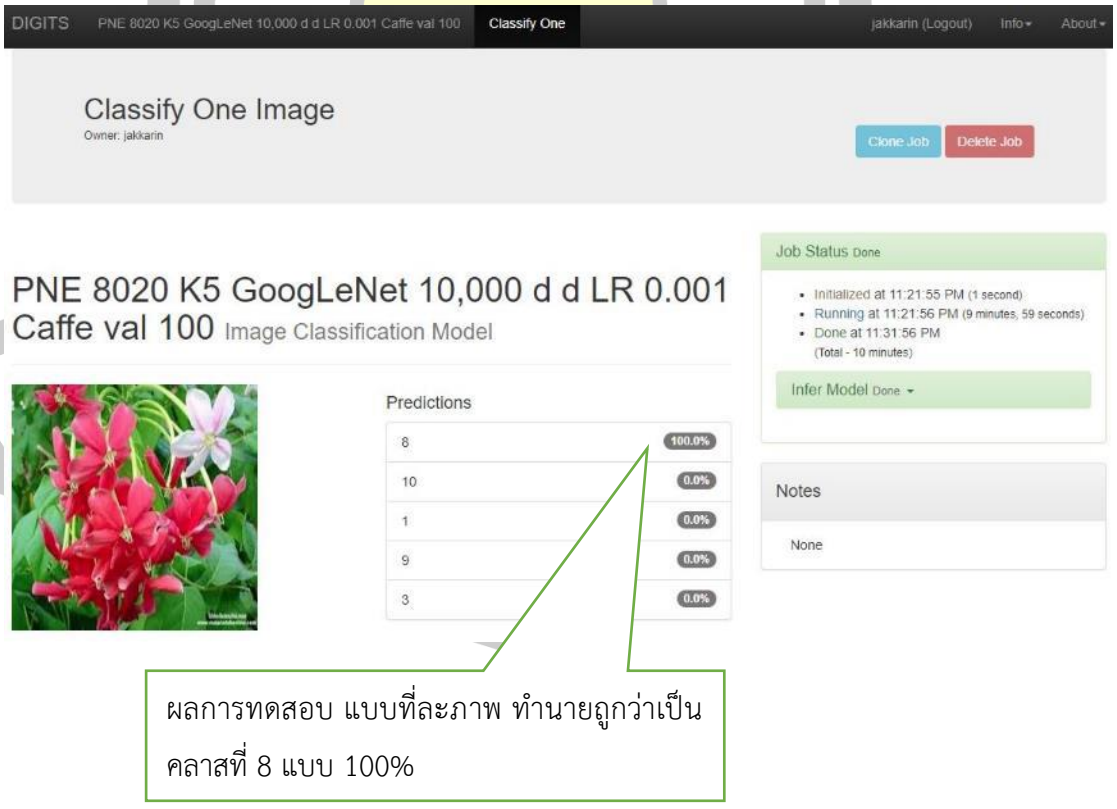
หลังจากนั้นให้ทำการตั้งชื่อโดเดลและกลุ่มของโมเดลระบบทำงานตามลำดับขั้นโดยสามารถสร้างโมเดลเอาไว้ได้ เมื่อทำการเรียนรู้เสร็จก็ทำงานเรียนรู้โมเดลที่กำหนดไว้ในลำดับต่อไปเรื่อยๆ

ส่วนที่ 3 จะเป็นส่วนของการทดสอบประสิทธิภาพของโมเดล

ในการวิจัยครั้งนี้ ยกตัวอย่างโมเดลที่ได้ทำการทดสอบแล้วมีประสิทธิภาพสูงที่สุด คือโมเดล GoogoLeNet

