

การพัฒนาระบบจำแนกยาสามัญและยาต้นแบบ โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน

ภัทรภณ เล่าหกิจวิฑูร

TNII

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

บัณฑิตศึกษา สาขาเทคโนโลยีวิศวกรรม

สถาบันเทคโนโลยีไทย-ญี่ปุ่น

ปีการศึกษา 2566

DEVELOPMENT OF GENERIC DRUGS AND ORIGINAL DRUGS CLASSIFICATION SYSTEM
USING CONVOLUTIONAL NEURAL NETWORK

Patpon Laohakijviton

TNII

A Thesis Submitted in Partial Fulfillment of Requirements
for The Degree of Master of Engineering Program in Engineering Technology

Graduate Studies

Thai-Nichi Institute of Technology

Academic Year 2023

หัวข้อวิทยานิพนธ์

การพัฒนาระบบจำแนกยาสามัญและยาต้นแบบ โดยใช้

โครงข่ายประสาทเทียมแบบคอนโวลูชัน

โดย

ภัทรภณ เล่าหกิจจิฑูร

สาขาวิชา

เทคโนโลยีวิศวกรรม

อาจารย์ที่ปรึกษาวิทยานิพนธ์

ดร.ศิวพงษ์ นิลวงษ์

บัณฑิตศึกษา สถาบันเทคโนโลยีไทย-ญี่ปุ่น อนุมัติให้บัณฑิตวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

.....รองอธิการบดีฝ่ายวิชาการ

(รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์)

วันที่.....เดือน.....พ.ศ.....

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ

(ดร.ณัฐกิตติ์ จิตรเอื้อตระกูล)

.....กรรมการ

(รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์)

.....กรรมการ

(ดร.ชัชชัย วรรณบูรณ์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์

(ดร.ศิวพงษ์ นิลวงษ์)

ภัทรภณ เลหาภิจิวิฑูร : การพัฒนาระบบจำแนกยาสามัญและยาต้นแบบ โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน. อาจารย์ที่ปรึกษา : ดร.ศิวพงษ์ นิลวงษ์, 74 หน้า.

ในโครงสร้างสาธารณสุขของประเทศไทย ปัญหาการจ่ายยาโดยไม่ระบุชื่อยาหรือการให้ข้อมูลที่จำเป็นไม่เพียงพอบนซองยาจากระบบบริการสุขภาพ เป็นหนึ่งในสาเหตุที่สำคัญของการเกิดเหตุการณ์ที่ไม่พึงประสงค์จากการใช้ยา (Adverse Drug Events) จนส่งผลกระทบต่ออาจทำให้การรักษาโรคไม่มีประสิทธิผล เกิดอาการถดถอยหรือมีอันตรายถึงชีวิต หนึ่งในกลวิธีที่ผู้ประกอบการมักใช้ประโยชน์จากปัญหาเหล่านี้ คือ การเลือกจ่ายยาสามัญแทนยาต้นแบบที่รู้ลักษณะภายนอกที่คล้ายคลึงกันไม่สามารถแยกแยะได้โดยง่าย แต่มีต้นทุนที่ราคาถูก มีคุณภาพในการรักษาที่ด้อยกว่า ซึ่งอาจเกิดจากวัตถุดิบที่มาจากต่างแหล่งแม้จะมีตัวยาสำคัญที่เป็นชนิดเดียวกัน หรือกรรมวิธีการผลิตที่ยาสามัญบางตัวไม่มีแม้ขั้นตอนการตรวจสอบประสิทธิภาพการรักษาหรือผลข้างเคียงที่เกิดจากการใช้ยา ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์ในการสร้างเครื่องมือต้นแบบที่ใช้จำแนกระหว่างยาต้นแบบและยาสามัญโดยใช้โครงข่ายประสาทเทียมแบบ Convolutional Neural Network (CNN) ที่กำลังเป็นที่นิยมอันเนื่องมาจากความสามารถในการจดจำ และสกัดคุณลักษณะสำคัญ (Feature Extraction) จากวัตถุในรูปภาพที่สามารถนำมาประยุกต์ใช้กับงานด้านต่างๆ โดยเฉพาะในสาขาคอมพิวเตอร์วิทัศน์ (Computer Vision) มาสร้างเป็นโมเดลการเรียนรู้เชิงลึกเพื่อใช้ในการจำแนกยาสามัญและยาต้นแบบพร้อมทำการทดสอบเปรียบประสิทธิภาพในการใช้งานกับ โมเดล CNN ที่สร้างจากเทคนิค Transfer Learning จากโมเดลที่มีอยู่เดิม 3 ชนิดได้แก่ VGG16 InceptionV3 และ Resnet50v2 เพื่อหาโมเดล CNN ที่มีประสิทธิภาพการจำแนกสูงสุด โดยใช้ข้อมูลรูปภาพที่ประกอบไปด้วยยาต้นแบบและยาสามัญจำนวน 26 ชนิด ผลลัพธ์ที่ได้จากการทดลองโมเดล CNN ที่สร้างโดยเทคนิค Transfer learning จากโมเดล ResNet50-v2 สามารถทำคะแนนการวัดประสิทธิภาพการจำแนกได้สูงที่สุดกว่า 85% รองลงมาคือ โมเดลที่สร้างจาก Inceptionv3, โมเดล CNN ที่พัฒนาขึ้นโดยไม่ใช้ Transfer Learning, และโมเดลที่สร้างจาก VGG16 ตามลำดับ

บัณฑิตศึกษา

สาขาวิชา เทคโนโลยีวิศวกรรม

ปีการศึกษา 2566

ลายมือชื่อนักศึกษา

ลายมือชื่ออาจารย์ที่ปรึกษา

PATPON LAOHAKJVITOON : DEVELOPMENT OF GENERIC DRUGS AND ORIGINAL DRUGS CLASSIFICATION SYSTEM USING CONVOLUTION NEURAL NETWORK.
 ADVISOR : DR. SIVAPONG NILWONG, 74 PP.

In Thailand's public health structure, there are problems in dispensing medicine without specifying the name of the drug or insufficient information on required medicines on medicine packets from the health service system. It is one of the important causes of Adverse Drug Events (ADEs) that may cause the treatment to be ineffective, worsening symptoms or be life-threatening. One tactic that entrepreneurs often use to take advantage of these problems is to prescribe generic drugs over original drugs whose appearance is not easily distinguishable. It has a low cost but has inferior treatment quality as well. This may be caused by raw materials that come from different sources, even though they contain the same type of important drug, or production methods that some generic drugs do not even have a step to check for treatment efficacy or side effects caused by the drug. Therefore, this research aims to create a prototype tool for discriminating between original drugs and generic drugs using Convolutional Neural Network that are becoming popular due to their ability to recognize and feature extraction that can be used that can be applied to works, especially in the field of computer vision, to create deep learning models for use in classifying generic drugs and original drugs, along with testing to compare efficiency in use with another three models are created using VGG16, InceptionV3 and Resnet50v2 architectures with transfer learning technique to find the CNN model with the highest classification performance using image dataset containing 26 original and generic drugs. As a result, ResNet50-v2 model built with Transfer Learning technique achieved the highest Classification performance measurement score of over 85%, followed by Inceptionv3 model, the CNN model developed without transfer learning, and VGG16 model respectively.

Graduate Studies

Student's Signature.....

Field of Engineering of Technology

Advisor's Signature.....

Academic Year 2023

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี อันเนื่องมาจากความกรุณาของ ดร.ศิวพงษ์ นิลวงษ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่กรุณาสละเวลาอันมีค่าในการให้คำแนะนำ ให้คำปรึกษาและแนวทางการพัฒนาในการทำวิจัย ตลอดจนสอนขั้นตอนการดำเนินการทำวิจัยจนงานวิจัยสำเร็จลุล่วงไปได้ด้วยดี

ขอขอบพระคุณ ดร.ณัฐกิตติ จิตรเอื้อตระกูล (มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ) รองศาสตราจารย์ ดร.วรากร ศรีเชวงทรัพย์ และ ดร.ชัชไชย วรรณบุรณ์ (สถาบันเทคโนโลยีไทย-ญี่ปุ่น) ที่กรุณาสละเวลา เป็นประธานและคณะกรรมการสอบวิทยานิพนธ์ในครั้งนี้

ขอขอบพระคุณ ดร.อัทนา เชนโต้ะ (สถาบันเทคโนโลยีไทย-ญี่ปุ่น) ที่ให้คำแนะนำในการเขียนเอกสารวิทยานิพนธ์ และระเบียบวิธีวิจัย อีกทั้งยังแนะนำแนวทางในการดำเนินการศึกษาวิจัยส่วนเพิ่ม เพื่อต่อยอดแนวทางการดำเนินการศึกษาและวิจัยเดิมที่ได้ดำเนินการมา

ขอขอบพระคุณ คุณพ่อ คุณแม่ และครอบครัว ที่ให้โอกาสในการศึกษา และขอขอบคุณเพื่อนๆ อาจารย์ และผู้มีพระคุณทุกท่าน ทั้งผู้ที่กล่าวนามและไม่ได้กล่าวนาม ณ ที่นี้ ที่คอยให้กำลังใจให้คำปรึกษาและสนับสนุนในการทำงาน จนดำเนินงานสำเร็จลุล่วงไปได้ด้วยดี ซึ่งผู้วิจัยมีความซาบซึ้งในความกรุณาอันดีจากทุกท่าน และหวังเป็นอย่างยิ่งว่า งานวิจัยฉบับนี้จะเป็นประโยชน์ต่อผู้ศึกษาไม่มากนักน้อย หากมีสิ่งใดขาดตกบกพร่อง ผู้วิจัยขออภัยมา ณ ที่นี้

ภัทรภณ เลาทกิจวิฑูร

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฌ
สารบัญรูป.....	ญ
บทที่	
1 บทนำ.....	1
1.1 ที่มาและความสำคัญ.....	1
1.2 วัตถุประสงค์ของการศึกษา.....	4
1.3 ขอบเขตการวิจัย.....	4
1.4 ขั้นตอนการดำเนินงาน.....	5
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	5
1.6 แผนงานและระยะเวลาดำเนินการ.....	6
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ประเภทที่วางจำหน่ายในประเทศไทย.....	7
2.2 ทฤษฎีที่เกี่ยวข้องกับโครงข่ายประสาทเทียม.....	8
2.3 ฟังก์ชันกระตุ้น (Activation Function).....	10
2.4 ฟังก์ชันสูญเสีย (Loss Function).....	11
2.5 การหาค่าที่ดีที่สุด (Optimization).....	12
2.6 การเรียนรู้เชิงลึก (Deep Learning).....	13
2.7 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network). ..	15
2.8 ฟังก์ชันซอฟต์แวร์แมกซ์ (Softmax function).....	23
2.9 การเรียนรู้เชิงถ่ายโอนข้อมูล (Transfer Learning).....	23
2.10 งานวิจัยที่เกี่ยวข้อง.....	28

สารบัญ (ต่อ)

บทที่	หน้า
3	ระเบียบวิธีการวิจัย 33
3.1	ขั้นตอนการดำเนินงาน..... 33
3.2	กรอบแนวคิดการวิจัย..... 34
3.3	ขั้นตอนการทดลอง 35
3.4	การเตรียมชุดข้อมูล (Data Processing)..... 36
3.5	อุปกรณ์ในการทดลอง และเครื่องมือที่ใช้ในงานวิจัย 40
3.6	การสร้างแบบจำลอง..... 43
4	ผลการวิจัยและการวิเคราะห์ผลการทดลอง 46
4.1	ผลลัพธ์ของระบบ..... 46
4.2	การวัดประสิทธิภาพ..... 48
4.3	อภิปรายผลลัพธ์ของโมเดลแต่ละชนิด 49
4.4	สรุปผล..... 55
5	สรุปผลและข้อเสนอแนะ..... 56
5.1	สรุปผลการวิจัย..... 57
5.2	ข้อเสนอแนะเพื่อการพัฒนางานวิจัย 57
บรรณานุกรม	 61
ภาคผนวก	 63
ภาคผนวก ก. ตารางบันทึกผลการจำแนกของโมเดล CNN ทั้งหมดที่ใช้ใน การทดลอง	 64
ภาคผนวก ข. Conference Paper.....	68
ประวัติย่อผู้วิจัย.....	74

สารบัญตาราง

ตาราง	หน้า
1.1 แผนงานและระยะเวลาในการดำเนินการศึกษาวิจัย ปี พ.ศ. 2561-2562	6
2.1 สรุปผลงานวิจัยที่เกี่ยวข้อง	28
3.1 การเปรียบเทียบการจับคู่ของยาทั้ง 2 ชนิดระหว่างยาต้นแบบ และยาสามัญ.....	36
3.2 ชุดคำสั่งภาษา Python ที่ใช้ในงานวิจัย.....	42
3.3 กลุ่มเลเยอร์ Feature learning ของโมเดล CNN ที่สร้างขึ้น.....	43
3.4 กลุ่มเลเยอร์ Fully connected layers ของโมเดล CNN ที่สร้างขึ้น.....	43
3.5 สรุป Training setting และ Hyperparameters ของโมเดล CNN.....	44
4.1 ผลลัพธ์ของโมเดล 4 อันดับแรกที่ใช้สถาปัตยกรรม InceptionV3 ที่ทำคะแนนการวัด ประสิทธิภาพได้สูงสุดจากทั้งหมด 20 โมเดล ในการจำแนกยาเม็ด 2 ชนิด	51
4.2 ผลลัพธ์ของโมเดล 4 อันดับแรกที่ใช้สถาปัตยกรรม VGG16 ที่ทำคะแนนการวัด ประสิทธิภาพได้สูงสุดจากทั้งหมด 20 โมเดล ในการจำแนกยาเม็ด 2 ชนิด	52
4.3 ผลลัพธ์ของโมเดล 4 อันดับแรกที่ใช้สถาปัตยกรรม ResNet50-v2 ที่ทำคะแนนการวัด ประสิทธิภาพได้สูงสุดจากทั้งหมด 20 โมเดล ในการจำแนกยาเม็ด 2 ชนิด	52
4.4 การเปรียบเทียบโมเดลโครงข่ายประสาทเทียม Convolution ที่ใช้ในการทดลองกับ โมเดล Transfer learning อีก 3 ชนิด ที่ใช้ในการเปรียบเทียบ	55

TNI

THAI - NICHU INSTITUTE OF TECHNOLOGY

สารบัญรูป

รูป	หน้า
2.1 ภาพประกอบหลักการทำงานของ โครงข่ายประสาทเทียม	9
2.2 โครงสร้างการทำงานของ Machine Learning	13
2.3 โครงสร้างการทำงานของ การเรียนรู้เชิงลึก (Deep Learning)	14
2.4 หลักการทำงานของโครงสร้างโครงข่ายประสาทเทียมแบบ Convolution	16
2.5 รูปแบบการทำงานของเลเยอร์คอนโวลูชัน (Convolution layer).....	17
2.6 อิทธิพลของจำนวน Kernel ในการทำ convolution ต่อ feature map.....	18
2.7 ความแตกต่างของ feature map จากการทำ convolution ระหว่าง stride เท่ากับ 1 (บน) และ stride เท่ากับ 2 (ล่าง).....	19
2.8 หลักการทำงานของ Zero-Padding	19
2.9 หลักการทำงานของ Max Pooling (ซ้าย) และ Average Pooling (ขวา) โดย ใช้ขนาด pool size 2 x 2	21
2.10 หลักการทำงานของพูลลิ่งเลเยอร์ (Pooling layer)	21
2.11 หลักการทำงานของ Fully-Connected Layer ของโครงข่ายประสาทเทียมใน ขั้นตอน Classification	22
2.12 โครงสร้างสถาปัตยกรรมโครงข่ายประสาทเทียม Inception-v3.....	24
2.13 โครงสร้างโมดูล Inception ที่ประกอบด้วยเลเยอร์ convolution และ maxpooling	25
2.14 การทำ feature map กับรูปภาพ RGB โดยใช้เลเยอร์ convolution ขนาด 1 x 1 pixels	25
2.15 ภาพโครงสร้างสถาปัตยกรรมโครงข่ายประสาทเทียม VGG16.....	26
2.16 สถาปัตยกรรมโมเดล ResNet50-v2	27
3.1 แผนผังขั้นตอนการดำเนินงาน	33
3.2 กรอบแนวคิดการวิจัย.....	34
3.3 ตัวอย่างข้อมูลรูปภาพเม็ดยาที่ใช้ในการทดลองระหว่าง (บน) ยาสามัญ และ (ล่าง) ยาต้นแบบ	37
3.4 แผนผังขั้นตอนการฝึกฝน (training) ชุดข้อมูล training set.....	38
3.5 แผนผังขั้นตอนการประเมินผล (evaluation) โดยใช้ชุดข้อมูล validation set.....	39

สารบัญรูป (ต่อ)

รูป	หน้า
3.6 คอมพิวเตอร์โน้ตบุ๊ก Asus TUF Gaming ที่ใช้สำหรับการเขียนโมเดล และ ดำเนินการทดลอง ขั้นตอนต่างๆ.....	40
3.7 โทรศัพท์มือถือ Huawei nova 3i ที่ใช้สำหรับการถ่ายภาพ เพื่อสร้างชุดข้อมูล แต่ละชนิดในงานทดลอง.....	40
3.8 โปรแกรม PyCharm Community Edition สำหรับการเขียนโมเดล CNN และดำเนินการ ทดลองทดลองด้วยภาษา Python.....	41
3.9 โลโก้โปรแกรม PyCharm Community Edition	41
3.10 โครงสร้างโมเดลโครงข่ายประสาทเทียม Convolution ที่ใช้ในการทดลอง	44
3.11 ตัวอย่างการสร้างโมเดล CNN ผ่านกระบวนการ transfer learning	45
4.1 แผนผังขั้นตอนกระบวนการการคาดการณ์ (prediction) ของโมเดล CNN เพื่อทำ การวัดประสิทธิภาพ.....	46
4.2 ตัวอย่างการทำ Classification ของ โมเดล CNN.....	47
4.3 ตาราง Confusion Matrix ในการวัดความสามารถของ machine learning ในการแก้ปัญหา classification	48
4.4 ตาราง Confusion Matrix ของโครงข่ายประสาทเทียม Convolution ที่ใช้ใน การทดลอง	50
4.5 ตาราง Confusion Matrix สำหรับโมเดล InceptionV3_6	51
4.6 ตาราง Confusion Matrix สำหรับโมเดล VGG16_17.....	52
4.7 ตาราง Confusion Matrix สำหรับโมเดล ResNet50V2_5	53
4.8 กราฟอัตราความแม่นยำในการตรวจจำยาเม็ด 26 ตัวจากโมเดล CNN ทั้งหมด ที่ใช้ในการทดลอง.....	54

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญ

ในโครงสร้างสาธารณสุขของประเทศไทยปัจจุบัน คลินิกเอกชนและร้านขายยา เป็นกำลังที่สำคัญประเภทหนึ่งของระบบบริการสุขภาพที่อยู่ใกล้ชิดประชาชนมากที่สุด มีบทบาทเปรียบเสมือนเป็นที่พึ่งด้านสุขภาพของชุมชนท้องถิ่นทั้งการดูแลรักษาโรคและการจ่ายยาเป็นอีกทางเลือกหนึ่งของประชาชนในการใช้บริการเมื่อมีอาการเจ็บป่วยเบื้องต้นไม่รุนแรงสามารถเข้าถึงการรักษาได้สะดวก รวดเร็ว [1]

แต่อย่างไรก็ตาม การเลือกใช้ระบบบริการสุขภาพประเภทนี้ ยังคงพบปัญหาในด้านต่างๆ เช่น ปัญหาด้านคุณสมบัติของบุคลากรทางการแพทย์ ปัญหาการจ่ายยาโดยผู้ที่ไม่ใช่เภสัชกร และปัญหาสำคัญที่เกิดขึ้นและถูกละเลยบ่อยครั้ง คือ ปัญหาการจ่ายยาโดยไม่ระบุชื่อยา หรือการให้ข้อมูลยาที่จำเป็นไม่เพียงพอ บนซองยาจากคลินิกเอกชนและร้านขายยาหลายแห่ง อันนำไปสู่ ปัญหาการเกิดเหตุการณ์ที่ไม่พึงประสงค์จากการใช้ยา (Adverse Drug Events) และปัญหาการใช้ยาไม่ได้ผล ซึ่งส่งผลกระทบต่อการรักษาโรคของผู้ป่วยที่อาจทำให้เกิดอาการถดถอยลงหรือมีอันตรายถึงชีวิต รวมไปถึงสิ้นเปลืองค่าใช้จ่ายที่เกินจำเป็น [2] คลินิกเอกชนและร้านขายยาบางแห่ง มักจะหลีกเลี่ยงการระบุชื่อ รวมทั้งสรรพคุณข้อบ่งใช้บนหน้าซองยา โดยเขียนเพียงวิธีรับประทานเท่านั้น อีกทั้งการจ่ายยาเม็ด มักจ่ายในรูปแบบไม่อยู่ในแผงยา การกระทำเช่นนี้ก็เนื่องจากเหตุผลการสร้างรายได้เปรียบทางธุรกิจ เพื่อให้ผู้ป่วยที่ซื้อหรือรับยาไปนั้น หากต้องการได้ยาตัวเดิมมาใช้อีก จะไม่สามารถนำยาไปถามหรือหาซื้อจากร้านขายยาหรือคลินิกอื่นได้ เพราะไม่สามารถระบุได้ชัดเจนว่าเป็นยาอะไร ทำให้จำเป็นต้องวนกลับมาซื้อที่ร้านขายยาหรือคลินิกเดิม [2] นอกจากนี้ ประเด็นสำคัญที่ทำให้เพิ่มปัญหาการใช้ยา โดยผู้ประกอบการคลินิกเอกชนและร้านขายยาที่มักใช้กลวิธีคือ การเลือกจ่ายยาสามัญแทนยาต้นแบบ ซึ่งยาสามัญโดยทั่วไปจะถูกผลิตให้มีรูปลักษณ์ที่คล้ายกับยาต้นแบบเพื่อให้ผู้ป่วยไม่สามารถสังเกตความแตกต่างได้ และด้วยต้นทุนการผลิตและราคาที่ต่ำกว่ามากของยาสามัญ ทำให้สร้างผลกำไรได้มากขึ้น แต่หากพิจารณาปัญหาการใช้ยาสามัญ ซึ่งถึงแม้จะมีชื่อตัวยาสำคัญหรือสารออกฤทธิ์เดียวกันกับยาต้นแบบ แต่โดยข้อเท็จจริงนั้น ยาสามัญมีข้อแตกต่างจากยาต้นแบบหลายประการ เช่น แหล่งที่มาของตัวยาสำคัญและสารองค์ประกอบอื่นในตำรับเป็นคนละแหล่ง เทคโนโลยีและกรรมวิธีในการผลิต ระบบการควบคุมคุณภาพแตกต่างกัน ซึ่งด้วยการลงทุนที่ต่ำมาก ทำให้มาตรฐานประสิทธิภาพและคุณภาพของยาสามัญบางตัวด้อยกว่ายาต้นแบบ ด้วยประการเหล่านี้ ทำให้การรักษาโรคด้วยยาสามัญอาจได้ผลแตกต่างจากยาต้นแบบ อีกทั้งยาสามัญอาจทำให้เกิดผล

ข้างเคียงกับผู้ป่วยได้แตกต่างไปจากยาต้นแบบ ทำให้การควบคุมเฝ้าระวังการใช้ยาสามัญทำได้ยากกว่ายาต้นแบบที่รับทราบข้อควรระวังการใช้อยู่แล้ว ซึ่งในบางกรณี ก็อาจเหนือความสามารถที่จะแก้ไขเหตุการณ์ที่ไม่พึงประสงค์จากการใช้ยาสามัญได้ [3] [4] จึงกล่าวได้ว่า ยาสามัญไม่ใช่จะสามารถใช้แทนยาต้นแบบได้เสมอ เพราะการรักษาโรคที่มีประสิทธิภาพและปลอดภัยที่สุดนั้น ยังต้องพึ่งพาการเลือกใช้ยาที่เป็นยาต้นแบบ ที่ผ่านการศึกษาวิจัยรักษาโรคนั้นและได้ติดตามความปลอดภัยในการใช้ยาเป็นอย่างดีแล้ว และด้วยปัญหาจากความคล้ายคลึงในรูปลักษณะของยาต้นแบบกับยาสามัญ ทำให้เมื่อเกิดกรณีผลการรักษาโรคเบี่ยงเบนไปจากที่ควรเป็น หรือเกิดผลข้างเคียงที่ไม่สัมพันธ์กับยา จำเป็นต้องตรวจสอบยาที่ผู้ป่วยได้รับว่าคืออะไร เป็นยาต้นแบบหรือยาสามัญ เพื่อเป็นข้อมูลสำคัญในการประเมินผลการรักษาและการแก้ไขเหตุการณ์ไม่พึงประสงค์ที่เกิดขึ้น ซึ่งความคล้ายคลึงในรูปลักษณะของยาต้นแบบกับยาสามัญนี้ มักเป็นปัญหาที่ทำให้เภสัชกรหรือบุคลากรทางการแพทย์ ไม่สามารถแยกแยะออกได้โดยง่าย และต้องใช้เวลามากในการวิเคราะห์จำแนก จนบางครั้งอาจไม่สอดคล้องกับภาวะฉุกเฉินหรือช่วงเวลาที่ต้องการได้ ดังนั้นการสร้างเครื่องมือที่ใช้จำแนกระหว่างยาทั้ง 2 ชนิดจึงเป็นสิ่งที่จำเป็นเพื่อให้บุคลากรทางการแพทย์สามารถจำแนก และระบุเมื่อยาว่าเป็นผลิตภัณฑ์ยาต้นแบบหรือยาสามัญได้อย่างถูกต้อง ซึ่งปัจจุบันในประเทศไทย มีเครื่องมือที่ใช้ตรวจสอบเมื่อยาที่ให้บริการบางส่วน อาทิ เว็บไซต์ฐานข้อมูลการพิสูจน์เอกลักษณ์ยาเม็ดและแคปซูลในประเทศไทย โดยคณะเภสัชศาสตร์ มหาวิทยาลัยอุบลราชธานี แต่ก็เป็นเครื่องมือที่ไม่สมบูรณ์ เนื่องจากเป็นระบบการป้อนข้อมูลที่มีความซับซ้อนเข้าใจยาก และการประมวลผลไม่ค่อยมีประสิทธิภาพอันเนื่องจากฐานข้อมูลเมื่อยาไม่เพียงพอ ไม่มีการอัปเดต อีกทั้ง ไม่ได้ตอบสนองการจำแนกเมื่อยาที่มีรูปลักษณะคล้ายคลึงกันว่าเป็นยาต้นแบบหรือยาสามัญ จึงไม่ตรงกับเป้าหมายที่ต้องการ

ในปัจจุบัน การจำแนกภาพ (Image Classification) ด้วยโครงข่ายประสาทเทียม Convolutional Neural Network กำลังเป็นที่นิยมกันอย่างแพร่หลาย อันเนื่องมาจากความสามารถในการ การจดจำคุณลักษณะที่สามารถนำมาประยุกต์ใช้กับงานด้านต่างๆ ในสาขา Computer vision ซึ่งได้มีการพัฒนาอย่างต่อเนื่อง และก้าวหน้าไปอย่างรวดเร็วจนกลายเป็นวิธีการเรียนรู้เชิงลึก (Deep learning) ที่เป็นที่รู้จัก โดยเฉพาะโครงสร้างโมเดลที่สามารถการสกัดคุณลักษณะสำคัญ (Feature extraction) ของวัตถุจากข้อมูลรูปภาพที่ใช้ในการเรียนรู้ แล้วจึงนำคุณลักษณะเหล่านั้นมาประมวลผลรวมกันเพื่อทำการคาดการณ์ (Prediction) และจำแนกวัตถุจากภาพถ่าย หรือกล้องวิดีโอได้อย่างมีประสิทธิภาพ ปัจจุบันมีงานวิจัยอยู่หลากหลายฉบับ ที่ใช้เทคโนโลยีคอมพิวเตอร์วิทัศน์ในการจดจำเพื่อจำแนกตรวจจำวัตถุแต่ละชนิด โดยประยุกต์โครงข่ายประสาทเทียมคอนโวลูมาใช้ในการทดลอง เช่น งานวิจัยการจำแนกภาพขวดเซตเปิดด้วยโครงข่ายประสาทเทียม CNN เพื่อจำแนกภาพขวดเครื่องดื่มออกเป็นประเภทที่กำหนด และสามารถระบุว่าเป็นประเภทที่ไม่รู้จัก พร้อมศึกษาประสิทธิภาพการทำงานในแง่ความแม่นยำเมื่อใช้กับปัญหาการจำแนกแบบเซตเปิด (open-set) ซึ่ง

