โครงการ

การเรียนการสอนเพื่อส่งเสริมประสบการณ์

ชื่อโครงการ	แบบจำลองระบบแนะนำไทยอาหารโดยใช้ Deep Learning และ Recurrent Neural Network	
ผู้ดำเนินงาน	ชื่อ นาย สมเกียรติ เลี้ยงหิรัญ	เลขประจำตัว 5932641823
	ชื่อ นางสาว ชนิดาภา ธรรมธัชวงศ์	เลขประจำตัว 5932603023
ภาควิชา	เทคโนโลยีทางภาพและการพิมพ์	
ปีการศึกษา	2562	

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การเรียนการสอนเพื่อส่งเสริมประสบการณ์

ปีการศึกษา 2562

แบบจำลองระบบแนะนำไทยอาหารโดยใช้ Deep Learning และ Recurrent Neural Network

โดย

นาย สมเกียรติ เลี้ยงหิรัญ

นางสาว ชนิดาภา ธรรมธัชวงศ์

รายงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตร์บัณฑิต

ภาควิชาเทคโนโลยีทางภาพและการพิมพ์ คณะวิทยาศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2562

เนื่องจากสถานการณ์ระบาดของโรคไวรัส COVID-19

ในช่วงภาคปลายของปีการศึกษา 2562

จึงส่งผลให้ไม่สามารถดำเนินการได้ครบถ้วนสมบูรณ์ตามวัตถุประสงค์ของโครงงาน

หัวข้อโครงงาน	แบบจำลองระบบแนะนำไทยอาหารโดยใช้ Deep Learning และ Recurrent Neural Network
นิสิตผู้ดำเนินโครงการ	นาย สมเกียรติ เลี้ยงหิรัญ
	นางสาว ชนิดาภา ธรรมธัชวงศ์
สาขาวิชา	ภาควิชาเทคโนโลยีทางภาพและการพิมพ์
อาจารย์ที่ปรึกษา	ร.ศ.ดร. อรัญ หาญสืบสาย

ภาควิชาเทคโนโลยีทางภาพและการพิมพ์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ยอมรับรายงานวิทยาศาสตร์นี้ เป็นส่วนหนึ่งของการศึกษาระดับปริญญาตรี

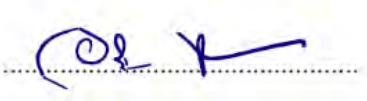
วันที่ส่ง.....

วันที่อนุมัติ.....

.....

หัวหน้าภาควิชาฯ

(รองศาสตราจารย์ดร.พิชญดา เกตุเมฆ)



อาจารย์ที่ปรึกษาโครงการ

(ร.ศ.ดร. อรัญ หาญสืบสาย)

ผู้วิจัย นาย สมเกียรติ เลี้ยงหิรัญ รหัสนิสิต 5932641823
นางสาว ชนิดาภา ธรรมธัชวงศ์ รหัสนิสิต 5932603023

ชื่อเรื่อง แบบจำลองระบบแนะนำไทยอาหารโดยใช้ Deep Learning และ Recurrent Neural Network

อาจารย์ที่ปรึกษา ร.ศ.ดร. อรัญ หาญสืบสาย

บทคัดย่อ

การทดลองนี้เป็นการศึกษาและทดลองการสร้างระบบการแนะนำอาหารไทยโดยใช้โครงข่ายประสาทเทียมแบบ Deep Learning และ Recurrent Neural Network ระบบดังกล่าวแบ่งเป็น 2 ส่วนปฏิบัติการ ได้แก่ 1) ระบบเรียนรู้จดจำภาพอาหารไทย และ 2) ระบบแนะนำอาหารไทย โดยทำการคัดเลือกอาหารไทยและเครื่องดื่มที่มีความเป็นสามัญและได้รับความนิยม จำนวนรวมทั้งสิ้น 25 ชนิด นำเข้าเป็นฐานข้อมูลรูปภาพของระบบ เพื่อให้ระบบโครงข่ายประสาทเทียมเชิงลึกทำการเรียนรู้และจดจำภาพอาหาร สร้างความสามารถในการระบุชนิดของอาหารนั้น ๆ ให้แก่โครงข่ายประสาทเทียมเชิงลึก โดยนำโครงข่ายประสาทเทียมเชิงลึกที่เกิดการเรียนรู้ไปใช้ร่วมกับ Recurrent Neural Network ทำการสร้างระบบจำลองการจับคู่แนะนำอาหารจากชนิดอาหารที่มีอยู่ในฐานข้อมูล ผลการทดลองที่ได้ทำให้ทราบถึงค่าร้อยละความแม่นยำในการระบุและแนะนำอาหารไทยของโครงข่ายประสาทเทียมที่เกิดการเรียนรู้ และทราบถึงแนวทางในการลดค่าความผิดพลาดที่เกิดขึ้น

ภาควิชาเทคโนโลยีทางภาพและการพิมพ์ ลายมือชื่อนิสิตผู้ดำเนินงาน.....

คณะวิทยาศาสตร์ ลายมือชื่อนิสิตผู้ดำเนินงาน.....

ปีการศึกษา 2562 ลายมือชื่ออาจารย์ที่ปรึกษาหลัก.....

Research Student	Mr. Somkiet leanghirun	ID No. 5932641823
	Ms. Chanidapha Dhammatouchwong	ID No. 5932603023
Title	Thai food recommendation model using Deep Learning and Recurrent Neural Network	
Project Advisor	Associate Professor Dr. Aran Hansuebsai	

Abstract

The idea of this study is to research and to creates Thai food recommendation model using Deep Learning and Recurrent Neural Network. The model is consisted of two operational systems which are 1) Thai food Images recognition system and 2) Thai food recommendation system. Model's image database is inputted with 25 types of commonly known and popular Thai foods and drinks. Deep learning model generated by MatLab coding program is trained using the model's image database in order to be learns each image' s attributes and able to identify each specific food item. The trained recognition model is used in collaborate with Recurrent Neural Network to creates Thai food recommendation model. Due to model's image database, the recommendation model is capable of selecting food item compatible the one inputted into the system. The results of model's testing show accuracy percentages of food item identification and food item recommendation which suggest the way for developing the model for practical use by reducing founded errors.

Department of Imaging and Printing Technology

Academic Year 2019

Student Signature.....

Student Signature.....

Project Advisor Signature.....

กิตติกรรมประกาศ

ขอขอบพระคุณ ร.ศ.ดร. อรัญ หาญสืบสาย อาจารย์ที่ปรึกษาโครงการผู้ซึ่งคอยให้ความช่วยเหลือดูแลและให้คำแนะนำที่มีค่าต่าง ๆ ในโครงการนี้เสมอมา จึงทำให้โครงการนี้ประสบผลจนลุล่วงไปได้ด้วยดี

ขอขอบพระคุณ ร้านอาหารภายในรั้วจุฬาลงกรณ์มหาวิทยาลัย โดยเฉพาะร้านอาหารราดแกงร้านสี่ประจำหอพักนิสิตจุฬาฯ ผู้ให้ความอนุเคราะห์ในการเก็บภาพถ่ายอาหารประเภทต่าง ๆ เพื่อใช้ในการรวบรวมข้อมูลและคัดเลือกภาพอาหารไทยได้รับความนิยมและมีความเป็นสามัญ รวมถึงขอขอบพระคุณทุกผู้ที่ให้การอนุเคราะห์แบ่งปันภาพถ่ายอาหารที่ใช้ในโครงการนี้

ผู้วิจัยขอขอบพระคุณครอบครัวของผู้วิจัยที่คอยให้โอกาสและให้กำลังใจเสมอมา

ผู้วิจัย

สารบัญ

บทคัดย่อ.....5
Abstract.....6
กิตติกรรมประกาศ.....7
บทที่ 1.....9
1.1 ความเป็นมาและมูลเหตุจูงใจในการเสนอโครงการ.....9
1.2 วัตถุประสงค์ของโครงการ.....10
1.3 ขอบข่ายการศึกษา.....10
1.4 ผลที่คาดว่าจะได้รับ.....10
บทที่ 2.....11
2.1 ทฤษฎี.....12
2.1.1 Deep learning.....12
2.1.2 Recommendation system.....14
2.1.3 ระบบสี.....16
2.2 งานวิจัยที่เกี่ยวข้อง.....21
2.2.1 งานวิจัยที่เกี่ยวข้อง (1).....21
2.2.2 งานวิจัยที่เกี่ยวข้อง (2).....22
บทที่ 3.....25
3.1 วัสดุอุปกรณ์.....25
3.2 วิธีการดำเนินงาน.....25
บทที่ 4.....29
4.1 ผลการทดลองค่าความแม่นยำของการจำแนกรูปภาพ.....29
4.2 ผลค่าความแม่นยำในการแนะนำอาหาร.....30
4.3 วิเคราะห์ผลการทดลอง.....33
บทที่ 5.....35
5.1 สรุปผลการทดลอง.....35
บรรณานุกรม.....36

บทที่ 1

บทนำ

1.1 ความเป็นมาและมูลเหตุจูงใจในการเสนอโครงการ

ปัจจุบันมีนักท่องเที่ยวจำนวนมากที่ไม่รู้ว่าอาหารไทยชนิดไหนที่ได้รับความนิยมและควรทานร่วมกัน เนื่องจากวัฒนธรรมอาหารไทยเป็นหนึ่งในประเทศที่มีวัฒนธรรมการรับประทานอาหารที่ทานอาหารร่วมกันหลายชนิด ยกตัวอย่าง เช่น น้ำพริกกับปลาทู กระเพราไก่กับไข่ดาว ซึ่งในมุมมองคนไทยอาจมองว่าเป็นเรื่องปกติที่จะรับประทานอาหารเหล่านี้ร่วมกัน อย่างไรก็ตามสำหรับชาวต่างชาติซึ่งไม่คุ้นเคยกับวัฒนธรรมไทยมาก่อนจะมองว่าเป็นเรื่องแปลกใหม่ และอาจเกิดความงุนงงในการเลือกคู่อาหารเพื่อรับประทาน สิ่งนี้เองจึงเป็นแรงจูงใจให้ผู้วิจัยเกิดแนวคิดที่จะอำนวยความสะดวกในการเลือกอาหาร จึงเกิดเป็นโครงการการสร้างแบบจำลองระบบอาหารไทยนี้ขึ้น

ผู้จัดทำพยายามสร้างแบบจำลองระบบแนะนำอาหาร เพื่อเป็นแนวทางในการพัฒนาเพื่อใช้ในงานบริการ ไม่เพียงแต่สำหรับชาวต่างชาติที่ไม่คุ้นเคยกับอาหารไทยดั้งเดิม แต่รวมไปถึงทุกบุคคลที่ต้องการการแนะนำอาหารในเบื้องต้น ในหลากหลายโอกาส อาทิ เช่น เมื่อมีโอกาสรับประทานอาหารในภัตตาคารใหม่ที่ไม่คุ้นชิน หรือการได้เลือกชิมอาหารพื้นเมืองบางประเภทในเขตภูมิภาคต่าง ๆ ที่มีหลากหลายและแตกต่างกันไปในแต่ละภาคส่วนของประเทศไทย

ในปัจจุบันมีวิธีหลากหลายในการรวบรวมข้อมูลดิบเพื่อนำไปวิเคราะห์ผลข้อมูลในด้านที่อยู่ในความสนใจ ไม่ว่าจะเป็น Point of Sale (PoS) คือ ระบบขายหน้าร้าน ซึ่งเป็นการเก็บข้อมูลการซื้อขายและเคลื่อนไหวของสินค้าจากระบบการจ่ายซื้อสินค้าหน้าร้านหรือระบบแคชเชียร์ หรืออีกวิธีการหนึ่งคือ การใช้ Internet of Thing (IOT) ในการเก็บข้อมูลด้านต่าง ๆ ซึ่งเป็นกระบวนการที่ได้รับความสนใจเป็นอย่างมากในช่วงหลายปีที่ผ่านมา จากความก้าวหน้าของระบบการสื่อสารแบบไร้สายและการเข้าถึงเทคโนโลยีอย่างทั่วถึง โทรศัพท์มือถือ คอมพิวเตอร์ และอินเตอร์เน็ต กลายเป็นส่วนหนึ่งในชีวิตประจำวันของผู้คนในสังคม การใช้แอฟลิเคชั่นซื้อขายของออนไลน์ แอฟลิเคชั่นบริการสั่งและส่งอาหาร หรือการใช้โซเชียลเน็ตเวิร์คถ่ายทอดเรื่องราวต่าง ๆ ในชีวิตประจำวัน ล้วนแล้วแต่เป็นการส่งถ่ายข้อมูลดิจิตอลซึ่งสามารถนำไปศึกษาหารูปแบบพฤติกรรมการเลือกซื้อสินค้าและบริการของผู้คนได้ทั้งสิ้น

