

## การเปรียบเทียบวิธีการวิเคราะห์โรคใบมะนาวด้วยการเรียนรู้เชิงลึก

### Comparison of Lime Leaf Disease Analysis using Deep Learning

อุมาภรณ์ สายแสงจันทร์<sup>1</sup> รพีพร ชำชอง<sup>1\*</sup> อรรถพล สุวรรณษา<sup>2</sup>

Umaporn Saisangchan<sup>1</sup>, Rapeeporn Chamchong<sup>1\*</sup> and Atthapol Suwannasa<sup>2</sup>

<sup>1</sup> Polar Laboratory, Faculty of Informatics, Maharakham University, Thailand

<sup>2</sup> ISAN Laboratory, Faculty of Informatics, Maharakham University, Thailand

\* Corresponding Author: rapeeporn.c@msu.ac.th

#### Received:

26 March 2022

#### Revised:

1 May 2022

#### Accepted:

11 May 2022

#### คำสำคัญ:

การเรียนรู้เชิงลึก, คอนโวลูชัน  
นิวรอลเน็ตเวิร์ค, การจำแนก  
ประเภท, โรคพืช

#### Keywords:

Deep Learning, Convolution  
Neural Networks, Classification,  
Plant Decease

**บทคัดย่อ:** ปัจจุบันมะนาวเป็นพืชชนิดหนึ่งที่ได้รับการนิยมปลูกกันจำนวนมาก เนื่องจากมะนาวใช้ในการประกอบอาหาร นอกจากนั้นมะนาวยังมีคุณสมบัติเป็นพืชสมุนไพร ประเทศไทยได้มีการนำมาทำเครื่องดื่มหรือนำมาทำเป็นยาสมุนไพร เพื่อใช้ในการบำรุงสุขภาพ เป็นต้น จากความนิยมดังกล่าว ทำให้การปลูกมะนาวจึงแพร่หลายภายในฟาร์มและเริ่มมีการปลูกตามบ้านมากยิ่งขึ้น ซึ่งปัจจุบันการปลูกมะนาวสามารถควบคุมผลผลิตให้ออกผลผลิตได้นอกฤดูกาล แต่มะนาวยังเป็นพืชที่เป็นโรคต่างๆ ได้ง่าย หากขาดการดูแลที่ดี ดังนั้นในงานวิจัยนี้จึงได้ศึกษาวิธีการวิเคราะห์ความเป็นโรคจากใบมะนาวโดยอาศัยการเรียนรู้เชิงลึก โดยเทคนิคที่ใช้คือการจำแนกใบมะนาวโดยใช้คอนโวลูชันนิวรอลเน็ตเวิร์ค ซึ่งสถาปัตยกรรมที่ศึกษาในที่นี้ได้แก่ LeNet-5 VGG16 ResNet-50 และสถาปัตยกรรมที่นำเสนอบนพื้นฐานของ VGGNet โดยการปรับปรุง Hyperparameter การปรับปรุงสถาปัตยกรรมนี้เป็นการลดขนาดของชั้นการเรียนรู้ และช่วยลดจำนวนพารามิเตอร์ของโครงข่าย ซึ่งจะทำการประมวลผลเร็วขึ้นแต่ประสิทธิภาพยังคงเดิม จากการทดลองเก็บรวบรวมข้อมูลภาพถ่ายใบมะนาวเดี่ยว จำนวน 5,710 ภาพ เป็นภาพสี RGB โดยแบ่งเป็นใบมะนาวที่เป็นโรคและไม่เป็นโรคเท่ากัน และทำการแบ่งชุดข้อมูลเป็นชุดสอนร้อยละ 80 และชุดทดสอบร้อยละ 20 พบว่าประสิทธิภาพการวิเคราะห์ด้วย LeNet-5 มีความถูกต้องร้อยละ 78.90 ซึ่งมีประสิทธิภาพต่ำสุด ในขณะที่สถาปัตยกรรมที่นำเสนอมีความถูกต้องร้อยละ 89.06 ซึ่งมีค่าสูงสุด แต่ไม่แตกต่างกับ ResNet-50 มากนัก

**Abstract:** Currently, lime is a type of plant that has been cultivated in many places. Because limes are used in cooking, and their properties are herb. In Thailand, limes are used for drinking and medical herb to use in health care. From this popularity, the cultivation of limes became widespread and began to be planted more on farms and more at home. Nowadays, lime cultivation can be controlled to produce off-season produce. However, lime is also prone to diseases without proper care. For these reasons, this research is to study the analysis of disease from lime leaves by using deep learning. The Convolutional Neural Networks (CNNs) of deep learning is used to classify the disease. In this study, LeNet-5, VGG16, and ResNet-50 architectures of CNNs are to the proposed architecture based on VGGNet by adjusting hyperparameter. This improvement in the proposed architecture is to reduce the size of the learning layer and can decrease the number of network parameters. This will make processing faster but the performance remains the same. The total number of single lime leaf images is 5,710. The input images are RGB color. The normal and disease lime leaves are separated equally. Training and test sets are 80 and 20 percent, respectively. The evaluation result is found that LeNet-5 has the lowest accuracy at 78.90 percent, while the proposed architecture has the highest accuracy at 89.06 percent but it is not different from ResNet-50.

## 1. บทนำ

ในปัจจุบัน มีผู้สนใจปลูกมะนาวเป็นจำนวนมากในประเทศไทย ซึ่งโดยปกติแล้วทุกครัวเรือนจะใช้มะนาวในเป็นเครื่องปรุงแต่งรสในอาหาร และมีความต้องการสูงในร้านอาหารเพื่อใช้ในการปรุงรส และในปัจจุบันมะนาวเป็นผลผลิตที่ให้ประโยชน์ต่อสุขภาพ โดยการขมมะนาวผสมเครื่องดื่มเพื่อบำรุงสุขภาพ และจัดเป็นผลที่ให้วิตามินซีสูง แต่เดิมนั้นมะนาวเป็นพืชที่ให้ผลผลิตได้เฉพาะฤดูกาล แต่ในปัจจุบันเริ่มมีการตัดแปลงการปลูกมะนาวในวงบ่อซีเมนต์เพื่อให้ได้ผลผลิตตลอดปี ซึ่งมะนาววิธีการดังกล่าวจะช่วยให้ผลผลิตได้ตลอดปี ดังนั้นในปัจจุบันทั้งเกษตรกรและคนทั่วไปได้ให้ความสนใจในศึกษา และหันมาปลูกมะนาวเป็นจำนวนมาก แต่อย่างไรก็ตาม มะนาวเป็นพืชที่พบโรคได้หลากหลาย หากไม่ได้รับการดูแลรักษาที่ดี ซึ่งโรคต่างๆ นั้น เกิดมาจากศัตรูพืช ซึ่งส่งผลให้มะนาวไม่เจริญเติบโต หรือให้ผลผลิตได้ตามต้องการ นอกจากนี้ หากผู้สนใจปลูกมะนาวใน

ระยะแรกเริ่มจะยังไม่ทราบลักษณะของโรคที่เกิดขึ้น และอาจปล่อยให้โรคดังกล่าวลุกลามอาจทำให้เกิดความเสียหายเป็นจำนวนมากได้ และส่วนใหญ่การเกิดโรคระยะแรกเริ่มนั้นเกิดที่ใบของมะนาว แล้วจึงลุกลามไปยังส่วนต่างๆ ของมะนาวต่อไป

จากการศึกษางานวิจัยที่เกี่ยวข้อง Geoffrey Hinton and Ruslan Salakhutdinov (Hinton & Salakhutdinov, 2006) ได้นำเสนอการเรียนรู้เชิงลึก (Deep Learning) ซึ่งแสดงให้เห็นว่าแต่ละชั้นของ Feedforward Neural Network จะทำงานได้อย่างมีประสิทธิภาพโดยก่อนการสอนด้วย Unsupervised Restricted Boltzmann Machine และในขั้นต่อมาโดยการปรับการทำงานด้วย Supervised Back-propagation ซึ่งเรียกวิธีการดังกล่าวว่า Deep Belief Network (DBN) จากการนำเสนอครั้งนี้จึงเป็นจุดเริ่มต้นของงานวิจัยทางด้านการเรียนรู้เชิงลึกเป็นต้นมา Wicht (Wicht & Henneberty, 2015) ได้ประยุกต์ใช้การเรียนรู้เชิงลึกในการรู้จำ Sudoku ที่มี