เป็นสภาพแวดล้อมที่คลาสเป้าหมาย (target class) ไม่ปรากฏในขั้นตอนการฝึก ศุภณัฐ จินตวัฒน์สกุล [5], งานวิจัยการประมวลผลภาพสำหรับการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์ โดยการจำลองการมองเห็นของมนุษย์ด้วยวิธีการเรียนรู้เชิงลึกเพื่อจำแนกภาพภาพถ่ายมะม่วงพันธุ์โชคอนันต์ออกเป็น 4 ระดับ ซึ่งประกอบด้วยคุณภาพระดับเกรด A, เกรด B, เกรด C และเกรด D (คือมะม่วงเน่า) ตามลำดับ ด้วยอัลกอริทึมโครงข่ายประสาทเทียมแบบคอนโวลูชัน นพรุจ พัฒนสารและ ญัฐวุฒิ ศรีวิบูลย์ [6], งานวิจัยการตรวจหาต้นไม้เป็นโรคด้วยการเรียนรู้เชิงลึก โดยใช้โครงข่ายประสาทเทียมคอนโวลูชันกับภาพถ่ายมุมสูงจากโดรน เพื่อตรวจหาต้นไม้ที่เป็นโรคโดยใช้โมเดลแบบลำดับชั้น (Cascade Model) ในการตรวจจับ (Object detection) และจำแนกประเภท รัตนโชติ พันธุ์ไธ [7] และงานวิจัยการดึงข้อมูลภาพยาเม็ดแบบไฮบริดโดยใช้โมเดล CNN ด้วยวิธีการแบบ supervised ที่มีเทคนิคการจัดกลุ่ม k-mean แบบไม่เชิงเส้นเป็นเทคนิคในการแบ่งส่วนและมีตัวแยกประเภท KNN เพื่อใช้แก้ไขปัญหการจำแนก (Classification) และการถดถอย (Regression) โดยมีข้อมูลภาพยาเม็ดมากกว่า 7,000 อิมเมจ จากยาทั้งหมด 1,000 ชนิด E. A. Adewumi [8] ซึ่งงานวิจัยที่กล่าวมาข้างต้นต่างให้ผลลัพธ์ที่ค่อนข้างดี แสดงให้เห็นถึงประสิทธิภาพในการจดจำ (Recognition) และจำแนกประเภทของโครงข่ายประสาทเทียมคอนโวลูชันในการตรวจจับวัตถุ (Object detection) ไม่ว่าจะเป็นภาพนิ่งหรือภาพเคลื่อนไหว นอกจากนี้ด้วยเทคนิคการเรียนรู้ถ่ายโอนข้อมูล (Transfer learning) ก็ยังทำให้การฝึกจำแบบลองการเรียนรู้เชิงลึกที่มีทั้งขนาด และความซับซ้อนสามารถทำได้ง่ายขึ้น และลดระยะเวลาที่ใช้ในการฝึกฝนที่อาจกินเวลานานหลายสัปดาห์ ด้วยการนำบางส่วนของแบบจำลองที่เคยผ่านการ pre-train กับงานที่ใกล้เคียงกัน มาใช้เป็นส่วนหนึ่งของโมเดลใหม่ [9] ดังนั้น งานวิจัยนี้จึงมีวัตถุประสงค์ ในการใช้เทคโนโลยีการจำแนกภาพด้วยวิธีการเรียนรู้เชิงลึก โดยการประยุกต์โครงข่ายประสาทเทียมแบบคอนโวลูชันมาใช้ในการพัฒนาอัลกอริทึมเพื่อสร้างแบบจำลองที่มีประสิทธิภาพในการจำแนกยาเม็ดระหว่างยาต้นแบบและยาสามัญ โดยใช้ข้อมูลรูปภาพทั้งหมด 6,240 ภาพ ที่ประกอบไปด้วยยาต้นแบบ และยาสามัญจำนวน 26 ตัว โดยแบ่งเป็น 5,200 ภาพ สำหรับชุดข้อมูลการฝึกฝน (training set) และ 1,040 ภาพ สำหรับชุดข้อมูลการทดสอบ (test set) พร้อมทั้งทดลองเปรียบเทียบประสิทธิภาพในการจำแนกกับโมเดล CNN ทั้ง 3 แบบที่สร้างจากสถาปัตยกรรม CNN ที่แตกต่างและเคยผ่านการทดลองกับงานที่มีลักษณะใกล้เคียงกัน ได้แก่ VGG16, InceptionV3 และ ResNet50-v2 ด้วยเทคนิค Transfer learning เพื่อหาโมเดลที่มีประสิทธิภาพในการจำแนกสูงสุดที่จะนำไปพัฒนาต่อเป็นเครื่องมือช่วยสนับสนุนในการระบุสาเหตุการแพ้ยา หรือการใช้ยาที่ไม่ถูกต้องจากการใช้ยาสามัญแทนยาต้นแบบ สำหรับบุคลากรทางการแพทย์ในอนาคตโดยการตรวจสอบยาเม็ด และจำแนกได้อย่างถูกต้อง

1.2 วัตถุประสงค์ของการศึกษา

1.2.1 เพื่อสร้างโมเดลต้นแบบในการจำแนกยาเม็ด ระหว่างยาดัชนีแบบ และยาสามัญ บนพื้นฐานโครงข่ายประสาทเทียมแบบคอนโวลูชันที่มีความแม่นยำสูงซึ่งจะพัฒนาต่อไปในอนาคต

1.2.2 เพื่อหาสถาปัตยกรรมโมเดล CNN ที่ดีที่สุดและ วิธีที่เหมาะสมในการจำแนกยาเม็ด ระหว่างยาดัชนีแบบและยาสามัญตามที่กำหนดจากยาเม็ดทั้งหมด 26 ตัวที่ใช้ในงานทดลอง

1.2.3 เพื่อทราบถึงหลักการทำงานของโครงข่ายประสาทเทียมแบบคอนโวลูชันในการสร้างโมเดลที่ใช้ในการจำแนกยาเม็ดระหว่างยาดัชนีแบบและยาสามัญ

1.2.4 เพื่อเรียนรู้ขั้นตอนวิธีในการสร้างโมเดลผ่านเทคนิค Transfer learning ในการดึงบางส่วนของโมเดลอื่นๆ มาใช้เป็นส่วนหนึ่งของโมเดลใหม่ ในการจำแนกยาเม็ดระหว่างยาดัชนีแบบและยาสามัญ

1.2.5 เพื่อนำความรู้ที่ได้เรียนรู้ มาพัฒนาใช้ในงานเภสัชกรรม

1.3 ขอบเขตของการศึกษา

1.3.1 ภาษาที่ใช้ในการเขียนโปรแกรมคือภาษา Python version 3.9 บนคอมพิวเตอร์โน้ตบุ๊ก ASUS TUF Dash F15 พร้อมการ์ดจอ GeForce RTX™ 3070 และซีพียู Intel Core™ i7 โดยใช้โปรแกรม PyCharm Community Edition พร้อมไลบรารี tensorflow-gpu 2.6.0

1.3.2 ประเภทของข้อมูลที่นำมาใช้ในการทดลองคือ ไฟล์ภาพ JPEG ขนาด 4,608 x 3,456 pixels ที่ถ่ายด้วยโทรศัพท์มือถือจำนวน 6,240 ภาพ จากยาทั้งหมด 2 ชนิดซึ่งแบ่งเป็นยาดัชนีแบบ และยาสามัญอย่างละ 13 ตัว โดยเป็นภาพด้านหน้าและภาพด้านหลังของยาเป็นหลัก ที่ถ่ายในมุมต่างๆ บนแผ่นรองตัดในสภาพแวดล้อมที่ไม่มีการควบคุม

1.3.3 ยาดัชนีแบบ และยาสามัญ ที่ใช้ในการทดลอง คือยาเม็ดแบบแข็งที่ใช้กันทั่วไป และสามารถหาซื้อได้ตามร้านขายยา คลินิกเอกชน หรือสถานพยาบาล ในประเทศไทย ซึ่งจำกัดจำนวนอยู่ที่ 26 ตัว ที่ใช้ในการทดลอง

1.4 ขั้นตอนการดำเนินงาน

1.4.1 ศึกษาค้นคว้าข้อมูล และงานวิจัยที่เกี่ยวข้อง

1.4.2 เขียนเค้าโครงการ เสนอกับอาจารย์ที่ปรึกษา

1.4.3 ศึกษาโปรแกรมที่ใช้ ภาษาคอมพิวเตอร์ ไลบรารี และอุปกรณ์ที่เกี่ยวข้อง

1.4.4 ทดสอบขีดความสามารถ และการใช้งานอุปกรณ์ โปรแกรม ไลบรารี กล้อง และคอมพิวเตอร์

1.4.5 เก็บรวบรวมข้อมูลภาพถ่ายยาเม็ด ทั้งยาต้นแบบและยาสามัญที่ขายตาม คลินิก สถานพยาบาล และร้านขายยาทั่วไปในประเทศไทย

1.4.6 จัดเตรียมชุดข้อมูลรูปภาพแบ่งออกเป็น 2 ชุด คือชุดสำหรับการฝึกฝน (training set) และชุดสำหรับการทดสอบ (test set) พร้อมทำการจัดประเภท (labeling) เพื่อความสะดวกในการใช้งาน

1.4.7 พัฒนาโมเดล CNN ที่ออกแบบมาเพื่อใช้ในการทดลองโดยมีวัตถุประสงค์เพื่อใช้ในการจำแนกยาเม็ดทั้ง 2 ชนิดและโมเดล CNN อีก 3 ชนิดที่สร้างด้วยเทคนิค transfer learning โดยการดัดแปลงโมเดล CNN ที่เคยผ่านการ pre-train กับงานที่ใกล้เคียงกันมาสร้างเป็น โมเดลใหม่เพื่อใช้ในการเปรียบเทียบโดยใช้ชุดข้อมูล training set และทำการประเมินผล (evaluation) ในระหว่างนั้นด้วยชุดข้อมูล test set พร้อมกำหนดจำนวน Epoch และพารามิเตอร์อื่นๆ ที่มีความเหมาะสม

1.4.8 ตรวจสอบ Accuracy, loss และเวลาที่ใช้ในการฝึก (train)

1.4.9 วัดประสิทธิภาพการจำแนกของโมเดลแต่ละชนิด ด้วย Accuracy, Precision, Recall, F1-score พร้อมทำการเปรียบเทียบ

1.4.10 นำโมเดลไปทดลองใช้งานจริงด้วยการทดลองทำนาย (prediction) กับภาพถ่ายยาเม็ดทั้ง 2 ชนิดที่ถ่ายด้วยกล้องโทรศัพท์มือถือ

1.4.11 วิเคราะห์ข้อมูล สรุปผล และเขียนรายงาน

1.4.12 นำเสนอโครงการวิจัย

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 ได้รับโมเดล CNN ที่มีความแม่นยำสูง ซึ่งจะถูกนำไปพัฒนาต่อให้กลายเป็นเครื่องมือสนับสนุนเภสัชกรและบุคลากรทางการแพทย์ ในการหาสาเหตุที่เกิดการแพ้ยา การใช้ยาไม่ไ้ผล และปัญหาการใช้ยาไม่ถูกต้องที่เกิดจากผลข้างเคียงของยาสามัญในอนาคด

1.5.2 ได้รับเทคนิคความรู้ วิธีการจำแนกยาเม็ดระหว่างยาต้นแบบและยาสามัญ จากภาพถ่ายโดยใช้คุณลักษณะ (feature) ต่างๆ ที่สามารถนำมาประยุกต์และพัฒนาต่อไป เพื่อสร้างระบบที่สามารถตรวจสอบยาเม็ดได้จริงในอนาคต

1.5.3 ได้เรียนรู้หลักการทำงานของการเรียนรู้เชิงลึก ในการจำแนกวัตถุจากประมวลผลคุณลักษณะต่างๆ ที่อยู่ในภาพ

1.5.4 ได้เรียนรู้การใช้โครงข่ายประสาทเทียม Convolutional ในการจำแนกยาเม็ดระหว่างยาต้นแบบและยาสามัญ จาก feature ของยาแต่ละชนิด

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ประเภทยาที่วางจำหน่ายในประเทศ

2.1.1 ยาต้นแบบ (Original drug)

ยาต้นแบบ คือ ยาที่ถูกสังเคราะห์ วิจัย คิดค้นพัฒนาและได้รับอนุมัติรับรองประสิทธิภาพและความปลอดภัยจนสามารถผลิตนำมาใช้รักษาโรคเป็นครั้งแรก ซึ่งต้องผ่านกรรมวิธี มากมายตั้งแต่ขั้นตอนการคิดค้นสารตั้งต้นและสังเคราะห์ตัวยาสำคัญที่มีคุณสมบัติในการรักษาโรคขั้นตอนการศึกษาวินิจฉัยสร้างสูตรตำรับต่างๆ การนำสูตรตำรับแต่ละรูปแบบผ่านกระบวนการทดลองทางเภสัชวิทยาอย่างละเอียดทั้งการทดลองกับสัตว์และมนุษย์ [4] โดยจำนวนการทดลองต้องพิสูจน์ได้ทางสถิติที่เป็นมาตรฐานทางวิทยาศาสตร์ จนแน่ใจได้ว่า ยาสูตรตำรับที่วิจัยพัฒนานั้นมีฤทธิ์และประสิทธิภาพในการรักษาโรคเป้าหมายได้อย่างแท้จริง มีความปลอดภัยหรือมีผลข้างเคียงน้อยที่สุดจึงจะยื่นขอรับการอนุมัติรับรองประสิทธิภาพและความปลอดภัยจากองค์การอาหารและยาของประเทศผู้ผลิตรายแรกนั้น เพื่อสามารถผลิตยาต้นแบบนำมาใช้รักษาผู้ป่วยได้อย่างกว้างขวางต่อไป

ยาต้นแบบทั้งหลายที่คิดค้นพัฒนาจนผลิตนำมาใช้รักษาโรคได้สำเร็จนั้น ต้องใช้เงินทุนมหาศาล บุคลากรหลายสาขาอาชีพ และอุปกรณ์เครื่องมือเพื่อการศึกษาวินิจฉัยทดลองมากมาย รวมทั้งระยะเวลายาวนานมากจึงเป็นส่วนสำคัญที่ทำให้ตัวยามีต้นทุนการผลิตที่สูง ซึ่งส่งผลโดยตรงต่อราคายา และผู้ผลิตยาก็จะต้องยื่นจดสิทธิบัตรยาเพื่อคุ้มครองสิทธิในการผลิตและจำหน่ายโดยส่วนใหญ่ยาต้นแบบที่มีคุณภาพเป็นที่ยอมรับทั่วโลกจะเป็นยาที่มาจากบริษัทผู้ผลิตรายใหญ่ของประเทศในยุโรปหรือสหรัฐอเมริกา [3] [4]

2.1.2 ยาสามัญ (Generic drug)

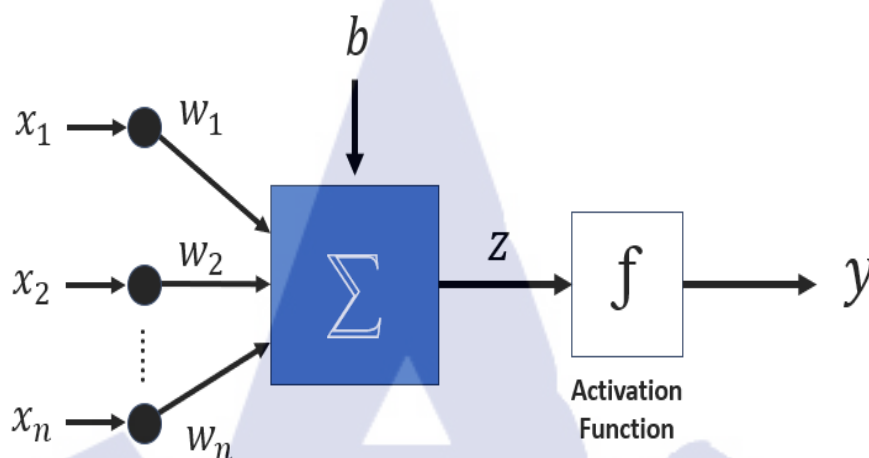
ยาสามัญ หรือ ยาเลียนแบบ (Local drug) คือ ยาที่ผลิตขึ้นภายใต้เครื่องหมายหรือชื่อการค้าที่ไม่ใช่เครื่องหมายหรือชื่อการค้าตามสิทธิของผู้ผลิตยาต้นแบบ ยาสามัญจะมีตัวยาสำคัญ ซึ่งเป็นข้อสารชนิดเดียวกันกับยาต้นแบบ โดยจะสามารถผลิตจำหน่ายได้หลังจากที่ยาต้นแบบได้รับการอนุมัติรับรองให้ใช้ในการรักษาโรคและหมดอายุสิทธิบัตรยาแล้ว[3]

ยาสามัญโดยทั่วไปจะมีต้นทุนการผลิตที่ต่ำกว่ายาต้นแบบมาก เนื่องจากยาสามัญไม่ได้มีขั้นตอนในการคิดค้นพัฒนาสูตรตำรับยา หรือการศึกษาวินิจฉัยทดลองสูตรตำรับยากับสัตว์และมนุษย์ ในช่วงกระบวนการทดลองทางเภสัชวิทยา ยาเลียนแบบจะมีเพียงขั้นตอนการทดสอบทางชีวสมมูล (Bioequivalence) คือ การประเมินคุณสมบัติของการแตกตัวและการกระจายของตัวยา การดูดซึม

และระดับยาในเลือด โดยกำหนดให้ค่าทางชีวสมมูลมีความแตกต่างกับยาต้นแบบได้ไม่เกิน 20% ซึ่งยาสามัญที่มีค่าทางชีวสมมูลเทียบเท่ายาต้นแบบจะถือสรุปว่ามีประสิทธิภาพไม่แตกต่างจากยาต้นแบบ [3] [4] แต่ในข้อเท็จจริงนั้น ยาสามัญ หรือยาเลียนแบบมีข้อแตกต่างจากยาต้นแบบ ได้แก่ แหล่งที่มาของวัตถุดิบหรือตัวยาสำคัญ ซึ่งส่วนใหญ่เป็นคนละแหล่ง แหล่งผลิตสารองค์ประกอบอื่นๆ ในตำรับที่ไม่ใช่ตัวยาสำคัญ เทคโนโลยีและกรรมวิธีในการผลิต ระบบการควบคุมคุณภาพ ฯลฯ ที่แตกต่างกัน ซึ่งสิ่งเหล่านี้เป็นปัจจัยที่ทำให้การรักษาด้วยยาสามัญ ได้ผลแตกต่างจากยาต้นแบบ แม้ว่าจะมีสารออกฤทธิ์เป็นสารชนิดเดียวกันก็ตาม รวมทั้งยาสามัญอาจทำให้เกิดผลข้างเคียงกับผู้ป่วยได้แตกต่างไปจากยาต้นแบบด้วยเช่นกัน ดังนั้น แม้ยาสามัญ หรือยาเลียนแบบมีความได้เปรียบกว่ายาต้นแบบในเชิงโครงสร้างต้นทุนและราคา การทำให้เพิ่มโอกาสการรักษาแก่ผู้ป่วย แต่ในขณะเดียวกัน คงต้องพึงระลึกเสมอว่าประสิทธิภาพและความปลอดภัยของการใช้ยาสามัญ อาจเบี่ยงเบนไปจากยาต้นแบบมากน้อยได้ ที่สำคัญยาสามัญหลายตัวที่ผลิตให้มีรูปลักษณะคล้ายคลึงยาต้นแบบ และนำมาจ่ายสลับแทนที่ยาต้นแบบ ก็เป็นสิ่งที่เภสัชกรหรือบุคลากรทางการแพทย์ต้องตรวจสอบให้ชัดเจนเพื่อประเมินการรักษาได้ถูกต้องด้วย

2.2 ทฤษฎีที่เกี่ยวข้องกับโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม หรือ Neural Networks คือ โมเดลทางคณิตศาสตร์ ที่มีการทำงานเลียนแบบสมองมนุษย์ในการประมวลผลสารสนเทศด้วยวิธีการคำนวณแบบ Connectionist โดยมีวัตถุดิบประสงค์เพื่อสร้างเครื่องมือที่ใช้ในการเรียนรู้ที่สามารถจดจำรูปแบบ (Pattern Recognition) และสร้างความรู้ใหม่ (Knowledge Extraction) [9] โดยอาศัยการใช้หน่วยประมวลผลจำนวนมาก และทำการเชื่อมต่อจนกลายเป็นโครงข่ายประสาทที่มีลักษณะการทำงานคล้ายคลึงกับเซลล์ประสาท(neurons) หรือระบบประสาทในสมองมนุษย์ โดยจะมีส่วนที่เรียกว่า ข้อมูลรับเข้า หรือข้อมูลอินพุต (input) ซึ่งจะผ่านการประมวลผลร่วมกัน กับค่าถ่วงน้ำหนัก (weight) ก่อนจะส่งไปยังฟังก์ชันการกระตุ้น (activation function) ซึ่งเป็นตัวกำหนดข้อมูลเอาต์พุต (output) ที่จะออกมาในลักษณะใดดังรูปภาพที่ 2.1



รูปที่ 2.1 ภาพประกอบหลักการทำงานของ โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมที่ทำงานเพียงแค่หนึ่งหน่วย ดังภาพประกอบที่ 2.1 จะถูกเรียกว่า เพอร์เซ็ปตรอน (Perceptron) ซึ่งเป็นหน่วยพื้นฐานที่เล็กที่สุดที่เกิดจากนิวรอน (neuron) หรือเซลล์ประสาทเทียมหลายตัวที่มาประกอบกันอย่างน้อยหนึ่งชั้นซึ่งสามารถเขียนได้ในรูปแบบของสมการดังสมการที่ 1 และสมการที่ 2 ตามลำดับ

โดย x_i คือ ข้อมูลอินพุต (Input) แต่ละตัวที่เข้าสู่โครงข่ายประสาทเทียม, w_i คือ ค่าถ่วงน้ำหนัก (weight), b คือ ค่าความเอนเอียง (bias), z คือ ผลลัพธ์เบื้องต้นที่เกิดจากการรวมผลของข้อมูลนำเข้าที่คูณกับค่าถ่วงน้ำหนัก, $f(.)$ คือ ฟังก์ชันการกระตุ้น (activation function) และ y คือ ข้อมูลเอาต์พุต (output)

$$z = \sum_{i=1}^N w_i x_i \quad (2.1)$$

$$y = f(b + z) \quad (2.2)$$

หลักการทำงานเมื่อนำผลรวมการคูณของข้อมูลนำเข้า (x_i) กับค่าถ่วงน้ำหนัก (w_i) ก็จะได้เป็นผลลัพธ์เบื้องต้น (z) ก่อนจะนำมาบวกด้วยค่า bias (b) แล้วเข้าสู่ฟังก์ชันการกระตุ้น เพื่อปรับค่าเป็นผลลัพธ์สุดท้ายโดยการเปรียบเทียบผลรวมเบื้องต้นกับค่า threshold ของฟังก์ชันการกระตุ้นก่อนจะออกมาเป็น ข้อมูลเอาต์พุต (y) ซึ่งจะกลายเป็นข้อมูลอินพุตในเพอร์เซ็ปตรอนถัดไปหากมีเซลล์ประสาทมากกว่า 1 เลเยอร์

สิ่งสำคัญต้องทราบในการทำโครงข่ายประสาทเทียม คือ ค่าถ่วงน้ำหนัก (weight) และค่า threshold ซึ่งเป็นสิ่งที่คอมพิวเตอร์ใช้ในการจดจำ และเป็นค่าที่ไม่แน่นอน แต่ก็สามารถกำหนดให้คอมพิวเตอร์สามารถปรับค่าเหล่านั้นได้โดยการใช้กระบวนการ back propagation ซึ่งเป็นกระบวนการย้อนกลับของการ recognize ในขั้นตอนการฝึก feed-forward โดยการสอนให้คอมพิวเตอร์รู้จัก pattern ของสิ่งที่ต้องการให้รู้จำ [9] ซึ่งคอมพิวเตอร์จะทำการปรับปรุง ค่าน้ำหนักคะแนนของเครือข่าย (network weight) หลังการใส่รูปแบบข้อมูลการฝึกให้แก่เครือข่ายในแต่ละครั้ง ซึ่งค่าเอาต์พุต จากเครือข่ายจะถูกนำไปเปรียบเทียบกับผลที่คาดหวัง แล้วทำการคำนวณหาค่าความผิดพลาด (Error) ซึ่งค่าความผิดพลาดที่ได้มานี้จะถูกส่งกลับไปยังเครือข่ายเพื่อใช้แก้ไขหาค่าน้ำหนักคะแนนต่อไป

2.3 ฟังก์ชันกระตุ้น (Activation Function)

ฟังก์ชันกระตุ้น หรือ Activation function คือ ฟังก์ชันที่เป็นตัวกำหนดว่า Perceptron จะส่งข้อมูลเอาต์พุตออกมาแบบใด[5] โดยการรับค่าผลรวมทั้งหมดจากทุกข้อมูลอินพุตที่คูณกับค่า weight ในทุกๆ dendrite ภายใน 1 นิวรอนมาปรับ เพื่อให้ได้ผลลัพธ์สุดท้ายโดยการเปรียบเทียบกับความถี่สัญญาณประสาทเอาต์พุต ซึ่งก็มีฟังก์ชันการกระตุ้นอยู่หลายรูปแบบโดยจะอธิบายในหัวข้อต่อไป

2.3.1 ฟังก์ชันซิกมอยด์ (Sigmoid function)

ฟังก์ชันซิกมอยด์ คือ ฟังก์ชันที่จะให้ผลลัพธ์ที่อยู่ในช่วง 0 ถึง 1 ดังสมการที่ 2 สามารถหา Derivative ได้ไม่ยาก เช่น 1 เท่ากับ Yes และ 0 เท่ากับ No แต่ก็มีข้อเสียเนื่องจากเกิดการ vanishing gradient หากอินพุตมีค่าน้อยกว่า -5 หรือมากกว่า 5 จนความชัน (slope) เข้าใกล้ 0 จะทำให้โมเดล train ไม่ไปไหน นอกจากนี้เอาต์พุตยังไม่มีค่าสมดุล และค่าเฉลี่ยไม่เท่ากับ 0 เพราะมีแต่ค่าเป็นบวกทำให้ Optimize ได้ยาก ดังนั้น Sigmoid Function จึงเสื่อมความนิยมลง และถูกแทนที่ด้วย ReLU Function ซึ่งจะอธิบายในหัวข้อ 2.3.3

$$\sigma(z) = \frac{1}{1 + \exp(-z)} = \frac{\exp(z)}{\exp(z) + 1} \quad (2.3)$$

2.3.2 ฟังก์ชันไฮเพอร์โบลิคแทนเจนต์ (Hyperbolic Tangent Function)

ฟังก์ชันไฮเพอร์โบลิคแทนเจนต์หรือ Tanh Function เป็นฟังก์ชันที่มีลักษณะคล้ายกับฟังก์ชัน Sigmoid แต่ให้ผลลัพธ์อยู่ในช่วง -1 ถึง 1 ดังสมการที่ 3 ทำให้ output มีความสมดุล มีค่าเฉลี่ยเท่ากับ 0

และสามารถทำ optimize ได้ง่าย จึงเหมาะสำหรับงาน Classification ระหว่าง 2 อย่างและเป็นตัว normalize ให้เป็น 1 Unit แต่ก็ไม่สามารถใช้เป็น probability ได้ยิ่งถ้าอินพุตมีค่าน้อยกว่า -3 หรือมีค่ามากกว่า 3 จน slope เข้าใกล้ 0 จะทำให้เกิดการ vanishing gradient ทำให้โมเดล train ไม่ไปไหนจึงไม่เหมาะจะนำมาเป็น activation function หลักของโมเดล CNN

$$\tanh(z) = \frac{\exp(z) - \exp(-z)}{\exp(z) + \exp(-z)} \quad (2.4)$$

2.3.3 ฟังก์ชันเรกติไฟด์เชิงเส้น (Rectified Linear Unit Function)

ฟังก์ชันเรกติไฟด์เชิงเส้นหรือ ReLU Function คือ ฟังก์ชันที่ให้ผลลัพธ์เป็น 0 หรือ z ดังสมการที่ 4 ซึ่งมีความเรียบง่ายกว่าทุก activation function แต่มีความทรงพลังมากกว่า เนื่องจากถ้าอินพุตเป็นค่าบวก และความชัน (slope) จะเท่ากับ 1 ตลอด ทำให้ไม่เกิดปัญหา vanishing gradient ส่งผลให้การเทรนโมเดลเร็วยิ่งขึ้นแม้ว่า output จะไม่สมดุล ทำให้การ optimize ทำได้ยาก และ output ไม่มี Limit จนจัดการได้ลำบากแต่ด้วยข้อดีที่ไม่เกิดปัญหา vanishing gradient จึงเป็น activation function ที่เหมาะสมจะนำมาใช้ในโมเดล CNN ที่สุด

$$\text{RelU}(z) = \max(0, z) \quad (2.5)$$

2.4 ฟังก์ชันสูญเสีย (Loss Function)

ในการฝึกฝนแบบจำลองโครงข่ายประสาทเทียม เพื่อให้เกิดการเรียนรู้ ค่าที่ได้จาก loss function คือ ตัวแปรที่ใช้ในการปรับค่าถ่วงน้ำหนัก (weight) เพื่อหาค่าที่ดีที่สุดที่ทำให้ค่าของ loss function ออกมาน้อยที่สุด [5] โดยปรับเปลี่ยนค่าถ่วงน้ำหนัก (weight) และค่า bias ในขั้นตอนการฝึก (train) ดังที่แสดงในสมการที่ 6 โดยค่า y คือ ค่าจริงค่า $\hat{Y}$ คือ ค่าที่ทำนายได้ และ $L(y, \hat{Y})$ คือ loss function

$$\mathbf{W}^* = \underset{\mathbf{W}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N L(y^{(i)}, \hat{\mathbf{Y}}^{(i)}) \quad (2.6)$$