จากพฤติกรรมการใช้เทคโนโลยีดังกล่าวในข้างต้น การวิเคราะห์รวบรวมข้อมูลดิจิตอล เช่น ข้อมูลจากไฟล์ภาพ จึงเป็นแนวทางที่ทำให้เก็บข้อมูลได้มาก รวดเร็ว และตรงตามพฤติกรรมการเลือกของผู้คน โดยแนวทางที่ผู้จัดทำโครงการเลือกใช้คือการใช้โครงข่ายประสาทเทียมเชิงลึก (Deep learning) ซึ่งเป็นปัญญาประดิษฐ์แขนงหนึ่ง ที่ได้รับการศึกษาแล้วว่ามีประสิทธิภาพในการจดจำข้อมูลอาหารโดยใช้เทคนิคการรับรู้จากภาพถ่ายได้ดี (Image recognition) [1] ผู้จัดทำโครงการจึงมุ่งหวังสร้างระบบโครงข่ายประสาทเทียมเชิงลึกที่สามารถทำการจดจำข้อมูลอาหารจากไฟล์ภาพถ่ายที่รวบรวบมาจากแหล่งดิจิตอลต่าง ๆ และสามารถทำการเลือกแนะนำอาหารที่ควรรับประทานร่วมกับข้อมูลภาพถ่ายที่ป้อนเข้าไปได้ โดยสร้างระบบกระบวนการ

แนะนำจาก Recurrent neural network ซึ่งเป็นโครงข่ายปัญญาประดิษฐ์อีกแขนงหนึ่งซึ่งมีความสามารถในการเลือกเรียงลำดับข้อมูลได้อย่างมีประสิทธิภาพ

1.2 วัตถุประสงค์ของโครงการ

เพื่อศึกษาเทคนิคการจดจำภาพถ่ายดวยการเรียนรู้แบบโครงข่ายประสาทเทียมเชิงลึก เพื่อสร้างและศึกษาระบบการแนะนำอาหารโดยใช้ RNN และเพื่อทดสอบความแม่นยำในการใช้งานระบบจำลองที่สร้างขึ้นสำหรับการนำไปประยุกต์ใช้งานจริง

1.3 ขอบข่ายการศึกษา

ฝึกโครงข่ายประสาทเทียม โดยทำการเรียนรู้จดจำภาพถ่ายและสร้างระบบแนะนำอาหารไทยจากข้อมูลอาหารไทยที่มีความเป็นสามัญและเป็นที่นิยมที่ได้จากการสำรวจร้านอาหารนิสิตในจุฬาลงกรณ์มหาวิทยาลัย จำนวนรวมทั้งสิ้น 25 ชนิด

1.4 ผลที่คาดว่าจะได้รับ

ได้แบบจำลองระบบแนะนำอาหารไทยโดยใช้ Deep Learning และ Recurrent Neural Network ทั้งยังได้ศึกษาการเขียนโปรแกรมเกี่ยวกับ Deep learning และ Recurrent Neural Network และได้ช่วยให้นักท่องเที่ยวตัดสินใจได้ง่ายขึ้นในการเลือกรับประทานอาหารไทย

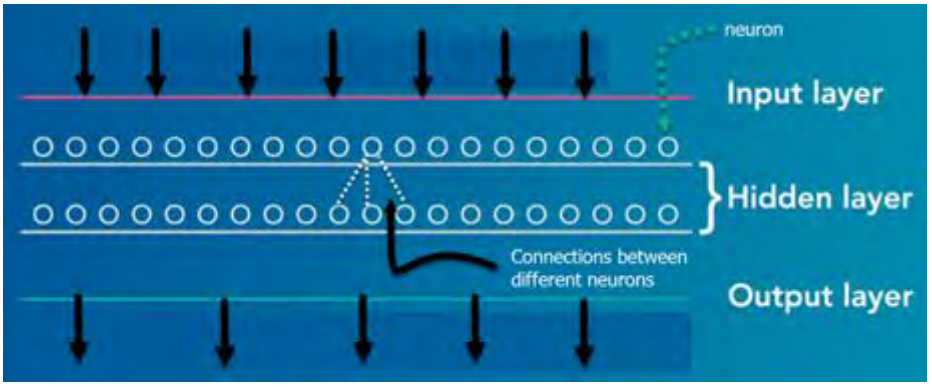
บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎี

2.1.1 Deep learning

Deep learning คือ ซอฟต์แวร์คอมพิวเตอร์ที่เลียนแบบการทำงานของระบบโครงข่ายประสาท (neurons) ในสมองมนุษย์ ซึ่งเป็นซับเซ็ตของ Machine learning นอกจากนี้หลักการทำงานของ Deep learning ยังถูกสร้างขึ้นมาจากการนำเอา Neural network หลายๆชั้นมาต่อกัน โดยชั้นแรกสุดจะทำหน้าที่ในการรับข้อมูล (Input layer) ชั้นสุดท้ายจะจะทำหน้าที่ในการส่งผลลัพธ์ของการประมวลผลออกมา (Output layer) ส่วนชั้นระหว่างชั้นแรกและชั้นสุดท้ายจะถูกเรียกว่า Hidden layer ซึ่ง Hidden layer ของแต่ละชั้นจะเปรียบว่าเสมือนสร้างจากเซลล์ประสาทจำนวนมาก ซึ่งมีหน้าที่ประมวลผลและรับข้อมูลจากชั้นที่อยู่สูงกว่าและส่งข้อมูลที่ถูกประมวลผลแล้วไปยังชั้นที่อยู่ต่ำกว่า ซึ่งข้อดีของการประมวลผลแบบนี้คือ ชั้นแต่ละชั้นสมามารถมีค่าถ่วงน้ำหนัก ค่าความเอนเอียง และวิธีประมวลผลทางคณิตศาสตร์ที่เป็นอิสระต่อกัน



รูป 2.1 แผนภาพการทำงานของ Deep learning

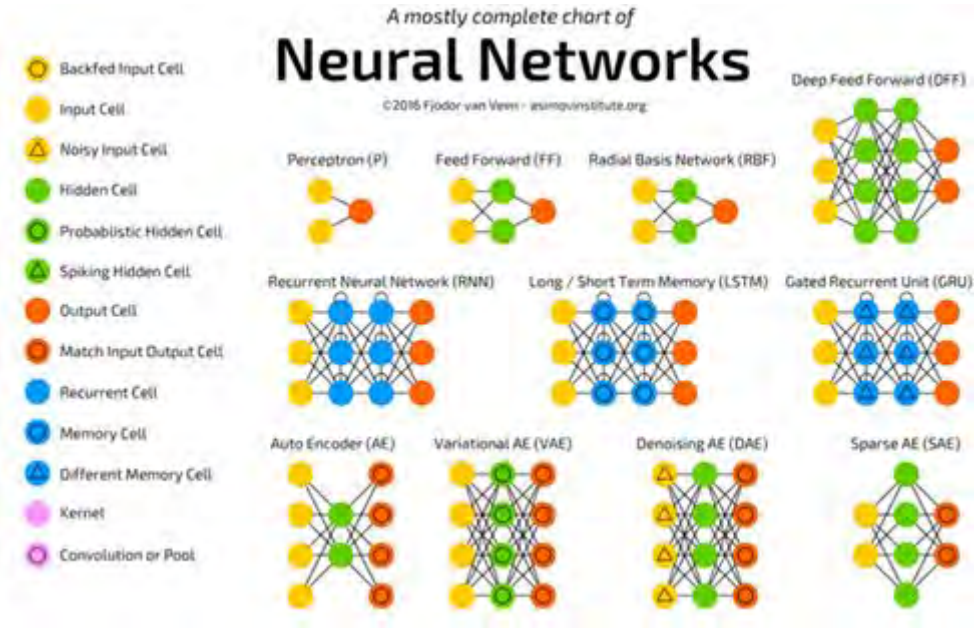
(ที่มา : https://www.thaiprogrammer.org/2018/12/deep-learning)

ลักษณะการเรียนรู้ของ Deep learning

การเรียนรู้ของ Deep learning แบ่งเป็นสองรอบ รอบที่หนึ่ง คือ การประยุกต์ใช้การแปลงแบบไม่เป็นเชิงเส้น (Nonlinear transformation) กับข้อมูลที่ได้รับ (Input) และได้ผลลัพธ์ (Output) ออกมาอยู่ในรูปของโมเดลทางสถิติ (Statistical model) รอบที่สอง คือ การนำโมเดลมาผ่านวิธีการทางคณิตศาสตร์ เช่น การหา Derivative โดยทั้งสองรอบนี้จะถูกทำซ้ำจนกว่าจะได้โมเดลที่ความแม่นยำ (Accuracy) ในระดับที่น่าพึงพอใจ ซึ่งการทำซ้ำของทั้งสองรอบนี้มีชื่อเรียกว่า Iteration โดยการทำซ้ำนี้เองยิ่งมีจำนวนการทำซ้ำที่เพิ่มมากขึ้นก็จะยิ่งทำให้ได้ผลลัพธ์ที่แม่นยำและเป็นที่พึงพอใจมากยิ่งขึ้น

ชนิดของโครงข่าย deep learning

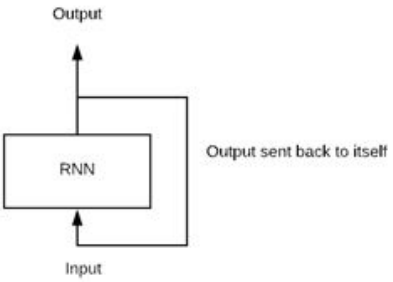
โครงข่ายประสาทแบบป้อนไปหน้า (Feed-forward neural networks) เป็นโมเดลที่มีโครงสร้างที่เรียบง่ายที่สุด เนื่องจากการดำเนินการของข้อมูลจะเป็นไปในทิศทางเดียว คือ รับข้อมูลจาก Input layer แล้วส่งไปต่อไปยัง Hidden layer เลื่อย ๆ จนกระทั้งถึง Output layer ก็จะหยุด ซึ่งจะสังเกตุได้ว่าจะไม่มีวงวนเกิดขึ้นเลย



รูปที่ 2.2 แผนภาพโครงข่ายประสาทชนิดต่าง ๆ

(ที่มา : https://www.guru99.com/deep-learning-tutorial.html)

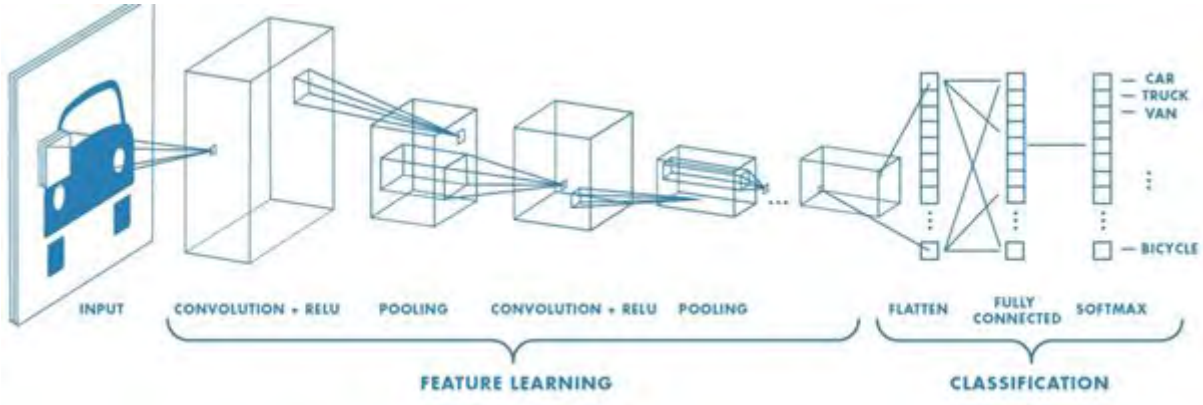
โครงข่ายแบบวนซ้ำ (Recurrent neural networks : RNN) คือ Neural networks แบบหลายชั้นที่สามารถเก็บข้อมูลไว้ที่ Node จึงทำให้มันสามารถรับข้อมูลเป็นแบบลำดับและให้ผลลัพธ์ออกเป็นลำดับของข้อมูลได้ กล่าวคือ RNN คือ Neural network ที่เชื่อมต่อกันหลาย ๆ อันและยังสามารถต่อกันเป็นวงวนได้ ดังนั้น RNN จึงเหมาะสมในการประมวลผลข้อมูลที่เป็นลำดับ



รูปที่ 2.3 แผนภาพกลไกการทำงานของโครงข่ายประสาทแบบวนซ้ำ

(ที่มา : https://www.guru99.com/deep-learning-tutorial.html)

โครงข่ายแบบคอนโวลูชัน (Convolutional neural networks : CNN) คือ โครงข่ายประสาทแบบหลายชั้นที่มีโครงสร้างเฉพาะตัว โดยถูกออกแบบมาเพื่อการเพิ่มความสามารถในการสกัดเอารูปแบบที่มีความซับซ้อนมากยิ่งขึ้นของข้อมูล ซึ่งนิยมใช้กับข้อมูลที่ไม่ค่อยเป็นระเบียบหรือไม่ได้มีโครงสร้างเป็นรูปแบบเฉพาะตัว ตัวอย่างเช่น รูปภาพ เป็นต้น ตัวอย่างการใช้งานโครงข่ายแบบคอนโวลูชัน เช่น การคิดคำบรรยายให้รูปภาพหรือการพยากรณ์รูปภาพเป็นต้น