ทั้งตัวเลขลายมือเขียนและพิมพ์จากภาพประกอบเก็บจาก  
กล้องมือถือ โดยใช้ Convolutional Deep Belief  
Network ในการจำแนก Feature จากภาพ แล้วจึง  
จำแนกโดยใช้ Support Vector Machine โดยข้อมูล  
นำเข้าเป็นภาพทั้งลายมือเขียนจากภาพ Sudoku  
จำนวน 200 ภาพประกอบได้จากกล้องมือถือที่อาจ  
จะมีการเคลื่อนไหวและแสงเงา ผลลัพธ์ที่ได้มีอัตรา  
การรู้จำเป็น 97.7% จะเห็นว่าการประยุกต์ใช้ Deep  
Belief Network สามารถช่วยในการสกัดคุณลักษณะ  
ได้จากภาพทั้งที่เป็นตัวพิมพ์และลายมือเขียน

งานวิจัยของ Jeon and Rhee (Yingjiu Guo,  
2018) ได้ศึกษา Plant Leaf Recognition Using  
a Convolution Neural Network เพื่อใช้ในการ  
จำแนกชนิดของต้นไม้โดยใช้ CNN โดยใช้โครงสร้าง  
ของ GoogLeNet พื้นฐาน (Model 1) และการ  
ประยุกต์ GoogLeNet (Model 2) กับ Flavia dataset  
โดยแบ่งรูปภาพเป็น 8 ชนิด (Lanceolate, Oval,  
Acicular, Linear, Oblong, Reniform (Kidney-shped),  
Cordate (Heart-shped) และ Palmate leaf) และ  
สุ่มเลือกรูปภาพชนิดละ 100 รูป ซึ่งรูปภาพดังกล่าว  
มีมุมในทิศต่างๆ คือ แนวตั้ง แนวนอน มุม 45 องศา  
และ 90 องศา พบว่าประสิทธิภาพความถูกต้อง  
(Accuracy) ของ Model 1 เป็น 99.6% (เวลา  
ประมวลผล 8 ชม. 43 นาที) และ Model 2 เป็น  
99.8% (เวลาประมวลผล 9 ชม. 18 นาที) ในสภาพ  
แวดล้อมเดียวกัน ซึ่ง Model 2 มีประสิทธิภาพดีกว่า  
Model 1 เล็กน้อย แต่เวลาในการประมวลผลเพิ่มขึ้น  
35 นาที อย่างไรก็ตามจะเห็นว่าประสิทธิภาพดังกล่าว  
ดีขึ้นแต่ไม่มีนัยสำคัญ ขณะที่เวลาก็เพิ่มขึ้นในการ  
เรียนรู้ แต่จากการทดลองผู้วิจัยได้ศึกษาเพิ่มเติม  
โดยการสร้างแปลงสีภาพถ้าหากภาพมีความเสียหายเกิด  
ขึ้นในรูปแบบต่างๆ โดยภาพประกอบมีสีผิดเพี้ยนจาก  
ภาพเริ่มต้น (Discolored image) 5%, 10%, 30%,  
50% และ 60% พบว่า Model ที่ 2 จะให้ประสิทธิภาพ  
ดีกว่าเล็กน้อยในทุกๆระดับ นอกจากนี้ ในงานวิจัย  
ดังกล่าวยังได้ทำการทดลองกับกรณีภาพประกอบมีความ

เสียหาย อาจเป็นรูปเกิดขึ้นเนื่องจากแมลงกัดแทะเป็นต้น  
โดยสร้างภาพประกอบมีรูปเกิดขึ้นที่ 5%, 10%, 15%  
และ 30% พบว่า Model 2 ดีกว่า Model 1 และจะ  
เห็นว่าถึงแม้ว่าการสร้างภาพประกอบมีสีผิดเพี้ยน หรือ  
ภาพเสียหาย ก็ยังสามารถรู้จำใบไม้ได้ถูกต้องมากกว่า  
98% และ 94% ตามลำดับ

Pawara และคณะ (Pawara, Okafor,  
Schomaker, & Wiering, 2017) ได้ศึกษาถึง Data  
Augmentation for Plant Classification โดยมี  
การทำ Data Augmentation (DA) กับภาพใน 6  
รูปแบบ ได้แก่ การหมุนภาพในทิศต่างๆ (Rotation)  
การเบลอภาพ (Blur) การสเกลรูปภาพขนาดต่างๆ  
กัน (Scaling) การปรับระดับความเข้มสี (Contrast)  
และการปรับแสง (Illumination) และการบิดภาพ  
ในลักษณะต่างๆ กัน (Projective) กับภาพพันธุ์พืช  
จากข้อมูลภาพ Folio, AgriPlant และ Swedish  
leaf โดยใช้การเรียนรู้เชิงลึกที่อยู่บนพื้นฐานของ CNN  
ได้แก่ สถาปัตยกรรมของ AlexNet, GoogLeNet โดย  
มีขนาดภาพนำเข้าเป็น 256x256 จุดภาพ พบว่าการ  
ทำ Augmentation นี้ช่วยเพิ่มประสิทธิภาพในการ  
รู้จำได้มากทั้งโครงสร้าง AlexNet และ GooLeNet  
เมื่อเทียบกับกรณีไม่ทำ Augmentation และหากมี  
การใช้ภาพประกอบมีการทำ Augmentation หลาก  
หลายวิธีร่วมกัน จะยิ่งเพิ่มประสิทธิภาพในการรู้จำได้  
ดีกว่ากรณีใช้ Augmentation แบบใดแบบหนึ่ง

Pawara และคณะ (Okafor, *et al.*, 2016)  
ได้ศึกษาถึงการจำแนกพันธุ์พืชโดยเปรียบเทียบการ  
ใช้ CNN กับการสกัดคุณลักษณะแบบโลคอล และถุง  
ของภาพ (Bags of Visual Words) ด้วยการเรียนรู้  
โดยใช้ KNN (K-nearest Neighborhoods) ซึ่ง  
พบว่าวิธีการจำแนกพันธุ์พืชด้วย CNN มีประสิทธิภาพ  
ประกอบดีในการใช้ถุงของภาพกับ KNN

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า การ  
เรียนรู้เชิงลึกด้วย CNN ได้รับความนิยมอย่างมาก  
ในการใช้จำแนกประเภทภาพถ่าย โดยสามารถสกัด

คุณลักษณะของภาพจากการเรียนรู้ด้วยโครงข่ายพร้อมทั้งจำแนกประเภท นอกจากนี้ CNN ยังมีหลากหลายสถาปัตยกรรม ดังนั้นในงานวิจัยนี้จึงได้นำเสนอแนวทางในการวิเคราะห์ว่ามะนาวเป็นโรคหรือไม่จากใบมะนาว ซึ่งการศึกษาครั้งนี้ อยู่บนพื้นฐานของการเรียนรู้เชิงลึก โดยใบมะนาวที่นำเข้าวิเคราะห์ต้องเป็นภาพใบมะนาวเดี่ยวเท่านั้น โดยมีวัตถุประสงค์ ดังนี้

1. เพื่อศึกษาเปรียบเทียบโมเดลการเรียนรู้เชิงลึก CNN ที่มีสถาปัตยกรรมแตกต่างกัน ในการจำแนกใบมะนาวที่เป็นโรคและไม่เป็นโรค
2. เพื่อปรับปรุงสถาปัตยกรรม CNN บนพื้นฐานของ VGGNet เพื่อให้การทำงานเร็วขึ้น แต่ยังคงประสิทธิภาพดั้งเดิม
3. ใช้เป็นแนวทางในการพัฒนาการวิเคราะห์โรคในพืชต่างๆ ให้กับเกษตรกรเข้าควบคุมได้ก่อนที่จะมีการลุกลามต่อไป