ซึ่งก็มี loss function อยู่หลากหลายประเภท โดยต้องเลือกใช้ให้เหมาะสมกับงาน และสามารถแก้ปัญหาได้ดีที่สุด เช่น Cross-entropy, MSE, Kullback Leibler ฯลฯ

2.5 การหาค่าที่ดีที่สุด (Optimization) [5]

การ Optimization มีเป้าหมายเพื่อหาพารามิเตอร์ หรือค่า weight ที่ดีที่สุดเพื่อปรับปรุงแบบจำลองให้มีประสิทธิภาพ โดยวิธีการเคลื่อนลงตามความชัน (Gradient Descent Algorithm) ดังในสมการที่ 8 ซึ่ง gradient หรือ $\nabla g(w)$ คือ เวกเตอร์ของอนุพันธ์ย่อย (Partial derivatives) ทั้งหมด ซึ่งจะบอกถึงความเปลี่ยนแปลงของ $g(w)$ เมื่อ w เกิดการเปลี่ยนค่า

$$\nabla g(w) = \left(\frac{\partial g}{\partial w_1}(w), \dots, \frac{\partial g}{\partial w_K}(w) \right) \quad (2.7)$$

ค่าถ่วงน้ำหนักใหม่ (w) จะหาได้จากสมการที่ 9 โดย α คือ อัตราการเรียนรู้ (learning rate) และ $\frac{\partial L}{\partial w}$ คือ gradient ของ cost function ที่เทียบกับค่า weight

$$W \leftarrow W - \alpha * \frac{\partial L}{\partial W} \quad (2.8)$$

2.5.1 ฟูลแบตช์เกรเดียนต์เดสเซนท์ (Full-Batch Gradient Descent)

Full-Batch Gradient Descent จะใช้ข้อมูลทุกจุดในการคำนวณ gradient ดังสมการที่ 10 ซึ่งแสดงให้เห็นถึง การคำนวณ gradient ตั้งแต่ข้อมูลตัวที่ 1 ถึง N

$$W \leftarrow W - \alpha * \sum_{i=1}^N \frac{\partial L(y_i, \hat{Y}^{(i)})}{\partial W} \quad (2.9)$$

2.5.2 สโตแคสติกเกรเดียนต์เดสเซนท์ (Stochastic Gradient Descent)

Stochastic Gradient Descent หรือ SGD จะใช้ข้อมูลเพียงจุดเดียวในการคำนวณ gradient แทนที่ใช้ทุกจุด แบบดังสมการที่ 11

$$W \leftarrow W - \alpha * \frac{\partial L(y_i, \hat{Y}^{(i)})}{\partial W} \quad (2.10)$$

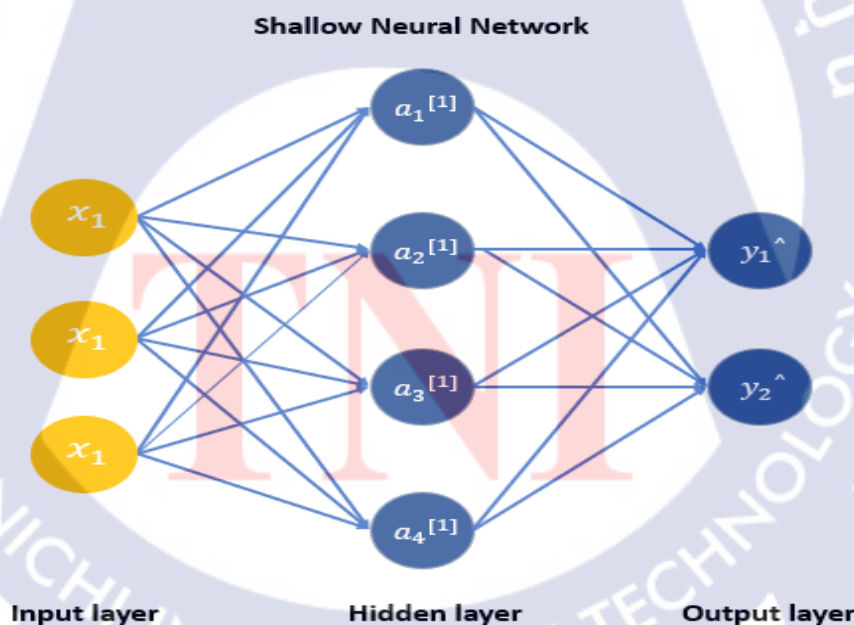
2.5.3 มินิแบตช์เกรเดียนต์เดสเซนท์ (Mini-Batch Gradient Descent)

Mini-Batch Gradient Descent จะใช้ข้อมูลบางส่วนในการคำนวณ gradient จากสมการที่ 12 แสดงให้เห็นการคำนวณ gradient จากข้อมูลตัวที่ k ถึง $k + m$

$$\mathbf{W} \leftarrow \mathbf{W} - \alpha * \sum_{i=k}^{k+m} \frac{\partial L(y_i, \hat{\mathbf{Y}}^{(i)})}{\partial \mathbf{W}} \quad (2.11)$$

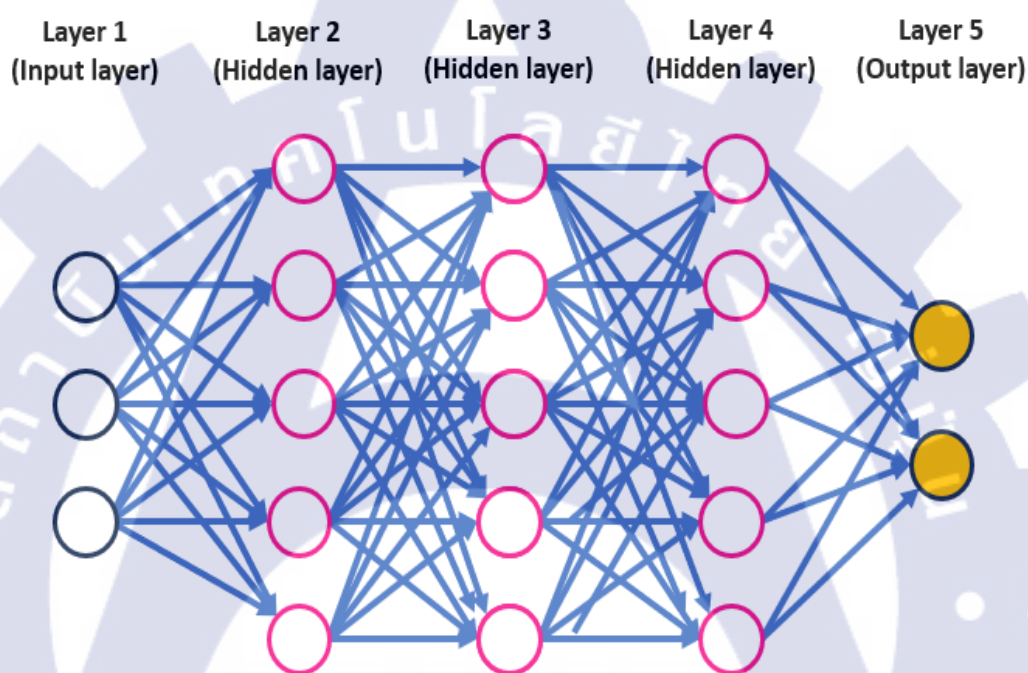
2.6 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก หรือ deep learning คือ หนึ่งในการทำ machine learning อย่างหนึ่ง เป็นเทคนิคที่ใช้ในการสร้างปัญญาประดิษฐ์ที่จำลองเลียนแบบการทำงานของเซลล์ neurons ภายในระบบประสาทของสมองมนุษย์ โดยใช้โครงข่ายประสาทเทียมแบบ machine learning ดังรูปภาพที่ 2.2 แต่จะมี hidden layer มากกว่า 2 ชั้นขึ้นไป เพื่อใช้ในการประมวลผลและตรวจสอบคุณสมบัติที่ดีขึ้น เป็นหนึ่งในเทคโนโลยีที่ทันสมัยและนิยมใช้กันมากในปัจจุบัน ทั้งในงานสาขาคอมพิวเตอร์วิทัศน์, ปัญญาประดิษฐ์ และการประมวลผลรูปภาพ



รูปที่ 2.2 โครงสร้างการทำงานของ Machine Learning

โดยข้อดีของการเรียนรู้เชิงลึกคือ เมื่อต้องการใช้งาน เช่น การประมวลผลภาพเพื่อแยกแยะ หรือทำการระบุ และตรวจสอบระหว่างคุณลักษณะ (feature) ก็สามารถทำได้เลย โดยไม่จำเป็นต้องให้ความรู้พื้นฐานกับระบบล่วงหน้าอันเนื่องมาจากการเรียนรู้เชิงลึกสามารถสร้างแบบจำลอง และหาคำตอบได้ด้วยการนำ layer ของ neural network หลายๆ ชั้นที่เรียกว่า hidden layer มาใช้วิเคราะห์ประมวลผล ในการหาคำตอบ ดังรูปภาพในรูปที่ 2.3 โดยจะมีเลเยอร์ input และเลเยอร์ output ทำหน้าเป็นตัวนำเข้า และแสดงผลลัพธ์จากข้อมูลที่ใส่เข้าไปตามลำดับ แบบเดียวกับ machine learning



รูปที่ 2.3 โครงสร้างการทำงานของ การเรียนรู้เชิงลึก (Deep learning)

คำว่า deep learning มาจากการใช้เลเยอร์ Neural Network มากกว่า 2 ชั้นขึ้นไป เพื่อให้เกิดการสร้างแบบจำลองที่สามารถเรียนรู้ได้ลึกขึ้น ดังนั้นจึงเทียบได้ว่าถ้าแต่ละชั้นของเลเยอร์ Neural Network ยิ่งใช้จำนวนมากเท่าใด ขั้นตอนการประมวลผลก็ยิ่งมีโครงสร้างและเรียนรู้ได้ลึกขึ้นเท่านั้น

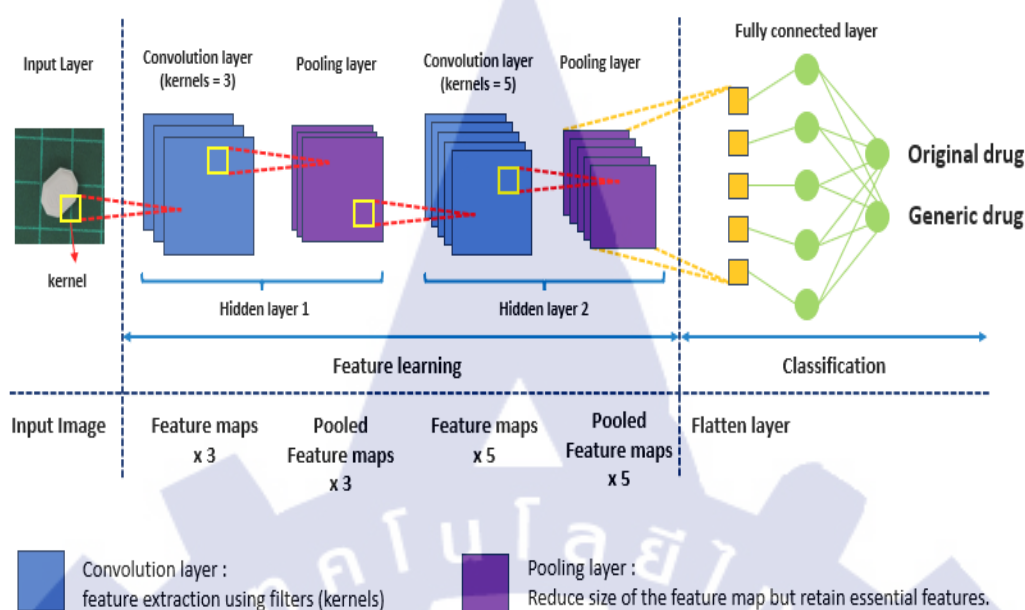
2.7 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network)

โครงข่ายประสาทเทียมแบบคอนโวลูชัน หรือ Convolutional Neural Network คือ หนึ่งในวิธีการ deep learning อย่างหนึ่งเป็นโครงข่ายประสาทเทียมที่ประกอบด้วยเซตตัวกรอง (kernel) ที่ใช้ในการ Convolution มีจุดเด่นด้านการสกัดคุณลักษณะ (feature extraction) จากภาพ โดยการปรับค่าเคอร์เนลให้ที่ความเหมาะสมและทำการจำแนก (classification) ซึ่งเป็นการจำลองการมองเห็นของมนุษย์ที่สามารถแยกแยะคุณลักษณะวัตถุผ่านการมองเห็น ก่อนจะนำมาประกอบวิเคราะห์ร่วมกัน และทำการระบุสิ่งที่มองเห็น โดยมีการคำนวณแบบ dot product ระหว่างพื้นที่ส่วนย่อยของภาพ (h) ที่แบ่งจากภาพต้นฉบับกลายเป็นพื้นที่ย่อยแต่ละอันกับตัวเคอร์เนล (Kernel : $F = k \times k$) เพื่อดึงคุณลักษณะเด่นจากภาพและทำการวิเคราะห์ โดยเคอร์เนลที่นำมาคำนวณจะเป็นเมทริกซ์จัตุรัส (Square Matrix) ที่มีขนาดเล็กกว่าพื้นที่ส่วนย่อย ซึ่งหากเป็นภาพสีจะมีขนาด 3×3 แต่ถ้าเป็นภาพขาวดำจะมีขนาดเท่ากับ 2×2 ซึ่งการคำนวณ Convolutional สามารถเขียนในรูปแบบสมการได้ดังสมการที่ 13 และสมการที่ 14 ตามลำดับ

$$G = h \times F \quad (2.12)$$

$$G[i,j] = \sum_{u=-k}^k \sum_{v=-k}^k h[u,v]F[i-u,j-v] \quad (2.13)$$

โครงข่ายประสาทเทียมแบบคอนโวลูชัน เป็นเทคนิคที่ได้รับความนิยมอย่างมากหลังจาก Alex Krizhevsky นักวิทยาศาสตร์คอมพิวเตอร์ได้นำเสนอ AlexNet สถาปัตยกรรมโมเดล CNN ซึ่งสามารถเอาชนะการแข่งขัน ImageNet Large Scale Visual Recognition Challenge หรือ ILSVRC ในปี 2012 โดยสามารถทำคะแนนได้สูงกว่าเทคนิคแบบดั้งเดิม และยังคงถูกพัฒนาอย่างต่อเนื่องจนติด 5 อันดับ error rate สูงสุดของชุดข้อมูล ImageNet ที่ลดจาก $\sim 25\%$ เป็น $\sim 2.25\%$ ภายใน 5 ปี หลังจากการนำเสนอ AlexNet ซึ่งในปัจจุบันโครงข่ายประสาทเทียม CNN ได้ถูกนำมาประยุกต์ใช้ในงานคอมพิวเตอร์วิทัศน์มากมายไม่ว่าจะเป็นการจำแนกรูปภาพ (Image Classification) การตรวจจับวัตถุ (Object Detection) การวิเคราะห์รูปภาพ และวิดีโอ (Image and Video Recognition) ฯลฯ [5]

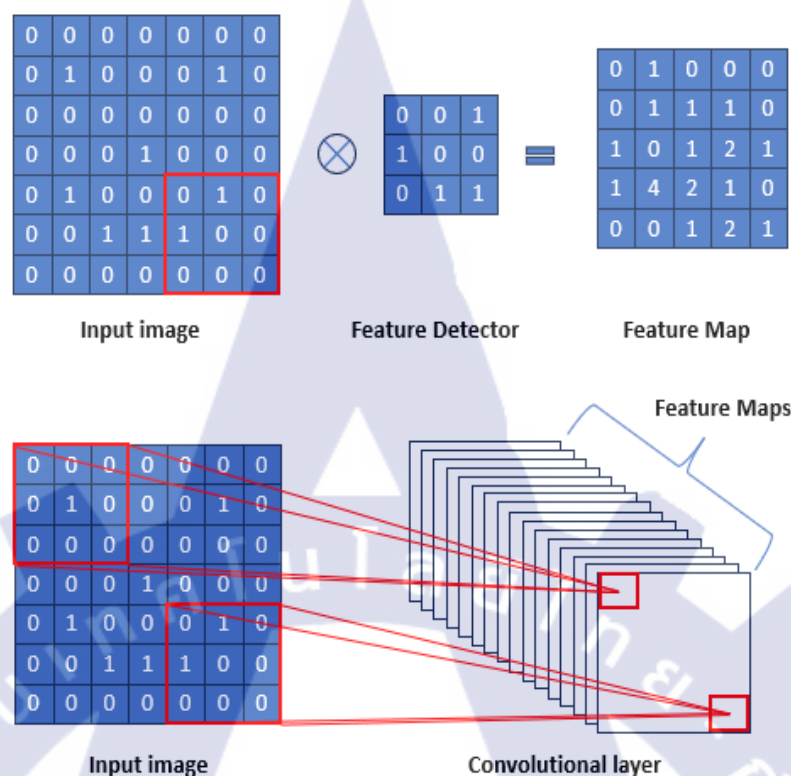


รูปที่ 2.4 หลักการทำงานของโครงสร้างโครงข่ายประสาทเทียมแบบ Convolution

โครงสร้างหลักของโครงข่ายประสาทเทียมแบบ Convolution จะประกอบด้วยเลเยอร์ 2 แบบ ที่ใช้ในการเรียนรู้คุณลักษณะ (feature learning) และจำแนก (classification) ดังรูปที่ 2.4 ถัดจากเลเยอร์อินพุต (Input Layer) ซึ่งทำหน้าที่รับข้อมูลนำเข้าสู่กระบวนการต่างๆ เพื่อการสกัดคุณลักษณะ และทำการประมวลผล โดยสามารถแบ่งออกได้เป็น เลเยอร์คอนโวลูชัน (Convolutional Layer), เลเยอร์พูลลิ่ง (Pooling Layer) และเลเยอร์เชื่อมต่อกันแบบสมบูรณ์ (Fully-Connected Layer) ซึ่งจะอธิบายในหัวข้อต่อไป

2.7.1 เลเยอร์คอนโวลูชัน (Convolutional Layer)

Convolutional Layer เป็นเลเยอร์ที่ใช้ในการทำการคอนโวลูชันระหว่างภาพอินพุต กับ ตัวเคอร์เนลเพื่อคำนวณหาผลลัพธ์ของ Neuron ที่เชื่อมต่อกับพื้นที่ส่วนย่อย (Local Region) ซึ่งเป็นการคำนวณแบบ dot product กับเคอร์เนลต่างๆ ที่มีขนาดเหมาะสมในการสกัดคุณลักษณะ (feature extraction) ของวัตถุที่อยู่ในภาพ ซึ่งจะมีการสุ่มและการปรับเปลี่ยนในขั้นตอนการเรียนรู้ก่อนจะออกมาเป็น แผนผังคุณลักษณะ (feature map) ซึ่งเป็นผลลัพธ์ที่ได้จากกระบวนการดังกล่าว ดังรูปที่ 2.5

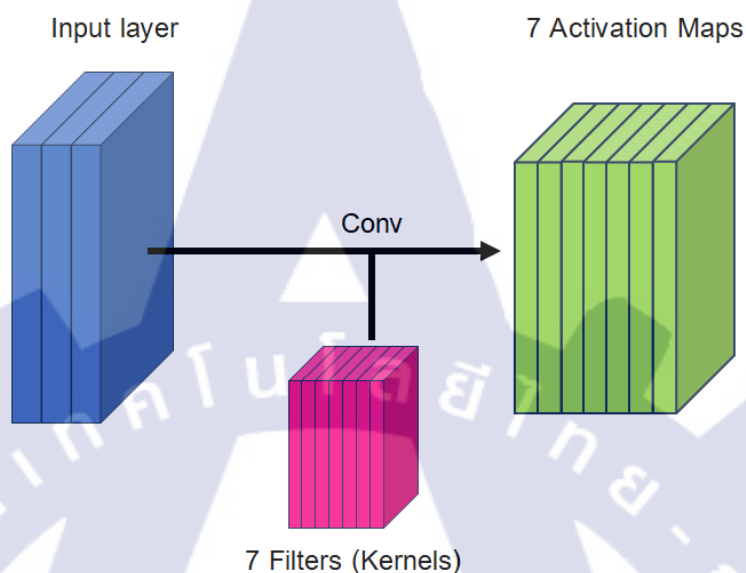


รูปที่ 2.5 รูปแบบการทำงานของเลเยอร์คอนโวลูชัน (Convolution layer)

Convolutional Layer เป็นเลเยอร์ที่ใช้ในการทำการคอนโวลูชันระหว่างภาพอินพุต กับ ตัวเคอร์เนลเพื่อคำนวณหาผลลัพธ์ของ Neuron ที่เชื่อมต่อกับพื้นที่ส่วนย่อย (Local Region) ซึ่งเป็นการ คำนวณแบบ dot product กับเคอร์เนลต่างๆ ที่มีขนาดเหมาะสมในการสกัดคุณลักษณะ (feature extraction) ของวัตถุที่อยู่ในภาพ ซึ่งจะมีการสุมและการปรับเปลี่ยนในขั้นตอนการเรียนรู้ก่อนจะออกมา เป็น แผนผังคุณลักษณะ (feature map) ซึ่งเป็นผลลัพธ์ที่ได้จากกระบวนการดังกล่าว ดังรูปที่ 2.5

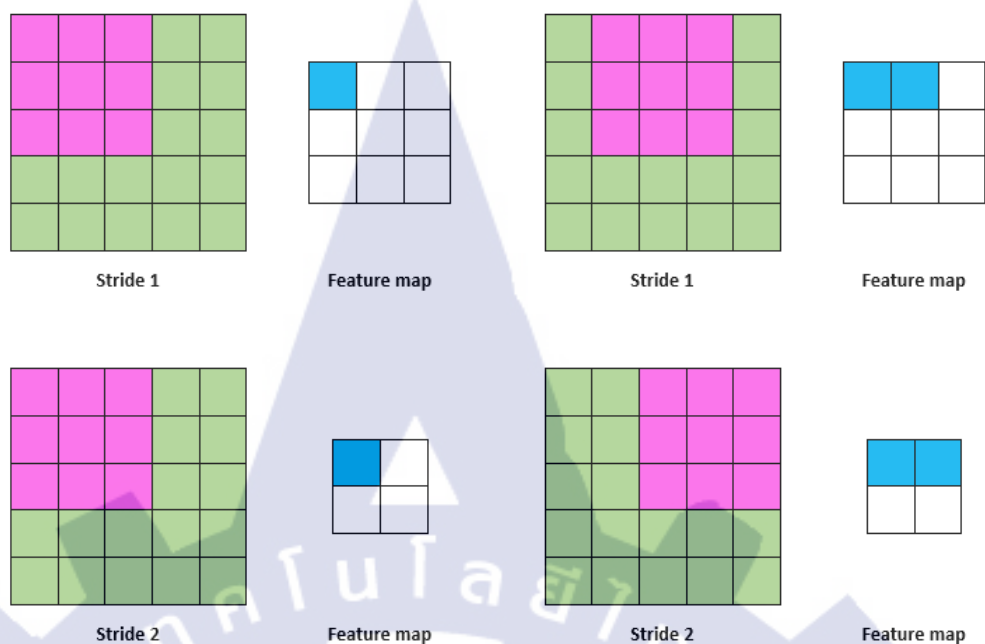
โดยขนาดอินพุต ($W_{input} \times H_{input} \times D_{input}$) ใน Convolutional Layer จะถูก กำหนดโดยความกว้าง ความสูงและ ความลึก ตามขนาดของรูปภาพอินพุต ซึ่งโดยปกติแล้ว ขนาด ความกว้าง (W_{input}) และความสูง (H_{input}) จะมีค่าเท่ากัน ส่วนของความลึกขนาดอินพุตจะมาจาก จำนวน Color depth ของรูปภาพแต่ละชนิด เช่นหากเป็นภาพสี RGB ความลึก (D_{input}) จะ เท่ากับ 3 แต่หากภาพเป็นภาพขาวดำความลึก (D_{input}) จะเท่ากับ 2 และในชั้น Hidden Layer ความลึก (D_{input}) จะเท่ากับจำนวนตัวกรองหรือ kernel ที่ใช้ในเลเยอร์ก่อนหน้า ส่วนพารามิเตอร์ สำหรับการสร้าง kernel ที่ใช้ในเลเยอร์ Convolution จะอธิบายในหัวข้อต่อไป

- จำนวนเคอร์เนล หรือ (K) คือ จำนวนตัวกรองที่ใช้ในการสกัดคุณลักษณะ ซึ่งส่งผลต่อความลึกของ feature map ตามจำนวน kernel ที่ทำการระบุหรือที่เรียกว่า Convolutional Layer ดังรูปภาพที่ 2.6



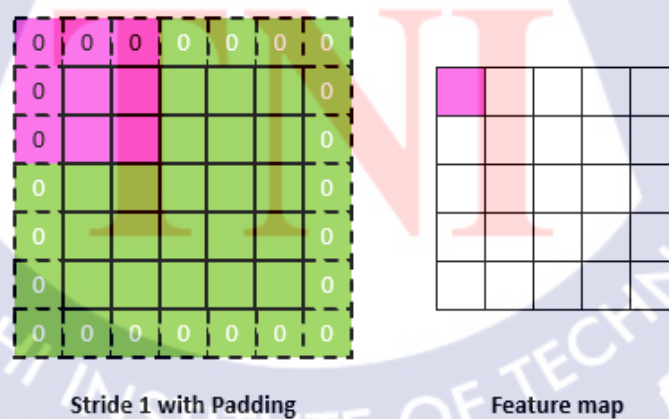
รูปที่ 2.6 อิทธิพลของจำนวน Kernel ในการทำ convoluion ต่อ feature map

- ขนาดเคอร์เนล (Kernel size) หรือ ($F \times F$) คือ ขนาดความกว้าง และความสูงของตัวเคอร์เนล ที่ใช้ในกระบวนการ convolution กับรูปภาพอินพุต เช่น เคอร์เนลขนาด 3×3 , 5×5 , 7×7 และ 11×11 เป็นต้น ตามความเหมาะสมในการใช้งาน
- สไตรด์ (Stride) หรือ (S) คือ จำนวนการเคลื่อนที่ของตัวเคอร์เนล ที่ใช้ในการทำ feature extraction กับรูปภาพอินพุตในกระบวนการ convolution ซึ่งสามารถกำหนดให้มากขึ้นได้ หากต้องการคำนวณหา feature ที่มีพื้นที่ทับซ้อนกันน้อยขึ้น แต่ก็ทำให้ feature map มีขนาดเล็กลงเช่นกัน ดังรูปภาพที่ 2.7



รูปที่ 2.7 ความแตกต่างของ feature map จากการทำ convolution ระหว่าง stride เท่ากับ 1 (บน) และ stride เท่ากับ 2 (ล่าง)

• ซีโรแพดดิ้ง (Zero-Padding) หรือ (P) คือการกำหนดให้มีการเติมค่า 0 เข้าไปที่บริเวณขอบของภาพอินพุตเพื่อเพิ่มขนาดของ feature map ดังรูปภาพที่ 2.8



รูปที่ 2.8 หลักการทำงานของ Zero-Padding

ซึ่งโดยปกติหากไม่มีการทำ padding ผลที่ได้จากการ convolution จะทำให้ได้ feature map มีขนาดเล็กกว่ารูปภาพอินพุต ดังนั้นเป้าหมายของการทำ padding คือการรักษาขนาดของ feature map ไม่ให้มีขนาดเล็กลงในขณะที่ทำการ feature extraction ซึ่งจำนวน padding สามารถหาได้จากสมการที่ 15 เมื่อต้องการรักษาขนาด feature map ให้เท่ากับภาพอินพุต โดย p คือ ขนาด pixel ของการทำ padding ในแต่ละด้าน และ f คือ ขนาดมิติตัวกรองที่ใช้ทำ convolution

$$p = \frac{f - 1}{2} \quad (2.14)$$

• เอาต์พุต (Output) หรือ แผนผังคุณลักษณะ (feature map) คือ ผลลัพธ์ที่ได้จากการทำ convolution มีขนาดเท่ากับ $(W_{output} \times H_{output} \times D_{output})$ ที่ได้จากการใช้ตัวกรอง kernel กับภาพอินพุต ผ่านกระบวนการ feature extraction ดังสมการที่ 16, 17 และ 18 ตามลำดับ

$$W_{output} = (W_{input} - F + 2P/S) + 1 \quad (2.15)$$

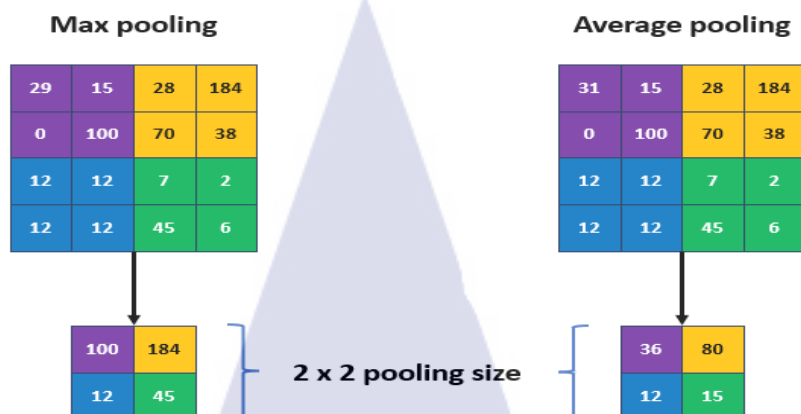
$$H_{output} = (H_{input} - F + 2P/S) + 1 \quad (2.16)$$

$$D_{output} = K \quad (2.17)$$

ผลที่ได้จากการทำ convolution หรือ feature map จะต้องผ่านฟังก์ชันกระตุ้น (activation function) เพื่อทำการปรับค่าให้เหมาะสม และส่งต่อไปเป็นอินพุตในเลเยอร์ถัดไป โดยฟังก์ชันกระตุ้นที่นิยมใช้ใน CNN มีหลายฟังก์ชัน ดังที่อธิบายในหัวข้อ 2.3