รูปที่ 2.4 แผนภาพการทำงานของโครงข่ายประสาทแบบคอนโวลูชัน

(ที่มา : https://www.guru99.com/deep-learning-tutorial.html)

ข้อจำกัดของ Deep learning

1) ต้องการระบุชื่อหรือจัดประเภทข้อมูล (Data labeling) ปัจจุบัน AI ส่วนมากอาศัยการ train โดยใช้หลักการของ supervise learning ซึ่งนั้นก็หมายความว่า มนุษย์จะต้องตั้งชื่อและจัดประเภทของข้อมูล (Data labeling) ด้วยตัวเองก่อนจะนำไป train ซึ่งข้อมูลเหล่านี้ก็มีจำนวนมากจึงอาจเกิดข้อผิดพลาดขึ้นได้ ตัวอย่างเช่น รถยนต์ขับเคลื่อนอัตโนมัติ (Self - driving car) จ้างคนหลายร้อยคนเพื่อที่จะ labeling วิดีโอที่จะใช้ในการ train ระบบ
2) ต้องการชุดข้อมูลจำนวนมาก (Obtain huge training datasets) deep learning นั้นต้องข้อมูลจำนวนมากในการ Train ข้อมูล ตัวอย่างเช่น ต้องการข้อมูลจำนวน 1000 ตัวอย่าง เพื่อจะทำให้ได้โมเดล Classification ที่มีประสทธิภาพและในบางกรณีก็ต้องการมากกว่า 1 ล้านตัวอย่างเพื่อที่จะทำให้ Model ทำงานได้อย่างมีประสิทธิภาพใกล้เคียงกับมนุษย์ แน่นอนว่า Deep learning นั้นเป็นที่นิยมในบริษัทเทคโนโลยีชั้นนำทั่วโลก ซึ่งพวกเขาใช้ Big data ที่สะสมข้อมูลที่มีเนื้อที่ไม่น้อยกว่า Petabytes (1 ล้าน Gigabytes) นั้นเป็นเหตุผลว่าทำไม บริษัทเทคโนโลยีชั้นนำจึงสามารถสร้าง Deep learning model ที่มีประสิทธิภาพและมีความแม่นยำสูง

3) ปัญหาในการอธิบาย (Explain a problem) ไม่ใช่เรื่องง่ายที่มนุษย์เราจะอธิบาย AI model ที่ซับซ้อนได้ ตัวอย่างเช่น คำถามว่าที่ทำไม Model ถึงเกิดการตัดสินใจแบบนั้น ซึ่งเป็นเรื่องยากที่มนุษย์จะอธิบายการตัดสินใจของ AI ได้ ซึ่งก็เป็นอีกเหตุผลหนึ่งที่ทำให้ AI ได้รับการยอมรับช้า

2.1.2 Recommendation system

Recommendation system หรือ ระบบช่วยแนะนำ คือ ระบบที่ช่วยแนะนำในสิ่งที่เราชอบมาให้ โดยอ้างอิงจากผู้ใช้งานที่คล้ายคลึงกับเราและประวัติการใช้งานของเราเป็นพื้นฐาน โดยระบบช่วยแนะนำนี้จะใช้ Rating เป็นตัววัดความพึ่งพอใจของลูกค้าเพื่อมาประมวลผลการแนะนำสินค้า โดย Rating แบ่งได้สองประเภท คือ Explicit Rating และ Implicit Rating

Explicit Rating เป็นการที่ผู้ใช้บอกเราตรง ๆ เลยว่าชอบหรือไม่ชอบอะไร โดยส่วนมากที่มักจะพบเห็น คือ การให้คะแนนรีวิวต่าง ๆ เช่น ผู้ใช้ A ให้คะแนนรีวิวสินค้า B ด้วยคะแนน 8/10 คะแนน ซึ่งทำให้สามารถรู้ได้ว่าผู้ใช้ A นั้นค่อนข้างชอบสินค้า B หรือ ผู้ใช้ A ให้คะแนนสินค้า C เพียง 1/10 คะแนน ก็จะทำให้พอทราบได้ว่าผู้ใช้ A ไม่ชอบสินค้า C เป็นต้น

ข้อดี

- ผู้ใช้บอกได้โดยตรงว่าชอบหรือไม่
- สามารถบอกระดับความชอบหรือไม่ชอบจากผู้ใช้ได้

ข้อเสีย

- ผู้ใช้ส่วนใหญ่ไม่ได้ซื้อหรือใช้สินค้าทั้งหมดในระบบ โดยเฉพาะหากเป็นเว็บไซต์ที่มีสินค้าจำนวนมาก ๆ หรือบางครั้งผู้ใช้ได้ลองใช้สินค้าแล้วก็อาจจะไม่ได้เข้ามาให้คะแนนเป็นต้น
- ทำให้ข้อมูลในส่วนของ rating นี้มีจำนวนน้อยมาก ๆ

Implicit Rating นอกเหนือจาก explicit rating ที่ผู้ใช้บอกโดยตรงแล้ว ก็ยังมีช่องทางอื่นอีกที่พอจะสามารถเดา ได้ว่าผู้ใช้ชอบหรือมีความสนใจในสินค้าหรือไม่ ตัวอย่างเช่น การคลิกเข้าไปดูสินค้านั้น หรือการตัดสินใจที่จะซื้อสินค้านั้น หรือการที่ผู้ใช้กด favorite เป็นต้น ซึ่ง rating ประเภทนี้นั้น จะมีปริมาณข้อมูลที่เยอะมากกว่า explicit rating รวมไปถึงเมื่อมีผู้ใช้ใหม่เข้ามา ก็จะสามารถเก็บข้อมูลนี้ได้เลยทันทีเมื่อเค้าเริ่มคลิก แต่ feedback ที่ได้นั้นก็จะไม่ชัดเจนเท่า explicit rating

ข้อดี

- มีปริมาณข้อมูลมากเมื่อเทียบกับ explicit rating
- สามารถใช้ได้กับผู้ใช้ใหม่ที่เข้ามา ถึงแม้ว่าเค้าจะยังไม่เคยมีประวัติการซื้อมาก่อน

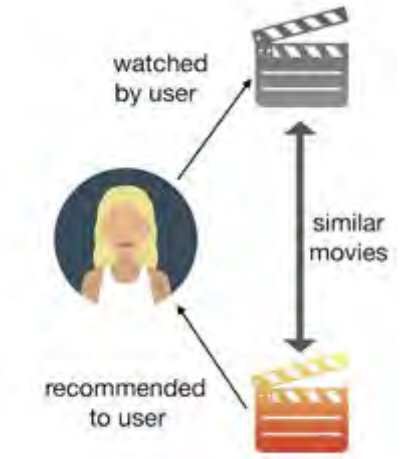
ข้อเสีย

- สามารถเดาได้แค่เพียงว่าผู้ใช้ชอบสิ่งไหน แต่ไม่มี negative feedback กลับมา เช่น หากผู้ใช้ไม่ได้คลิกสินค้า B ผู้ใช้อาจไม่ได้ชอบสินค้า B แต่เป็นเพราะว่าผู้ใช้ไม่ได้เห็นสินค้า B เป็นต้น

- ข้อมูลค่อนข้างแปรปรวน เช่น การที่ผู้ใช้ตัดสินใจที่จะซื้อสินค้า B กับเรา อาจจะเป็นเพราะว่าผู้ใช้นั้นซื้อให้ผู้อื่น ไม่ได้ซื้อให้ตัวเอง จึงบอกไม่ได้ว่าผู้ใช้ชอบสินค้า B หรือไม่ หรือการที่ผู้ใช้ไม่ได้ซื้อสินค้า C ก็ไม่ได้แปลว่าผู้ใช้จะไม่ชอบสินค้า C ผู้ใช้อาจจะชอบสินค้า C เหมือนกัน แต่มีสินค้า C อยู่แล้วที่บ้าน หรือผู้ใช้บางคนอาจจะคลิกไปมั่ว ๆ เล่น ๆ แต่ไม่ได้มีความสนใจในสินค้านั้นจริง ๆ ก็เป็นได้
- บอกปริมาณความชอบของผู้ใช้ไม่ได้แตกต่างกับ explicit rating เช่น การให้คะแนนรีวิวที่สามารถบอกได้ว่าผู้ใช้ชอบมากหรือน้อยเพียงใดจากระดับคะแนนรีวิวที่ให้มา

ประเภทของโมเดลที่ใช้ในการทำ recommendation system

1. Content-based Filtering คือ การนำประวัติของผู้ใช้คนนั้น มาใช้ในการนำระบบช่วยแนะนำ ยกตัวอย่างเช่น ผู้ใช้ A ชอบดูภาพยนตร์ตลกระบบก็จะแนะนำภาพยนตร์ตลกไปให้

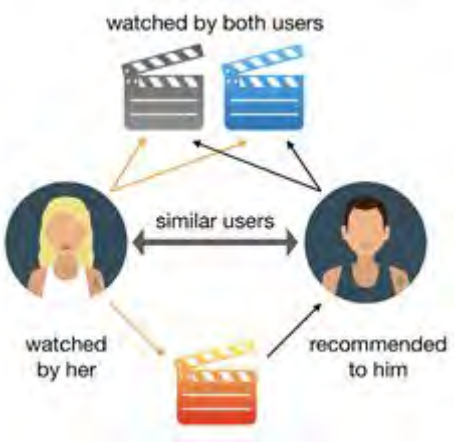


รูปที่ 2.5 แผนภาพการทำงานของโมเดลชนิด Content-based Filtering

(ที่มา :
https://towardsdatascience.com/how-to-build-from-scratch-a-content-based-movie-recommender-with-natural-language-processing-25ad400eb243)

2. Collaborative Filtering คือ การพยายามเอาข้อมูลของคนหมู่มากหรือผู้ใช้คนอื่นมาช่วยในการ พยากรว่าผู้ใช้คนนี้จะชอบอะไร ซึ่งสามารถทำได้หลายวิธีดังนี้

   1. Memory-based วิธีนี้จะมุ่งเน้นไปที่ข้อมูลแล้วหาความสัมพันธ์หรือความคล้ายระหว่างผู้ใช้หรือสินค้าจากข้อมูลโดยตรง ยกตัวอย่างเช่น ผู้ใช้ A ชอบดูเกราะกายสิทธิ์เหมือนกับผู้ใช้ B นอกจากนี้ผู้ใช้ A ก็ชอบดูแก้วหน้าม้าด้วย ดังนั้นระบบช่วยแนะนำจึงแนะนำแก้วหน้าม้าให้กับผู้ใช้ B ด้วยเป็นต้น

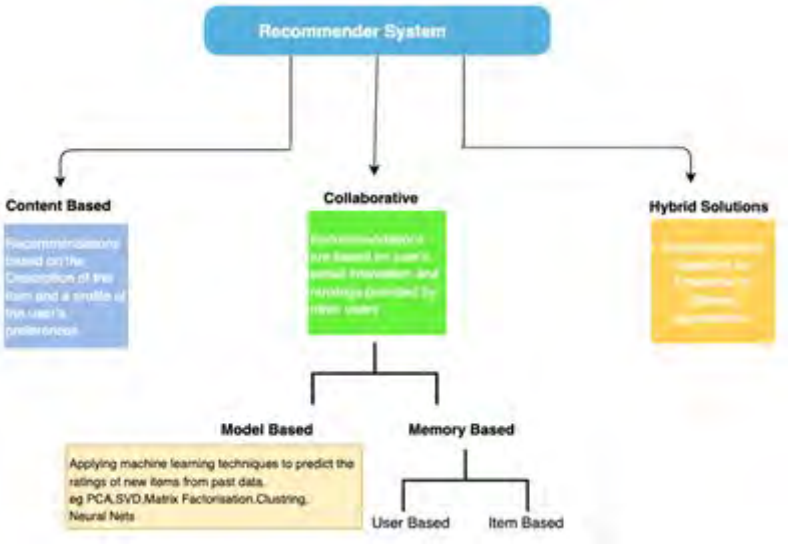


รูปที่ 2.6 แผนภาพการทำงานแบบ Memory-base

(ที่มา : https://towardsdatascience.com)

2. Model-based วิธีนี้จะมีการใช้เทคนิคของ Machine learning เพื่อหา User embedding และ Item embedding มาทำนาย Rating ที่ผู้ใช้จะให้กับสินค้า

3. Hybrid วิธีนี้จะใช้หลาย ๆ วิธีการมาทำงานร่วมกัน



รูปที่ 2.7 แผนภาพประเภทของระบบช่วยแนะนำ

(ที่มา : https://towardsdatascience.com)

2.1.3 ระบบสี

ระบบสี RGB

เป็นระบบสีที่เกิดจากการรวมกันของสามแสงสี แสงสีแดง เขียวและน้ำเงินโดยมีการรวมกันแบบ การผสมสีแบบวก หรือ Additive ซึ่งโดยปกติจะนำไปใช้ในจอภาพแบบ CRT (Cathode ray tube) ในการใช้งานระบบสี RGB มีการสร้างมาตรฐานที่แตกต่างกันออกไปที่นิยมใช้งาน เช่น RGB_{CIE} และ RGB_{NTSC} สำหรับระบบสีแบบ RGB ของ CIE เป็นระบบสีที่ถูกพัฒนาโดย CIE (Commission International l 'Eclairage) ซึ่งอ้างอิงสี