## 2. ทฤษฎีที่เกี่ยวข้อง

### 2.1 การจำแนกประเภท (Classification) โดยการเรียนรู้เชิงลึก

โดยทั่วไปการจำแนกประเภท (Classification) จะแบ่งขั้นตอนของการจำแนกออกเป็น 3 ขั้นตอนหลักๆ ในการทำงานได้ดังภาพประกอบ 1

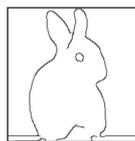
จากขั้นตอนการทำงานของการทำงานจำแนกประเภทดังกล่าวนี้ จะเห็นว่าจะเป็นการยากที่จะต้อง

สกัดคุณลักษณะที่สำคัญของวัตถุออกมา โดยเฉพาะตัวหนังสือ ตั้งแต่ปี คศ. 2006 เป็นต้นมานั้น การเรียนรู้โครงสร้างชั้นลึก (Deep structured learning) หรือเรียกว่า การเรียนรู้ชั้นลึก (Deep Learning) หรือการเรียนรู้แบบลำดับชั้น (Hierarchical Learning) ได้กลายมาเป็นงานวิจัยใหม่ทางด้าน Machine Learning (Bengio, 2009) ซึ่งงานวิจัยดังกล่าวนี้มีพื้นฐานมาจากงานวิจัยทางด้านโครงข่ายประสาทเทียม (Artificial Neural Network, ANN) โดยอาศัยโครงสร้างการทำงานมาจาก Multi-layer Perceptron (MLP) โดยรู้จักกันในชื่อ Back-propagation (BP) แต่พบว่ามี Computation cost ที่ค่อนข้างสูง ดังนั้นในช่วงระหว่าง คศ. 1990-2000 พบว่าเทคนิคที่ถูกนำมาใช้ในการเรียนรู้ที่ประสบความสำเร็จคือ Support Vector Machines (SVMs) แต่วิธีการดังกล่าวนี้ยังคงมีปัญหาเมื่อนำไปประยุกต์ใช้งานกับข้อมูลจริง

อย่างไรก็ตาม ภายหลังจากการนำเสนอการเรียนรู้เชิงลึก โดย Geoffrey Hinton and Ruslan Salakhutdinov (Hinton & Salakhutdinov, 2006) ซึ่งแสดงให้เห็นว่าแต่ละชั้นของ Feedforward Neural Network จะทำงานได้อย่างมีประสิทธิภาพ โดยก่อนการสอนด้วย Unsupervised Restricted Boltzmann Machine และในขั้นต่อมาโดยการปรับการทำงานด้วย Supervised Back-propagation ซึ่งเรียกวิธีการดังกล่าวว่า Deep Belief Network (DBN) ซึ่งจากการนำเสนอครั้งนี้จึงเป็นจุดเริ่มต้นของงานวิจัยทางด้านการเรียนรู้เชิงลึก เป็นต้นมา



Input Image



คุณลักษณะในระดับต่ำ  
(edges, shain code, etc)



กระต่าย

Classification  
ด้วยการเรียนรู้ของเครื่อง

ภาพประกอบ 1 ขั้นตอนการทำงานของการทำงานจำแนกประเภท

สำหรับหลักการของการเรียนรู้เชิงลึก นั้น ได้เริ่มมีการนำมาประยุกต์ใช้กับงานวิจัยทางด้าน Computer Vision และ Image Processing โดยเรียนรู้ภาพจากข้อมูลนำเข้าเป็นพิกเซล (Pixel) ด้วยตัวจำแนกประเภท (Classifier) โดยแต่ละลำดับชั้นของการทำงานจะจำแนกคุณลักษณะออกมาเป็น Output ในแต่ละชั้นและนำมาเป็น Input ของลำดับชั้นถัดไป ดังภาพประกอบ 2

สำหรับขั้นตอนการเรียนรู้ด้วยการเรียนรู้เชิงลึก โดยทั่วไปแบ่งออกเป็น 3 ประเภทคือ

1. โครงข่ายชั้นลึกแบบมีผู้สอน (Supervised Deep Network) ซึ่งทำงานได้ดีเมื่อจำนวนผลเฉลย (Label) มีขนาดใหญ่ โครงสร้างของโมเดล (Model) มีส่วนสำคัญมาก เช่น โครงสร้างแบบ Convolution ได้แก่ Deep Neural Network (DNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN)

2. โครงข่ายชั้นลึกแบบไม่มีผู้สอน (Unsupervised Deep Network) เป็นการเรียนรู้จากโครงสร้างทางสถิติหรือขึ้นกับข้อมูลที่ยังไม่มีผลเฉลย เหมาะสมกับกรณีที่จำนวนผลเฉลยมีไม่มากนัก ได้แก่ Deep

Autoencoder, Deep Belief Network (DBN), Restricted Boltzmann Machines (RBM) เป็นต้น

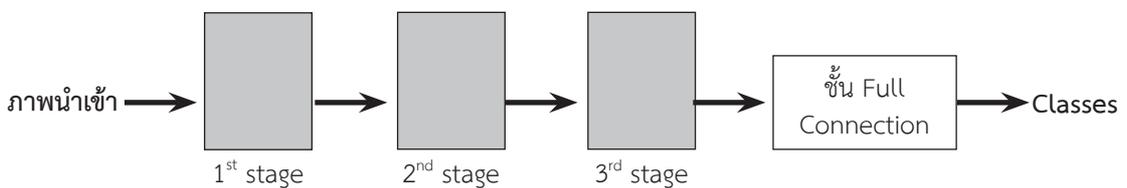
3. โครงข่ายชั้นลึกแบบผสม (Hybrid Deep Network) เป็นการผสมผสานขั้นตอนการทำงานก่อนการสอนในแต่ละชั้นด้วยการใช้ Unsupervised Deep Network ก่อน แล้วจึงใช้การสอนในขั้นต่อมาด้วย Supervised Deep Network ซึ่งปัจจุบันวิธีแบบผสมผสานนี้ได้รับความนิยมขึ้นมาก เช่นการใช้ DBN แล้วตามด้วย DNN หรือ CNN เป็นต้น

## 2.2 โครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Network, CNN)

โครงข่ายประสาทคอนโวลูชัน (Deng & Yu, 2013) มีพื้นฐานมาจากโครงข่ายประสาทแบบหลายชั้น (Multi-layer Perceptron, MLP). หลักการทำงานของ CNN เป็นโครงข่ายแบบหลายชั้นที่มีชั้นโครงสร้างการเชื่อมพิเศษ การเรียงต่อกันของตัวจำแนกคุณลักษณะหลายชั้น การคำนวณให้มีคุณลักษณะแบบโกลบอลมากขึ้นและทำให้คุณลักษณะมีความคงที่มากขึ้นในลำดับชั้นที่สูงขึ้น ซึ่งโครงสร้างของ CNN แสดงดังภาพประกอบ 3