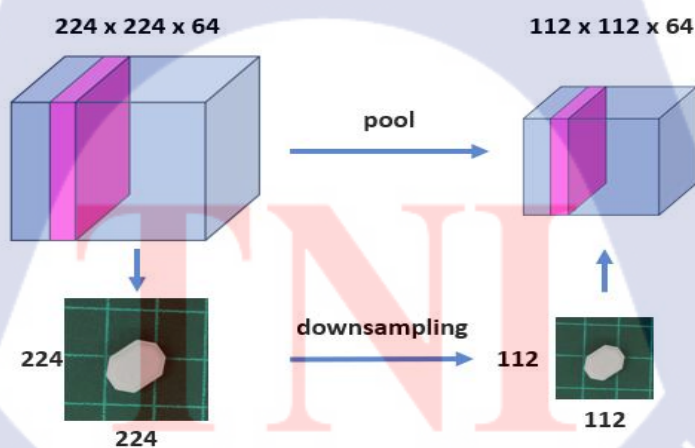
2.7.2 เลเยอร์พูลลิ่ง (Pooling layer)

Pooling layer คือ เลเยอร์ที่มีกว้างต่อจากเลเยอร์ Convolution โดยมีวัตถุประสงค์เพื่อลดขนาด feature map ที่ได้จากกระบวนการ convolution ให้มีขนาดเล็กลงโดยยังคงคุณลักษณะที่จำเป็นก่อนจะส่งไปยังเลเยอร์ถัดไป โดยมี 2 วิธีที่นิยมใช้คือ การพูลค่ามากที่สุด (Max Pooling) และ การพูลค่าเฉลี่ย (Average Pooling) ดังรูปที่ 2.9



รูปที่ 2.9 หลักการทำงานของ Max Pooling (ซ้าย) และ Average Pooling (ขวา) โดยใช้ขนาด pool size 2 x 2

ในการลดขนาดของ feature map จะเป็นการกำหนดขนาดของ pool size (F×F) และจำนวน stride (S) เพื่อให้ข้อมูลมีขนาดเล็กลงพร้อมกับลดพารามิเตอร์ไปพร้อมกันดังที่แสดงในรูป 2.10



รูปที่ 2.10 หลักการทำงานของพูลลิ่งเลเยอร์ (Pooling layer)

โดยขนาด feature map ($W_{input} \times H_{input} \times D_{input}$) ที่เข้าสู่กระบวนการ pooling ก่อนจะออกมาเป็นข้อมูลเอาต์พุตที่มีขนาดเล็กลง เพื่อนำไปใช้ในเอน์ดถัดไปจะมีขนาดเท่ากับ ($W_{output} \times H_{output} \times D_{output}$) ที่ได้จากสมการที่ 19, 20 และ 21 ตามลำดับ

$$W_{output} = ((W_{input} - F)/S) + 1 \quad (2.18)$$

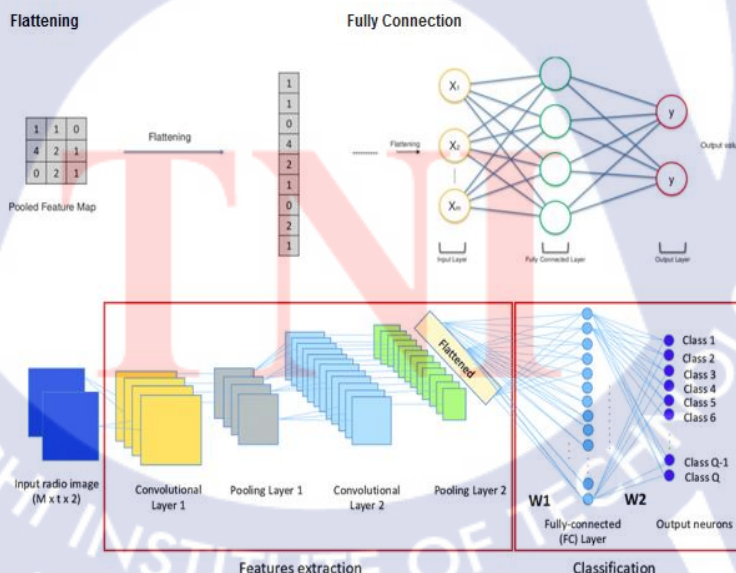
$$H_{output} = ((H_{input} - F)/S) + 1 \quad (2.19)$$

$$D_{output} = D_{input} \quad (2.20)$$

จุดประสงค์ของ Convolution layer และ Pooling layer คือ การสกัดคุณลักษณะ (feature extraction) จากภาพอินพุต โดยใน hidden layer ชั้นแรกๆ จะได้ feature ระดับล่าง เช่น เส้น ส่วนในชั้น hidden layer ที่ลึกขึ้นจะเป็นการ feature extraction ที่ซับซ้อนขึ้น ซึ่งจำนวนชั้น Convolution layer และ Pooling layer ที่เพิ่มขึ้นจะเป็นตัวกำหนดความลึกของโครงข่ายประสาทเทียมที่เพิ่มมากขึ้นตามจำนวนชั้นของเลเยอร์ทั้ง 2 [5]

2.7.3 เลเยอร์เชื่อมต่อกันแบบสมบูรณ์ (Fully-Connected Layer)

Fully-Connected Layer คือ เลเยอร์ที่เชื่อมโยงระหว่าง feature map และตัวเอาต์พุตซึ่งอาจมี 1 หรือ 2 เลเยอร์ ก่อนจะส่งข้อมูลเข้าสู่ฟังก์ชัน Softmax โดยจะอยู่ตรงส่วนท้ายของตัวโครงข่ายซึ่งหมายถึงทุกๆ neuron ที่อยู่ในเลเยอร์สุดท้ายของ feature map จะถูกนำไปปรับขนาดใหม่ (Reshape) หรือที่เรียกว่า Flatten ให้กลายเป็นคอลัมน์เดี่ยวเพื่อส่งไปคำนวณต่อในเลเยอร์ Fully-Connected Layer เพื่อให้สามารถวิเคราะห์ข้อมูลได้อย่างความสะดวกดังรูปภาพที่ 2.11



รูปที่ 2.11 หลักการทำงานของ Fully-Connected Layer ของโครงข่ายประสาทเทียม
ใน ขั้นตอน classification

ในการทำ Full Connection คือ การนำข้อมูลเอาต์พุตที่ผ่านการ Flattening จากชั้นตอนก่อนหน้าเข้าสู่โมเดลเชิงลึก เพื่อทำการประมวลผลพร้อมกำหนดข้อมูลนำออก (output) ในเอาต์พุตเลเยอร์ เพื่อทำการ prediction ตามจำนวนผลลัพธ์ที่กำหนดไว้

2.8 ฟังก์ชันซอฟต์แวร์แมกซ์ (Softmax function) [5]

ฟังก์ชัน Softmax คือ ฟังก์ชันที่ถูกวางไว้ในเลเยอร์ชั้นสุดท้ายเพื่อหาความน่าจะเป็นของแต่ละคลาสที่ถูกกำหนด โดยคลาสที่มีความน่าจะเป็นสูงสุด คือ ผลจากการ prediction โดยดูจากคะแนนความน่าจะเป็นของแต่ละคลาสที่ได้จากสมการที่ 22

$$\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_{k=1}^K \exp(z_k)} \quad (2.21)$$

โดยค่า z_i คือคะแนน logit score ในเวกเตอร์ z ซึ่งความน่าจะเป็นของแต่ละคลาสสามารถหาได้จาก exponential ของ logit หารด้วยผลรวมของ exponential ของ logit ทุกตัวก่อนจะออกมาเป็นคะแนนความน่าจะเป็นของแต่ละคลาส โดยมีผลรวมเท่ากับ 1

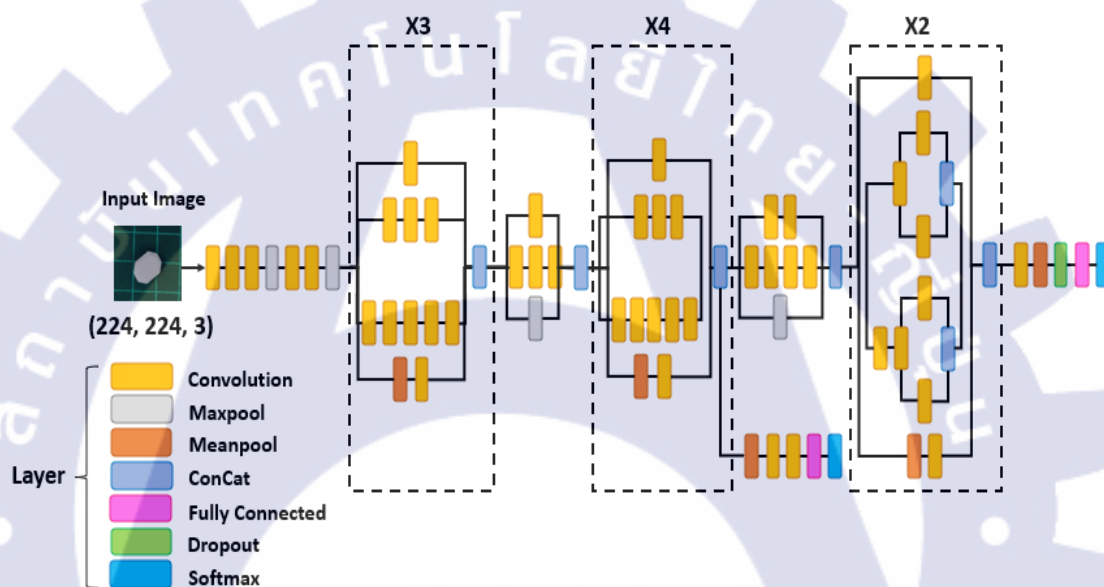
2.9 การเรียนรู้เชิงถ่ายโอนข้อมูล (Transfer Learning)

โดยทั่วไปการจำแนกรูปภาพโดยใช้แบบจำลองการเรียนรู้เชิงลึกมักจะประสบปัญหาเรื่องการใช้เวลาในการ training ที่นานเกินไปนับตั้งแต่ต้นจนจบกระบวนการ อันเนื่องมาจากความซับซ้อนและความลึกของตัวโมเดลตามจำนวนชั้นเลเยอร์ที่เพิ่มขึ้นพร้อมกับพารามิเตอร์ และค่าถ่วงน้ำหนักที่มีอาจจำนวนมากมายหลายล้านจนจำเป็นต้องใช้พลังในการประมวลผลมหาศาล พร้อมกับชุดข้อมูลขนาดใหญ่ และเวลาในการประมวลผลที่อาจกินเวลาหลายวัน หรืออาจกินหลายสัปดาห์

การเรียนรู้เชิงถ่ายโอนข้อมูล หรือ Transfer Learning คือ เทคนิคที่ช่วยลดเวลาในการฝึก โดยการนำส่วนที่ใช้ในการประมวลผลของสถาปัตยกรรมโมเดลที่เคยผ่านการ pre-train กับงานที่ใกล้เคียงกัน มาใช้เป็นส่วนหนึ่งของโมเดลใหม่ โดยไม่จำเป็นต้อง random initialize ค่า weight และสามารถเรียนรู้จากข้อมูลปริมาณไม่มากได้ เมื่อเทียบกับการเรียนรู้ตั้งแต่ศูนย์ ซึ่งก็ขึ้นอยู่กับความแตกต่างกันระหว่าง domain ของข้อมูลที่ทำการศึกษาและ domain ที่ต้องการเรียนรู้

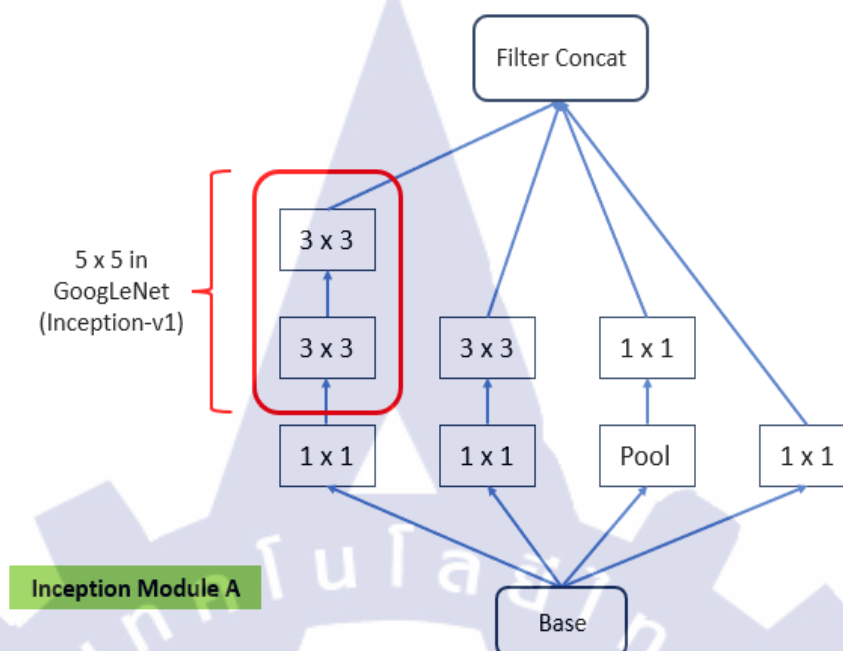
2.9.1 Inception-V3

Inception-V3 คือ สถาปัตยกรรมโครงข่ายประสาทเทียมที่ใช้ในการ Image analysis และ Object detection ได้รับการพัฒนาโดย Google เปิดตัวครั้งแรกในงาน ImageNet Recognition Challenge โดยเป็นโครงข่ายประสาทเทียมที่ถูกต่อยอดมาจาก Inception-V2 ที่ลดโครงสร้างภายในออกเป็น 5 Step ได้แก่ Inception Module A จำนวน 5 Module , Grid Size of Reduction Step1 จำนวน 1 Module, Inception Module B จำนวน 4 Module, Grid Size of Reduction Step2 จำนวน 1 Module, Inception Module C จำนวน 2 Module พร้อมกับส่วน Head (8x8x2048) ที่สามารถแยก output ได้ถึง 1,000 คลาส ดังรูปที่ 2.12



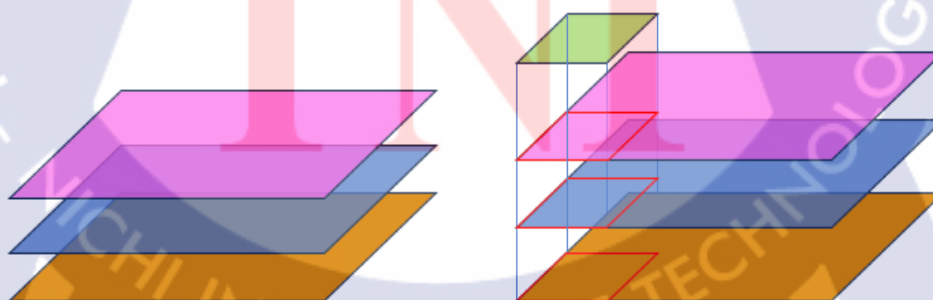
รูปที่ 2.12 โครงสร้างสถาปัตยกรรมโครงข่ายประสาทเทียม Inception-v3

ทำให้โมเดลมีพารามิเตอร์ลดลงจากเดิมแต่ก็ยังคงมีประสิทธิภาพ โดยการปรับปรุง convolutions และ maxpooling ที่จากเดิม 5x5 เหลือ 3x3 pixels และ 3x3 ลดลงเหลือ 2x2 pixels ตามลำดับ โดยวัตถุประสงค์ในการออกแบบ 1x1 convolution คือ ต้องการให้ขนาด (shape) ข้อมูลขาออก (outputs) เป็น tensor เช่นกำหนด (N, F, H, W) ซึ่ง N คือ batch size, F คือ จำนวนของ convolution filters ส่วน H, W คือ spatial ของมิติ ดังรูปที่ 2.13



รูปที่ 2.13 โครงสร้างโมดูล Inception ที่ประกอบด้วยเลเยอร์ convolution และ maxpooling

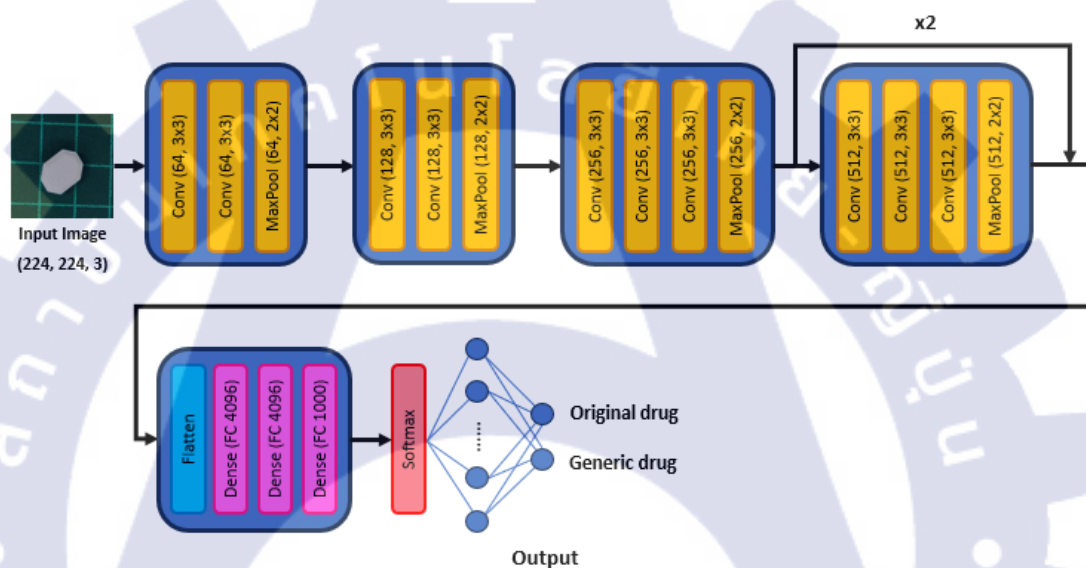
ภายในโมดูล inception กรณีที่ข้อมูลนำเข้าเป็นภาพ RGB เราสามารถเปรียบเทียบได้กับการทำ feature map ดังรูปด้านซ้าย ซึ่งสมมุติว่า ในกรณีที่ต้องการผลลัพธ์ feature map เพียง 1 คำนั้ถูกทำได้โดยใช้ convolution ขนาด 1x1 pixels ที่มีการขยับที่ 1 stride ดังรูปที่ 2.14 โดยค่า weight ที่คูณในแต่ละชั้นของ RGB นั้นจะมีค่าที่ไม่เท่ากัน ส่งผลให้ค่าที่ได้ในแต่ละพิกเซล ของ feature map หลังการทำ convolution นี้จะมีที่เป็นลักษณะเฉพาะ



รูปที่ 2.14 การทำ feature map กับรูปภาพ RGB โดยใช้เลเยอร์ convolution ขนาด 1x1 pixels

2.9.2 VGG16

VGG16 หรือ Visual Geometry Group 16 คือ สถาปัตยกรรมโครงข่ายประสาทเทียมที่ถูกพัฒนาขึ้นโดยกลุ่มนักวิจัยจากมหาวิทยาลัยออกซฟอร์ด (Oxford University) ซึ่งได้รับความสนใจมาจากการแข่งขัน ILSVR ในปี 2014 และเป็นที่นิยมจนถึงปัจจุบัน โดยมีจุดเด่นคือ การแทนที่ hyperparameter จำนวนมากที่เน้นไปที่การออกแบบ เลเยอร์ conv2D ขนาด 3x3 pixels พร้อมกำหนด stride เท่ากับ 1 และการใช้ same padding และ maxpooling ขนาด 2x2 pixels พร้อมกำหนด stride เท่ากับ 2 แบบเดียวกันตลอดทั้งโครงสร้าง โดยชื่อของ VGG16 หมายถึงโมเดลที่มี 16 ชั้นเลเยอร์พร้อมกับค่าน้ำหนัก ซึ่งเป็นเครือข่ายขนาดใหญ่ และมีพารามิเตอร์ประมาณ 138 ล้าน ดังรูปที่ 2.15

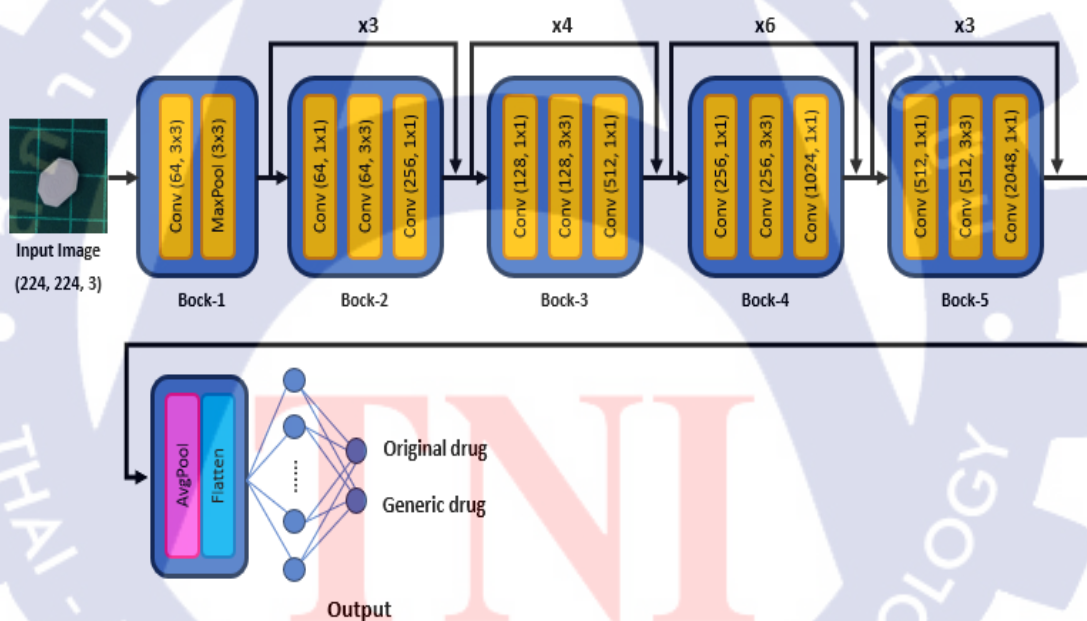


รูปที่ 2.15 ภาพโครงสร้างสถาปัตยกรรมโครงข่ายประสาทเทียม VGG16

2.9.3 ResNet50-v2

ResNet50-v2 คือ สถาปัตยกรรมโครงข่ายประสาทเทียมที่ถูกพัฒนาต่อยอดมาจาก ResNet50 มีเลเยอร์ทั้งหมด 50 เลเยอร์ โดยถูกแบ่งออกเป็น 4 Block ได้แก่ conv 2, conv 3, conv 4 และ conv 5 ตามลำดับดังรูปที่ 2.16 ซึ่งภายใน Block แต่ละตัวจะประกอบไปด้วย เลเยอร์ Convolution จำนวน 3, 4, 6 และ 3 ชั้น ตามลำดับซึ่งรวมถึงชั้นเลเยอร์ Convolutional ที่ติดกับข้อมูลนำเข้า (input) และข้อมูลนำออก (output) เพื่อใช้ในการดึงคุณลักษณะทำให้มิติภาพสูงขึ้นรวมเป็น 2 ชั้น โดยชื่อของ ResNet50 หมายถึง จำนวนเลเยอร์ทั้งหมด 50 ชั้นที่อยู่ในโครงข่ายประสาทเทียม โดยมีจำนวนเลเยอร์ทั้งสิ้น $(3 + 4 + 6 + 3)$ คูณ 3 เท่ากับ 48 และบวกกับอีก 2 ชั้นเลเยอร์นำเข้า และเลเยอร์นำออก จึงเท่ากับเลเยอร์ 50 ชั้น

โดยมีวัตถุประสงค์การออกแบบ คือการแก้ปัญหา vanishing gradient ซึ่งมักเกิดขึ้นกับโครงข่ายประสาทเทียมที่มีความลึกค่อนข้างมาก ซึ่งมีจำนวนชั้นเลเยอร์มากถึง 152 เลเยอร์ หรือเท่ากับ 8 เท่าของ VGG16 โดยใช้เทคนิคการออกแบบ module ที่มีลักษณะเป็นทางลัดลง



รูปที่ 2.16 สถาปัตยกรรมโมเดล ResNet50-v2

2.10 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการใช้โครงข่ายประสาทเทียมแบบ Convolution หรือ Deep learning มาประยุกต์เป็นเครื่องมือจำแนกวัตถุประเภทยาเม็ด โดยมีการนำเสนอวิธีการ และผลลัพธ์ เพื่อใช้เป็นแนวอ้างอิงปรับปรุงวิธีการทดลองให้มีประสิทธิภาพมากยิ่งขึ้นสามารถสรุปได้ในตารางที่ 2.1

ตารางที่ 2.1 สรุปผลงานวิจัยที่เกี่ยวข้อง

คณะผู้วิจัย	หัวข้อวิจัย	วิธีการวิจัย
E. A. Adewumi [8]	A Convolutional Neural Network (CNN) based Pill Image Retrieval System	ได้นำเสนอโมเดล CNN การดึงข้อมูลภาพยาเม็ดแบบไฮบริดโดยใช้วิธี supervised ที่มีเทคนิคการจัดกลุ่ม k-mean แบบไม่เชิงเส้น แล้วยกประเภท kNN เพื่อแก้ปัญหการจำแนกประเภทและการถดถอย
Y. Ou et al. [10]	Automatic Drug Pills Detection based on Convolution Neural Network	นำเสนอระบบ visual system สำหรับตรวจจับเม็ดยาอัดโนมิตี ซึ่งประกอบด้วย 2 ขั้นตอนหลัก คือ การตรวจจับ และจัดหมวดหมู่โดยใช้ภาพยาเม็ด 1,680 สำหรับการระบุตำแหน่ง 470,000 ภาพสำหรับการจัดประเภทและอีก 400 ภาพสำหรับการตรวจสอบ
Y. Ou et al. [11]	Automatic drug pills detection based on enhanced feature pyramid network and convolution neural networks	นำเสนอเครื่องตรวจจับเม็ดยาที่ประกอบด้วย โครงข่ายพีระมิดคุณลักษณะ CNN ที่ได้รับการปรับแต่ง (EFPN) ที่ใช้ในการระบุตำแหน่ง และโมเดล Inception-ResNetv2 สำหรับการจำแนก โดยใช้ชุดข้อมูลภาพยาทั้งหมด 612 ชนิด
J. J. Caban et al. [12]	Automatic identification of prescription drugs using shape distribution models	นำเสนอเทคนิคการจำแนกยาตามใบสั่งแพทย์จากภาพเม็ดยาอัดโนมิตีทั้ง 568 ชนิดที่ใช้ในสหรัฐ โดยใช้เทคนิคการกระจายรูปร่างที่ได้รับการดัดแปลงเพื่อตรวจสอบรูปร่าง สี รอยประทับ พร้อมสร้างตัวบอกค่าคงที่ที่สามารถ recognize ยาชนิดเดียวกันในมุมมองอื่น

ตารางที่ 2.1 สรุปผลงานวิจัยที่เกี่ยวข้อง (ต่อ)