ด้วยสีแดงที่ 700 nm สีเขียวที่ 546.1 nm และสีน้ำเงินที่ 435.8 nm สำหรับระบบสีแบบ RGB ของ NTSC เป็นระบบที่พัฒนาโดย NTSC (National Television System Committee) เพื่อใช้สำหรับการแสดงภาพของจอภาพแบบ CRT เป็นมาตรฐานสำหรับผู้ผลิตแบบ CRT ให้มีลักษณะเดียวกัน

ระบบสี HSV

ระบบสี HSV (Hue Saturation Value) เป็นการพิจารณาสีโดยใช้ Hue Saturation และ Value ซึ่ง Hue คือค่าสีของสีหลักซึ่งคือ สีแดง เขียวและน้ำเงิน มีค่าอยู่ระหว่าง 0 และ 255 ซึ่งถ้า Hue มีค่าเท่ากับ 0 จะแทนสีแดงและเมื่อ Hue มีค่าเพิ่มขึ้นเรื่อย ๆ สีก็จะเปลี่ยนแปลงไปตามสเปกตรัมของสีจนถึง 256 จึงจะกลับมาเป็นสีแดงอีกครั้ง ซึ่งสามารถแทนให้อยู่ในรูปขององศาได้ ดังนี้คือ สีแดง = 0 องศา สีเขียวเท่ากับ 120 องศา สีน้ำเงินเท่ากับ 240 องศา ในระบบ Hue พบว่าจะมีค่าอย่างน้อยหนึ่งค่าที่จะเท่ากับ 0 แต่ถ้ามีสองค่าเท่ากับ 0 แล้ว hue จะเป็นมุมของค่าสีมีค่าเป็นไปตามสีที่สามและถ้าทั้งสามสีมีค่าเท่ากับ 0 แล้วจะทำให้ไม่มีค่าของ Hue หรือสีที่ได้จะมีค่าเท่ากับสีขาวนั่นเอง ตัวอย่างเช่น จอภาพขาว-ดำ ถ้าเกิดมีสีใดสีหนึ่งมีค่าเท่ากับ 0 จะทำให้ค่าสีที่ได้เป็นไปตามสีที่เหลือ การให้น้ำหนักในการพิจารณาเมื่อสีแดงมีค่าเท่ากับ 0 [11]

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 งานวิจัยที่เกี่ยวข้อง (1)

Human activity classification based on sound recognition and residual convolutional neural network โดย Minhyuk Junga and Seokho Chib ตีพิมพ์ในวารสาร Elsevier ฉบับ Automation in Construction (2020)

งานวิจัยศึกษาเกี่ยวกับการจัดหมวดหมู่กิจกรรมของมนุษย์โดยใช้การจดจำเสียงด้วยระบบโครงข่ายประสาทเทียม(Residual convolutional neural) ข้อมูลเสียงจากการทำกิจกรรมต่าง ๆ ของมนุษย์จะถูกบันทึกและแปลงเก็บไว้ในรูปของคลื่นเสียง เนื่องจากคลื่อนเสียงมีคุณลักษณะที่เป็นรูปแบบ (Pattern) และสามารถถูกนำมาเชื่อมโยงเพื่อใช้ในการเรียนรู้และระบุรูปแบบของเสียงนั้น ๆ โดยใช้โครงข่ายประสาทเทียมได้

การวิจัย

1. กระบวนการการเก็บข้อมูล ข้อมูลเสียงถูกรวบรวมผ่าน open-source ออนไลน์ เช่น Youtube-8M และ AudioSet ถูกนำไปสร้างรูปแบบข้อมูลเสียงการทำกิจกรรมในชีวิตประจำวันเป็น 10 หมวดหมู่ โดยต้องคัดเลือกให้ข้อมูลที่จัดเก็บมีความหลากหลายและอยู่ภายใตเงื่อนไขการจัดเก็บเดียวกัน เพื่อไม่ใช้เกิดการชี้นำตัวจัดแบ่งหมวดหมู่(Classifier) ตัวแปรควบคุมในส่วนของข้อมูลเหล่านี้ คือ ความยาวของบันทึกข้อมูลเสียง ระยะเวลา 6 นาที และสภาพแวดล้อมการจัดเก็บเสียง ซึ่งในที่ นี้ คือ การจัดเก็บเสียงการทำกิจกรรมเฉพาะที่อยู่ภายในอาคาร(Indoor-environment) หากไม่ควบคุมปัจจัยทั้งสองอย่างนี้ตัวแปรของเสียงที่โครงข่ายประสาทเทียมเลือกใช้เป็นเกณฑ์(Criteria) ในการแบ่งหมวดหมู่อาจเกิดการตกหล่นได้

Table 1
Human activity classes and relevant sound events.

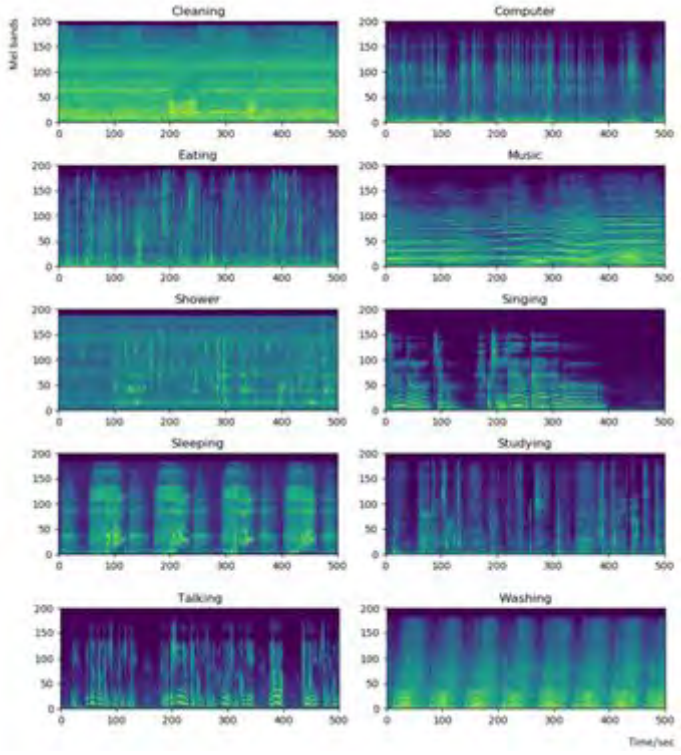
Human activities	Sound events
Cleaning	Engine sound of vacuum cleaner.
Computer	Clicking sound of keyboard and mouse
Eating	Eating foods with chopsticks, forks, and spoon and drinking
Music	Sound from classical, jazz, rock, hip-hop and pop music
Showering	Falling water from showerheads and faucets and washing hair, face, and body
Singing	Sounds of humming and singing without playing an instrument
Sleeping	Sounds of breathing and snoring
Studying	Sounds of turning over a book page and writing with pen and pencil
Talking	Monologue, group conversation, telephone conversation
Washing (laundry)	Sounds of washing machine's watering, draining, and spinning

Table 2
Descriptions of the sound dataset after the data collection and preprocessing.

Activity class name	Number of recordings	Number of audio data	Length of sound data (sec)
Cleaning	28	864	8640
Computer	28	864	8640
Eating	28	864	8640
Music	28	864	8640
Showering	28	864	8640
Singing	28	864	8640
Sleeping	36	864	8640
Studying	28	864	8640
Talking	28	864	8640
Washing	98	864	8640

ตารางที่ 1 และ 2 แสดงการแบ่งกลุ่มข้อมูลเสียงและรูปแบบการเก็บข้อมูลเสียงตามลำดับ

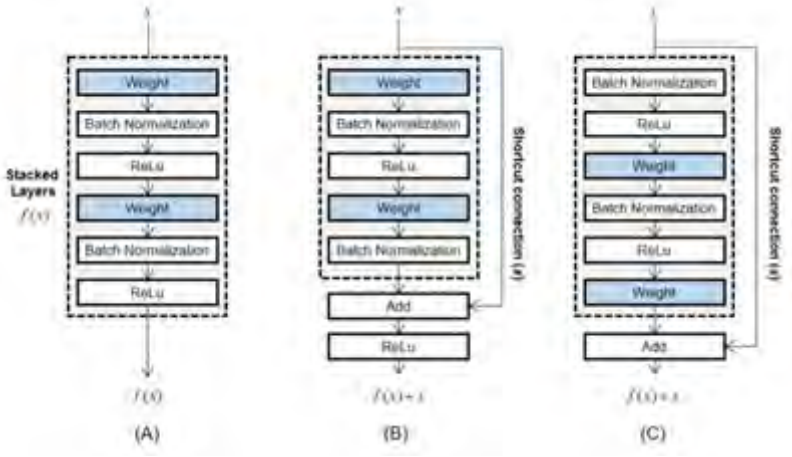
2. การแยกคุณลักษณะของเสียงจากไฟล์ข้อมูลเพื่อใช้ในการวิเคราะห์ (Feature extraction) ข้อมูลเสียงถูกแยกออกมาด้วย Log Mel-filter bank energies ข้อมูลเสียงที่ได้แสดงรูปแบบได้ดังภาพที่ 2.8 จะเห็นว่าข้อมูลเสียงของแต่ละกิจกรรมมีรูปแบบที่แตกต่างกัน



รูปที่ 2.8. ข้อมูลคลื่นเสียงใน spectrogram ที่ได้จากการแยกคุณลักษณะจากข้อมูลเสียง

3. การสร้างโครงข่ายประสาทเทียม (Neural network architecture) ข้อมูลที่ได้จากการแยกคุณลักษณะเสียง จะเห็นว่าข้อมูลดังกล่าวมีลักษณะที่เป็นรูปแบบและเกิดการซ้ำ ดังนั้นโครงข่ายประสาทเทียมจะสามารถแบ่งหมวดหมู่ข้อมูลเสียงเหล่านี้ได้อย่างถูกต้อง หากโครงข่ายประสาทเทียมสามารถจดจำและระลึกได้ว่ารูปแบบของคลื่นที่แสดงใน spectrogram นี้ เป็นรูปแบบของคลื่นเสียงจากกิจกรรมประเภทไหน โครงข่ายประสาทเทียมแบบคอนโวลูชันหรือ Convolutional Neural Network (CNN) มีความเหมาะสมอย่างมากในการนำมาใช้คัดแยกประเภท ของข้อมูลที่มีมิติข้อมูลมากกว่า 1 มิติ เพราะให้ผลลัพธ์ที่มีความแม่นยำ จากที่ได้รับการพิสูจน์แล้วจากผลการวิจัยด้านโครง