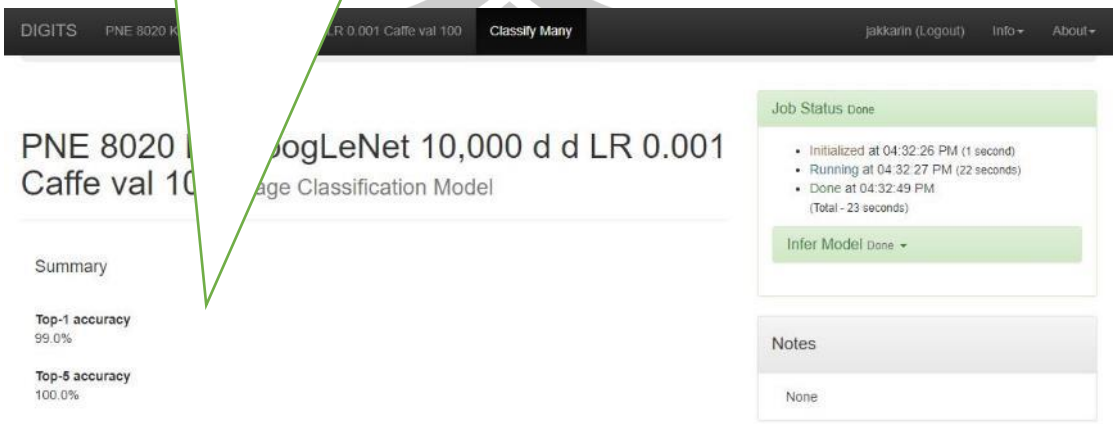


การทดสอบประสิทธิภาพสามารถทำได้ 2 แบบ คือ ทดสอบทีละภาพดังตัวอย่างที่ 1 และสามารถทดสอบแบบรวมทุกคลาสดังต่ออย่างที่ 2



ผลการทดสอบ แบบที่ละภาพ ทำนายถูกว่าเป็นคลาสที่ 8 แบบ 100%

ผลการทดสอบ แบบรวมทุกคลาสมีค่า Top1-Accurancy อยู่ที่ 99.00%



Confusion matrix

	1	10	2	3	4	5	6	7	8	9	Per-class accuracy
1	59	0	0	0	0	0	1	0	0	0	98.33%
10	0	58	0	0	0	0	1	0	1	0	96.67%
2	0	0	60	0	0	0	0	0	0	0	100.0%
3	0	0	0	60	0	0	0	0	0	0	100.0%
4	0	0	0	0	60	0	0	0	0	0	100.0%
5	0	0	0	0	0	60	0	0	0	0	100.0%
6	0	1	0	0	0	0	59	0	0	0	98.33%
7	0	0	0	0	0	0	0	60	0	0	100.0%
8	0	0	0	1	0	0	0	0	59	0	98.33%
9	0	0	0	0	0	0	0	0	1	59	98.33%

พหุน ปณ ฑิโต ชีเว

ประวัติผู้เขียน

ชื่อ	นายจักรินทร์ สนุกแสน
วันเกิด	วันที่ 16 สิงหาคม พ.ศ. 2531
สถานที่เกิด	อำเภอสังขะ จังหวัดสุรินทร์
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 90 หมู่ 1 ตำบลสังขะ อำเภอสังขะ จังหวัดสุรินทร์ รหัสไปรษณีย์ 32150
ตำแหน่งหน้าที่การงาน	นักวิชาการคอมพิวเตอร์
สถานที่ทำงานปัจจุบัน	มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์ 145 หมู่15 ตำบลนอกเมือง อำเภอเมืองสุรินทร์ จังหวัดสุรินทร์ รหัสไปรษณีย์ 32000
ประวัติการศึกษา	พ.ศ. 2546 มัธยมศึกษาตอนต้น โรงเรียนสังขะ จังหวัดสุรินทร์ พ.ศ. 2549 มัธยมศึกษาตอนปลาย โรงเรียนสังขะ จังหวัดสุรินทร์ พ.ศ. 2554 ปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยราชภัฏสุรินทร์ พ.ศ. 2562 ปริญญาวิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม
ทุนวิจัย	ทุนการศึกษาบุคลากรสายสนับสนุน มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์
ผลงานวิจัย	J. Sanuksan and O. Surinta (2019). Deep Convolutional Neural Networks for Plant Recognition in Natural Environment, Journal of Science and Technology Mahasarakham University, 38(2), 113-124. (TCI 1)