คณะผู้วิจัย	หัวข้อวิจัย	วิธีการวิจัย
S. Suntronsuk และ S. Ratanotayanon [13]	Automatic Text Imprint Analysis from Pill Images	นำเสนอวิธีการสกัดข้อความรอยประทับจากเม็ดยา โดยอาศัยชุดกฎเกณฑ์ (set or rules) เพื่อหาตำแหน่งรอยประทับและเทคนิคการจัดสัญญาณรบกวนเพื่อลบพิกเซลที่มีปัญหาจากภาพไบนารีที่ได้จาก OTSU's thresholding
L. Tan et al. [14]	Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification	ได้ทำการทดลองเปรียบเทียบประสิทธิภาพ ระหว่าง object detection โมเดลทั้ง 3 ได้แก่ RetinaNet, SSD และ YOLO v3 ในการตรวจจับยาเม็ดแบบเรียลไทม์ เพื่อหาโมเดลที่ดีที่สุด โดยทำการ train แต่ละอัลกอริทึมกับชุดข้อมูลภาพยาเม็ด
L. Tan et al. [15]	Comparison of YOLO v3, Faster R-CNN, and SSD for Real-Time Pill Identification	ได้ทำการทดลองเปรียบเทียบประสิทธิภาพเพื่อหาโมเดลที่ดีที่สุดระหว่าง object detection โมเดลทั้ง 3 ได้แก่ YOLO v3, Faster R-CNN และ SSD ในการตรวจจับ และระบุยาเม็ดแบบเรียลไทม์ โดยทำการ train แต่ละอัลกอริทึมด้วยชุดข้อมูลภาพเม็ดยาทั้งหมด
Y. F. Wong et al. [16]	Development of fine-grained pill identification algorithm using deep convolutional network	พัฒนาวิธีการระบุและตรวจสอบยาเม็ดอัตโนมัติ ด้วย Deep Convolutional Network หรือ DCN โดยใช้ภาพเม็ดยาในมุมมองต่างๆ ที่ถ่ายจากโทรศัพท์มือถือในสภาพแวดล้อมที่แตกต่าง พร้อมทำการเปรียบเทียบกับ 2 วิธีพื้นฐานที่มี
N. Usuyama et al. [17]	ePillID Dataset: A Low-Shot Fine-Grained Benchmark for Pill Identification	ทดสอบ ePillID ซึ่งเป็นเกณฑ์มาตรฐานความละเอียดต่ำ ที่ใหญ่ที่สุดและเป็นสาธารณะในการทำ image recognition ซึ่งประกอบด้วยชุดข้อมูลภาพเม็ดยา 8,184 ภาพจากยาทั้งหมด 4,092 ชนิดที่ถ่ายไว้ 2 ด้าน เพื่อนำไปใช้พัฒนาระบบการระบุตัวยาแบบ One-Short โดยการประเมินแบบจำลอง

ตารางที่ 2.1 สรุปผลงานวิจัยที่เกี่ยวข้อง (ต่อ)

คณะผู้วิจัย	หัวข้อวิจัย	วิธีการวิจัย
		พื้นฐาน baseline ต่างๆ ด้วยเกณฑ์เปรียบเทียบเชิงประจักษ์
P. Chpawa และ K. Kanjanawanishkul [18]	Pill Identification with Imprints Using a Neural Network	พัฒนาฐานข้อมูลยาที่ใช้กันทั่วไปในประเทศไทย พร้อมสร้างระบบช่วยเหลือเภสัชกรในการระบุยาเม็ดที่ไม่รู้จักแบบเรียลไทม์ซึ่งจะเน้นไปที่การใช้รอยประทับบนยาเม็ดที่มีสีรูปร่างที่เกือบเหมือนกัน และนำเสนออัลกอริทึมแบบใหม่เพื่อใช้แยก feature vector ออกจากตัวรอยประทับโดยไม่มี ความแปรผันต่อการหมุน ก่อนป้อนเข้าสู่โครงข่ายประสาทเทียมโดยจำกัดจำนวนเม็ดยาไว้ที่ 6 ชนิด
S. Ling et al. [19]	Few-Shot Pill Recognition	นำเสนอชุดข้อมูลภาพยา CURE แบบใหม่ที่มี ตัวอย่างต่อคลาสเพิ่มขึ้น, W2-net สำหรับจัดการ ปัญหาการ recognition ยาเม็ดไม่กี่ครั้งและ สถาปัตยกรรมเชิงลึกแบบ multi-stream พร้อมกล ยุทธ์การเรียนรู้แบบ 2 stage เพื่อใช้ประโยชน์จาก ข้อมูลที่เกี่ยวข้องกับโดเมนในกรณีที่มีข้อมูลมีขนาด เล็กได้ดียิ่งขึ้น
U. Patel [20]	Machine Learning-based Pharmaceutical Tablet Inspection and Recognition Techniques – A Review	ทบทวนวิเคราะห์บทความและวิธีการต่างๆ ที่ใช้ในการวิเคราะห์ตัวยา Tablet ผ่าน deep learning และ machine learning เพื่อตรวจสอบและระบุตัวยา โดยอ้างอิงจากพารามิเตอร์ต่างๆ พร้อมทั้งทำ การเปรียบเทียบ ระหว่างเทคนิค machine learning กับ deep learning สมัยใหม่ และวิธีการ วิเคราะห์ต่างๆ เพื่อหาการประยุกต์ใช้งานใน อุตสาหกรรม หรือการตรวจสอบยาโดยผู้ป่วย

ตารางที่ 2.1 สรุปผลงานวิจัยที่เกี่ยวข้อง (ต่อ)

คณะผู้วิจัย	หัวข้อวิจัย	วิธีการวิจัย
X. Zeng et al. [21]	MobileDeepPill: A Small-Footprint Mobile Deep Learning System for Recognizing Unconstrained Pill Images	นำเสนอระบบจดจำภาพเม็ดยาบนมือถืออัตโนมัติ (MobileDeepPill) ซึ่งเป็น vision system mobile ขนาดเล็กที่ประกอบด้วยสถาปัตยกรรม multi-CNNs แบบใหม่ที่ได้รับการออกแบบพร้อม triplet loss function และ เฟรมเวิร์กการบีบอัดโมเดลเชิงลึกที่อิงการกลั่นความรู้ ในการจดจำภาพยาเม็ดแบบไม่มีข้อจำกัดมาใช้ในการทดลอง กับชุดข้อมูลรูปภาพเม็ดยาจากหอสมุดแพทยศาสตร์แห่งชาติ สหรัฐ
L. Sousa et al. [22]	Pill Image Classification using Machine Learning	นำเสนอระบบการจำแนกประเภทยาเม็ดอัตโนมัติที่อิงจากรูปร่าง และสีของเม็ดยาที่อยู่ในภาพ โดยใช้เทคนิคการประมวลผลภาพในการระบุ กับชุดข้อมูลย่อยของชุดข้อมูล NLM PIR ที่จัดทำโดยหอสมุดแพทยศาสตร์แห่งชาติพร้อมชุดแอตทริบิวต์ SVM และตัวแยกประเภท Multilayer Perceptron classifier
A. Mehmood et al. [23]	MobilePill: Accurate Pill Image Classification via Deep Learning on Mobile	ได้พัฒนาแนวทางการใช้โมเดล CNN แบบไฮบริดในการจำแนกภาพเม็ดยาผ่านโทรศัพท์มือถือโดยใช้วิธีการจำแนกคุณลักษณะ (feature) ที่แตกต่างกันของรูปยาเม็ดที่มีความคล้ายคลึงกัน ซึ่งโมเดล CNN แบบ hybrid จะประกอบด้วย โมเดลอิสระแต่ละแบบที่แตกต่างกันพร้อมชุดข้อมูลการฝึกฝน (training set) ที่มีความหลากหลาย ตามรูปแบบเภสัชภัณฑ์ (dosage forms) ของแต่ละคนใช้ เพื่อแก้ไขปัญหาความแม่นยำ และความลำบากในการให้ข้อมูลยาจากรูปผ่านโทรศัพท์มือถือ

ตารางที่ 2.1 สรุปผลงานวิจัยที่เกี่ยวข้อง (ต่อ)

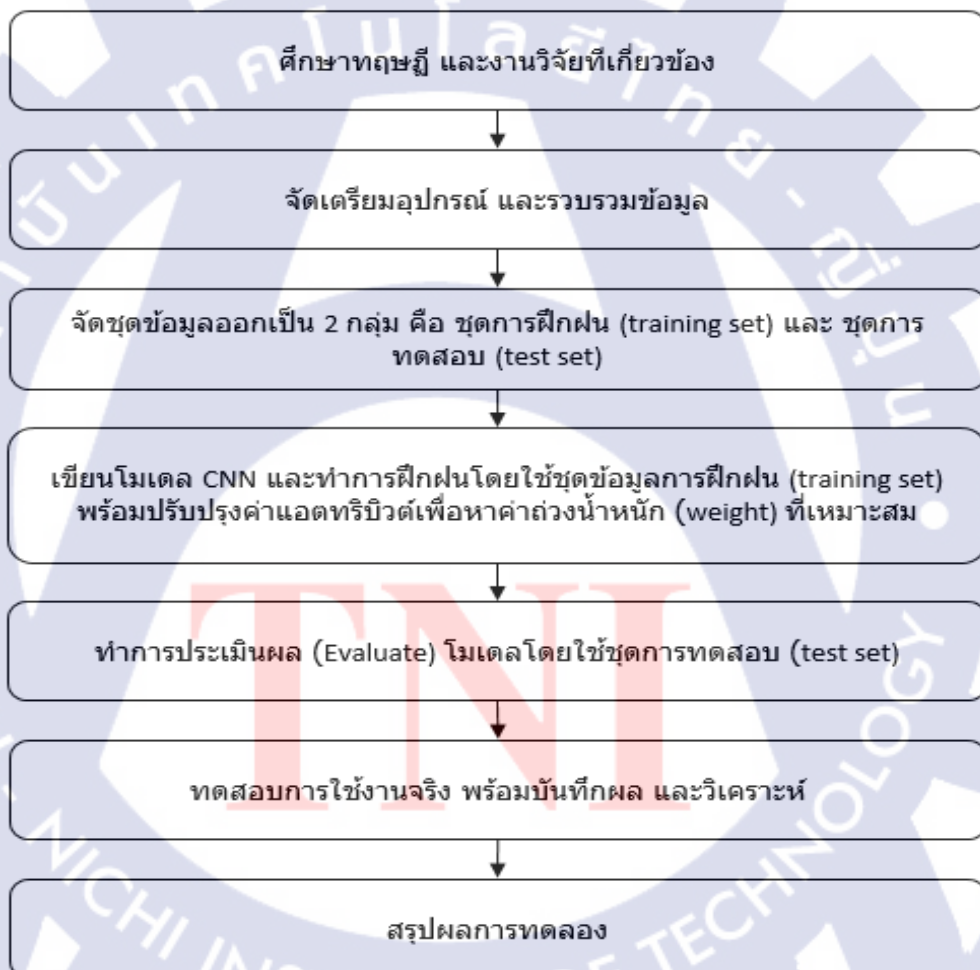
คณะผู้วิจัย	หัวข้อวิจัย	วิธีการวิจัย
Y. Wang et al [24]	Pill Recognition Using Minimal Labeled Data	ได้นำเสนอวิธีการระบุยาเม็ดอัตโนมัติ CNN แบบใหม่จากภาพแบบเดียวกับชุดข้อมูลสาธารณะโดยสถาบันสุขภาพแห่งชาติซึ่งมียามากถึง 1,000 ชนิด โดยขั้นตอนที่ 1 จะเป็นการจำกัดวง และระบุตำแหน่งตัวยาในภาพโดยการตรวจจับบริเวณที่มีความเข้มขอบสูงสุดเพื่อแก้ปัญหาชุดข้อมูลการฝึกฝน (training set) พร้อมป้ายกำกับ (label) ที่มีอยู่น้อย ส่วนขั้นตอนที่ 2 จะเป็นการเปลี่ยน domain จากภาพที่ใช้ในการ train ซึ่งถ่ายในห้องปฏิบัติการไปเป็นภาพผู้บริโภคที่ถ่ายภายใต้สภาพธรรมชาติที่ไม่มีการควบคุม โดยมีเทคนิคการเสริมข้อมูลในการสร้างภาพยาเม็ดสังเคราะห์ที่ใช้การฝึก และมี GoogLeNet Inception Network เป็นตัวแยกประเภทหลัก
W. Swastika et al. [25]	Preliminary Study of Multi Convolution Neural Network-Based Model To Identify Pills Image Using Classification Rules	พัฒนาโมเดล multi CNN สำหรับการระบุยาเม็ดที่ประกอบด้วยโมเดล CNN ทั้ง 3 ที่ได้รับการพัฒนาโดยใช้กฎการจำแนกเพื่อใช้ในการระบุคุณลักษณะหลักของยาเม็ด 3 อย่าง ได้แก่ สี, รูปร่าง และรอยประทับ โดยใช้ชุดข้อมูลภาพยาทั้ง 24,000 แบ่งเป็น training set 95% และ test set 5% พร้อม Optimizers Adadelata และขนาดอินพุต 64x64 ซึ่งผลลัพธ์ของแต่ละโมเดลจะถูกประมวลผลด้วยกฎที่กำหนดไว้ล่วงหน้าเพื่อสร้างคลาสของยาเม็ด ยาเม็ดแต่ละเม็ด

บทที่ 3

ระเบียบวิธีการวิจัย

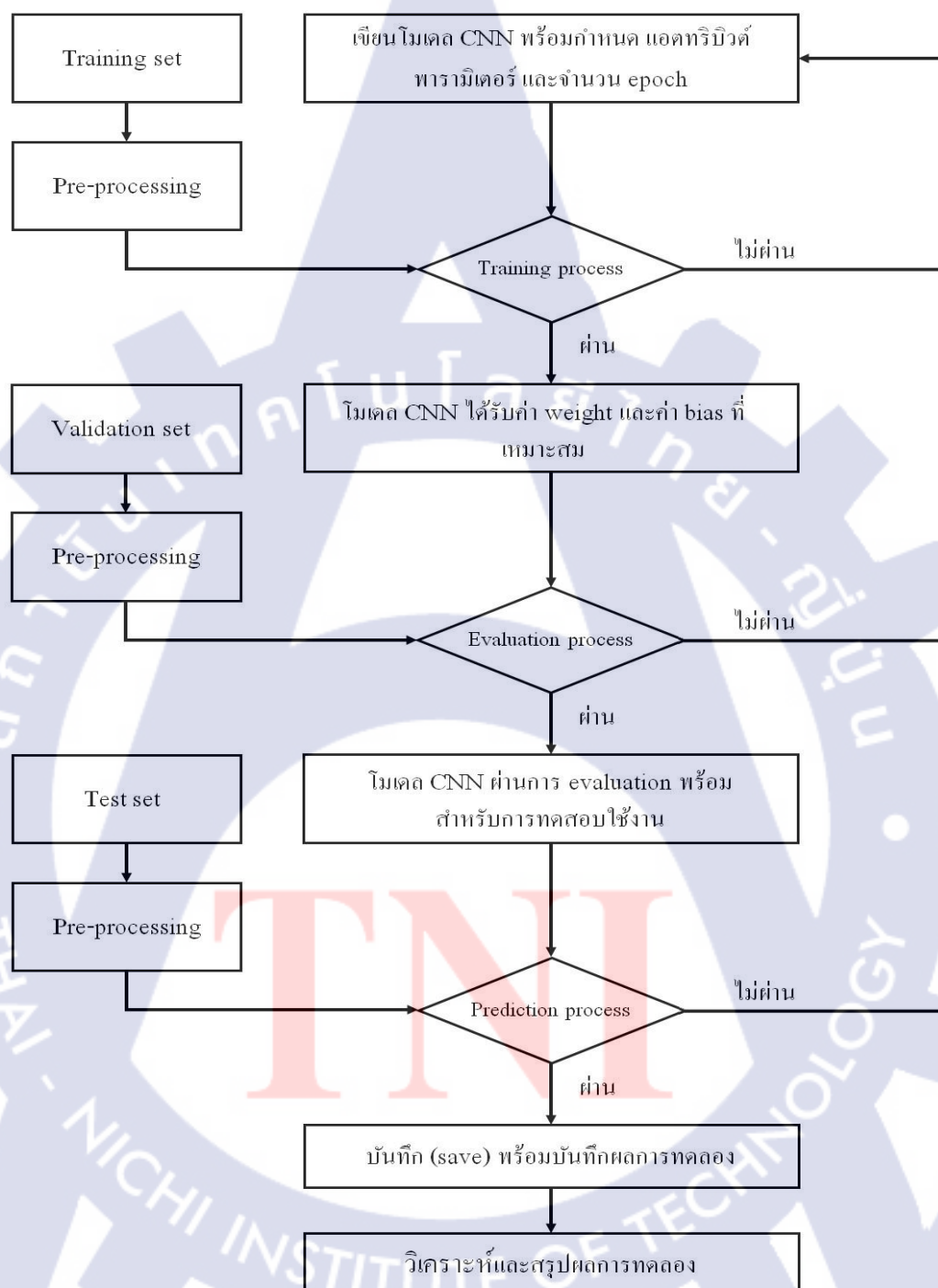
3.1 ขั้นตอนการดำเนินงาน

ในเนื้อหาบทนี้กล่าวถึงระเบียบวิธีวิจัยเกี่ยวกับการดำเนินการงานทดลอง โดยเริ่มจากการศึกษาทฤษฎี และงานวิจัยที่เกี่ยวข้อง จากนั้นจึงเตรียมอุปกรณ์ รวบรวมข้อมูล ทำการจัดกลุ่มก่อนสร้างโมเดล และทำการฝึกฝน พร้อมกับวัดประสิทธิภาพในการทดสอบ และบันทึกผลตามลำดับดังแผนผังการดำเนินงานวิจัยที่แสดงในรูป 3.1



รูปที่ 3.1 แผนผังขั้นตอนการดำเนินงาน

3.2 กรอบแนวคิดการวิจัย



รูปที่ 3.2 กรอบแนวคิดการวิจัย

3.3 ขั้นตอนการทดลอง

จากรูปภาพที่ 3.2 ขั้นตอนการดำเนินงานการทดลอง จากกรอบแนวคิดการวิจัยจะถูกอธิบายดังต่อไปนี้

1. ชุดข้อมูลรูปภาพทั้งยาต้นแบบ และยาสามัญที่ถ่ายจากโทรศัพท์มือถือจะถูกรวบรวมจัดกลุ่ม และทำการจัดประเภท (labeling) ล่วงหน้า เพื่อความสะดวกในการใช้งาน ก่อนจะแบ่งออกเป็น 2 ชุดคือ ชุดข้อมูล training set และ test set เพื่อใช้ในการฝึกฝน (training) และประเมินผล (evaluate) ตามลำดับ

2. เขียนโมเดล CNN ที่ใช้ในการทดลอง และโมเดล CNN อีก 3 ชนิดที่ใช้ในการเปรียบเทียบกับเทคนิค transfer learning ซึ่งสร้างจากสถาปัตยกรรม CNN ที่แตกต่างกันผ่านภาษา python บนโปรแกรม Pycharm โดยใช้ไลบรารี Keras และ Tensorflow พร้อมกำหนดขนาดอินพุต (224, 224, 3) และจำนวน kernel ใน Dense (128) โดยมีฟังก์ชันการกระตุ้น Relu, Sigmoid และ adam เป็น optimizer เพื่อใช้ในกระบวนการ convolution และ classification ตามลำดับพร้อมกำหนดจำนวน epoch หรือจำนวนรอบที่ใช้ในการ training

3. ทำการประมวลผลข้อมูลล่วงหน้า (Pre-processing) ทั้งชุดข้อมูล training set และ validation set โดยใช้คำสั่ง ImageDataGenerator จากไลบรารี Keras เพื่อทำการ Normalization โดยการปรับขนาด (rescale) รูปภาพใหม่หาค่าพิกเซลข้อมูลทั้งหมดด้วย 255 และใช้ flow_from_directory เพื่อหาตำแหน่งไฟล์ที่เก็บข้อมูล พร้อมทำการกำหนด target_size, batch_size ขนาด 64 และ class_mode เพื่อเตรียมข้อมูลให้พร้อมก่อนป้อนเข้าสู่โมเดล

4. ป้อนชุดข้อมูล training set เข้าสู่โมเดล CNN เพื่อทำการ training ตามจำนวน epoch ที่ได้กำหนดเพื่อให้โมเดลทำการเรียนรู้ และหาค่า weight และ bias ที่เหมาะสมที่ใช้ในการจำแนก แต่หากเกิดความผิดพลาด (Error) เช่น เกิดปัญหา overfitting หรือใช้เวลา train นานเกินไปจะถือว่าการทดลองไม่ผ่าน และทำการหยุดกระบวนการ พร้อมจดค่าพารามิเตอร์ แล้วกลับไปแก้ไขในขั้นตอนที่ 2 และ 3 อีกครั้ง ก่อนจะดำเนินการตามข้อ 4

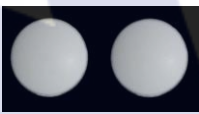
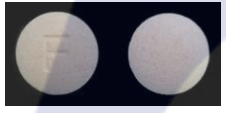
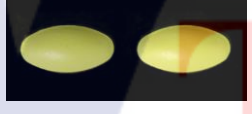
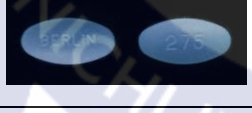
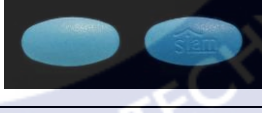
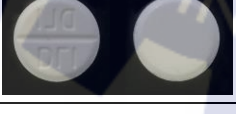
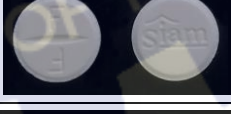
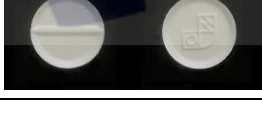
5. หากโมเดลผ่านการทดสอบในข้อที่ 4 ให้ทำการ evaluate โดยใช้ชุดข้อมูล validation set ที่ผ่านกระบวนการ Pre-processing เพื่อทำการวัดประสิทธิภาพ ซึ่งหากเกิดความผิดพลาด (Error) หรือ ค่า accuracy ไม่ดีเท่าที่ควร ให้กลับไปแก้ไขในขั้นตอนที่ 2 และ 3 แล้วทำเหมือนขั้นตอนก่อนหน้านี้

6. บันทึก (Save) ตัวโมเดล และทำการทดสอบประสิทธิภาพในการจำแนก ผ่านการใช้งานจริง โดยการนำภาพเม็ดยาทั้ง 2 ชนิดจำนวนหนึ่ง มาใช้ในการ prediction พร้อมดูผลลัพธ์ในการจำแนก

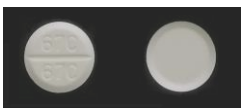
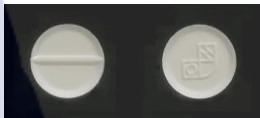
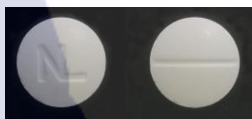
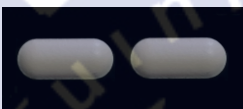
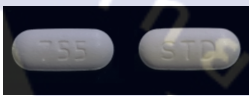
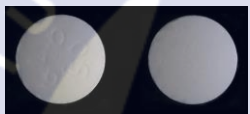
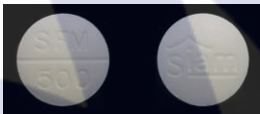
3.4 การเตรียมชุดข้อมูล (Data Processing)

ในส่วนการเตรียมชุดข้อมูล ชุดข้อมูลที่นำมาใช้ในงานทดลองประกอบไปด้วยรูปภาพเม็ดยาที่มีขายตามร้านยาทั่วไป หรือสถานพยาบาลเอกชนต่างๆ ทั้งหมด 6,240 ภาพ จากยาทั้งหมด 26 ตัว โดยแบ่งออกเป็น 2 หมวดหมู่ คือ ยาต้นแบบและ ยาสามัญอย่างละ 13 ตัว หรือคิดเป็น 240 ภาพ ตามการจัดคู่ดังตารางที่ 3.1

ตารางที่ 3.1 การเปรียบเทียบการจับคู่ของยาทั้ง 2 ชนิดระหว่างยาต้นแบบ และยาสามัญ

ตารางการเปรียบเทียบลักษณะรูปร่างจากการจับคู่ยาทั้ง 2 ชนิดที่ใช้ในงานทดลอง				
คู่ที่	ยาต้นแบบ (Original drugs)		ยาสามัญ (Generic drugs)	
1		NORVASC		Amlopine
2		Dramamine		Dimenhydrinate
3		Methycobal		Methyl-B
4		Flemex		Muflex
5		Ponstan 500 mg		Mefamed
6		MODURETIC		AMILOZIDE
7		Soproxen		Annoxen-S
8		Lasix 40 mg		Furetic 40 mg
9		Mucosolvan		Mucolid

ตารางที่ 3.1 การเปรียบเทียบการจับคู่ของยาทั้ง 2 ชนิดระหว่างยาดั้งเดิม และยาสามัญ (ต่อ)

ตารางการเปรียบเทียบลักษณะรูปร่างจากการจับคู่ยาทั้ง 2 ชนิดที่ใช้ในงานทดลอง				
คู่ที่	ยาดั้งเดิม (Original drugs)		ยาสามัญ (Generic drugs)	
9		Mucosolvan		Mucolid
10		NORGESIC		Myogesic
11		Lyrica 75 mg		Precius
12		DUOCETZ		Utraphen
13		Glucophage 500 mg		Siamformet 500 mg

ชุดข้อมูล (Dataset) ทั้งหมดจะถูกถ่ายจากโทรศัพท์มือถือ Huawei nova 3i พร้อมความละเอียด 16 ล้านพิกเซล และระบบประมวลผลที่ช่วยในการถ่ายภาพในสภาพแวดล้อมแบบปิด โดยเป็นภาพขนาด 4,608 x 3,456 พิกเซล ที่ถ่ายจากมุมด้านหน้าและด้านหลังเป็นหลักดังตัวอย่างข้อมูลดังรูปที่ 3.3

Genetic drugs



Original drugs

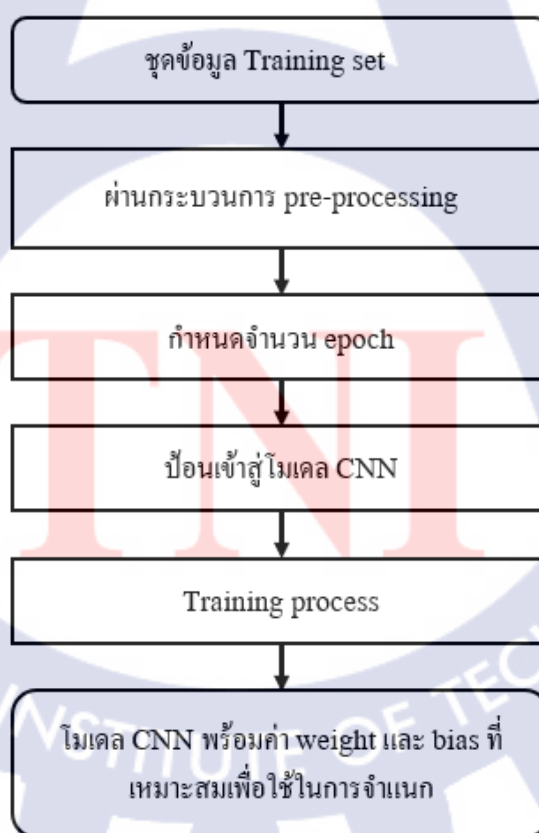


รูปที่ 3.3 ตัวอย่างข้อมูลรูปภาพเม็ดยาที่ใช้ในการทดลองระหว่าง (บน) ยาสามัญ และ (ล่าง) ยาดั้งเดิม

อันเนื่องมาจากมัมเหล่านี้ เป็นมัมที่แสดงคุณลักษณะเฉพาะของยาเม็ดแต่ละชนิดได้ชัดเจนที่สุด โดยเฉพาะรอยประทับ (imprint) ที่อยู่บนเม็ดซึ่งเป็น feature เพียงอย่างเดียวที่เป็นความแตกต่างระหว่างยาต้นแบบและ ยาสามัญที่โมเดลใช้ในการจดจำ (recognize) พร้อมทำการ labeling ข้อมูลล่วงหน้า เพื่อความสะดวกในการใช้งาน โดยมีบุคลากรทางการแพทย์ และเภสัชกรที่มีความเชี่ยวชาญเป็นผู้ให้คำปรึกษา ก่อนจะแบ่งออกเป็น 2 กลุ่มคือ ชุดข้อมูลการฝึกฝน (training set) และชุดการตรวจสอบ (validation set) ที่ใช้สำหรับการเรียนรู้และ evaluate ตามลำดับ ซึ่งขั้นตอนในการใช้งานของชุดข้อมูลทั้ง 2 ชนิดจะถูกอธิบายในหัวข้อต่อไป

3.4.1 ชุดข้อมูล training set

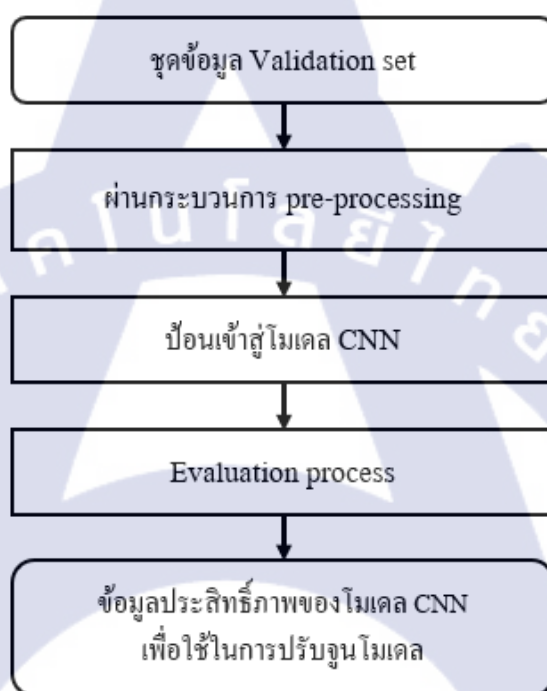
ชุดข้อมูล training set คือ ชุดข้อมูลที่โมเดลใช้ในการประมวลผลเพื่อให้เกิดการเรียนรู้ พร้อมคำนวณค่า weight และ bias ที่เหมาะสมเพื่อใช้ในการ classification โดยงานวิจัยนี้ชุดข้อมูล 83% จากทั้งหมดหรือคิดเป็น 5,200 ภาพ จะถูกนำมาใช้ในการเรียนรู้ก่อนทำการ pre-processing แล้วป้อนเข้าสู่โมเดล และทำการ training ตามแผนผังขั้นตอนดังรูปที่ 3.4