ภาพประกอบ 2 ขั้นตอนการจำแนกประเภทโดยใช้การเรียนรู้แบบลำดับชั้น



ภาพประกอบ 3 โครงสร้างทั่วไปของ CNN

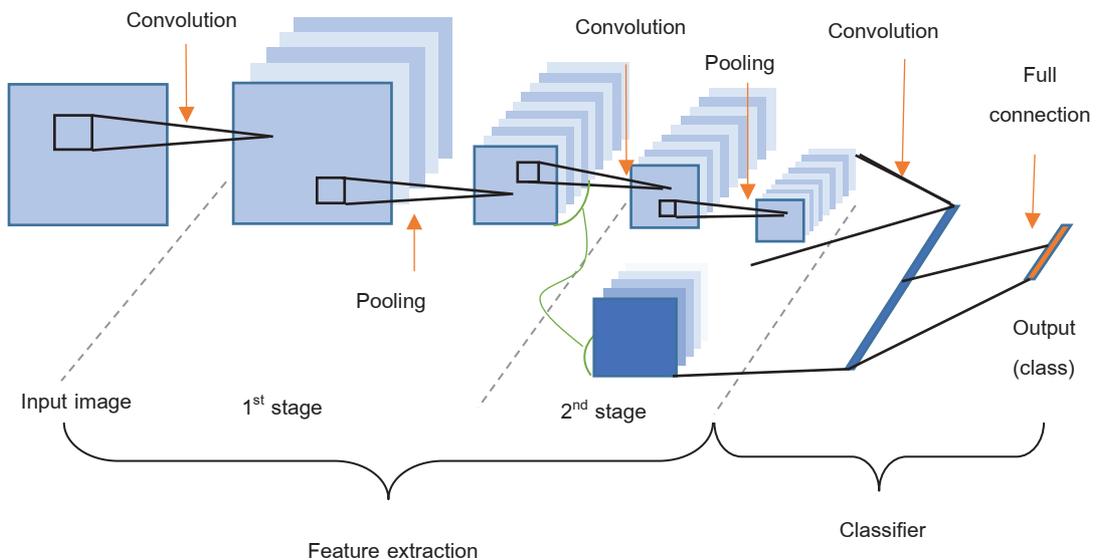
จากภาพประกอบ 3 ในแต่ละ Stage ของ Convolution จะเชื่อมต่อกันเป็นแบบ Stack ที่ต่อเนื่องกันไป โครงสร้างของ CNN เป็นการต่อกันหลายชั้นในการจำแนกคุณลักษณะโดยใช้ Feed-forward มีขั้นตอนดังต่อไปนี้ โดยค่า Input ในชั้น Convolution จะเรียนรู้จากตัวกรองเกรเดียน (Gradient filter) การทำ Pooling และการ Normalization ซึ่ง Feature map เหล่านี้จะถูกเรียนรู้แบบมีผู้สอน (Supervised learning) โดยใช้การจำแนก Back-propagation

ในแต่ละชั้น Feature input มาจากจุดรอบๆ ข้างของชั้นก่อนหน้าซึ่งเรียกว่า Receptive field โดยการจำแนกคุณลักษณะ การหาขอบ (Edge) มุม (Corners) ความโค้ง (Curve) หรือจุดปลาย (End-points) ซึ่งจะถูกรวมกันใช้ชั้นที่สูงขึ้น ซึ่งในการหาคุณลักษณะแบบพื้นฐานที่ไม่ใช้การเรียนรู้นี้จะพบคุณลักษณะบางอย่างอาจถูกกำจัดหายไปได้ ซึ่ง Kernel weight ของตัวกรอง Convolution จะถูกดำเนินการอยู่ในขั้นตอนของขบวนการเรียนรู้

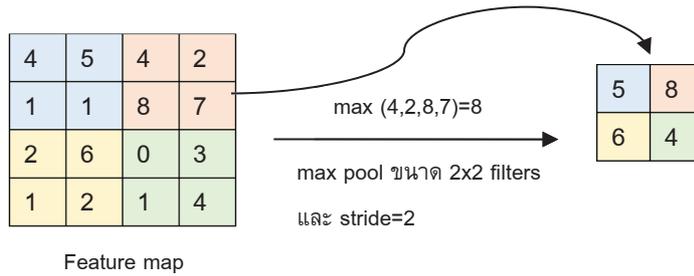
CNN สามารถนำมาใช้ในงานวิจัยได้หลากหลายสาขา ได้แก่ การรู้จำภาพ (Image recognition) การรู้จำเสียง (Speech recognition) การประมวลผลภาษาธรรมชาติ (Natural language processing) ซึ่ง CNN ได้กลายมาเป็นงานที่มีความสำคัญมาก เพราะ Weight ของชั้น Convolutional นั้นได้นำมาใช้ในการจำแนกคุณลักษณะ ขณะเดียวกันในชั้น Full connection ก็จะนำมาใช้ในการจำแนกประเภท (Classification) ในระหว่างการสอน โครงสร้างของ CNN ที่ได้รับการปรับปรุงในปัจจุบันนั้นยังช่วยให้ประหยัดหน่วยความจำ เวลาในการคำนวณ และเพิ่มประสิทธิภาพให้กับโปรแกรม

จากภาพประกอบ 4 แสดงรายละเอียดของสถาปัตยกรรม CNN 2 ระดับ ซึ่งประกอบด้วย

- ภาพนำเข้า (Input image) ในอยู่ในรูปของเมทริกซ์ 4 มิติ ที่ประกอบไปด้วย จำนวนของรูปภาพ (Samples) ความสูงของภาพ (Height of an image) ความกว้างของภาพ (Width of an image) และจำนวนของมิติสี (Channels of an image)



ภาพประกอบ 4 สถาปัตยกรรม CNN แบบ 2 ระดับ (LeCun & Bengio, 1998)



ภาพประกอบ 5 การดำเนินการ Pooling

- ชั้น Convolution เป็นส่วนที่เชื่อมต่อระหว่างข้อมูลนำเข้า (Input data) ไปยังชั้นซ่อน (Hidden layer) โดยการประยุกต์จำนวนของมาส์กตัวกรอง เช่น ขนาด 3x3 or 5x5 ซึ่งตัวกรองมาส์กเป็นการแปลงภาพนำเข้าและสร้าง Feature map.
- ชั้น Pool (Pooling layer) เป็นการหาตัวอย่างย่อย (sub-sampling) เพื่อลดขนาดเมทริกซ์ของ Feature map โดยการประยุกต์การ Stride และหาค่ามากที่สุด หรือหาค่าเฉลี่ย หาในแต่ละ Stride ดังตัวอย่างการหาค่า Max pooling ที่มีขนาดของ window=2x2 และ stride=2 สามารถดำเนินการได้ดัง
- ชั้น Full connection เป็นชั้นตอนสุดท้ายในการทำนายผล Softmax output ที่ได้จาก Feature map จากชั้น Convolution และ ชั้น Pooling

### 2.3-สถาปัตยกรรมโครงข่ายประสาทคอนโวลูชัน (Convolutional Neural Network Architecture)

ในปี 1998 LeNet-5 ได้ถูกนำเสนอโดย LeCun และคณะ (LeCun, Bottou, Bengio, & Haffner, 1998) ซึ่งเป็น CNN พื้นฐานที่เป็นที่รู้จัก ซึ่ง LeNet-5 ประกอบไปด้วยคอนโวลูชัน 7 ชั้น เพื่อใช้ในการจำแนกตัวเลข ซึ่งได้ประยุกต์ใช้กับเซ็คชันการที่มีภาพนำเข้าเป็นภาพระดับเทาที่มีขนาดนำเข้า 32x32 พิกเซล ภาพหลังจากที่ CNN ได้มีการนำเสนอโครงการ ImageNet ก็ได้ก่อตั้งขึ้นเพื่อจัดการแข่งขัน

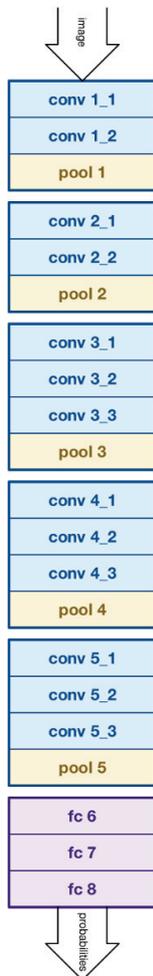
การรู้จำภาพจากฐานข้อมูลขนาดใหญ่ โดยโครงการนี้ได้มีการจัดแข่งขันทุกๆ ปีตั้งแต่ปี 2012 เป็นต้นมา ในการแข่งขัน CNN ปีแรก (2012) นั้น AlexNet (Krizhevsky, Sutskever, & Hinton, ImageNet classification with deep convolutional neural networks, May 2017) เป็นผู้ชนะ ซึ่งสถาปัตยกรรม AlexNet จะมีความคล้ายคลึงกับ LeNet-5 แต่มีจำนวนชั้นมากกว่า และจำนวน Filter ในแต่ละชั้นมากกว่า โดยมีทั้งหมด 8 ชั้น โดยการเรียนรู้ทั้งหมด 6 วันโดยใช้ Nvidia Geforce GTX 580 GPU จำนวน 2 ตัว สถาปัตยกรรมนี้นำเสนอ โดย SuperVision group ( (Krizhevsky, Sutskever, & Hinton, 2012 AlexNet, 2012))