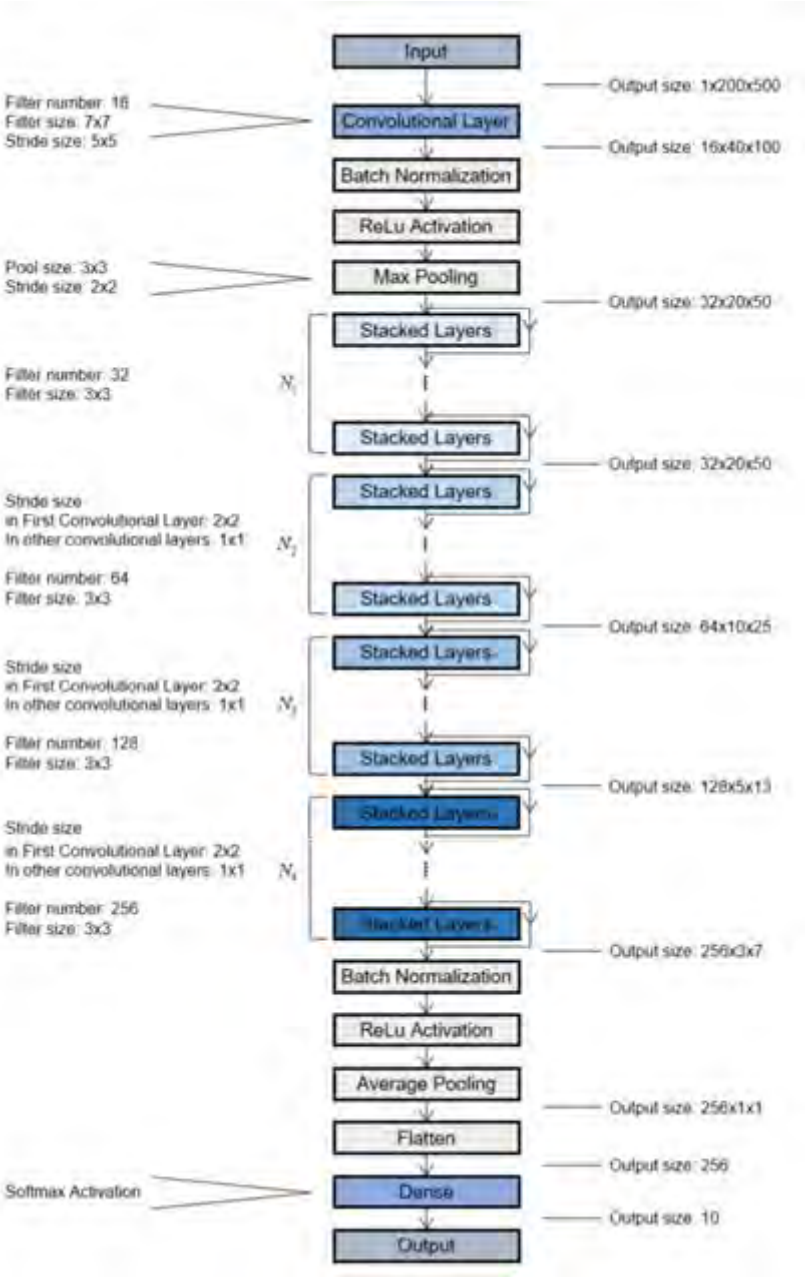
ข่ายประสาทเทียมอื่น ๆ ประสิทธิภาพการทำงานของ CNN จะขึ้นอยู่กับ จำนวนชั้นและความซับซ้อนของโครงข่ายประสาท (Depths of network) การเพิ่มจำนวนชั้นและความซับซ้อนของโครงข่ายประสาทจะทำให้ CNN สามารถเรียนรู้รูปแบบและคุณลักษณะของข้อมูลที่ซับซ้อนและหลากหลายมากขึ้นได้ อย่างไรก็ตามหากเพิ่มจำนวนคุณลักษณะ (Features) ที่ใช้เป็นเกณฑ์มากเกินไปอาจเกิดเป็นอุปสรรคในการสร้างโครงข่ายประสาทเทียมได้ เพราะทำให้ CNN จดจำข้อมูลสำหรับตัวชี้วัดที่ไม่สามารถนำไปใช้ประโยชน์ได้ ตัวชี้วัดที่ไม่จำเป็นต้องใช้ในการเรียนรู้ของโครงข่ายประสาทนี้ เรียกว่า “Free parameter” ซึ่งในระหว่างกระบวนการป้อนข้อมูลให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ หรือในระหว่างกระบวนการ “Training” โครงข่ายประสาทเทียมจะทำการตัด Free parameter ออกจากกระบวนการเรียนรู้ ในงานวิจัยเมื่อศึกษาต่อพบว่านอกจากการตัด Free parameter ออกแล้ว พบว่าการสร้างทางลัด (Short-cut) ในการส่งข้อมูลจะเป็นการตัวเร่งกระตุ้นให้การส่งกระจายข้อมูลทั่วทั้งระบบเป็นไปอย่างรวดเร็วมากขึ้น และเสริมให้ประสิทธิภาพการทำงานของระบบดีขึ้น โดยผังแสดงการออกแบบโครงข่ายประสาทจะเป็นไปดังรูปที่ 2.9



รูปภาพที่ 2.9 ความแตกต่างของ การเรียนรู้แบบโครงข่ายประสาทเทียมปกติ(A) แบบโครงข่ายประสาทเทียมที่มีการสร้างทางเชื่อมต่อลัด(B) และแบบโครงข่ายประสาทเทียมที่ยังไม่สร้างทางเชื่อมต่อลัด(C)

โดยการศึกษาข้างต้นมีต้นแบบมากจากโมเดลของ ResNet ซึ่งถูกพัฒนาโดย He et al ในส่วนต่อไปนี้จะกล่าวถึงโครงสร้างชั้นต้นของโครงค่ายประสาทที่เรียกว่า “Stacked layers” ซึ่งประกอบไปด้วย 1) Weight ซึ่งเป็นเกณฑ์วัด 2) Batch Normalization 3) ReLU หรือ Rectified linear units ซึ่งเป็น “Activation function” ทำหน้าที่แปลงค่าข้อมูลนำเข้า (Input) ด้วยสมการ f(x)=x+=max(0,x) เพื่อให้ข้อมูลอยู่ในสถานะ “activated” และ 4) ชั้นคอนโวลูชัน (Convolution layer) ซึ่งประกอบด้วยตัวกรอง (Filter) ขนาด 3 x 3 โดยรูปแบบของ stacked layer ปรากฏในรูปที่ 2.9 รูปที่ 2.10 แสดงการทำงานของโครงข่ายประสาทเทียมทั้งระบบ ขั้นตอนการพัฒนาโครงข่ายประสาทเทียมเริ่มจากชั้นคอนโวลูชันที่มีตัวกรอง ซึ่งจะขอเรียกว่า filter ขนาด 7 x 7 และ stride ขนาด 5 x 5 โดย stride คือตัวกำหนดระยะ step ในการเลื่อน filter จากนั้นข้อมูลจะถูกส่งผ่านไปยังชั้น batch normalization และถูกส่งต่อไปอีกยังชั้น ReLU จากนั้นจึงผ่านไปยังชั้นที่มีการทำ max pooling ขนาด 3 x 3 และ stride ขนาด 2 x 2 ดังแสดงในรูปภาพที่ 2.10 กระบวนการจากถูกแบ่งเป็น 4

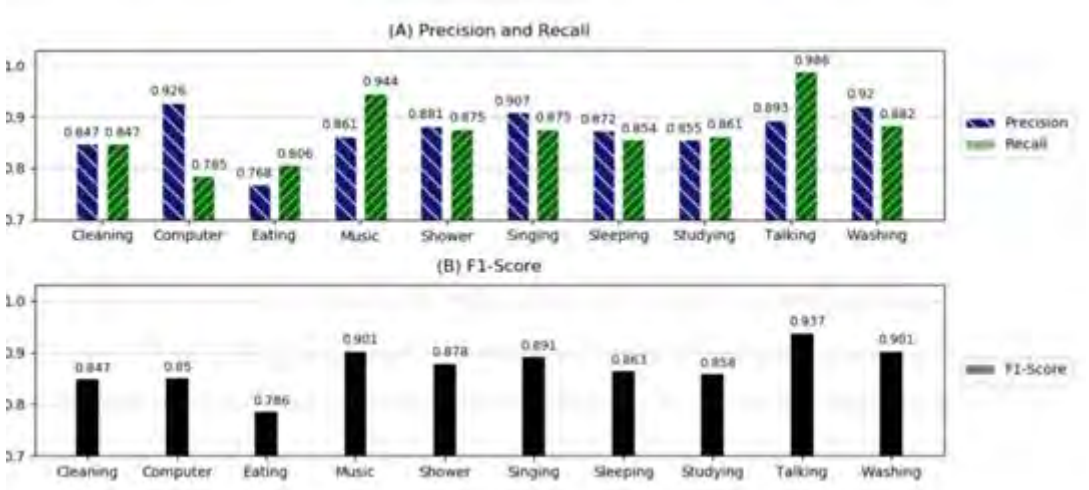
กระบวนการในช่วงเดียวกันจะมีขนาดผังคุณลักษะ (feature map) และจำนวนของคุณลักษณะ ที่เท่ากัน เมื่อผ่านกระบวนการเหล่านี้ข้อมูลนำเข้าลดขนาด feature map จากการผ่าน filter และกระบวนการ max pooling ทำให้ข้อมูลนำเข้าถูก flatten คือการถูกลดมิติของข้อมูล และจะได้ข้อมูลลัพธ์ (Out put) เป็น array 1 มิติ ซึ่งมีขนาดเท่ากับหมวดหมู่ข้อมูลที่ต้องการแยกประเภท



รูปภาพที่ 2.10 แสดงลำดับการทำงานของโครงข่ายประสาทเทียม

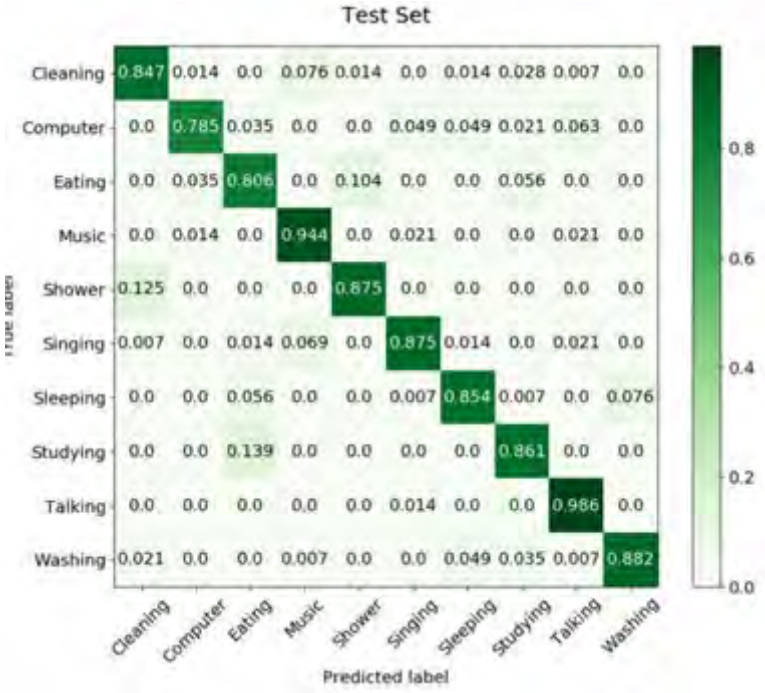
4. การฝึกตัวโมเดลและการพยากรณ์ (Model training and predication) มีการใช้ kernel initializer adaptive moment estimation (Adam) เป็น โดยทำการเทรนด์โมเดลเป็นจำนวน 200 รอบ (epochs) และเพื่อไม่ให้เกิดการ overfitting รูปแบบโมเดลที่มี weight ซึ่งให้ประสิทธิภาพการทำงานของโครงข่ายประสาทดีที่สุดจะถูกเลือกเป็นโมเดลในขั้นสมบูรณ์

สรุปผลการวิจัย



รูปภาพที่ 2.11 ผลความแม่นยำในการพยากรณ์ข้อมูลเสียง

จากผลที่ได้ในรูปที่ 2.11 ค่าความแม่นยำของการพยากรณ์อยู่ที่ 87.2% และหากพิจารณาจากความแม่นยำในแต่ละหมวดหมู่ หมวดหมู่กิจกรรมการพูดให้ความแม่นยำที่สุด ที่ค่าความแม่นยำเท่ากับ 93.7% ในส่วนของประสิทธิภาพการพยากรณ์ที่แตกต่างกันในแต่ละหมวดหมู่นั้น หากพิจารณาจาก confusion matrix ในรูปภาพที่ 2.12 จะสังเกตได้ว่า ส่วนที่เกิดความคลาดเคลื่อนในการพยากรณ์จะมากจากข้อมูลของเสียงที่มีคุณลักษณะรูปแบบของข้อมูลที่ใกล้เคียงกัน



รูปภาพที่ 2.12 แสดง Confusion Matrix ของข้อมูล

จากงานวิจัยที่เกี่ยวข้องนี้สามารถนำมาประยุกต์ใช้กับโครงงานวิจัยได้ ในเรื่องของการออกแบบโครงข่ายประสาทเทียมเพื่อให้สามารถคัดแยกและแบ่งหมวดหมู่ข้อมูลนำเข้าได้อย่างถูกต้องแม่นยำและนำไปสร้างเป็นระบบแนะนำได้

2.2.2 งานวิจัยที่เกี่ยวข้อง (2)

Food object recognition using a mobile device: Evaluation of currently implemented systems โดย Simon Knez และ Luka Sajn ตีพิมพ์ในวารสาร Elsevier ฉบับ Trends in Food Science & Technology

งานวิจัยศึกษาเกี่ยวกับการจำแนกภาพอาหารด้วยอุปกรณ์แบบพกพา โดยมีกระบวนการทั้งสิ้น 6 ขั้นตอน ดังนี้

1. การรวบรวมรูปภาพ (image acquisition)
2. การประมวลรูปภาพ (image processing)
3. การแยกส่วนรูปภาพ (image segmentation)
4. การแยกคุณลักษณะของรูปภาพ (feature extraction)
5. การจัดหมวดหมู่รูปภาพ (image classification)
6. การวิเคราะห์ปริมาณ (volume estimation)

งานวิจัยจัดทำขึ้นเพื่อศึกษาภาพรวมของวิธีการจำแนกภาพอาหาร และอภิบายกระบวนการโดยย่อของแต่ละระบบวิธี เพื่อนำไปประยุกต์ใช้ในด้านต่าง ๆ อาทิ เช่น นำไปสร้างเป็นฐานข้อมูลในระบบแนะนำอาหาร ในระบบรวบรวมข้อมูล หรือในระบบตอบรับอัตโนมัติ รวมทั้งเป็นการศึกษาและให้ข้อมูลเบื้องต้นเกี่ยวกับแนวทางการเก็บรวบรวมข้อมูลและศึกษาวิธีการที่มีการใช้อย่างแพร่หลายในปัจจุบัน โดยเรียกการวิเคราะห์ข้อมูลจากแนวทางเหล่านี้ว่า “Automatic analysis and recognition”