รูปที่ 3.4 แผนผังขั้นตอนการฝึกฝน (training) ชุดข้อมูล training set

3.4.2 ชุดข้อมูล validation set

ชุดข้อมูล validation set คือ ชุดข้อมูลที่ใช้ในการประเมินผล (evaluation) เพื่อทำการประวัตประสิทธิภาพ loss และ accuracy ของตัวโมเดล CNN โดยในงานวิจัยนี้ชุดข้อมูล 17% จากชุดข้อมูลทั้งหมดหรือคิดเป็น 1,040 ภาพ จะถูกนำมาใช้ในการ evaluate โดยผ่านการทำ pre-processing เช่นเดียวกับชุดข้อมูล training set แล้วป้อนเข้าสู่โมเดล ตามแผนผังขั้นตอนดังรูปที่ 3.5



รูปที่ 3.5 แผนผังขั้นตอนการประเมินผล (evaluation) โดยใช้ชุดข้อมูล validation set

3.5 อุปกรณ์ในการทดลอง และเครื่องมือที่ใช้ในงานวิจัย

อุปกรณ์และเครื่องมือที่ใช้ในงานวิจัยและดำเนินการทดลองในขั้นตอนงานวิจัยจากหัวข้อที่ 3.3 จะถูกอธิบายในข้อต่อไปซึ่งประกอบด้วย

3.5.1 คอมพิวเตอร์

คอมพิวเตอร์โน้ตบุ๊ก Asus TUF Gaming พร้อมการ์ดจอ Nvidia RTX 3060 ที่ใช้สำหรับการเขียนโปรแกรม ออกแบบ สร้างโมเดล และทำการทดสอบ ดังรูปที่ 3.6 ซึ่งมีประสิทธิภาพในการประมวลผลใช้งานโมเดล CNN และกล้องจากโทรศัพท์มือถือ Huawei nova 3i ที่ใช้ในการถ่ายภาพเพื่อสร้างชุดข้อมูล ดังรูปที่ 3.7



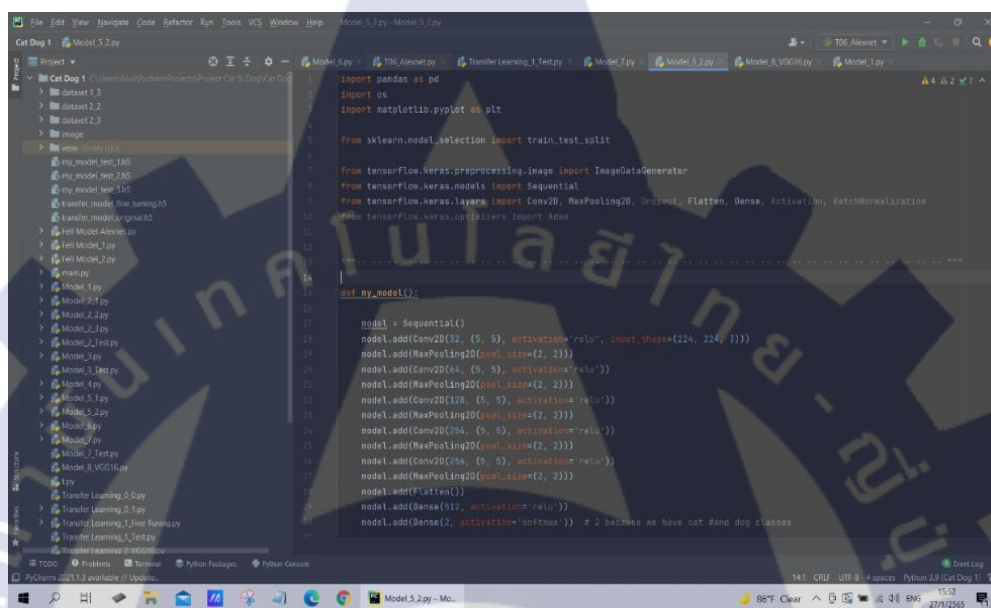
รูปที่ 3.6 คอมพิวเตอร์โน้ตบุ๊ก Asus TUF Gaming ที่ใช้สำหรับการเขียนโมเดล และการดำเนินการทดลองขั้นตอนต่างๆ



รูปที่ 3.7 โทรศัพท์มือถือ Huawei nova 3i ที่ใช้สำหรับการถ่ายภาพ เพื่อสร้างชุดข้อมูลแต่ละชนิดในงานทดลอง

3.5.2 โปรแกรม

โปรแกรม PyCharm Community Edition คือ โปรแกรมคอมพิวเตอร์ที่ใช้ในงานทดลอง จากหัวข้อที่ 3.3 ซึ่งใช้ในการเขียนโมเดล CNN และดำเนินการทดลองขั้นตอนต่างๆ ตั้งแต่การจัดเตรียมชุดข้อมูล, training, evaluation, prediction และทำการวัดประสิทธิภาพการใช้งานของโมเดล CNN โดยใช้ภาษา Python ดังรูปที่ 3.8



รูปที่ 3.8 โปรแกรม PyCharm Community Edition สำหรับการเขียนโมเดล CNN และดำเนินการทดลองด้วยภาษา Python



รูปที่ 3.9 โลโก้โปรแกรม PyCharm Community Edition

3.5.3 ชุดคำสั่งและไลบรารี

ชุดคำสั่งที่ใช้ในงานวิจัยจากหัวข้อที่ 3.3 จะแสดงอยู่ในตารางที่ 3.2 ซึ่งเป็นไลบรารี (library) ที่ใช้คู่กับภาษา Python เพื่อใช้ในการเขียนโมเดล CNN และดำเนินงานวิจัยในขั้นตอนต่างๆ

ตารางที่ 3.2 ชุดคำสั่งภาษา Python ที่ใช้ในงานวิจัย

โลโก้	ชื่อ	หน้าและวิธีการใช้งาน
	TensorFlow	เป็นไลบรารีที่ใช้ในการพัฒนา Machine Learning โดยเฉพาะการสร้างโมเดลโครงข่ายประสาทเทียม เพื่อใช้ในการวิจัย
	Keras	เป็น Deep Learning Library ในภาษา Python ที่ใช้ในการพัฒนา Machine Learning เช่นเดียวกับ TensorFlow
	NumPy	เป็นโมดูลส่วนเสริมของ Python ที่มีฟังก์ชันเกี่ยวกับคณิตศาสตร์ และการคำนวณ
	Pandas	คือ ไลบรารี ที่ใช้ในการจัดการกับ dataframe และข้อมูลรูปภาพด้วยภาษา Python ในการทดลองโมเดล CNN
	OpenCV	คือ ไลบรารีที่ใช้ในการประมวลผลภาพขั้นพื้นฐาน เช่น การอ่าน การปรับ การเบลอ การผสม และการเพิ่มคุณภาพของภาพ

3.6 การสร้างแบบจำลอง

3.6.1 โมเดลโครงข่ายประสาทเทียม Convolution ที่สร้างขึ้น

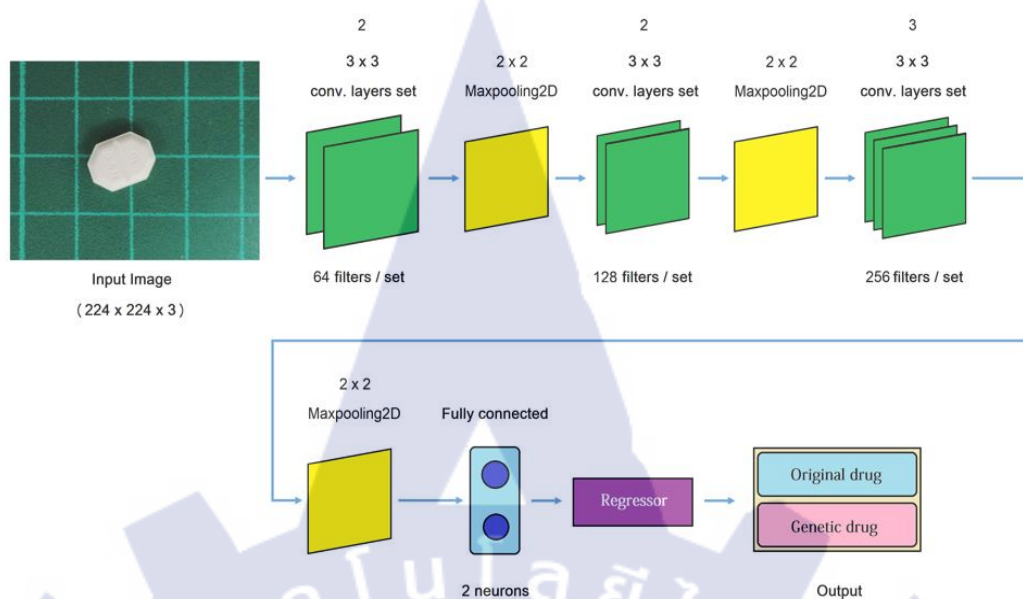
การสร้างแบบจำลอง CNN ที่ใช้ในการจำแนกยาเม็ดทั้ง 2 ชนิดจะเป็นการเขียนโมเดลใหม่ตั้งแต่ศูนย์ โดยใช้ไลบรารี Keras และ TensorFlow ในการสร้างเลเยอร์แต่ละชั้นซึ่งเป็นองค์ประกอบหลักของโครงสร้างโมเดล CNN ตั้งแต่กลุ่มเลเยอร์ Feature learning ที่ใช้ในการสกัดคุณลักษณะจากภาพอินพุต (224, 224, 3) ไปจนถึงกลุ่มเลเยอร์ Fully connected ที่ใช้ในการจำแนกรูปภาพให้ออกเป็นผลลัพธ์ตามจำนวนที่ระบุ พร้อมกับกำหนดจำนวนตัวกรอง ขนาด และพารามิเตอร์อื่นๆ ที่เหมาะสมดังตารางที่ 3.3 และ 3.4 ตามลำดับ

ตารางที่ 3.3 กลุ่มเลเยอร์ Feature learning ของโมเดล CNN ที่สร้างขึ้น

องค์ประกอบโครงสร้างของโมเดล CNN ที่ใช้ในการทดลอง				
Layer type	Size Kernel	Number of neurons	Activation	Num.
Conv2D	3 x 3	64	relu	2
Maxpooling2D	2 x 2	-	-	1
Conv2D	3 x 3	128	relu	2
Maxpooling2D	2 x 2	-	-	1
Conv2D	3 x 3	256	relu	3
Maxpooling2D	2 x 2	-	-	1

ตารางที่ 3.4 กลุ่มเลเยอร์ Fully connected layers ของโมเดล CNN ที่สร้างขึ้น

องค์ประกอบโครงสร้างของโมเดล CNN ที่ใช้ในการทดลอง			
Layer type	Units	Rate	Activation
Flatten	-	-	-
Dense	1536	-	relu
Dropout	-	0.2	-
Dense	1536	-	relu
Dropout	-	0.2	-
Dense	2	-	-



รูปที่ 3.10 โครงสร้างโมเดลโครงข่ายประสาทเทียม Convolution ที่ใช้ในการทดลอง

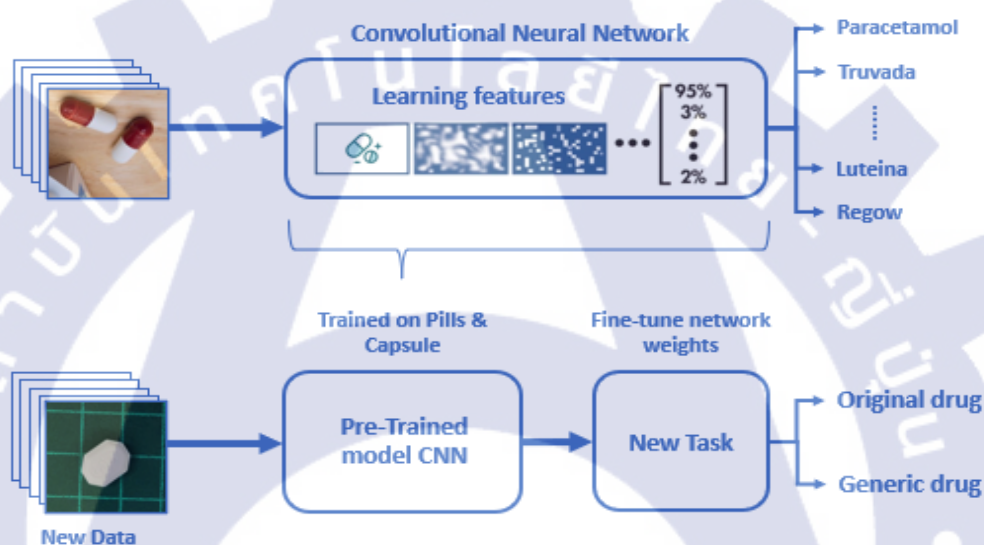
ในส่วนของการออกแบบ model.compile กำหนดให้ SGD เป็น optimizer พร้อมกำหนดอัตราการเรียนรู้ (learning rate) เท่ากับ 0.001 และ batch size เท่ากับ 64 โดยมีการคิด loss แบบ sparse categorical crossentropy (from_logits = True) พร้อมกำหนด metrics เป็น accuracy เพื่อทำการการวัดประสิทธิภาพแต่ละรอบระหว่างกระบวนการฝึกฝน (training) และเพื่อให้ค่า output เป็น onehot encoding

ตารางที่ 3.5 สรุป Training setting และ Hyperparameters ของโมเดล CNN

Training setting and Hyperparameters	
Platform	Tensorflow-gpu ver. 2.6.2
Device	Nvidia RTX 3060
Optimizer	SGD
Learning rate	0.001
Epoch	635
Batch size	64
Loss	sparse categorical crossentropy (from_logits = True)
Metrics	accuracy

3.6.2 โมเดล transfer learning ที่ใช้ในการเปรียบเทียบ

การสร้างโมเดล transfer learning ทั้ง 3 ชนิดที่ใช้ในเปรียบเทียบจะเป็นการใช้ไลบรารี Keras และ TensorFlow ในการดั่งสถาปัตยกรรมแบบจำลอง CNN ที่มีอยู่ใน keras applications และเคยผ่านการฝึกฝนกับงานที่ใกล้เคียงกันด้วยเทคนิค transfer learning ได้แก่ ResNet50-v2, VGG16 และ InceptionV3 ซึ่งส่วน feature learning จะถูกตรึง (Freeze) เอาไว้พร้อมกำหนดขนาดเลเยอร์นำเข้า (224, 224, 3) และ Fully connected layers ก่อนจะทำการ run เพื่อสร้างโมเดล ซึ่งชุดข้อมูล training set และ validation set ที่ผ่านการ pre-processing จะถูกป้อนเข้าสู่โมเดลเพื่อการ training และ evaluate ตามลำดับ ดังตัวอย่างกระบวนการในรูปภาพที่ 3.10



รูปที่ 3.11 ตัวอย่างการสร้างโมเดล CNN ผ่านกระบวนการ transfer learning

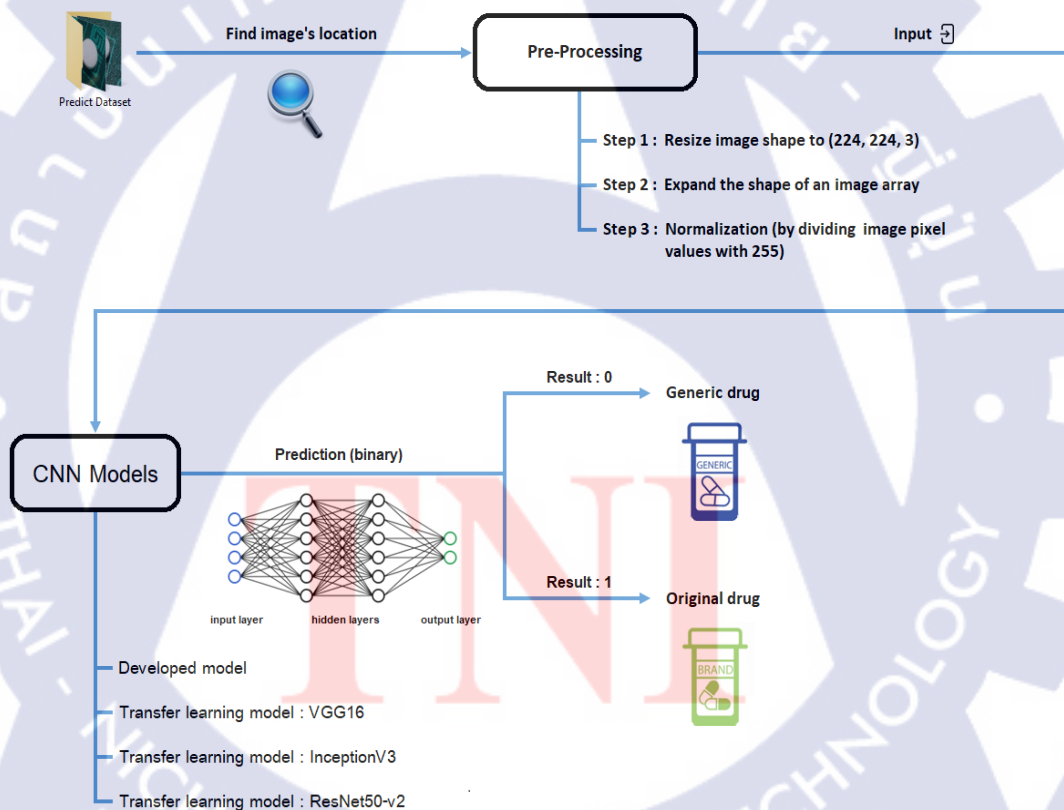
ในการออกแบบ Fully connected layers ชุดข้อมูลที่ผ่านการ feature extraction จะถูกส่งไปยัง `flatten()` เพื่อเปลี่ยนชนิดข้อมูลให้กลายเป็นเวกเตอร์ array ขนาด 1 มิติ และส่งเข้าโครงข่าย โดยการออกแบบเลเยอร์แรกเป็น dense 128 โดยมีฟังก์ชันกระตุ้น `relu` พร้อมกำหนด `he_uniform` เป็น `kernel_initializer` และ output layer เป็น dense 2 โดยใช้ ฟังก์ชันกระตุ้น `sigmoid` กำหนดให้ `adam` เป็น optimizer และ batch size เท่ากับ 64 โดยมีการคิด loss แบบ `sparse categorical_crossentropy` ส่วน `compile model` เพื่อให้ค่า output เป็น onehot encoding จึงกำหนดให้ metrics เป็น accuracy เพื่อทำการการวัดประสิทธิภาพแต่ละรอบระหว่างการ training

บทที่ 4

ผลการวิจัยและการวิเคราะห์ผลการทดลอง

4.1 ผลลัพธ์ของระบบ

ในเนื้อหาบทนี้จะกล่าวถึงผลลัพธ์ที่ได้จากการทดลองตามกระบวนการการทดลองในบทที่ 3 เพื่อหาโมเดลที่มีประสิทธิภาพสูงสุดในการจำแนกยาเม็ดระหว่างยาต้นแบบ และยาสามัญจากโมเดลโครงข่ายประสาทเทียม Convolution ที่เขียนขึ้นเพื่อใช้ในการจำแนกยาทั้ง 2 ชนิด โดยเฉพาะกับโมเดล CNN อีก 3 ชนิดที่ใช้ในการเปรียบเทียบซึ่งสร้างจากการดัดแปลงโมเดล CNN ที่เคยผ่านการ pre-train กับงานที่ใกล้เคียงกัน ได้แก่ VGG16, InceptionV3 และ ResNet50-v2 ด้วยเทคนิค transfer learning ตามขั้นตอนที่ใช้ในการคาดการณ์ (prediction) ดังรูปภาพที่ 4.1



รูปที่ 4.1 แผนผังขั้นตอนกระบวนการการคาดการณ์ (prediction) ของโมเดล CNN เพื่อทำการภาพวัดประสิทธิภาพ

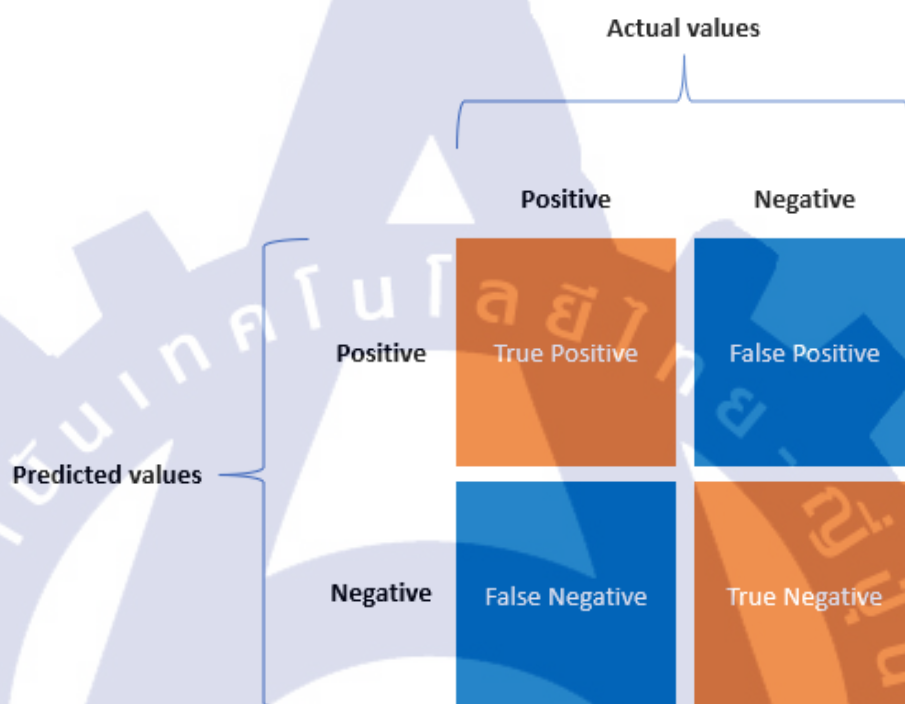
โดยชุดข้อมูล test set ที่ใช้สำหรับการ prediction จะถูกดึงจาก folder ไปยัง กระบวนการ pre-processing ด้วยชุดคำสั่ง os ก่อนจะทำการปรับขนาดรูปภาพใหม่เป็น (224, 224, 3) พร้อมขยาย (expand) ขนาด array รูปภาพและ ทำการ normalization โดยการหารค่าพิกเซล ของตัวรูปภาพด้วย 255.0 เพื่อปรับค่าข้อมูลให้อยู่ในช่วง 0-1 ตามขอบเขตที่กำหนด ก่อนจะป้อนเข้าสู่โมเดลทั้งหมดที่ใช้ในการทดลองแล้วจึงออกมาเป็นผลลัพธ์ในรูปแบบของข้อมูล binary โดย 0 เท่ากับ ยาสามัญและ 1 เท่ากับ ยาต้นแบบ ดังตัวอย่างการทำ Classification ในรูปภาพที่ 4.2



รูปที่ 4.2 ตัวอย่างการทำ Classification ของ โมเดล CNN

ในกรณีผลลัพธ์ที่ได้จากการ prediction มีค่าเท่ากับ 0 ซึ่งตรงกับชนิดรูปภาพยาสามัญที่ป้อนเข้าสู่โมเดล ค่า True Positive (TP) จะบวกเพิ่มเป็น 1 แต่หากไม่ตรงกันค่า False Positive (FP) จะบวกเพิ่มแทน เช่นเดียวกันกับยาต้นแบบ หากผลลัพธ์ที่ได้จากการ prediction มีค่าเท่ากับ 1 ซึ่งตรงกับชนิดของภาพยาต้นแบบที่ป้อนเข้าสู่โมเดล ค่า True Negative (TN) จะบวกเพิ่ม 1 แต่หากไม่ตรงกันค่า False Negative (FN) จะบวกเพิ่มแทนซึ่งผลรวมของค่า TP, FP, TN และ FN ที่ได้จากการ prediction กับรูปภาพชุดข้อมูล dataset ทั้งหมดที่ใช้สำหรับการคาดการณ์จะถูกสรุปอยู่ในตาราง Confusion Matrix ดังรูปภาพ 4.3 ซึ่งเป็นเครื่องมือในการวัดความสามารถของ machine learning

ในการแก้ปัญหา classification ของโมเดล CNN ซึ่งค่าที่ได้เหล่านี้จะถูกนำไปคำนวณเพื่อหาค่า Accuracy, Precision, Recall และ F1 score ดังสมการที่ 23, 24, 25 และ 26 ตามลำดับ เพื่อใช้ในการประเมินประสิทธิภาพการทำงานของตัวโมเดล CNN ในการจำแนกยาเม็ดทั้ง 2 ชนิด ระหว่างยาต้นแบบและ ยาสามัญ



รูปที่ 4.3 ตาราง Confusion Matrix ในการวัดความสามารถของ machine learning ในการแก้ปัญหา classification

4.2 การวัดประสิทธิภาพ

การวัดประสิทธิภาพของโมเดล CNN ที่สร้างขึ้นโดยเทคนิคต่างๆ ดังกล่าวในบทที่ 3.6 เพื่อดูความถูกต้องในการจำแนกชนิดยา หรือ classification และการทำนายผลยา หรือ prediction ผู้วิจัยได้กำหนดเกณฑ์ในการประเมิน โดยใช้ Accuracy score, Precision score, Recall score และ F-1 score มาใช้ในการเปรียบเทียบโมเดลแต่ละชนิดที่ใช้โครงสร้างสถาปัตยกรรม CNN และจำนวนรอบ epoch ที่แตกต่างกัน ซึ่ง Accuracy, Precision, Recall และ F-1 Scores สามารถคำนวณค่าได้ จากสมการที่ 23, 24, 25 และ 26 ตามลำดับ

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4.1)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (4.2)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (4.3)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4.4)$$

โดย TP (True Positive) คือ ค่าที่พยากรณ์ถูกต้องเชิงบวก, TN (True Negative) คือ ค่าที่พยากรณ์ถูกต้องเชิงลบ, FP (False Positive) คือ ค่าที่พยากรณ์ผิดพลาดเชิงบวก และ FN (False Negative) คือ ค่าที่พยากรณ์ผิดพลาดเชิงลบ

4.3 อภิปรายผลลัพธ์ของโมเดลแต่ละชนิด

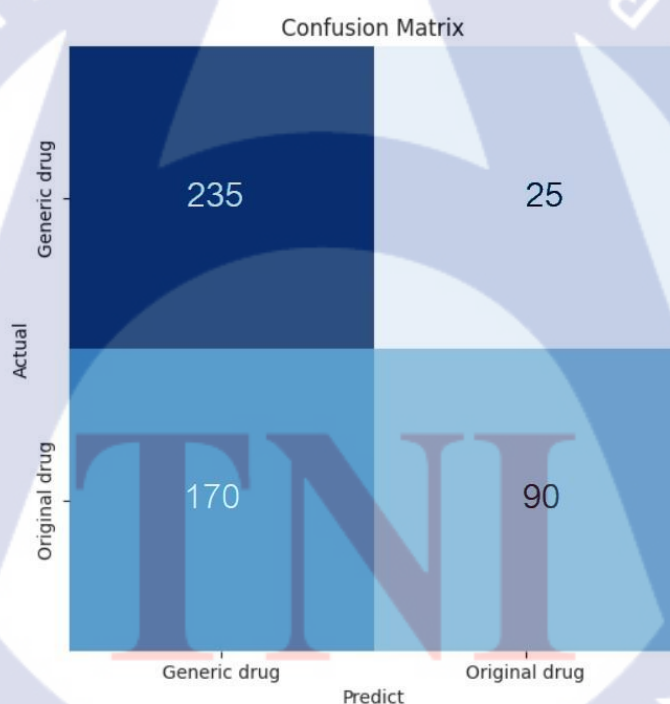
ในส่วนของการอภิปรายผลลัพธ์ที่ได้จากการทดลองจะเป็นสรุปผลการเปรียบเทียบคะแนนการวัดประสิทธิภาพการจำแนกแหว่งโมเดลโครงข่ายประสาทเทียม Convolution ที่ถูกออกแบบมาโดยมีวัตถุประสงค์เพื่อใช้จำแนกยาเม็ดทั้ง 2 ชนิดระหว่างยาต้นแบบและยาสามัญ กับโมเดลโครงข่ายประสาทเทียม 3 ชนิดที่ใช้ในการเปรียบเทียบซึ่งสร้างการจากเทคนิค transfer learning โดยการดัดแปลงสถาปัตยกรรมโมเดล CNN ที่มีอยู่ในไลบรารี keras.applications และเคยผ่านการใช้การทดลองกับงานที่ใกล้เคียงกันมาสร้างเป็นโมเดลใหม่ พร้อมทำการปรับแต่งเลเยอร์ Fully Connected และกำหนดขนาดของข้อมูลนำเข้าตามหัวข้อที่ 3.6 เพื่อหาโมเดลที่มีประสิทธิภาพในการจำแนกสูงสุด โดยเป็นการพิจารณาจากคะแนน Accuracy, Precision, Recall และ F1 score ที่ได้จากการคำนวณสมการที่ 23, 24, 25 และ 26 ตามลำดับ โดยใช้ค่า True Positive (TP), True Negative (TN), False Positive (FP) และ False Negative (FN) ที่ได้จากตาราง Confusion Matrix ตามหัวข้อที่ 4.1 ซึ่งเป็นผลลัพธ์ที่ได้จากการ prediction ของโมเดล CNN กับชุดข้อมูล test set ที่ประกอบด้วยรูปภาพ 130 ภาพที่ถ่ายจากความสูง 18.5 เซนติเมตร พร้อมทำการหมุน 90° 180° และ 270° องศา เพื่อเพิ่มปริมาณชุดข้อมูลกลายเป็น 520 โดยแบ่งออกเป็น 260 ภาพสำหรับยาสามัญและ ยาต้นแบบมาใช้ในการคำนวณ