ในปี 2013 ZFNet (Zeiler & Fergus, 2014) เป็นผู้ชนะการแข่งขันสถาปัตยกรรม CNN โดยมีจำนวน 8 ชั้น และปรับ hyper-parameters ของ AlexNet ต่อมาผู้ชนะการแข่งขันในปี 2014 ได้แก่สถาปัตยกรรม GoogLeNet (Inception V1) จาก Google team ประสิทธิภาพความถูกต้องของการรู้จำนี้ใกล้เคียงกับประสิทธิภาพการจำแนกของมนุษย์ สถาปัตยกรรมมีทั้งหมด 22 ชั้น ซึ่งประกอบด้วย Convolution ขนาดเล็กเพื่อลดจำนวนของพารามิเตอร์ ซึ่งสามารถลดจำนวนพารามิเตอร์จาก AlexNet ซึ่งมีขนาด 60 ล้านพารามิเตอร์ลงเหลือเพียง 4 ล้านพารามิเตอร์ GoogLeNet (Szegedy, et al., 2015) มีพื้นฐานมาจาก LeNet-5 (LeCun, Bottou, Bengio, & Haffner, 1998) โดยนำเสนอ

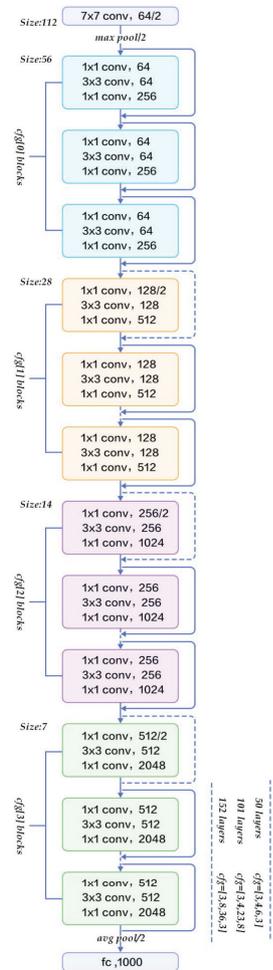
โมเดลใหม่ซึ่งเรียกว่า Inception ซึ่งโมเดลนี้ใช้ Batch normalization การทำภาพ distortion และค่า RMSprop ซึ่งโครงข่ายนี้จะเป็นโมเดลขนาดเล็กที่ดำเนินการซ้ำๆ ภายในโครงข่าย ในขณะที่เดียวกันในปี 2014 ยังมีสถาปัตยกรรม CNN อีกประเภทชื่อว่า VGGNet (Simonyan & Zisserman, 2014) ซึ่งอยู่ใน 5 อันดับแรกของการแข่งขัน และได้รับความนิยมในการนำมาประยุกต์ใช้อย่างแพร่หลาย โดยประกอบด้วย Convolution 16 ชั้น และมีโครงสร้างเป็น Uniform ซึ่งวิธีนี้ไม่ต่างจาก AlexNet มากนัก โดยมีมาส์กขนาด 3x3 ด้วยตัวกรองจำนวนมาก VGGNet ถูกเรียนรู้อยู่บน GPU จำนวน 4 ตัวเป็นเวลา 2-3

สัปดาห์ ซึ่งเหมาะสำหรับการใช้ในการจำแนกคุณลักษณะจากภาพ ซึ่งโมเดลนี้มีจำนวน 138 ล้านพารามิเตอร์ สถาปัตยกรรม VGGNet แสดงดังภาพประกอบ 6

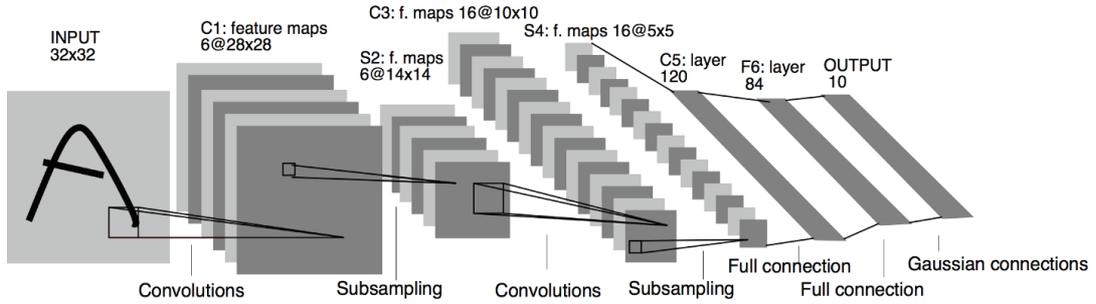
สำหรับผู้ชนะในปี 2015 ของ ImageNet ได้แก่ ResNet จาก Microsoft (He, Zhang, Ren, & Sun, 2016) โดยได้นำเสนอสถาปัตยกรรมใหม่โดย “skip connections” และทำ Batch normalization เพิ่มขึ้น ซึ่งจำนวนชั้นทั้งหมด 152 ชั้นกับการเชื่อมแบบ residual connection ซึ่งมีความซับซ้อนกว่า VGGNet ทำงานบน 8 GPU เป็นเวลา 2-3 สัปดาห์ ซึ่งสถาปัตยกรรม ResNet แสดงดังภาพประกอบ 7



ภาพประกอบ 6 VGGNet (Das, 2017)



ภาพประกอบ 7 ResNet (Das, 2017)



ภาพประกอบ 8 สถาปัตยกรรม LeNet-5 (LeCun, Bottou, Bengio, & Haffner, 1998)

สำหรับในงานวิจัยนี้ นำเสนอการเปรียบเทียบวิธีการเรียนรู้เชิงลึกด้วย CNN จากสถาปัตยกรรม 3 ประเภท ได้แก่ สถาปัตยกรรม LeNet-5 VGGNet และ ResNet ซึ่งมีรายละเอียดดังนี้

### 2.3.1 สถาปัตยกรรม LeNet-5

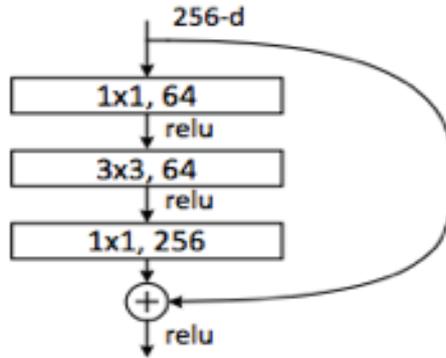
LeNet-5 (LeCun & Bengio, 1998) เป็นสถาปัตยกรรมพื้นฐานของ CNN ซึ่งมีรายละเอียดของสถาปัตยกรรมดังภาพประกอบ 8 โดยมีชั้น Convolution ทั้งหมด 5 ชั้น (C1-C5) ชั้น Full connection 1 ชั้น และ ชั้น Output อีก 1 ชั้น รวมเป็น 7 ชั้น

### 2.3.2 สถาปัตยกรรม VGG16

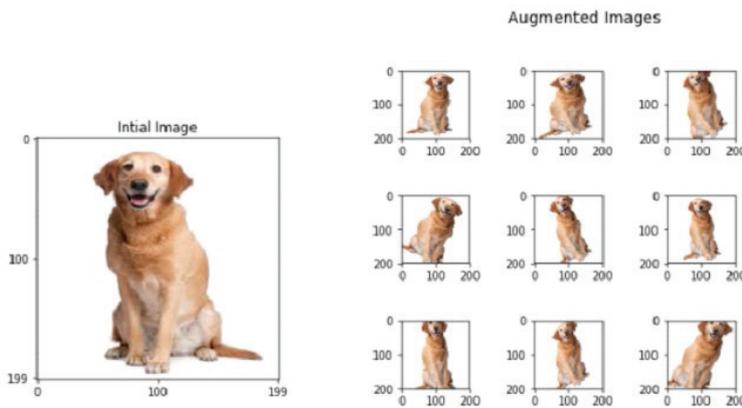
สำหรับสถาปัตยกรรม VGGNet ในที่นี้ใช้ VGG16 (Simonyan & Zisserman, 2014) ซึ่งเป็นสถาปัตยกรรมของ VGGNet ที่มีผลการแข่งขันดีที่สุด โดยมี Convolution ทั้งหมด 16 ชั้น ซึ่งโครงสร้างของ VGGNet จะเป็นลักษณะแบบ uniform โดยมีมาส์กฟิลเตอร์ขนาด maxpooling 2x2 และ stride=2 ฟังก์ชัน Activation ในแต่ละชั้นของ Convolution คือ 'ReLU' โดยที่โมเดลนี้มีจำนวนพารามิเตอร์ 138 ล้านพารามิเตอร์ สถาปัตยกรรมของ VGGNet ดังภาพประกอบ 9 ในการศึกษาครั้งนี้การฟังก์ชัน Activation ในชั้นตอนสุดท้ายเป็น SGD