กระบวนการรวบรวมรูปภาพ จะทำการเลือกเก็บข้อมูลการบริโภคจากเว็บไซต์ต่าง ๆ เช่น foodlog.jp โดยเก็บข้อมูลภาพเหล่านี้เพื่อนำไปป้อนให้ระบบปัญญาประดิษฐ์ที่พัฒนาเกิดการเรียนรู้ เป็นการเก็บข้อมูลลักษณะที่ได้จากการมองเห็น (visual information) เพียงอย่างเดียว โดยโมเดลระบบในรูปแบบนี้ ต้องมีการจำกัดขอบข่ายข้อมูลและประสิทธิภาพของฮาร์ตแวร์ที่ใช้เพราะล้วนแต่เป็นปัจจัยต่อประสิทธิภาพในการเก็บที่ต้องนำมาคำนวณ

ในขั้นที่สองคือ การประมวลรูปภาพ ภาพที่ได้จะถูกทำการปรับ pixel ทำ frame normalization และผ่านกระบวนการ colour coversion techniques เพื่อให้ข้อมูลภาพที่ผ่านกระบวนการเหล่านี้ถูกลด noise ลง และถูกปรับให้มีขนาดภาพเท่ากันทุกภาพ (uniform size) รวมทั้งถูกปรับช่องสี (color channels) ให้เป็นช่องที่ต้องการ การปรับเหล่านี้ทำให้สามารถลดตัวแปรด้านความสว่างของแสงที่มีอิทธิพลต่อรูปภาพได้ ส่งผลให้การแบ่งส่วนภาพ (image segmentation) ได้มีประสิทธิภาพมากยิ่งขึ้น

ในขั้นของการแยกส่วนรูปภาพเป็นขั้นตอนที่เกิดการแยก "ส่วนภาพที่สนใจ" หรือ ROI (region of interest) ออกจากส่วนอื่น ๆ ของภาพซึ่งอยู่นอกเหนือความสนใจ ซึ่งนิยามของ ROI คือส่วนของภาพที่มีวัตถุที่สนใจปรากฏอยู่ เช่น หากต้องการแยกภาพผลส้มออกจากภาพตระกร้าผลไม้ ROI คือส่วนของภาพที่มีผลส้มปรากฏอยู่ โดยหลักการพื้นฐานของ ROI คือแปลงส่วนภาพที่สนใจให้เป็น "Foreground" และส่วนภาพที่ไม่สนใจเป็น "Background" เพื่อกำจัดข้อมูลของภาพพื้นหลังที่ไม่มีความสำคัญ เช่น สภาวะแสง และมุมมองของภาพ คงเหลือไว้แต่ข้อมูลที่เป็นประโยชน์เพื่อลดเวลาในการประมวลรูปภาพ

การแยกคุณลักษณะของรูปภาพ เป็นขั้นตอนที่จะแตกต่างกันตามขั้นตอนระบบ (algorithm) ที่ใช้ ระบบจะทำการแยกข้อมูลลักษณะของภาพโดยเลียนแบบระบบการแยกแยะทางสายตาของมนุษย์และเลือก

จากข้อมูลที่มีความสำคัญมากที่สุดก่อน โดยสำหรับภาพอาหาร ข้อมูลที่เป็นปัจจัยสำคัญได้แก่ สี รูปร่าง และลักษณะผิว Deep learning มีขั้นตอนการปฏิบัติการที่คล้ายคลึงกับการทำงานของระบบประสาทในสมอง สามารถวิเคราะห์รูปแบบของวัตถุได้ จึงเหมาะสมแก่การใช้ในการทำงานรูปแบบนี้

ในขั้นการจัดหมวดหมู่รูปภาพและวิเคราะห์ปริมาณจะทำการรวบรวมลักษณะข้อมูลของภาพที่ได้และถูกนำไปจัดหมวดหมู่ ซึ่งปัจจัยสำคัญที่ส่งต่อความถูกต้องของข้อมูลคือ กลุ่มข้อมูล (data-set) ที่ป้อนเข้าเพื่อใช้ในการเรียนรู้ของระบบ

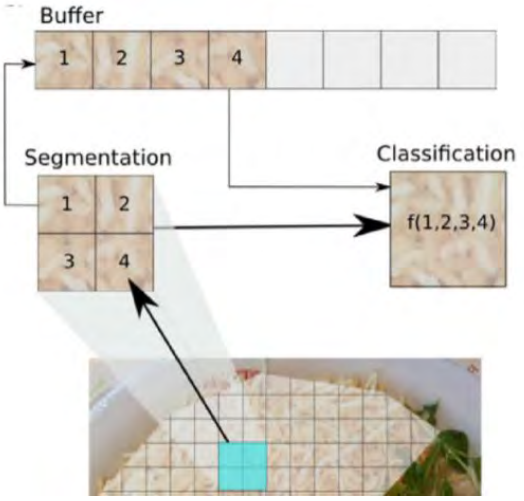
ตัวอย่างระบบ เช่น ระบบ Aizawa et al., - 2013 [7] ทำการรวบรวมข้อมูลภาพถ่ายจากเว็บไซต์ http://www.foodlog.jp/ ทำการประมวลรูปภาพโดยปรับขนาดให้มีขนาดเท่ากันทุกรูปที่ 320 x 240 pixels ในขั้นตอนการแยกส่วนรูปภาพทำการแตกรูปภาพออกเป็นบล๊อคสี่เหลี่ยมเพื่อเตรียมทำการแยกคุณลักษณะของรูปภาพ ในขั้นการแยกคุณลักษณะของรูปภาพ คุณสมบัติหลายประการเช่น สี จากการอ่านค่าสี RGB และค่า HSV จะถูกนำมาแบ่งแยกรูปภาพออกเป็น 5 หมวดรูปภาพอาหาร ได้แก่ หมวดธัญพืช หมวดผัก หมวดเนื้อสัตว์ หมวดนม และหมวดผลไม้ และ 1 หมวดที่ไม่ใช่อาหาร ตามค่าสีและค่าความถี่จากการอ่านข้อมูลในแต่ละบล๊อค โดยมีการแปลงข้อมูลทั้งหมดเป็นเวกเตอร์ และทำการปรับค่าความแม่นยำด้วย สมการแบบ Naïve Bayesian ซึ่งเป็นการทำเหมืองข้อมูลเพื่อวิเคราะห์หาความน่าจะเป็นของสิ่ง ๆ หนึ่ง จากสิ่งที่เกิดขึ้นมาก่อน [8] ซึ่งสมการนี้ทำให้สามารถแบ่งแยกข้อมูลรูปภาพได้อย่างเม่นยำมากขึ้น

รูปภาพที่ 2.13 แสดงการแยกหมวดหมู่รูปภาพอาหารตามการทดลอง ระบบ Aizawa et al., - 2013 [7]

Oliveira et al., - 2014 [9] เป็นอีกหนึ่งระบบตัวอย่าง เป็นระบบแอฟลิเคชั่นในอุปกรณ์มือถือแบบสำเร็จในตัว ทำการเก็บข้อมูลจากรูปภาพอาหารที่ถ่ายจากมุมมองด้านบน ในระบบนี้ไม่มีการระบุขั้นตอนการประมวลรูป ในส่วนของขั้นตอนการแยกส่วนรูปภาพ ระบบทำการระบุส่วนที่สนใจและใช้สมการ Region growing ซึ่งเป็นกระบวนการขยายกลุ่มของ pixel ให้มีขนาดใหญ่ขึ้นตามเกณฑ์ที่กำหนด [10] จากนั้นทำการคำนวณกำหนดเกณฑ์การแบ่งส่วนรูปภาพจากค่า HSV ค่า SVC texture และค่า LAB เมื่อสิ้นสุดทั้งสามการคำนวณระบบจะทำการสร้าง ROI ขึ้นสามค่า ค่าสี RGB และค่าสี LAB ของแต่ละ ROI ซึ่งมีลักษณะเป็นเวกเตอร์ 12 มิติ จะถูกแยกออกมาวิเคราะห์ในขั้นการแยกคุณลักษณะของรูปภาพ ดังรูปภาพที่ 2.14 เมื่อสิ้นสุดกระบวนการแยกส่วน รูปภาพจะถูกนำมาจัดหมวดหมู่โดยวิเคราะห์จากคุณลักษณะด้วย โครงข่ายประสาทเทียม GoogleLeNet รูปแบบ CNN ในการจัดหมวดหมู่ หมวดหมู่ของอาหารถูกอ้างอิงฐานข้อมูล (Data-set) ดังต่อไปนี้ Restaurant, Food-101, Food-201 และ Gfood-3D

Waltner et al., - 2017 [12] ระบบนี้เก็บข้อมูลภาพถ่ายอาหารจากการแสกนรูปภาพ เช่นจากระบบเครื่อง Scanner ในระบบนี้ไม่มีการระบุขั้นตอนการประมวลรูป ในส่วนของขั้นตอนการแยกส่วนรูปภาพระบบทำการวาด Bounding box [13] ซึ่งเป็นการสร้างขอบเขตส่วนของภาพที่สนใจ คล้ายกับการสร้าง ROI

ทำการแยกคุณลักษณะของรูปภาพ โดยสนใจเพียงค่าสีคือ ค่า RGB ค่า HSV และค่า LAB ทำการรวมค่าเหล่านี้และสร้างเป็นเวกเตอร์ ทำการเทรนด์ (Train) ระบบโดยป้อนฐานข้อมูลอาหาร 23 หมวด



รูปภาพที่ 2.14 แสดงการแยกคุณลักษณะของภาพจาก ROI ขนาด 2 x 2 ระบบ Oliveira et al., - 2014 [9]

Mezgec and Seljak - 2017 [14] เป็นระบบที่สร้างขึ้นเพื่อแนะนำอาหารที่เหมาะสมกับผู้ป่วยโรคพากินสัน (Parkinson’s disease) ระบบการทำงานคือการที่ผู้ใช้ถ่ายรูปอาหารเพื่อป้อนข้อมูลให้ระบบ โดยระบบมีการติดตั้งโครงข่ายประสาทเทียมแบบ CNN ที่ถูกพัฒนามาเป็นพิเศษ เรียกว่า NutrieNet ทำการให้ข้อมูลอาหารที่ผู้ป่วยควรรับประทาน 5 อันดับ ในระบบนี้ไม่มีการระบุขั้นตอนการประมวลรูป ในส่วนของขั้นตอนการแยกส่วนรูปภาพใช้โครงข่ายประสาทเทียม NutrieNet ซึ่งถูกดัดแปลงมาจาก AlexNet ในขั้นตอนนี้มีการลดมิติของรูปภาพ คือการลดคุณลักษณะของภาพลง เพื่อให้ประมวลได้ง่ายขึ้น โดยคุณลักษณะที่ถูกแยกออกมาขึ้นกับสมการที่ใช้ใน NutrieNet โดยระบบนี้มีการสร้างโครงข่ายประสาทที่ซับซ้อนทำให้มีความละเอียดอ่อนสามารถแยกรูปภาพอาหารได้ 520 หมวดหมู่

สรุปผลจากงานวิจัย โดยภาพรวมแล้วระบบจดจำและจำแนกอาหารประการไปด้วยขั้นตอน 6 ขั้นตอนดังที่กล่าวในข้างต้น โดยแต่ละระบบมีรายละเอียดของขั้นตอนและฐานข้อมูลที่เลือกใช้แตกต่างกันไปตามการใช้งาน ซึ่งจากตัวอย่างระบบที่กล่าวมาทำให้เห็นว่าการใช้โครงข่ายประสาทเทียมในการจำแนกและจัดหมวดหมู่รูปภาพอาหารเป็นวิธีที่มีประสิทธิภาพ

บทที่ 3

วิธีการดำเนินงานวิจัย

3.1 วัสดุอุปกรณ์

3.1.1 คอมพิวเตอร์

3.1.2 VGA รุ่น GTX 1660 Ti

3.1.3 โปรแกรม

- MATLAB
- Photos Snapping Tools

3.2 วิธีการดำเนินงาน

3.2.1 ระบบเรียนรู้จดจำภาพอาหารไทย

1. การเก็บรวบรวมรูปภาพ

เนื่องจากอาหารไทยเป็นชนิดอาหารที่ยังไม่มีการเก็บภาพตัวอย่าง หรือมีการสร้าง Data-set อย่างจริงจังจึงไม่สามารถดึงข้อมูลจากฐานข้อมูลที่เป็นที่รู้จักกันดีอย่าง Food-101 หรือ Food-201 ได้ ด้วยจุดประสงค์ในการสร้างแบบจำลองระบบแนะนำอาหารไทย เราจึงกำหนดขอบเขตของการทดลอง โดยการใช้ชนิดอาหารไทยที่ที่มีความเป็นสามัญและเป็นที่นิยมในโรงอาหารนิสิตของจุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 3.1แสดงตัวอย่างการสำรวจภาพอาหาร

โดยจากการสำรวจพฤติกรรมการสั่งอาหารจากร้านอาหารข้าวร้านแกงในโรงอาหารในมหาวิทยาลัย ได้เลือกอาหารจำนวน 25 อย่าง แบ่งออกเป็น 6 หมวด เพื่อทำการสร้างฐานข้อมูลอาหารไทย โดยอาหารทั้ง 25 ชนิด แสดงดังตารางที่ 3.1 นี้