โดยโมเดล CNN ที่ใช้ในการเปรียบเทียบจะประกอบไปด้วยโมเดลทั้งหมด 60 โมเดลซึ่งแบ่งออกเป็น 3 กลุ่มตามสถาปัตยกรรมแต่ละชนิดที่ใช้งานทดลอง กลุ่มละ 20 โมเดล ซึ่งโมเดลแต่ละกลุ่มประกอบไปด้วย โมเดล CNN ที่สร้างจากจำนวน epoch ที่แตกต่างกัน 4 ชนิด ได้แก่ 20, 15, 10 และ 5 epochs ตามลำดับ เพื่อจำกัดการรอบ epoch การฝึกฝน (training) ที่เหมาะสมสำหรับ

สถาปัตยกรรม CNN แต่ละชนิด พร้อมหา ค่า weight และ bias ที่โมเดลแต่ละชนิดสามารถทำการจำแนกได้อย่างมีประสิทธิภาพสูงสุด ก่อนทำการเปรียบเทียบกับโมเดลโครงข่ายประสาทเทียมที่ใช้ในการทดลองซึ่งผลลัพธ์และรายละเอียดที่ได้จากการทดลองแต่ละส่วนจะถูกอธิบายในหัวข้อต่อไป

4.3.1 โมเดลโครงข่ายประสาทเทียม Convolution ที่สร้างขึ้นเพื่อใช้ในการทดลอง

จากผลลัพธ์การ prediction โมเดลโครงข่ายประสาทเทียม Convolution ที่เขียนขึ้นเพื่อใช้จำแนกยาเม็ดทั้ง 2 ชนิดระหว่างยาสามัญและยาต้นแบบ สามารถทำคะแนน Accuracy ได้ถึง 0.6250 พร้อม Precision, Recall และ F1 score เท่ากับ 0.9038, 0.5802 และ 0.7068 ตามลำดับจาก ค่า True Positive (TP) = 235, True Negative (TN) = 90, False Positive (FP) = 25 และ False Negative (FN) = 170 จากตาราง Confusion Matrix ในรูปภาพที่ 4.4 โดยมีอัตราความสำเร็จในการตรวจจำยาสามัญได้อย่างถูกต้องมากกว่ายาต้นแบบ 61.70% ด้วยจำนวนรอบที่ใช้ในการฝึกฝน 635 epoch



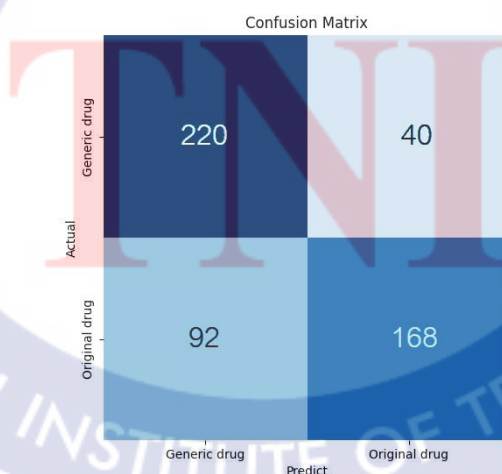
รูปที่ 4.4 ตาราง Confusion Matrix ของโครงข่ายประสาทเทียม Convolution ที่ใช้ในการทดลอง

4.3.2 InceptionV3

ตารางที่ 4.1 ผลลัพธ์ของโมเดล 4 อันดับแรกที่ใช้สถาปัตยกรรม InceptionV3 ที่ทำคะแนนการวัดประสิทธิภาพได้สูงสุดจากทั้งหมด 20 โมเดล ในการจำแนกยาเม็ด 2 ชนิด

Top	Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
1	InceptionV3_14	220	168	40	92	0.8462	0.7051	0.7462	0.7692	10
2	InceptionV3_17	201	187	59	73	0.7731	0.7336	0.7462	0.7528	5
3	InceptionV3_18	193	193	67	67	0.7423	0.7423	0.7423	0.7423	5
4	InceptionV3_16	224	159	36	101	0.8615	0.6892	0.7365	0.7658	15

จากผลลัพธ์ที่ได้จากการ prediction ของโมเดล CNN ที่ใช้ในการเปรียบเทียบ ซึ่งสร้างจากสถาปัตยกรรม InceptionV3 ดังตารางที่ 4.2 โมเดลที่มีประสิทธิภาพสูงสุดจากทั้งหมด 20 โมเดลที่ใช้สถาปัตยกรรมเดียวกัน คือโมเดล InceptionV3_14 ที่ใช้จำนวนรอบการฝึกฝน 10 epoch ซึ่งสามารถคะแนน Accuracy ได้สูงสุดถึง 0.7462 พร้อมคะแนน Precision = 0.8462, Recall = 0.7051 และ F1-score = 0.7692 จากค่า True Positive (TP) = 220, True Negative (TN) = 168, False Positive (FP) = 40 และ False Negative (FN) = 92 ดังตาราง Confusion Matrix ในรูปภาพที่ 4.5 โดยมีอัตราความสำเร็จในการตรวจจำยาสามัญได้อย่างถูกต้องมากกว่ายาต้นแบบ 23.64%



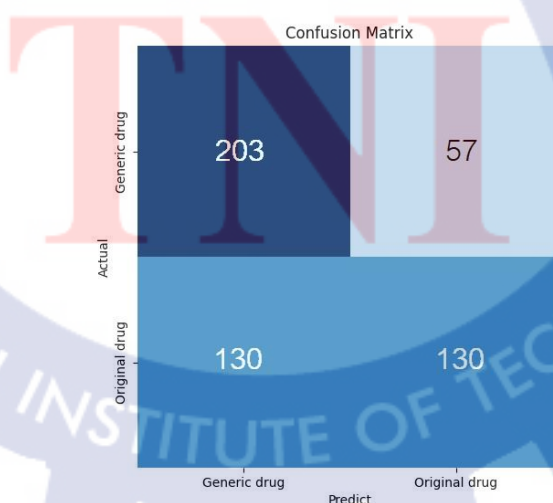
รูปที่ 4.5 ตาราง Confusion Matrix สำหรับโมเดล InceptionV3_14

4.3.3 VGG16

ตารางที่ 4.2 ผลลัพธ์ของโมเดล 4 อันดับแรกที่ใช้สถาปัตยกรรม VGG16 ที่ทำคะแนนการวัดประสิทธิภาพได้สูงสุดจากทั้งหมด 20 โมเดล ในการจำแนกยาเม็ด 2 ชนิด

Top	Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
1	VGG16_13	203	130	57	130	0.7808	0.6096	0.6404	0.6847	10
2	VGG16_15	200	133	60	127	0.7692	0.6116	0.6404	0.6814	10
3	VGG16_9	197	135	63	125	0.7577	0.6118	0.6385	0.6770	15
4	VGG16_14	173	159	87	101	0.6654	0.6314	0.6385	0.6479	10

จากผลลัพธ์ที่ได้จากการ prediction ของโมเดล CNN ที่ใช้ในการเปรียบเทียบ ซึ่งสร้างจากสถาปัตยกรรม VGG16 ดังตารางที่ 4.2 โมเดลที่มีประสิทธิภาพสูงสุดในการจำแนกยาเม็ดทั้ง 2 ชนิดจากทั้งหมด 20 โมเดลที่ใช้สถาปัตยกรรมเดียวกัน คือโมเดล VGG16_13 ที่ใช้จำนวนรอบในการฝึกฝน 10 epoch ซึ่งสามารถคะแนน Accuracy ได้สูงสุดถึง 0.6404 พร้อมคะแนน Precision = 0.7808, Recall = 0.6096 และ F1-score = 0.6847 จากค่า True Positive (TP) = 203, True Negative (TN) = 130, False Positive (FP) = 57 และ False Negative (FN) = 130 จากตาราง Confusion Matrix ในรูปภาพที่ 4.6 โดยมีอัตราความสำเร็จในการตรวจจับยาสามัญได้อย่างถูกต้องมากกว่ายาต้นแบบ 35.96%



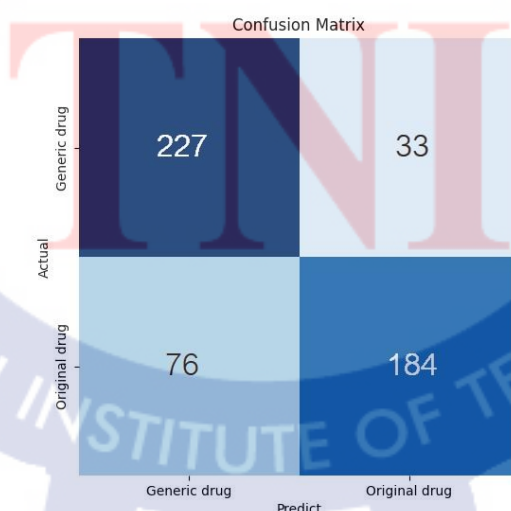
รูปที่ 4.6 ตาราง Confusion Matrix สำหรับโมเดล VGG16_13

4.3.4 ResNet50-v2

ตารางที่ 4.3 ผลลัพธ์ของโมเดล 4 อันดับแรกที่ใช้สถาปัตยกรรม ResNet50-v2 ที่ทำคะแนนการวัดประสิทธิภาพได้สูงสุดจากทั้งหมด 20 โมเดล ในการจำแนกยาเม็ด 2 ชนิด

Top	Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
1	ResNet50V2_18	227	184	33	76	0.8731	0.7492	0.7904	0.8064	5
2	ResNet50V2_13	206	196	54	64	0.7923	0.7630	0.7731	0.7774	10
3	ResNet50V2_2	194	207	66	53	0.7462	0.7854	0.7712	0.7653	10
4	ResNet50V2_4	195	202	65	58	0.7500	0.7708	0.7635	0.7602	20

จากผลลัพธ์ที่ได้จากการ prediction ของโมเดล CNN ที่ใช้ในการเปรียบเทียบ ซึ่งสร้างจากสถาปัตยกรรม ResNet50-v2 ดังตารางที่ 4.3 โมเดลที่มีประสิทธิภาพสูงสุดในการจำแนกยาเม็ดทั้ง 2 ชนิดจากทั้งหมด 20 โมเดลที่ใช้สถาปัตยกรรมเดียวกัน คือโมเดล ResNet50V2_18 ที่ใช้จำนวนรอบในการฝึกฝน 5 epoch ซึ่งสามารถคะแนน Accuracy ได้สูงสุดถึง 0.7904 พร้อมคะแนน Precision = 0.8731, Recall = 0.7492 และ F1-score = 0.8064 จากค่า True Positive (TP) = 227, True Negative (TN) = 184, False Positive (FP) = 33 และ False Negative (FN) = 76 จากตาราง Confusion Matrix ดังรูปภาพที่ 4.7 โดยมีอัตราความสำเร็จในการตรวจจำยาสามัญได้อย่างถูกต้องมากกว่ายาต้นแบบ 18.94%

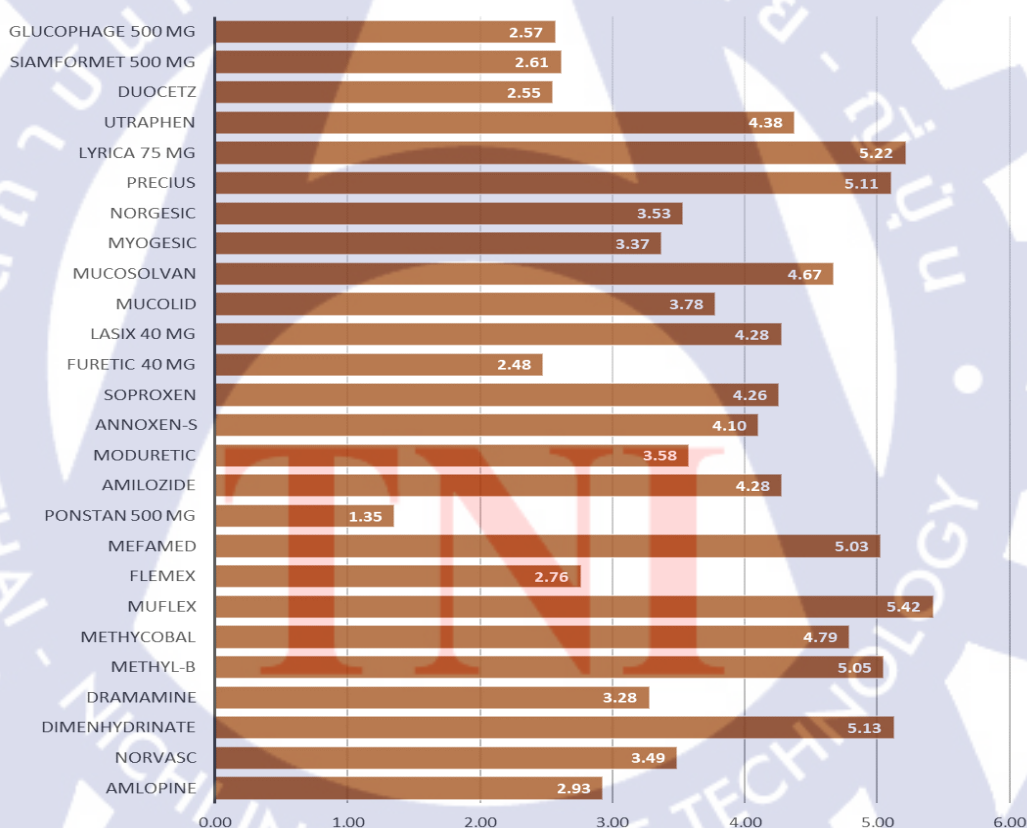


รูปที่ 4.7 ตาราง Confusion Matrix สำหรับโมเดล ResNet50V2_18

4.3.5 อภิปรายผลลัพธ์ของโมเดลแยกตามชนิดของยาเม็ด

จากการทดลองตรวจจับยาเม็ดทั้งหมด 26 ตัวกับโมเดล CNN ทั้งหมดที่ใช้ในการทดลองตัวยามีอัตราความสำเร็จในการตรวจจับได้อย่างถูกต้องต่ำที่สุด คือยาต้นแบบ Ponstan 500 mg ที่มีอัตราความแม่นยำในการจำแนกเพียง 1.35% ส่วนยาที่มีอัตราความสำเร็จในการตรวจจับได้อย่างถูกต้องมากที่สุด คือยาสามัญ Muflex ที่มีอัตราความแม่นยำในการจำแนกสูงถึง 5.42% ดังกราฟเปรียบเทียบอัตราความแม่นยำระหว่างยาทั้ง 26 ตัวในรูปภาพที่ 4.8 ที่ได้จากการคำนวณโดยใช้ค่า Ture position และ Ture negative ซึ่งเป็นค่าการจำแนกที่ถูกต้องที่ได้จากกระบวนการ prediction ของยาสามัญและยาต้นแบบแต่ละชนิดตามลำดับกับค่าผลรวมของการจำแนกที่ถูกต้อง (Ture) ทั้งหมดมาคำนวณคิดเป็นอัตราส่วนร้อยละความถูกต้องในการตรวจจับของยาแต่ละเม็ดทั้งหมด 26 ตัวที่ใช้ในการทดลอง

อัตราความแม่นยำในการตรวจจับของยาเม็ด 26 ตัวจากโมเดล CNN ทั้งหมดที่ใช้ในการทดลอง



รูปที่ 4.8 กราฟอัตราความแม่นยำในการตรวจจับยาเม็ด 26 ตัวจากโมเดล CNN ทั้งหมดที่ใช้ในการทดลอง

4.4 สรุปผล

ตารางที่ 4.4 การเปรียบเทียบโมเดลโครงข่ายประสาทเทียม Convolution ที่ใช้ในการทดลองกับ
โมเดล Transfer learning อีก 3 ชนิด ที่ใช้ในการเปรียบเทียบ

Top	Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
1	ResNet50V2_18	227	184	33	76	0.8731	0.7492	0.7904	0.8064	5
2	InceptionV3_14	220	168	40	92	0.8462	0.7051	0.7462	0.7692	10
3	VGG16_13	203	130	57	130	0.7808	0.6096	0.6404	0.6847	10
4	Model_CNN	235	90	25	170	0.9038	0.5802	0.6250	0.7068	635

จากการเปรียบเทียบประสิทธิภาพในการจำแนกระหว่างโมเดลโครงข่ายประสาทเทียม Convolution ที่ออกแบบมาเพื่อใช้ในการจำแนกยาเม็ด 2 ชนิดระหว่างยาต้นแบบและยาสามัญ กับโมเดล CNN อีก 3 ชนิด ที่ใช้สถาปัตยกรรมที่แตกต่างกันผ่านกระบวนการ transfer learning ดังตารางที่ 4.1 โมเดล CNN ที่สามารถทำคะแนนการวัดประสิทธิภาพได้สูงสุดซึ่งมีความแม่นยำในการจำแนกยาเม็ดทั้ง 2 ชนิดที่ไม่สามารถแยกแยะได้โดยง่ายเนื่องจากลักษณะ (feature) ภายนอกที่มีความคล้ายคลึงทั้ง ขนาด (size), สี (color) และ รูปร่าง (shape) ทั้ง 26 ตัวที่ใช้ในงานทดลอง และมีเปอร์เซ็นต์ความแตกต่างในการตรวจจำระหว่างยาทั้ง 2 ชนิดต่ำที่สุดคือ โมเดล ResNet50V2_18 ที่ใช้สถาปัตยกรรม ResNet50-v2 ผ่านเทคนิค transfer learning พร้อมจำนวนรอบในการฝึกฝน (training) 5 epoch ด้วยคะแนน Accuracy สูงสุดถึง 0.7904 พร้อมคะแนน Precision = 0.8731, Recall = 0.7492 และ F1-score = 0.8064 จากโมเดลทั้งหมด 61 โมเดลที่ใช้ในการทดลอง และเป็นโมเดลที่ทำคะแนนได้สูงสุดจากโมเดลทั้งหมด 20 โมเดลที่ใช้สถาปัตยกรรม ResNet50-v2 ซึ่งมีประสิทธิภาพมากกว่าโมเดลที่ใช้สถาปัตยกรรม VGG16, InceptionV3 และโมเดลโครงข่ายประสาทเทียม Convolution ที่ใช้ในการทดลอง โดยมีอัตราความแตกต่างในการตรวจจำยาสามัญได้ถูกต้องแม่นยำมากกว่ายาต้นแบบเพียง 18.94% ซึ่งน้อยกว่าโมเดล CNN อีก 3 ชนิดที่ใช้ในการทดลอง ดังนั้น โมเดลที่ใช้สถาปัตยกรรม ResNet50-v2 ผ่านเทคนิค transfer learning จึงเป็นโมเดล CNN ที่เหมาะสมจะเป็นโมเดลต้นแบบที่จะถูกนำไปพัฒนาต่อยอดในงานเภสัชกรรม กลายเป็นสนับสนุนเครื่องมือช่วยเหลือสำหรับเภสัชกรและบุคลากรทางการแพทย์ในการหาสาเหตุที่เกิดการแพ้ยา การใช้ยาไม่ได้ผลและปัญหาการใช้ยาไม่ถูกต้องที่เกิดจากผลข้างเคียงของยาสามัญ ซึ่งจะสามารถช่วยชีวิตผู้ป่วยได้รวดเร็วยิ่งขึ้นในอนาคต

บทที่ 5

สรุปผลการทดลอง

งานวิจัยนี้มีวัตถุประสงค์ เพื่อหาวิธีการที่เหมาะสมในการสร้างโมเดล CNN ต้นแบบที่จะนำไปพัฒนาต่อไปในอนาคตเป็นเครื่องมือสำหรับงานเภสัชกรรม และบุคลากรทางการแพทย์ ในการจำแนกยาเม็ดระหว่างยาต้นแบบ และยาสามัญ ที่ไม่สามารถแยกแยะออกได้โดยง่ายอันเนื่องมาจากลักษณะ (feature) ที่มีความคล้ายคลึง ทั้งขนาด (size), รูปร่าง (shape) และสี (color) ยกเว้นรอยประทับ (imprint) ที่สลักอยู่บนตัวเม็ดทั้งด้านหน้าและหลัง อันเป็นลักษณะเพียงอย่างเดียวที่สามารถใช้แยกแยะ เพื่อลดเวลาวินิจฉัยในการระบุสาเหตุการแพ้ยา หรือการใช้ยาที่ไม่ถูกต้องที่เกิดจากการใช้ยาสามัญแทนยาต้นแบบ และสามารถช่วยชีวิตผู้ป่วยให้พ้นขีดอันตรายได้เร็วขึ้น โดยการหาโมเดล Deep learning ที่มีประสิทธิภาพในการจำแนกสูงสุดที่สร้างจาก Convolution Neural Network และทำการเปรียบเทียบกับโมเดลทั้ง 3 ชนิดที่สร้างจากการดัดสถาปัตยกรรมโมเดล CNN ที่เคยผ่านการใช้งาน และมีรับประกันประสิทธิภาพจากงานวิจัยหลายฉบับด้วยเทคนิค Transfer learning ได้แก่ InceptionV3, VGG16 และ ResNet50-v2 ซึ่งเคยผ่านการใช้งานกับงานที่ใกล้เคียงกันมาสร้างเป็นโมเดลใหม่ ก่อนจะนำมาปรับแต่งชั้น Fully Connected layer พร้อมกำหนดขนาดข้อมูลนำเข้า (input size) โดยใช้โปรแกรม PyCharm Community Edition บนคอมพิวเตอร์โน้ตบุ๊ก ASUS TUF Dash F15 ผ่านภาษา ภาษา Python version 3.9 โดยใช้ภาพเม็ดยาทั้งหมด 6,240 ภาพ ซึ่งแบ่งเป็น 5,200 ภาพสำหรับชุดข้อมูลการฝึกฝน (training set) และอีก 1,040 ภาพ สำหรับชุดข้อมูลการตรวจสอบ (validation set) จากยาเม็ดทั้งสิ้น 26 ตัวซึ่งเป็นยาเม็ดแบบแข็งที่ใช้กันทั่วไป และสามารถหาซื้อได้ตามร้านขายยา คลินิกเอกชน หรือสถานพยาบาลต่างๆ ในประเทศไทยที่ถ่ายจากกล้องโทรศัพท์มือถือพร้อมทำการ label ข้อมูลล่วงหน้าเพื่อความสะดวกในการใช้งาน โดยมีบุคลากรทางการแพทย์ และเภสัชกรที่มีความเชี่ยวชาญเป็นผู้ให้คำปรึกษา มาใช้ในการทดลองซึ่งได้นำเสนอผลการทดลองแล้วในบทที่ 4 โดยการเปรียบเทียบค่า Accuracy, Precision, Recall และ F1-score ที่ใช้ในวัดประสิทธิภาพในการจำแนกของโมเดล CNN ที่ได้จากการ prediction กับชุดข้อมูล test set ที่ประกอบไปด้วยภาพที่ใช้ในการทดสอบจำนวน 520 ภาพ โดยพิจารณาจากโมเดลที่ทำคะแนน Accuracy ได้สูงที่สุด ซึ่งเป็นค่าความแม่นยำในการจำแนก และประสิทธิภาพของตัวโมเดล และจะสรุปผลการวิจัยและข้อเสนอแนะเพื่อนำไปพัฒนาในการวิจัยต่อไป

5.1 สรุปผลการวิจัย

จากการเปรียบเทียบโมเดล CNN ที่ใช้ในการทดลอง กับโมเดลอีก 3 ชนิดที่ใช้ในการเปรียบเทียบซึ่งสร้างจากเทคนิค Transfer learning โมเดล CNN ที่สามารถทำคะแนนการวัดประสิทธิภาพการจำแนกได้สูงที่สุด และมีเปอร์เซ็นต์ความแตกต่างในการตรวจจำระหว่างยาทั้ง 2 ชนิดต่ำที่สุดคือ โมเดล ResNet50V2_18 ที่ใช้ที่สถาปัตยกรรม ResNet50-v2 พร้อมจำนวนรอบการฝึกฝน 5 epoch ซึ่งสามารถทำคะแนนทั้ง Accuracy Recall และ F-1 Score ได้สูงที่สุดถึง 79.04%, 74.92% และ 80.64% ตามลำดับ มีเพียงคะแนน Precision เท่านั้นที่ได้ 87.31% ซึ่งเป็นอันดับสามรองจากโมเดลโครงข่ายประสาทเทียม Convolution ที่เขียนขึ้นเองที่ได้คะแนน Recall สูงถึง 88.85% เท่ากันจากโมเดลทั้งหมด 61 โมเดลที่ใช้ในการทดลอง

โดยสาเหตุที่โมเดลโครงข่ายประสาทเทียม Convolution ที่ที่เขียนขึ้นซึ่งถูกออกแบบมาเพื่อใช้ในการจำแนกยาเม็ดทั้ง 2 ชนิดโดยใช้รอบการฝึกฝน 635 epoch สามารถทำคะแนนการวัดประสิทธิภาพการจำแนกได้น้อยกว่าโมเดล CNN อีก 3 ชนิดที่เหลือซึ่งสร้างจากสถาปัตยกรรม ResNet50-v2 และ InceptionV3 และ VGG16 ด้วยเทคนิค transfer learning ดังตารางที่ 4.1 อาจเกิดจากคุณภาพของชุดข้อมูล dataset ที่ยังขาดการปรับแต่งที่เหมาะสมมากพอเพื่อให้โมเดลสามารถสกัดและ ดึงคุณลักษณะของตัวเม็ดยาที่อยู่ในภาพได้อย่างมีประสิทธิภาพ และทรัพยากรที่ใช้ในการคำนวณ หรือสเปคของคอมพิวเตอร์ที่ใช้ในการทดลองที่ไม่มาก หรืออาจไม่มีประสิทธิภาพเพียงพอที่จะใช้ในการสร้างโมเดล Deep learning ตั้งแต่ 0 ที่มีความซับซ้อน, พารามิเตอร์จำนวนมาก หรือมีชั้นเลเยอร์ค่อนข้างลึกมากได้ ทำให้การจะรันโมเดลแต่ละตัวจำเป็นต้องใช้เวลาที่นานหลายวัน และ จำนวน epoch ที่มากกว่าโมเดล CNN ที่สร้างด้วยเทคนิค transfer learning หลายเท่า

5.2 ข้อเสนอแนะเพื่อการพัฒนาวิจัย

5.2.1 เพิ่มจำนวนชุดข้อมูลรูปภาพเม็ดยา และชนิดให้มีความหลากหลายมากขึ้น เพื่อขยายขอบเขตการทดลองให้ครอบคลุมยิ่งขึ้น

5.2.2 เพิ่มขั้นตอนการทดลองเปรียบเทียบอิทธิพลของขนาดข้อมูลรูปภาพนำเข้าที่มีต่อประสิทธิภาพในการจำแนกของโมเดล CNN เพื่อหาขนาดของข้อมูลรูปภาพนำเข้าที่มีความเหมาะสมที่สุด

5.2.3 ลองเปลี่ยนสถาปัตยกรรม หรือเพิ่มจำนวนโมเดลที่ใช้ในการเปรียบเทียบให้มากขึ้น พร้อมกำหนดจำนวน epoch ให้มีความหลากหลายเพื่อจำกัดหาจำนวน epoch, ค่า weight และ bias ที่มีความเหมาะสมกับโมเดลแต่ละชนิดให้มีความแม่นยำมากขึ้น

5.2.4 ควรมีการทดลองในโปรแกรมหรือ platform อื่นๆ นอกเหนือจากโปรแกรม Pycham เพื่อทำการเปรียบเทียบหาความแตกต่างทั้งระยะเวลาในการรันโมเดล ผลลัพธ์ และ ประสิทธิภาพการใช้งานของแต่ละโปรแกรม



บรรณานุกรม

TNI

บรรณานุกรม

- [1] แสงสุข พิทยานุกุล และศิริ ชะระอำ, “การตัดสินใจเลือกใช้บริการร้านขายยาของผู้บริโภคไทย,” *วารสารวิชาการบริหารธุรกิจ*, ปีที่ 6, ฉบับที่ 15, หน้า 135-145, มกราคม-ธันวาคม 2550.
- [2] สำนักงานคณะกรรมการคุ้มครองผู้บริโภค, “ชื่อยา ต้องมีชื่อยา,” [Online]. Available: https://www.ocpb.go.th/mobile_detail.php?cid=3&nid=647&fbclid=IwAR1ZdOTjnS QnYeRzVj2KlasaM3Rwjs_9_0GjMXAV1DPJMcKp-DsWQD1Eax0. [Accessed: 18 พฤศจิกายน 2564].
- [3] นพพร ชื่นกลิ่น, “กว่าจะมาเป็นยา องค์การเภสัชกรรมผลิตยาตามสูตรตำรับ และมาตรฐาน คุณภาพทุกขั้นตอน,” [Online]. Available: [http://wongkarnpat.com/upfi leya/GPO\[1\].pdf](http://wongkarnpat.com/upfi leya/GPO[1].pdf). [Accessed: 19 พฤศจิกายน 2564].
- [4] สมศักดิ์ เทียมเก่า, “ประเภทยา: ยาชื่อสามัญและยาดันแบบ (Generic and Original Drugs),” [Online]. Available: <https://haamor.com/ยาชื่อสามัญและยาดันแบบ>. [Accessed: 5 ธันวาคม 2564].
- [5] ศุภณัฐ จินตวัฒน์สกุล, “การจำแนกภาพขวดแบบเซตเปิดด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน,” *ปริญญานิพนธ์ วท.ม. (เทคโนโลยีสารสนเทศ)*, มหาวิทยาลัยศรีนครินทรวิโรฒ, กรุงเทพฯ, 2562.
- [6] นพรุจ พัฒนสาร และ ณัฐวุฒิ ศรีวิบูลย์, “การประมวลผลภาพสำหรับการจำแนกคุณภาพมะม่วงพันธุ์โชคอนันต์โดยการจำลองการมองเห็นของมนุษย์ด้วยวิธีการเรียนรู้เชิงลึก,” *JIST Journal of Information Science and Technology*, ปีที่ 10, ฉบับที่ 1, หน้า 24-29, มกราคม-มิถุนายน 2563.
- [7] รัตน์โชติ พันธุ์วิไล, “การตรวจหาต้นไม้มันโรดโดยอัตโนมัติด้วยภาพถ่ายมุมสูงจากโดรนและวิธีการเรียนรู้เชิงลึก,” *วิทยานิพนธ์ วท.ม. (วิศวกรรมข้อมูลขนาดใหญ่)*, มหาวิทยาลัยธุรกิจบัณฑิต, กรุงเทพฯ, 2562.
- [8] E. A. Adewumi, “A Convolutional Neural Network (CNN) based Pill Image Retrieval System,” M.S. Thesis (Data Analytics), RIT Dubai, Dubai, UAE, 2020.
- [9] สุริยะ ชยะธรรมกุล, “การจำแนกผลึกน้ำตาลด้วยวิธีการเรียนรู้เชิงลึก,” *สารนิพนธ์ วท.ม. (วิศวกรรมซอฟต์แวร์)*, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ, 2563.