Input 32x32
3x3 conv, 64
3x3 conv, 64
2x2 maxpool/2
3x3 conv, 128
3x3 conv, 128
2x2 maxpool/2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
2x2 maxpool/2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
2x2 maxpool/2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
2x2 maxpool/2
FC, 4096
FC, 4096
dropout, 0.2
FC, classes
Output

ภาพประกอบ 9 สถาปัตยกรรม VGG16



ภาพประกอบ 10 โครงสร้าง Residual block ของสถาปัตยกรรม ResNet-50 (He, Zhang, Ren, & Sun, 2016)



ภาพประกอบ 11 ภาพแสดงการทำ Augmentation

ที่มา: <https://towardsdatascience.com/image-augmentation-14a0aafd0498>

### 2.3.3 สถาปัตยกรรม ResNet-50

สำหรับสถาปัตยกรรม ResNet อยู่หลากหลายสถาปัตยกรรมซึ่งในที่นี้เลือกใช้ ResNet-50 (He, Zhang, Ren, & Sun, 2016) โดยสถาปัตยกรรม ResNet-50 ได้ประยุกต์มาจาก VGG-19 ซึ่งโครงสร้างของสถาปัตยกรรม ResNet-50 ที่เลือกใช้จะมีโครงสร้างซึ่งแสดงไว้ดังภาพประกอบ 7 และมีโครงสร้าง Residual block จำนวน 3 layer ดังภาพประกอบ 10

### 2.4 การทำ Augmentation

การทำ Augmentation (Pawara, Okafor, Schomaker, & Wiering, 2017) ให้กับชุดข้อมูลเพื่อนำมาเรียนรู้ทำเพื่อเพิ่มความหลากหลายของชุดข้อมูลให้มีความหลากหลายเพื่อที่ Model จะได้เรียนรู้ได้ประสิทธิภาพได้ดียิ่งขึ้น โดยการทำ Augment ได้แก่ การหมุนภาพ (Rotation) ในมุมต่างๆ การเลื่อนภาพ (Shift) การบิดภาพ (Shearing) การซูม (Zooming) การปรับแสงเงา (Shading) เป็นต้น เพื่อให้ได้ภาพในลักษณะต่างๆ กัน สำหรับตัวอย่างการทำ Augment สามารถแสดงได้ดังภาพประกอบ 11

ตารางที่ 1 Confusion Matrix

ลักษณะของเหตุการณ์ (Actual Condition)	ผลจากจำแนก (Predicted Result)		ผลรวมการจำแนก
	Negative	Positive	
Negative	TN	FP	TN+FP
Positive	FN	TP	FN+TP
ผลรวมการจำแนก	TN+FN	FP+TP	TN+FN+FP+TP

## 2.5 การวัดประสิทธิภาพการจำแนก

การวัดประสิทธิภาพของการจำแนกกลุ่ม (Classification) หรือการจำแนกนั้น มีหลายวิธีที่ใช้ในการวัดประสิทธิภาพ เช่น อัตราการตรวจจับถูกต้อง (True Positive Rate, Recall) อัตราการจำแนกผิดพลาด (False Positive Rate) เป็นต้น โดยทั่วไปผลลัพธ์ที่ได้จากการจำแนกจะจัดอยู่ในรูปแบบของตาราง Confusion Matrix (Fawcett, 2006) ซึ่งเป็นการวัดผลตามลักษณะจริงที่ได้ถูกจำแนกที่ถูกต้องไว้แล้วและสรุปผลการจำแนกที่ได้จากการทำนายด้วยเทคนิคหรืออัลกอริทึมต่างๆ ที่นำมาใช้งาน ซึ่ง Confusion Matrix แสดงได้ดังตารางที่ 1

โดยที่

TN คือผลจากการจำแนกถูกต้องเมื่อข้อมูลเป็น Negative (True Negative)

FP คือผลจากการจำแนกผิดพลาดเมื่อข้อมูลเป็น Negative (False Positive)

FN คือผลจากการจำแนกผิดพลาดเมื่อข้อมูลเป็น Positive (False Negative)

TP คือผลจากการจำแนกถูกต้องเมื่อข้อมูลเป็น Positive (True Positive) ซึ่งค่าจากตาราง Confusion Matrix สามารถนำมาคำนวณการวัดประสิทธิภาพมาตรฐานต่างๆ ในการจำแนกการทำนาย ได้ดังนี้

1. ค่าความถูกต้องของการจำแนก (Accuracy) คือสัดส่วนของผลการทำนายที่ถูกต้องจากข้อมูลภาพทั้งหมดที่ใช้ในการทดสอบ คำนวณได้ดังสมการ

$$Accuracy = \frac{TP + TN}{TN + FN + FP + TP} \quad (1)$$

2. อัตราการจำแนกถูกต้องเมื่อข้อมูลเป็นจริง (True Positive Rate : TPR หรือ Recall) คือสัดส่วนของผลการจำแนกกลุ่มได้ถูกต้อง ซึ่งคำนวณได้จากสมการ

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

ซึ่งค่าดังกล่าวนี้เป็นการวัดประสิทธิภาพเพื่อบ่งบอกว่าระบบสามารถวิเคราะห์ภาพได้ถูกต้องมากน้อยแค่ไหน

3. ค่าความแม่นยำ (Precision) สามารถคำนวณได้ดังนี้

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

### 3. วิธีดำเนินการวิจัย

ในงานวิจัยนี้เป็นการประยุกต์ใช้การเรียนรู้เชิงลึก (Deep Learning) ในการจำแนกประเภทโรคมะนาวจากใบมะนาวเดี่ยวซึ่งมีพื้นฐานอยู่บนโครงข่ายคอนโวลูชัน โดยเปรียบเทียบประสิทธิภาพกับการจำแนกประเภท ซึ่งจะเปรียบเทียบ 4 สถาปัตยกรรมดังนี้ LeNet-5, VGG16, ResNet-50 และสถาปัตยกรรมที่นำเสนอที่อยู่บนพื้นฐานของ VGGNet ดังนี้

#### 3.1 สถาปัตยกรรมที่นำเสนอบนพื้นฐานของ VGGNet

วิธีการที่นำเสนอในงานวิจัยนี้อยู่บนพื้นฐานของ VGGNet โดยปรับปรุง Hyperparameter ของสถาปัตยกรรมดังกล่าว ซึ่งได้แก่ ขนาดของ Filter (Filtering size) ของแต่ละชั้น เช่นใน VGG16 นั้นขนาดของ Filter ของ Convolution แต่ละชั้นเป็น 64, 128, 256, 512 และ แต่ในที่นี่เป็น 32, 64, 128 และ 256 นอกจากนี้ลดจำนวนชั้น Convolution ในแต่ละ Block เหลือเพียง 1 ชั้น 512 ส่วน Full Connection มี Hidden Node เพียง 1024 โดย VGG16 มี Hidden Node เป็น 4096 ซึ่งจะสามารถลดขนาดของจำนวนพารามิเตอร์ลงได้

Input 32x32
3x3 conv, 32
2x2 maxpool/2
3x3 conv, 64
2x2 maxpool/2
3x3 conv, 128
2x2 maxpool/2
3x3 conv, 256
2x2 maxpool/2
FC, 1024
FC, 1024
dropout, 0.2
FC, classes
Output