ประเภทอาหาร: 6 หมวด											
ข้าว		ผัด		ต้ม/แกง		เครื่องเคียง		น้ำพริก		เครื่องดื่ม	
1	ข้าวสวย	1	ผัดผักรวม	1	แกงเขียวหวาน	1	ไข่ดาว	1	พริกอ่อง	1	น้ำใบเตย
2	ข้าวผัด	2	คะน้าหมูกรอบ	2	ต้มจืด	2	ไข่เจียว	2	พริกหนุ่ม	2	น้ำจับเลี้ยง
3	เส้นก๋วยเตี๋ยว	3	ผัดเต้าหู้	3	ต้มแซ่บ	3	ไข่มะตูม			3	เก๊กฮวย
		4	กระเพรา	4	พะแนง	4	หมูทอด			4	กระเจี๊ยบ
		5	ลาบ	5	แกงเลียง	5	ปลาทู				
						6	แคบหมู				

ตารางที่ 3.1 แสดงการแบ่งหมวดอาหารทั้ง 6 หมวด

จากการสำรวจการสั่งอาหาร แนวโน้มการสั่งอาหารเป็นดังนี้ อาหารประเภทผัดสามารถแบ่งได้เป็นรสจัดและรสอ่อน ผัดรสจัด ได้แก่ กระเพรา และลาบ ผัดรสอ่อน ได้แก่ ผัดผักรวม คะน้าหมูกรอบ และผัดเต้าหู้ หมวดแกงแบ่งเป็น แกงรสจัด และแกงรสอ่อน เช่นกัน โดยแกงรสจัดได้แก่ ต้มแซ่บ พะแนง และแกงเลียง แกงรสอ่อน ได้แก่ แกงเขียวหวาน และต้มจืด

เมื่อกำหนดประเภทอาหาร เราทำการสร้างฐานข้อมูล (Data-set) สำหรับเทรนด์ (Train) ของอาหารทั้ง 25 ชนิด โดยค้นรูปจากแหล่งข้อมูลดิจิตอล ป้อนข้อมูลรูปภาพของอาหารแต่ละชนิด ชนิดละ 100 รูปภาพ โดยรวมฐานข้อมูลใช้รูปทั้งหมด 2,500 รูปภาพ สำหรับการเก็บภาพตัวอย่างสำหรับทดสอบ (Test) ความ

แม่นยำของการระบุจดจำภาพถ่าย ใช้ภาพจากการเก็บรูปอาหารจากโรงอาหารในมหาวิทยาลัย ทำการแบ่ง Folder ไฟล์จำนวน 100 โฟลเดอร์ โดยตัดภาพอาหารให้เหลืออาหารเพียงชนิดเดียวในภาพและพอดีกับภาพ

รูปภาพที่ 3.2 ตัวอย่างรูปภาพที่ใช้ในการเทรนด์โครงข่ายประสาท

2. การประมวลรูปภาพและการแยกส่วนรูปภาพ

ขั้นตอนประมวลรูปภาพที่ใช้เทรนด์ จะทำการปรับขนาดรูปให้มีขนาด 150 x 150 pixels ซึ่งทุกรูปจะถูกป้อนเข้าให้เป็นค่าระบบสีแบบ RGB โดยฟังก์ชั่น Image processing: gray2rgb ใน MATLAB

3. การแยกคุณลักษณะของรูปภาพ

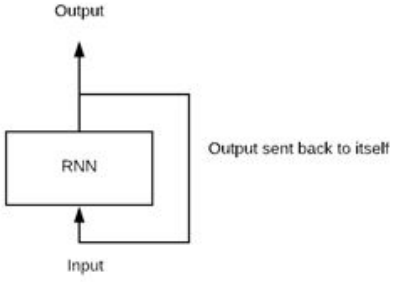
ทำการแยกคุณลักษณะของรูปภาพ (Extract features) ออกมา โดยใช้ MScratchNet ซึ่งเป็นโค้ดโปรแกรมที่เขียนขึ้นในโปรแกรม MATLAB โครงข่ายประสาทเทียมทำการ Epoch 20 รอบ เพื่อลดมิติคุณลักษณะที่ไม่มีประโยชน์ต่อการจำแนกภาพออก

4. การจัดหมวดหมู่รูปภาพ

นำโครงข่ายประสาทเทียมที่ผ่านการเทรนด์แล้วอ่านข้อมูลรูปภาพ แล้วระบุประเภทอาหาร 6 หมวดหมู่

3.2.2 ระบบแนะนำอาหารไทย

1. สร้าง RNN ป้อนวนค่าผ่านการโค้ดในโปรแกรม MATLAB

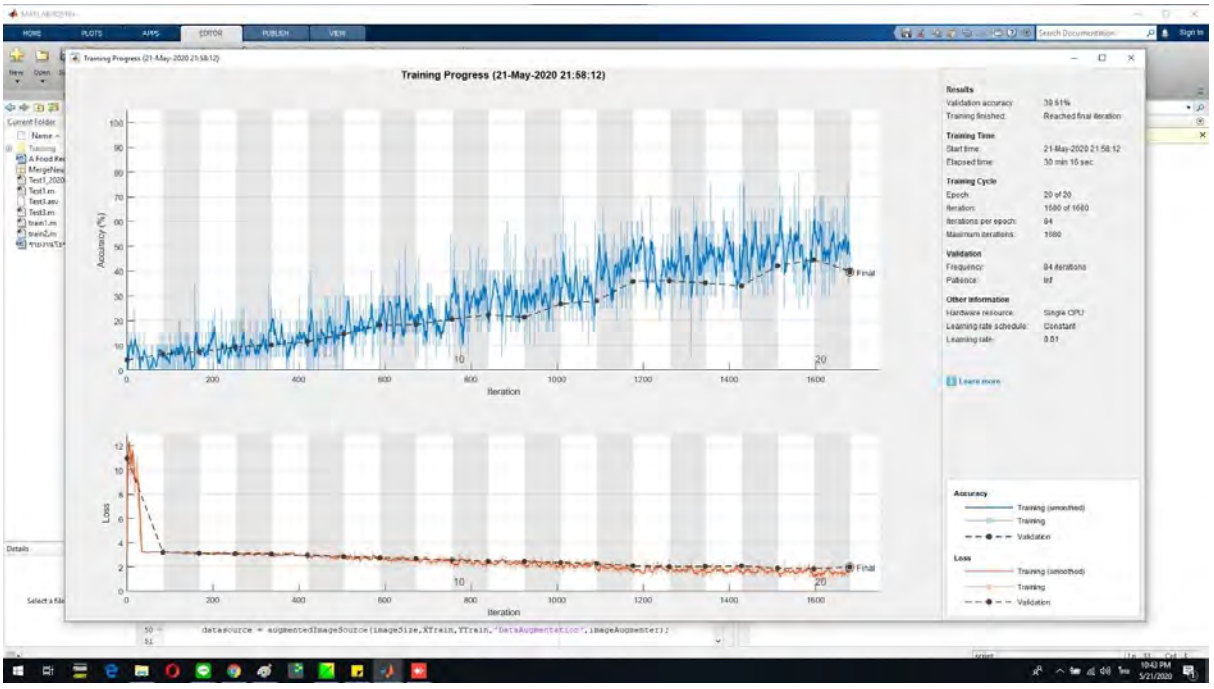


2. ทำการทดสอบ (Test) ชุดโปรแกรมโดยการใช้รูปอาหาร 100 โฟลเดอร์ที่สร้างขึ้น ทำการสุ่มรูปอาหาร แล้วดูผลความเหมาะสมของชนิดอาหารที่คืนค่ากลับมา บันทึกผลและวิเคราะห์ผล

บทที่ 4

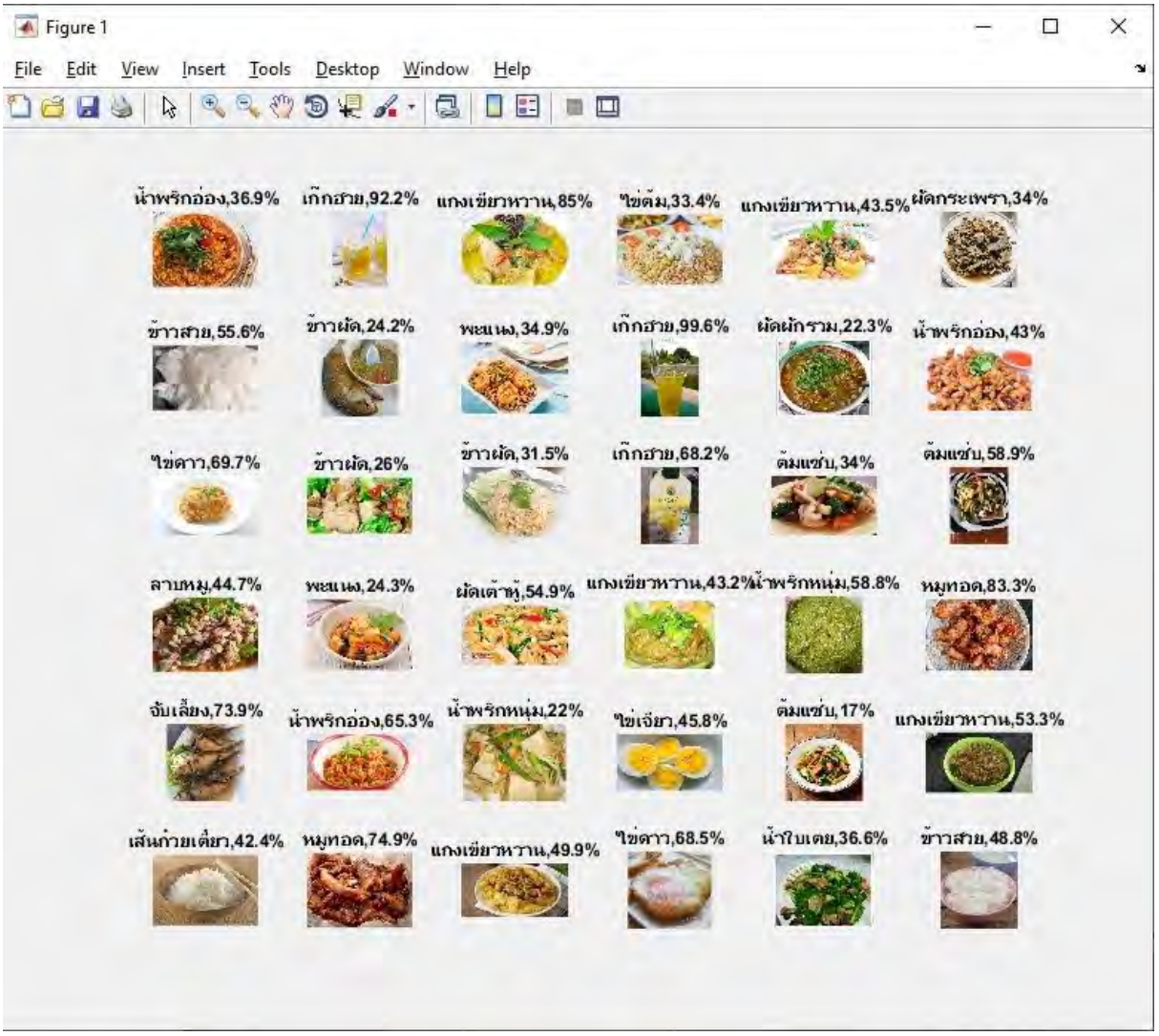
ผลการทดลองและอภิปรายผล

4.1 ผลการทดลองค่าความแม่นยำของการจำแนกรูปภาพ



รูปที่ 4.1 กราฟแสดงผลค่าความแม่นยำในการจำแนกจดจำภาพถ่าย

จากการทำการเทรนด์โครงข่ายประสาทเทียม โดยใช้ฮาร์ดแวร์แบบ Single CPU พบว่าได้ค่าความแม่นยำโดยรวม (Validation accuracy) 39.61% กระบวนการเทรนด์ถึงสิ้นสุดเป็นแบบ Reached final literation โดยทำการ Epoch ครบ 20 รอบ มีการ Literation 1,680 ครั้ง มีอัตราการเรียนรู้ 0.01