- [10] Y. Ou et al., "Automatic Drug Pills Detection based on Convolution Neural Network," Research Square [Online]. Available: https://www.researchgate.net/publication/333229352_Automatic_Drug_Pills_Detection_based_on_Convolution_Neural_Neural_Network. [Accessed: October 14, 2022].
- [11] Y. Ouet al., "Automatic drug pills detection based on enhanced feature pyramid network and convolution neural networks," *IET Computer Vision*, vol. 14, no. 1, pp. 9-17, February 2020.
- [12] J. J. Cabanet al., "Automatic identification of prescription drugs using shape distribution models," *19th IEEE International Conference on Image Processing*, Orlando, FL, USA, September 30 – October 3, 2012, pp. 1005-1008.
- [13] S. Suntronsuk and S. Ratanotayanon, "Automatic text imprint analysis from pill images," *9th International Conference on Knowledge and Smart Technology, KST 2017*, Chonburi, Thailand, February 1-4, 2017, pp. 288-293.
- [14] L. Tanet al., "Comparison of RetinaNet, SSD, and YOLO v3 for Real-Time Pill Identification," Research Square [Online]. Available: https://www.researchgate.net/publication/356449885_Comparison_of_RetinaNet_SSD_and_YOLO_v3_for_real-time_pill_identification. [Accessed: October 25, 2022].
- [15] L. Tanet al., "Comparison of YOLO v3, Faster R-CNN, and SSD for Real-Time Pill Identification," Research Square [Online]. Available: https://www.researchgate.net/publication/353590069_Comparison_of_YOLO_v3_Faster_RCNN_and_SSD_for_Real-Time_Pill_Identification. [Accessed: October 25, 2022].
- [16] Y. F. Wonget al., "Development of fine-grained pill identification algorithm using deep convolutional network," *Journal of Biomedical Informatics*, vol. 74, pp. 130-136, October 2017.
- [17] N. Usuyamaet al., "ePillID dataset: A low-shot fine-grained benchmark for pill identification," *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2020*, Seattle, WA, USA, June 14-19, 2020, pp. 3971-3977.
- [18] P. Chpawa and K. Kanjanawanishkul, "Pill identification with imprints using a neural network," *Maharakham International Journal of Engineering Technology*, vol. 1, no. 1, pp. 30-35, January - June 2015.

- [19] S. Linget et al., "Few-shot pill recognition," *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020*, Seattle, WA, USA, June 13-19, 2020, pp. 9789-3998.
- [20] U. Patel, "Machine learning-based pharmaceutical tablet inspection and recognition techniques – A Review," *2021 International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021*, Coimbatore, India, March 25-27, 2021, pp. 220-222.
- [21] X. Zeng et al., "MobileDeepPill: A small-footprint mobile deep learning system for recognizing unconstrained pill images," *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, MobiSys '17*, Niagara Falls, NY, USA, 19-23 June, 2017, pp. 56-67.
- [22] L. Sousa et al., "Pill image classification using machine learning," *8th Brazilian Conference on Intelligent Systems, BRACIS 2019*, Salvador, Brazil, October 15-18, 2019, pp. 556-561.
- [23] A. Mehmood et al., "MobilePill: Accurate pill image classification via deep learning on mobile," *International Conference on Information and Communication Technology Convergence, ICTC 2019*, Jeju, Korea (South), 2019, pp. 1362-1367.
- [24] Y. Wang et al., "Pill recognition using minimal labeled data," *IEEE Third International Conference on Multimedia Big Data, BigMM 2017*, Laguna Hills, CA, USA, April 19-21, 2017, pp. 346-353.
- [25] W. Swastika et al., "Preliminary study of multi convolution neural network-based model to identify pills image using classification rules," *International Seminar on Intelligent Technology and Its Applications, ISITIA 2019*, Surabaya, Indonesia, August 28-29, 2019, pp. 376-380.

ภาคผนวก

TNI

ภาคผนวก ก.

ตารางบันทึกผลการจำแนกของโมเดล CNN ทั้งหมดที่ใช้ในการทดลอง



ตารางที่ ก.1 ผลลัพธ์ประสิทธิภาพการจำแนกของโมเดล CNN ที่ใช้สถาปัตยกรรม InceptionV3 จำนวน 20 โมเดล เพื่อใช้ในการเปรียบเทียบ

Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
InceptionV3_1	227	143	33	117	0.8731	0.6599	0.7115	0.7517	20
InceptionV3_2	234	123	26	137	0.9000	0.6307	0.6865	0.7417	
InceptionV3_3	231	136	29	124	0.8885	0.6507	0.7058	0.7512	
InceptionV3_4	222	137	38	123	0.8538	0.6435	0.6904	0.7339	
InceptionV3_5	224	158	36	102	0.8615	0.6871	0.7346	0.7645	
InceptionV3_6	224	159	36	101	0.8615	0.6892	0.7365	0.7658	15
InceptionV3_7	219	161	41	99	0.8423	0.6887	0.7308	0.7578	
InceptionV3_8	233	129	27	131	0.8962	0.6401	0.6962	0.7468	
InceptionV3_9	224	157	36	103	0.8615	0.6850	0.7327	0.7632	
InceptionV3_10	232	137	28	123	0.8923	0.6535	0.7096	0.7545	
InceptionV3_11	252	74	8	186	0.9692	0.5753	0.6269	0.7221	10
InceptionV3_12	237	124	23	136	0.9115	0.6354	0.6942	0.7488	
InceptionV3_13	221	149	39	111	0.8500	0.6657	0.7115	0.7466	
InceptionV3_14	220	168	40	92	0.8462	0.7051	0.7462	0.7692	
InceptionV3_15	191	188	69	72	0.7346	0.7262	0.7288	0.7304	
InceptionV3_16	241	114	19	146	0.9269	0.6227	0.6827	0.7450	5
InceptionV3_17	201	187	59	73	0.7731	0.7336	0.7462	0.7528	
InceptionV3_18	193	193	67	67	0.7423	0.7423	0.7423	0.7423	
InceptionV3_19	211	131	49	129	0.8115	0.6206	0.6577	0.7033	
InceptionV3_20	220	137	40	123	0.8462	0.6414	0.6865	0.7297	

ตารางที่ ก.2 ผลลัพธ์ประสิทธิภาพการจำแนกของโมเดล CNN ที่ใช้สถาปัตยกรรม ResNet50v2 จำนวน 20 โมเดล เพื่อใช้ในการเปรียบเทียบ

Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
ResNet50V2_1	179	216	81	44	0.6885	0.8027	0.7596	0.7412	20
ResNet50V2_2	194	207	66	53	0.7462	0.7854	0.7712	0.7653	
ResNet50V2_3	193	201	67	59	0.7423	0.7659	0.7577	0.7539	
ResNet50V2_4	195	202	65	58	0.7500	0.7708	0.7635	0.7602	
ResNet50V2_5	189	206	71	54	0.7269	0.7778	0.7596	0.7515	
ResNet50V2_6	195	202	65	58	0.7500	0.7708	0.7635	0.7602	15
ResNet50V2_7	179	208	81	52	0.6885	0.7749	0.7442	0.7291	
ResNet50V2_8	172	188	88	72	0.6615	0.7049	0.6923	0.6825	
ResNet50V2_9	191	202	69	58	0.7346	0.7671	0.7558	0.7505	
ResNet50V2_10	186	211	74	49	0.7154	0.7915	0.7635	0.7515	
ResNet50V2_11	188	204	72	56	0.7231	0.7705	0.7538	0.7460	10
ResNet50V2_12	220	144	40	116	0.8462	0.6548	0.7000	0.7383	
ResNet50V2_13	206	196	54	64	0.7923	0.7630	0.7731	0.7774	
ResNet50V2_14	178	208	82	52	0.6846	0.7739	0.7423	0.7265	
ResNet50V2_15	185	201	75	59	0.7115	0.7582	0.7423	0.7341	
ResNet50V2_16	165	219	95	41	0.6346	0.8010	0.7385	0.7082	5
ResNet50V2_17	185	211	75	49	0.7115	0.7906	0.7615	0.7490	
ResNet50V2_18	227	184	33	76	0.8731	0.7492	0.7904	0.8064	
ResNet50V2_19	182	203	78	57	0.7000	0.7615	0.7404	0.7295	
ResNet50V2_20	190	196	70	64	0.7308	0.7480	0.7423	0.7393	

ตารางที่ ก.3 ผลลัพธ์ประสิทธิภาพการจำแนกของโมเดล CNN ที่ใช้สถาปัตยกรรม VGG16
จำนวน 20 โมเดล เพื่อใช้ในการเปรียบเทียบ

Model	TP	TN	FP	FN	Precision	Recall	Accuracy	F1	Epoch
VGG16_1	169	157	91	103	0.6500	0.6213	0.6269	0.6353	20
VGG16_2	171	148	89	112	0.6577	0.6042	0.6135	0.6298	
VGG16_3	175	154	85	106	0.6731	0.6228	0.6327	0.6470	
VGG16_4	163	162	97	98	0.6269	0.6245	0.6250	0.6257	
VGG16_5	149	171	111	89	0.5731	0.6261	0.6154	0.5984	
VGG16_6	171	156	89	104	0.6577	0.6218	0.6288	0.6393	15
VGG16_7	186	143	74	117	0.7154	0.6139	0.6327	0.6607	
VGG16_8	158	161	102	99	0.6077	0.6148	0.6135	0.6112	
VGG16_9	197	135	63	125	0.7577	0.6118	0.6385	0.6770	
VGG16_10	140	180	120	80	0.5385	0.6364	0.6154	0.5833	
VGG16_11	157	166	103	94	0.6038	0.6255	0.6212	0.6145	10
VGG16_12	156	168	104	92	0.6000	0.6290	0.6231	0.6142	
VGG16_13	203	130	57	130	0.7808	0.6096	0.6404	0.6847	
VGG16_14	173	159	87	101	0.6654	0.6314	0.6385	0.6479	
VGG16_15	200	133	60	127	0.7692	0.6116	0.6404	0.6814	
VGG16_16	160	155	100	105	0.6154	0.6038	0.6058	0.6095	5
VGG16_17	187	143	73	117	0.7192	0.6151	0.6346	0.6631	
VGG16_18	118	194	142	66	0.4538	0.6413	0.6000	0.5315	
VGG16_19	181	140	79	120	0.6962	0.6013	0.6173	0.6453	
VGG16_20	146	160	114	100	0.5615	0.5935	0.5885	0.5771	

ภาคผนวก ข.
Conference Paper



Vision-Based Drug Classification System using Transfer Learning Approach

Patpon Laohakijvitoon
Master of Engineering Program in Engineering Technology
Thai-Nichi Institute of Technology
Bangkok, Thailand
patpon23@gmail.com

Sivapong Nilwong
Department of Computer Engineering and Artificial Intelligence
Thai-Nichi Institute of Technology
Bangkok, Thailand
sivapong@tni.ac.th

Abstract—Dispensing generic drugs instead of original drugs from pharmacies may cause adverse drug events to patients. It is difficult to separate between generic drugs and originals due to their similarities, except for imprints on the tablets. In this paper, a vision-based medicine tablets classification system using Convolutional Neural Networks (CNN) with a transfer learning approach is proposed. The proposed classification system aims to help its user distinguish between the original drugs and generic drugs based on imprints on the medicine tablets. Multiple CNN models are created and tested as the classifier for the drug classification system. Three well-known CNN models are chosen as base models for creating CNNs with the transfer learning, including VGG16, Inception-V3, and ResNet50-V2. The developed CNN models are tested with images of medicine tablet taken from smartphones. Results from the experiments are evaluated through accuracies, precisions, recalls, and F-1 scores. Experimental results show satisfying performances of the proposed system in classifying original drugs and generic drugs using visual information.

Keywords—medicine tablet, convolutional neural network, transfer learning, image classification, deep learning

I. INTRODUCTION

In the present world, people in some countries can purchase medicines directly from pharmacies and private clinics [1]. Dispensing medicines to patients can have many possible errors such as missing or incorrect information, incorrect strength, and incorrect dosage [2]. These problems may result in varieties of problems which could be fatal. For the problem of missing information, some pharmacies can take an opportunity in dispensing the generic drugs instead of the originals. Generic drugs are almost the same as their original counterparts in both visuals and ingredients, though they are around 85 percents cheaper than original ones [3]. Despite containing the same ingredients and bioequivalent values of their originals [4], adverse effects may occur because some generic drugs are not tested for the therapeutic safety efficacy [5]. Generic drugs may also be contaminated during the manufacturing process, such as the incident in [6].

Recent research employed different approaches to detect or classify drugs in various applications. For example, the detection of counterfeit medicines in [7] and [8]. Deep learning algorithms were also employed for medicines detection and classification, such as the detection of traditional Chinese medicines using YOLOv8 [9]. Convolution Neural Network (CNN) is one of deep learning algorithms which have been widely utilized for vision-based tasks. Structures of

CNNs mimic the human visual system, allow them to process visual data better than other algorithms. Implementation of CNN-based algorithms also appeared in the field of medical, such as in [9]. Some other examples include the Vietnamese herbal medicine recognition using a CNN that combined from VGG16 and Xception models [10], the detection of lung cancer [11], and the COVID-19 detection from photographs of chest X-ray using the VGG16 model [12].

From the success in various applications, CNNs have promising potentials for classifying original drugs and generic drugs from their appearances. Although these two types of drugs are greatly similar to each other, there are some visual differences that can be observed. One of the differences is the imprint on the tablet, which may be difficult for some patients. Despite performances of CNN models, the creation of CNNs can consume time and computational power. However, there is a transfer learning technique that can ease the CNN creation by reusing existing CNN models [13]. Therefore, this paper presents a vision-based drug classification system based on the CNN with a transfer learning approach. The proposed system consists of a smartphone and a laptop with CNN inside. The smartphone is used as a camera to take a picture and send it to the laptop. The laptop classifies the acquired picture using the CNN as one of two classes: original drug and generic drug. Multiple CNN models were created through the transfer learning approach. Each CNN model used one of three existing models as the base, including VGG16 [14], Inception-V3 [15], and ResNet50-V2 [16]. The created CNN models were trained using images gathered from different medicine tablets. The classification system was evaluated using additional medicine images taken from the smartphone. Satisfying results were achieved from the experiments.

Remaining sections of this paper are divided into sections as follows: Section 2 describes the proposed generic drug classification system and its components, Section 3 explains the implemented transfer learning approach, Section 4 displays the experiments and the results, and Section 5 concludes the paper and discusses future works.

II. DRUG CLASSIFICATION SYSTEM

The proposed drug classification system aims to classify images of original drugs and generic drugs which are differed mostly by their imprints. The system hardware includes a smartphone and a laptop. The smartphone is employed as the camera and the laptop is used to compute the classification algorithm based on the convolutional neural networks.

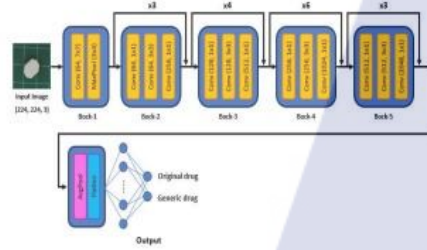


Fig. 5. Architecture of the implemented ResNet50-V2 model.

Inception-V3 is also one of the most used CNN architectures for image analysis and object detection. It was developed from Inception-V2, in which the internal structure was reduced to five steps, or five Inception modules [15]. Inception-V3 has fewer parameters than its predecessor, yet remains effective due to the improved size of convolutional layers and max-pooling layers. There is also an implementation of 1×1 convolutional filters in some layers to reshape the output sizes and dimensions. The Inception-V3 model is employed for the transfer learning approach with the input size of $224 \times 224 \times 3$, same as the VGG16. Architecture of the implemented Inception-V3 model is visualized in Fig. 4.

ResNet50-V2 is a CNN architecture developed from the ResNet50. Its design aims to solve the problem of vanishing gradients that often occur with relatively deep neural networks [16]. The number of layers in ResNet50-V2 is 50 layers, as the name suggested. Connections between some layers in ResNet50-V2 are applied with a design technique called skip connections, in which some levels in the network are skipped. The ResNet50-V2 model was applied for the transfer learning approach with the same input size as in the employed VGG16 and Inception-V3. The employed architecture of ResNet50-V2 is displayed as a diagram in Fig. 5.

III. IMPLEMENTED TRANSFER LEARNING APPROACH

Transfer learning is a technique that can reduce the time-consuming process of the CNN training. Typically, training a CNN model require a significant amount of time and computational power. These problems usually caused by the increased complexity and depth of the model, since the number of layers and parameters in CNN models are significantly large. Transfer learning can solve these problems by reusing parts of the CNN model that have been pre-trained for similar tasks, to create a new CNN model [13]. The new model can be trained using new data for its corresponding tasks, without the need for weight initialization, as described in Fig. 6. Transfer learning technique is also viable for training a CNN model with a small amount of data in the dataset.

The CNNs applied for the drug classification system was built and trained using the transfer learning technique. There are various transfer learning approaches available. The implemented transfer learning approach takes one of three selected models as described in section 2 of this paper, to create a new CNN. The parts in pre-trained model with convolutional layers for feature extraction were frozen. The trained parts were dense layers at the end of the network. Each of the models created through the applied approach was

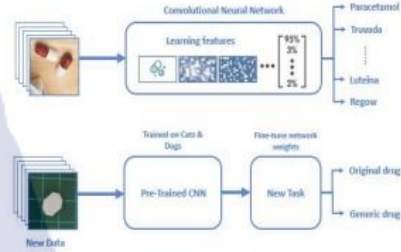


Fig. 6. Differences between, top: conventional training of CNNs, and bottom: implementation of the transfer learning.

redesigned to have an input layer with the size of $224 \times 224 \times 3$, and two dense layers, while other parts of the model remained the same as its pre-trained counterpart. The first dense layer contains 128 neurons, and the second dense layer contains 2 neurons. Two neurons in the second dense layer refer to two output classes: original drug, and generic drug.

The created CNNs were trained with the drug tablets image dataset, as mentioned in section 2. Hyperparameter settings applied the Adam optimizer, the batch size of 64, and using the sparse categorical cross entropy as the loss function. The number of epochs were varied among the created models. There were four variations of the number of epochs employed for the training, which were 5, 10, 15, and 20 epochs.

IV. EXPERIMENTAL RESULTS

There was a great number of CNN models created and trained through the implemented transfer learning approach. In order to measure the performances of these models, a separated image dataset of drug tablets was utilized for the experiment. The dataset for the experiments includes 130 colored images of drug tablets with different sizes and lightings. The experiments were conducted by having each CNN model classify all of 130 images in the dataset, then measure the performance through the classification results.

There were four measures implemented to evaluate performances from classification results of the CNN models, including accuracy, precision, recall, and F-1 score. In this work, the positive classification results refer to the results of the original drugs, while the negative results refer to results of the generic drugs. The highest possible value for all four applied evaluation measures is 1, and the lowest possible value is 0. Accuracy can measure the performance of a CNN model by calculating the ratio of correct results to total classification results. The accuracy can be calculated from the following equation:

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

where TP is the number of correct classification results for the original drug, TN is the number of correct results for the generic drug, FP is the number of wrong classification results that output as the original drugs, and FN is the number of wrong results that output as the generic drugs. The precision and recall values were also calculated from the same set of

TABLE I. RESULTS FROM DRUG CLASSIFICATION OF THE CNNs

Base Model	Epoch	Precision	Recall	Accuracy	F-1
VGG16	20	0.7231	0.6438	0.6615	0.6812
	15	0.6	0.6964	0.6692	0.6446
Inception-V3	15	0.9231	0.7229	0.7846	0.8108
	20	0.9077	0.7284	0.7846	0.8082
ResNet50-V2	5	0.9846	0.8205	0.8846	0.8951
	20	0.8462	0.9016	0.8769	0.8730

variables as in the calculation of the accuracy. Precision values were used to measure the ratio of the correct positive results to the total results of the total number of positive results. Precision can be calculated from the equation described mathematically as:

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (2)$$

where TP , TN , FP , and FN are the same as the accuracy equation in (1). The recalls measure the amount of correct positive results comparing to actual labels in the tested dataset. Recall can be calculated from the following equation:

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (3)$$

where TP , TN , FP , and FN are the same as in (1) and (2). Performances of the created CNN models were also evaluated through the F-1 scores. F-1 score is a metric for measuring classification performances, in which the F-1 represents both precision and recall in one metric. F-1 score can be calculated from the precision and recall as in the following equation:

$$F = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where F refers to the calculated F-1 score, $Precision$ is the result from (2), and $Recall$ is the result from (3). F-1 scores can be focused more during the result inspection than precisions and recalls, since they are used to calculate the F-1.

As mentioned, there was a great number of CNNs trained and tested. Therefore, the top two CNN models with best performances from each of three base pre-trained models are shown in the result section for better visuals instead. In total, there are six models displayed as the result in Table I. From results of the six models, results from CNN models which transferred from the ResNet50-V2 acquired the best performances among three base models. Among the results of ResNet50-V2-based models, the model that utilized the number of epochs of 5 gave the better results than the 20 epochs variation, with the F-1 score of 0.8951 and the accuracy of 0.8846. The worst of three base models was the VGG16, in which the best model that transferred from VGG16 achieved the F-1 score of 0.6812 and the accuracy of 0.6615.

It can be seen from the results that CNN models that transferred from ResNet50-V2 achieved the highest values in all measures. The results pointed that ResNet50-V2 is the most suitable CNN model for the applied transfer learning

with the drug classification task in this study. There is a notable point in the results where some results from the model with lower epoch settings were better than the higher ones. This was probably caused by the randomized weight during the initialization. However, the causes of this issue could be investigated further in future works.

CONCLUSION

This paper presented a drug classification system which can distinguish images of medicine tablets using visual information. The proposed system employed a CNN to classify medicine images into two classes: original drug and generic drug. A smartphone was used to acquire an image to be used as an input for the system, with slight preprocessing prior the classification with the CNN. Multiple CNN models were developed through the transfer learning approach, in which each model utilized one of three existing models as its base, including VGG16, Inception-V3, and ResNet50-V2. The developed CNN models were tested with an additional dataset gathered from smartphones. Satisfying results were achieved from the experiments, in which the models based on ResNet50-V2 performed the best. From the results, the vision-based system proved to be viable for classifying between generic drugs and original drugs using their imprints.

Future works can continue with various improvements and applications. Improvements of this work can include implementations of other CNN models built through different techniques, experiments with more diverse dataset, and the utilization of image processing algorithms to reduce effects of environmental conditions. This work can be further applied into different applications such as a mobile application for assisting patients in classifying the received medicine tablets from healthcare services, a system for detecting counterfeit drugs, and an application which can help organizing medicine tablets in pharmacies.

REFERENCES

- [1] "Hospital Operator THG in Pharmacy Push." Bangkokpost.com. <https://www.bangkokpost.com/business/general/2684289/hospital-operator-thg-in-pharmacy-push> (accessed Dec. 1, 2023).
- [2] I. S. Um, A. Clough, and E. C.K. Tan, "Dispensing error rates in pharmacy: A systematic review and meta-analysis," *Research in Social and Administrative Pharmacy*, vol. 20, issue 1, 2024, pp.1-9.
- [3] M. Ghoshal, "Generic vs. Brand-name Drugs: Is There a Difference?," Healthline.com. <https://www.healthline.com/health/drugs/generic-vs-brand> (accessed Dec. 1, 2023).
- [4] "Generic Drug Facts." FDA Homepage. <https://www.fda.gov/drugs/generic-drugs/generic-drug-facts> (accessed Nov. 28, 2023).
- [5] K. Uhl, J. R. Peters, "How the FDA Ensures High-Quality Generic Drugs," *Am Fam Physician*, vol. 97, issue 11, 2018, pp. 696-697.
- [6] D. Jimenez, "Generic drug safety: US regulators struggle to keep up with a global market." pharmaceutical-technology.com. <https://www.pharmaceutical-technology.com/features/generic-drug-safety-us-regulators-struggle-global-market> (accessed Nov. 24, 2023).
- [7] S. R. Shinde, K. Bhavsar, S. Kimbature, S. Khandelwal, A. Ghose and A. Pal, "Detection of Counterfeit Medicines Using Hyperspectral Sensing," *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, 2020, pp. 6155-6158.
- [8] M. Alsallal, M. S. Sharif, B. Al-Ghazawi and S. M. Mkat al Mutoki, "A Machine Learning Technique to Detect Counterfeit Medicine Based on X-Ray Fluorescence Analyser," *2018 International Conference on Computing, Electronics & Communications Engineering (iCECE)*, Southend, UK, 2018, pp. 118-122.

- 
- [9] Y. Su, B. Cheng and Y. Cai, "Detection and Recognition of Traditional Chinese Medicine Slice Based on YOLOv8," *2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT)*, Qingdao, China, 2023, pp. 214-217.
 - [10] T. N. Quoc and V. T. Hoang, "Traditional Vietnamese Herbal Medicine Image Recognition by CNN," *2023 15th International Conference on Knowledge and Smart Technology (KST)*, Phuket, Thailand, 2023, pp. 1-5.
 - [11] Tasmim et al., "A Lightweight-CNN Model for Efficient Lung Cancer Detection and Grad-CAM Visualization," *2023 International Conference on Information and Communication Technology for Sustainable Development (ICT4SD)*, Dhaka, Bangladesh, 2023, pp. 254-258.
 - [12] K. S. Prasad, S. Pasupathy, P. Chinnasamy and A. Kalaierasi, "An Approach to Detect COVID-19 Disease from CT Scan Images using CNN - VGG16 Model," *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2022, pp. 1-5.
 - [13] L. Torrey, and J. Shavlik, "Transfer Learning," *Transfer Learning. Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, IGI Global, 2010, pp. 242-264.
 - [14] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations (ICLR 2015)*, 2015, pp. 1-14.
 - [15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2818-2826.
 - [16] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," *Computer Vision – ECCV 2016*, vol. 9908, 2016, pp. 630-645.

ประวัติย่อผู้วิจัย



ชื่อ – สกุล

นายภัทรภณ เลาหกิจวิฑูร

วัน – เดือน – ปี เกิด

14 กุมภาพันธ์ พ.ศ. 2539

ที่อยู่ปัจจุบัน

95 ปากซอยรามคำแหง 203 ถนนรามคำแหง
แขวงมีนบุรี เขตมีนบุรี กรุงเทพมหานคร 10510

โทรศัพท์มือถือ : 086-045-0400

E-mail : la.patpon_st@tni.ac.th

ประวัติการศึกษา

พ.ศ. 2566

วิศวกรรมศาสตรมหาบัณฑิต สาขาเทคโนโลยีวิศวกรรม
สถาบันเทคโนโลยีไทย-ญี่ปุ่น

พ.ศ. 2562

วิศวกรรมศาสตร สาขาวิศวกรรมอุตสาหกรรม
สถาบันเทคโนโลยีไทย-ญี่ปุ่น

ประวัติการทำงาน

พ.ศ. 2566 – ปัจจุบัน

วิศวกรแผนก IE ฝ่ายเทคโนโลยีการผลิต
บริษัท วาโก้ศรีราชา จำกัด

พ.ศ. 2562

วิศวกรแผนกออกแบบ
บริษัท ไทย โคะอิโตะ จำกัด