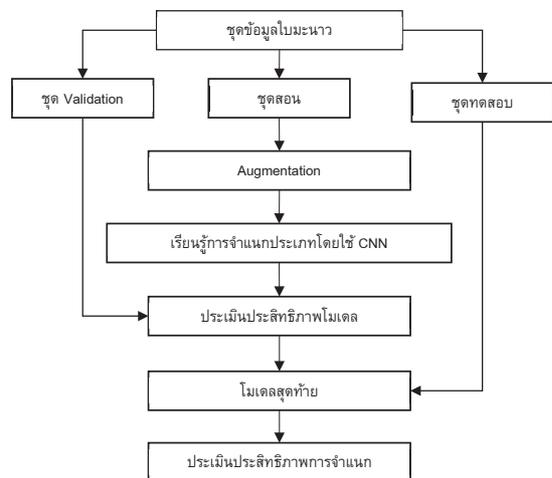
ภาพประกอบ 12 สถาปัตยกรรมที่นำเสนอ

### 3.2 ภาพรวมการประเมินประสิทธิภาพของโมเดล

สำหรับภาพรวมการประเมินประสิทธิภาพของโมเดลการเรียนรู้การจำแนก แสดงดังภาพประกอบ 13 โดยชุดข้อมูลโใบมะนาวจะถูกแบ่งเป็นชุดสอน (Training set) ร้อยละ 80 และชุดทดสอบ (Test set) ร้อยละ 20 ซึ่งโมเดลสุดท้ายของการจำแนกจะถูกประเมินประสิทธิภาพโดยใช้ชุดทดสอบเพื่อจำแนกประเภทโใบมะนาวที่เป็นโรค หรือไม่โรค

### 4. ผลการทดลอง

สำหรับการทดลองนี้ได้เก็บรวบรวมข้อมูลภาพถ่ายโใบไม้ที่เก็บได้จากต้นมะนาวที่เป็นโรค และไม่เป็นโรค ซึ่งโรคที่เกิดขึ้นได้แก่โรคหนอนขนใบ โรคราสีน้ำตาล โรคราดำ โรคใบหงิกงอ โรคจากแมลงกัดแทะ และโรคใบแก้ว ในงานวิจัยนี้เป็นการวิเคราะห์ว่าโใบมะนาวเป็นโรคหรือไม่เป็นโรค โดยไม่มีการจำแนกชนิดของโรค ดังนั้นจำนวนคลาสที่ใช้ในการจำแนกมีเพียง 2 ชนิดคือ เป็นโรค และไม่เป็นโรค โดยการจำแนกโใบมะนาวว่าเป็นโรคหรือไม่เป็นโรค ได้ดำเนินการจำแนกโดยผู้เชี่ยวชาญทางด้านการเกษตร ซึ่งจำนวนภาพถ่าย

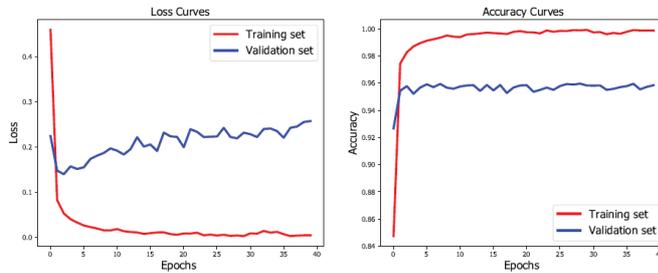


ภาพประกอบ 13 ภาพรวมขั้นตอนการเรียนรู้เพื่อการจำแนกโรคจากโใบมะนาว

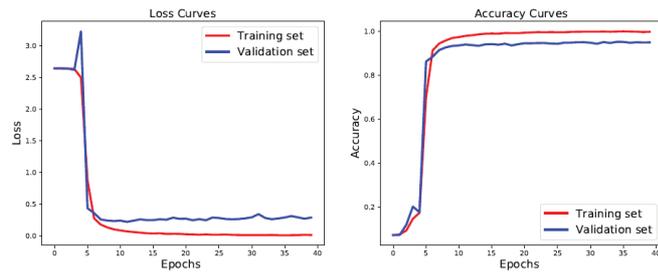
ที่เก็บรวบรวมได้มีจำนวน 5,710 ภาพ เป็นมะนาวไม่  
เป็นโรค 2,855 ภาพ และมะนาวเป็นโรค 2,855 ภาพ  
และแบ่งข้อมูลเป็นชุดสอน และชุดทดสอบ (80:20)  
ขนาดของภาพนำเข้ามีขนาด 224x224 พิกเซล  
โดยการนำเข้าภาพจะทำการตัดภาพถ่ายให้เหลือเฉพาะ  
ส่วนของภาพไปไม้เดียวในรูปแบบภาพสี RGB โดย  
การทดลองเปรียบเทียบผลการประเมินประสิทธิภาพ  
ด้วยสถาปัตยกรรมต่างๆ ของ CNN ดังนี้ LeNet-5,

VGG16, ResNet-50 และวิธีการที่นำเสนอบน  
พื้นฐานของ VGGNet

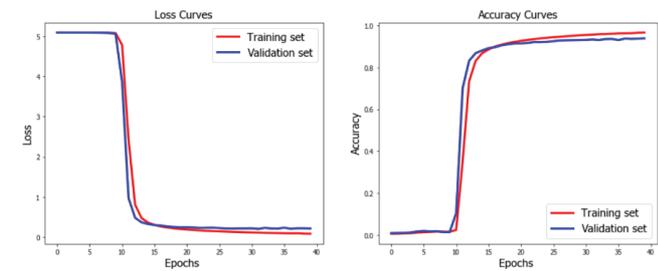
จากการเรียนรู้เพื่อสร้างโมเดลสำหรับการ  
วิเคราะห์โรคจากภาพถ่ายใบมะนาว ด้วยสถาปัตยกรรม  
แต่ละแบบ โดยจำนวนรอบของการเรียนรู้กำหนดไว้  
ที่ 40 รอบ ซึ่งสามารถประสิทธิภาพของการเรียนรู้  
ด้วยค่า Loss และค่า Accuracy ได้ตั้งภาพประกอบ 14  
ถึงภาพประกอบ 17



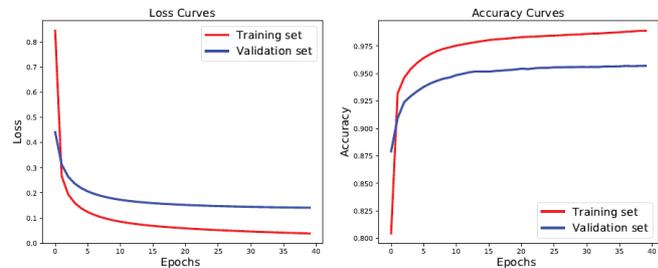
ภาพประกอบ 14 แสดง Loss (ซ้าย) และ Accuracy (ขวา) ของสถาปัตยกรรม LeNet-5



ภาพประกอบ 15 แสดง Loss (ซ้าย) และ Accuracy (ขวา) ของสถาปัตยกรรม VGG16



ภาพประกอบ 16 แสดง Loss (ซ้าย) และ Accuracy (ขวา) ของสถาปัตยกรรม ResNet-50



ภาพประกอบ 17 ภาพแสดง Loss (ซ้าย) และ Accuracy (ขวา) ของสถาปัตยกรรมที่นำเสนอ

**ตารางที่ 2** ผลการประเมินประสิทธิภาพกับชุดทดสอบ

สถาปัตยกรรม CNN	เปอร์เซ็นต์		
	Accuracy	Precision	Recall
LeNet-5	78.90	79.23	79.11
VGG16	86.30	86.64	86.31
ResNet-50	88.26	88.42	88.29
วิธีที่นำเสนอ	89.06	89.32	89.08

จากภาพประกอบ 14 ถึง ภาพประกอบ 17 เป็นการเปรียบเทียบค่า Loss และ ค่า Accuracy ของ ข้อมูลชุดสอน (Training set) และชุดประเมินประสิทธิภาพ โมเดล (Validation set) ซึ่งพบว่าสถาปัตยกรรม LeNet-5 นั้นมีความแตกต่างของค่า Loss และ Accuracy ของ ข้อมูลชุดสอนและชุดประเมินมาก และประสิทธิภาพ ของการเรียนรู้ยังมีกว่าแกว่งตัวมากกว่าสถาปัตยกรรม อื่นๆ ส่วนสถาปัตยกรรมอื่นๆ นั้นค่า Loss และ Accuracy ของข้อมูลชุดสอนและชุดประเมินไม่แตกต่างกันมากนักและเริ่มคงที่ตั้งแต่รอบที่ 20 ส่วนสถาปัตยกรรม ที่นำเสนอ เริ่มคงที่ตั้งแต่รอบที่ 25 เป็นต้นไป

จากตารางที่ 2 ผลการประเมินประสิทธิภาพ การวิเคราะห์โรคจากใบมะนาว พบว่าวิธีที่นำเสนอ โดย ทำการปรับสถาปัตยกรรมจาก VGGNet มาประยุกต์ใช้ จะให้ประสิทธิภาพในการจำแนกโรคจากภาพถ่ายใบ มะนาวด้วยค่าความถูกต้อง (Accuracy) ค่า Precision และค่า Recall สูงสุด แต่ไม่แตกต่างกับวิธี ResNet-50 มากนัก สำหรับวิธีที่นำเสนอนี้สามารถลดเวลาในการ ทำงานลงได้เมื่อเทียบกับ VGG16 เพราะมีจำนวน ชั้นน้อยกว่านั่นเอง ส่วน LeNet-5 ให้ประสิทธิภาพ แยกที่ต่ำในทุกๆ ด้าน แต่จะใช้เวลาน้อยที่สุดในการ ประมวลผลเนื่องจากมีจำนวนชั้นน้อยที่สุด และ พบว่า ResNet-50 ก็ให้ผลลัพธ์ที่มีประสิทธิภาพ สูงกว่า VGG16 และ LeNet-5 ตามลำดับ

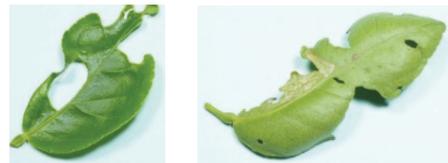
ตัวอย่างภาพถ่ายใบมะนาวแสดงดังนี้



ก) โรคราสีน้ำตาลหรือราดำ



ข) โรคหนอนขอนใบ



ค) โรคจากแมลงกัดแทะ



ง) โรคใบเหลืองและใบแก้ว



จ) ใบปกติ

ภาพประกอบ 18 ตัวอย่างภาพถ่ายใบมะนาว

## 5. สรุปผลและข้อเสนอแนะ

จากการทดลองพบว่าวิธีการวิเคราะห์โรคในมะนาวที่เป็นโรคและไม่เป็นโรคจากภาพถ่ายใบเดียว โดยเปรียบเทียบผลการทดสอบจากโมเดลที่ได้จากการเรียนรู้ด้วยวิธีการพื้นฐานทั้ง 3 วิธีได้แก่ LeNet-5 VGG16 และ ResNet-50 นั้นพบว่า ResNet-50 มีประสิทธิภาพโดยรวมที่ดีที่สุด ส่วน LeNet-5 ให้ประสิทธิภาพต่ำสุด แต่เนื่องจาก LeNet-5 มีจำนวนชั้นน้อยสุด จึงทำให้การเรียนรู้เป็นไปอย่างรวดเร็ว ซึ่งแตกต่างจาก VGG16 และ ResNet-50 ซึ่งมีจำนวนชั้นมาก จากการนำเสนอวิธีการโดยการปรับปรุงสถาปัตยกรรม VGGNet นำมาใช้ทดสอบเปรียบเทียบประสิทธิภาพกับวิธีทั้ง 3 ข้างต้น ให้ประสิทธิภาพประกอบดีกว่า แต่ไม่มากนัก เมื่อเทียบกับ ResNet-50

เนื่องจากการทดลองเป็นการเลือกใช้ Hold-out Method ในการแบ่งข้อมูลเพื่อการเรียนรู้และทดสอบ พบปัญหาจำนวนข้อมูลที่ค่อนข้างน้อย เนื่องจากการเกิดโรคไม่ได้มีจำนวนมากนัก และมีการเกิดโรคเป็นช่วงฤดูกาล จึงต้องทำ Augmentation เข้าช่วยในการเรียนรู้ให้มีประสิทธิภาพมากขึ้น แต่อย่างไรก็ตาม ข้อมูลที่ใช้ทดสอบยังมีข้อจำกัด จำนวนที่น้อยเกินไป ที่อาจไม่มีความหลากหลาย ทำให้ประสิทธิภาพการเรียนรู้ยังไม่ดีเพียงพอทั้งๆ ที่ Class ที่ใช้ในการจำแนกมีเพียง 2 ประเภทเท่านั้น

สำหรับการวิจัยต่อไป เพื่อปรับปรุงประสิทธิภาพการวิเคราะห์ให้ดีขึ้น ควรดำเนินการเก็บรวบรวมข้อมูลให้มากขึ้น นอกจากนี้ควรใช้วิธีการประเมินประสิทธิภาพโดยใช้วิธี Cross-validation ถ้าหากจำนวนข้อมูลมีจำนวนน้อยแต่ต้องใช้เครื่องประมวลผลที่มีประสิทธิภาพสูงขึ้น เนื่องจากภาพนำเข้ามีขนาดใหญ่ซึ่งใช้เวลาในการเรียนรู้นานซึ่งอาจจะเป็นข้อจำกัดในการทดลอง และปรับปรุงการจำแนกประเภทโดยสามารถระบุโรคแต่ละชนิดของใบมะนาวต่อไป

## 6. กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม ที่ให้ทุนสนับสนุนการดำเนินการวิจัยครั้งนี้ให้สำเร็จลุล่วงไปได้ด้วยดี

## 7. เอกสารอ้างอิง

- Bengio, Y. (2009). Learning Deep Architectures for AI. *Found. Trends Mach. Learn.*, 2, 1-127.
- Das, S. (2017, Nov. 16). *CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more*. Retrieved Nov. 15, 2018, from <https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5>
- Deng, L., and Yu, D. (2013). Deep learning: Method and application. In *Foundations and Trends in Signal Processing* (Vol. 7, pp. 197-387).
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 770-778).
- Hinton, G.E., and Salakhutdinov, R.R. (2006). Reducing the Dimensionality of Data with Neural Network. *Science*, 313, 504-507.

- Krizhevsky, A., Sutskever, I., and Hinton, E.G. (2012). 2012 AlexNet. *Adv. Neural Inf. Process. Syst*, 1-9.
- Krizhevsky, A., Sutskever, I., and Hinton, G.E. (May 2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 84-90. doi:10.1145/3065386
- LeCun, Y., and Bengio, Y. (1998). Convolutional Networks for Images, Speech, and Time Series. In M.A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks* (pp. 255-258). MIT Press.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86, pp. 2278-2324.
- Okafor, E., Pawara, P., Karaaba, F., Surinta, O., Codreanu, V., Schomaker, L., and Wiering, M. (2016). Comparative study between deep learning and bag of visual words for wild-animal recognition. *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, (pp. 1-8). Athens.
- Pawara, P., Okafor, E., Schomaker, L., and Wiering, M. (2017). Data augmentation for plant classification. *International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS 2017)*, (pp. 615-626).
- Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv: 1409.1556*. Retrieved from <http://arxiv.org/abs/1409.1556>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... Rabinovich, A. (2015). Going Deeper with Convolutions. *Computer Vision and Pattern Recognition (CVPR)*. Retrieved from <http://arxiv.org/abs/1409.4842>
- Wicht, B., and Hennebert, J. (2015). Mixed handwritten and printed digit recognition in Sudoku with Convolutional Deep Belief Network. *Proceedings of the 2015 13th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 861-865). IEEE Computer Society. doi:10.1109/ICDAR.2015.7333884
- Yingjiu Guo, D.W. (2018). Recognition of a plant leaf based on convolutional neural networks. *Proc. SPIE 10806, Tenth International Conference on Digital Image Processing (ICDIP 2018)*. Shanghai, China.
- Zeiler, M. D., and Fergus, R. (2014). Visualizing and understanding convolutional networks. *Computer Vision – ECCV 2014. Lecture Notes in Computer Science. 8689*. Springer, Cham.