รูปที่ 4.2 แสดงผลการทายและค่าความแม่นยำในการจำแนกจดจำภาพถ่ายอาหาร

4.2 ผลค่าความแม่นยำในการแนะนำอาหาร

ลำดับ/หมวด	1	2	3	4	5	6	
1	ข้าวสวย	ผัดผักรวม	ต้มแซ่บ	หมูทอด	น้ำใบเตย		1 แกงเขียวหวาน
2	ข้าวสวย	ผัดผักรวม	ต้มแซ่บ	หมูทอด	น้ำจับเลี้ยง		2 แกงเขียวหวาน
3	ข้าวสวย	ผัดผักรวม	ต้มแซ่บ	ปลาทู	น้ำใบเตย	น้ำพริกหนุ่ม	3 แกงเขียวหวาน
4	ข้าวสวย	ผัดผักรวม	ต้มแซ่บ	ปลาทู	น้ำจับเลี้ยง	น้ำพริกหนุ่ม	4 แกงเขียวหวาน
5	ข้าวสวย	ผัดผักรวม	ต้มแซ่บ	ไข่เจียว	น้ำใบเตย		5 แกงเขียวหวาน
6	ข้าวสวย	ผัดผักรวม	ต้มแซ่บ	ไข่เจียว	น้ำจับเลี้ยง		6 แกงเขียวหวาน
7	ข้าวสวย	ผัดผักรวม	พะแนง	หมูทอด	น้ำใบเตย		7 แกงเขียวหวาน
8	ข้าวสวย	ผัดผักรวม	พะแนง	หมูทอด	น้ำจับเลี้ยง		8 ข้าวสวย
9	ข้าวสวย	ผัดผักรวม	พะแนง	ปลาทู	น้ำใบเตย	น้ำพริกหนุ่ม	9 น้ำใบเตย

10	ข้าวสวย	ผัดผักรวม	พะแนง	ปลาทู	น้ำจับเลี้ยง	น้ำพริกหนุ่ม	10 แกงเขียวหวาน
11	ข้าวสวย	ผัดผักรวม	พะแนง	ไข่เจียว	น้ำใบเตย		11 พะแนง
12	ข้าวสวย	ผัดผักรวม	พะแนง	ไข่เจียว	น้ำจับเลี้ยง		12 พะแนง
13	ข้าวสวย	ผัดผักรวม	แกงเลียง	หมูทอด	น้ำใบเตย		13 แกงเขียวหวาน
14	ข้าวสวย	ผัดผักรวม	แกงเลียง	หมูทอด	น้ำจับเลี้ยง		14 แกงเขียวหวาน
15	ข้าวสวย	ผัดผักรวม	แกงเลียง	ปลาทู	น้ำใบเตย	น้ำพริกหนุ่ม	15 พะแนง
16	ข้าวสวย	ผัดผักรวม	แกงเลียง	ปลาทู	น้ำจับเลี้ยง	น้ำพริกหนุ่ม	16 พะแนง
17	ข้าวสวย	ผัดผักรวม	แกงเลียง	ไข่เจียว	น้ำใบเตย		17 พะแนง
18	ข้าวสวย	ผัดผักรวม	แกงเลียง	ไข่เจียว	น้ำจับเลี้ยง		18 แกงเขียวหวาน
19	ข้าวสวย	คะน้าหมูกรอบ	ต้มแซ่บ	ไข่ดาว	น้ำใบเตย		19 แกงเขียวหวาน
20	ข้าวสวย	คะน้าหมูกรอบ	ต้มแซ่บ	ไข่ดาว	น้ำจับเลี้ยง		20 แกงเขียวหวาน
21	ข้าวสวย	คะน้าหมูกรอบ	ต้มแซ่บ	ไข่เจียว	น้ำใบเตย		21 แกงเขียวหวาน
22	ข้าวสวย	คะน้าหมูกรอบ	ต้มแซ่บ	ไข่เจียว	น้ำจับเลี้ยง		22 น้ำจับเลี้ยง
23	ข้าวสวย	คะน้าหมูกรอบ	ต้มแซ่บ	แคบหมู	น้ำใบเตย	น้ำพริกอ่อง	23 แกงเขียวหวาน
24	ข้าวสวย	คะน้าหมูกรอบ	ต้มแซ่บ	แคบหมู	น้ำจับเลี้ยง	น้ำพริกอ่อง	24 แกงเขียวหวาน
25	ข้าวสวย	คะน้าหมูกรอบ	พะแนง	ไข่ดาว	น้ำใบเตย		25 แกงเขียวหวาน
26	ข้าวสวย	คะน้าหมูกรอบ	พะแนง	ไข่ดาว	น้ำจับเลี้ยง		26 แกงเขียวหวาน
27	ข้าวสวย	คะน้าหมูกรอบ	พะแนง	ไข่เจียว	น้ำใบเตย		27 แกงเขียวหวาน
28	ข้าวสวย	คะน้าหมูกรอบ	พะแนง	ไข่เจียว	น้ำจับเลี้ยง		28 พะแนง
29	ข้าวสวย	คะน้าหมูกรอบ	พะแนง	แคบหมู	น้ำใบเตย	น้ำพริกอ่อง	29 พะแนง
30	ข้าวสวย	คะน้าหมูกรอบ	พะแนง	แคบหมู	น้ำจับเลี้ยง	น้ำพริกอ่อง	30 พะแนง
31	ข้าวสวย	คะน้าหมูกรอบ	แกงเลียง	ไข่ดาว	น้ำใบเตย		31 แกงเขียวหวาน
32	ข้าวสวย	คะน้าหมูกรอบ	แกงเลียง	ไข่ดาว	น้ำจับเลี้ยง		32 แกงเขียวหวาน
33	ข้าวสวย	คะน้าหมูกรอบ	แกงเลียง	ไข่เจียว	น้ำใบเตย		33 น้ำใบเตย
34	ข้าวสวย	คะน้าหมูกรอบ	แกงเลียง	ไข่เจียว	น้ำจับเลี้ยง		34 ข้าวสวย
35	ข้าวสวย	คะน้าหมูกรอบ	แกงเลียง	แคบหมู	น้ำใบเตย	น้ำพริกอ่อง	35 แกงเขียวหวาน
36	ข้าวสวย	คะน้าหมูกรอบ	แกงเลียง	แคบหมู	น้ำจับเลี้ยง	น้ำพริกอ่อง	36 แกงเขียวหวาน
37	ข้าวสวย	ผัดเต้าหู้	ต้มแซ่บ	หมูทอด	น้ำใบเตย		37 แกงเขียวหวาน
38	ข้าวสวย	ผัดเต้าหู้	ต้มแซ่บ	หมูทอด	น้ำจับเลี้ยง		38 แกงเขียวหวาน

39	ข้าวสวย	ผัดเต้าหู้	ต้มแซ่บ	ไข่เจียว	น้ำใบเตย		39 น้ำใบเตย
40	ข้าวสวย	ผัดเต้าหู้	ต้มแซ่บ	ไข่เจียว	น้ำจับเลี้ยง		40 น้ำจับเลี้ยง
41	ข้าวสวย	ผัดเต้าหู้	ต้มแซ่บ	ปลาทู	น้ำใบเตย	น้ำพริกหนุ่ม	41 แกงเขียวหวาน
42	ข้าวสวย	ผัดเต้าหู้	ต้มแซ่บ	ปลาทู	น้ำจับเลี้ยง	น้ำพริกหนุ่ม	42 แกงเขียวหวาน
43	ข้าวสวย	ผัดเต้าหู้	พะแนง	หมูทอด	น้ำใบเตย		43 พะแนง
44	ข้าวสวย	ผัดเต้าหู้	พะแนง	หมูทอด	น้ำจับเลี้ยง		44 พะแนง
45	ข้าวสวย	ผัดเต้าหู้	พะแนง	ไข่เจียว	น้ำใบเตย		45 ข้าวสวย
46	ข้าวสวย	ผัดเต้าหู้	พะแนง	ไข่เจียว	น้ำจับเลี้ยง		46 ข้าวสวย
47	ข้าวสวย	ผัดเต้าหู้	พะแนง	ปลาทู	น้ำใบเตย	น้ำพริกหนุ่ม	47 พะแนง
48	ข้าวสวย	ผัดเต้าหู้	พะแนง	ปลาทู	น้ำจับเลี้ยง	น้ำพริกหนุ่ม	48 พะแนง
49	ข้าวสวย	ผัดเต้าหู้	แกงเลียง	หมูทอด	น้ำใบเตย		49 แกงเขียวหวาน
50	ข้าวสวย	ผัดเต้าหู้	แกงเลียง	หมูทอด	น้ำจับเลี้ยง		50 แกงเขียวหวาน
51	ข้าวสวย	ผัดเต้าหู้	แกงเลียง	ไข่เจียว	น้ำใบเตย		51 แกงเขียวหวาน
52	ข้าวสวย	ผัดเต้าหู้	แกงเลียง	ไข่เจียว	น้ำจับเลี้ยง		52 แกงเขียวหวาน
53	ข้าวสวย	ผัดเต้าหู้	แกงเลียง	ปลาทู	น้ำใบเตย	น้ำพริกหนุ่ม	53 แกงเขียวหวาน
54	ข้าวสวย	ผัดเต้าหู้	แกงเลียง	ปลาทู	น้ำจับเลี้ยง	น้ำพริกหนุ่ม	54 แกงเขียวหวาน
55	ข้าวสวย	กระเพรา	แกงเขียวหวาน	ไข่เจียว	เก๊กฮวย		55 แกงเขียวหวาน
56	ข้าวสวย	กระเพรา	แกงเขียวหวาน	ไข่เจียว	กระเจี๊ยบ		56 แกงเขียวหวาน
57	ข้าวสวย	กระเพรา	แกงเขียวหวาน	ไข่ดาว	เก๊กฮวย		57 แกงเขียวหวาน
58	ข้าวสวย	กระเพรา	แกงเขียวหวาน	ไข่ดาว	กระเจี๊ยบ		58 ข้าวสวย
59	ข้าวสวย	กระเพรา	แกงเขียวหวาน	หมูทอด	เก๊กฮวย		59 ข้าวสวย
60	ข้าวสวย	กระเพรา	แกงเขียวหวาน	หมูทอด	กระเจี๊ยบ		60 พะแนง
61	ข้าวสวย	กระเพรา	ต้มจืด	ไข่ดาว	เก๊กฮวย		61 แกงเขียวหวาน
62	ข้าวสวย	กระเพรา	ต้มจืด	ไข่ดาว	กระเจี๊ยบ		62 ไข่ดาว
63	ข้าวสวย	กระเพรา	ต้มจืด	ไข่เจียว	เก๊กฮวย		63 พะแนง
64	ข้าวสวย	กระเพรา	ต้มจืด	ไข่เจียว	กระเจี๊ยบ		64 พะแนง
65	ข้าวสวย	กระเพรา	ต้มจืด	หมูทอด	เก๊กฮวย		65 พะแนง
66	ข้าวสวย	กระเพรา	ต้มจืด	หมูทอด	กระเจี๊ยบ		66 แกงเขียวหวาน
67	ข้าวสวย	ลาบ	แกงเขียวหวาน	ไข่เจียว	เก๊กฮวย		67 แกงเขียวหวาน

68	ข้าวสวย	ลาบ	แกงเขียวหวาน	ไข่เจียว	กระเจี๊ยบ		68 แกงเขียวหวาน
69	ข้าวสวย	ลาบ	แกงเขียวหวาน	ไข่ดาว	เก็กฮวย		69 แกงเขียวหวาน
70	ข้าวสวย	ลาบ	แกงเขียวหวาน	ไข่ดาว	กระเจี๊ยบ		70 แกงเขียวหวาน
71	ข้าวสวย	ลาบ	แกงเขียวหวาน	หมูทอด	เก็กฮวย		71 แกงเขียวหวาน
72	ข้าวสวย	ลาบ	แกงเขียวหวาน	หมูทอด	กระเจี๊ยบ		72 ข้าวสวย
73	ข้าวสวย	ลาบ	ต้มจืด	ไข่ดาว	เก็กฮวย		73 แกงเขียวหวาน
74	ข้าวสวย	ลาบ	ต้มจืด	ไข่ดาว	กระเจี๊ยบ		74 แกงเขียวหวาน
75	ข้าวสวย	ลาบ	ต้มจืด	ไข่เจียว	เก็กฮวย		75 พะแนง
76	ข้าวสวย	ลาบ	ต้มจืด	ไข่เจียว	กระเจี๊ยบ		76 แกงเขียวหวาน
77	ข้าวสวย	ลาบ	ต้มจืด	หมูทอด	เก็กฮวย		77 แกงเขียวหวาน
78	ข้าวสวย	ลาบ	ต้มจืด	หมูทอด	กระเจี๊ยบ		78 ข้าวสวย
79	ข้าวผัด	ผัดผักรวม	ต้มแซ่บ	ไข่เจียว	น้ำใบเตย		79 ข้าวสวย
80	ข้าวผัด	ผัดผักรวม	ต้มแซ่บ	ไข่เจียว	น้ำจับเลี้ยง		80 ข้าวสวย
81	ข้าวผัด	ผัดผักรวม	พะแนง	ไข่เจียว	น้ำใบเตย		81 ไข่เจียว
82	ข้าวผัด	ผัดผักรวม	พะแนง	ไข่เจียว	น้ำจับเลี้ยง		82 ไข่เจียว
83	ข้าวผัด	กระเพรา	ต้มจืด	ไข่เจียว	เก็กฮวย		83 ข้าวสวย
84	ข้าวผัด	กระเพรา	ต้มจืด	ไข่เจียว	กระเจี๊ยบ		84 แกงเขียวหวาน
85	ข้าวผัด	กระเพรา	ต้มจืด	ไข่ดาว	เก็กฮวย		85 ข้าวสวย
86	ข้าวผัด	กระเพรา	ต้มจืด	ไข่ดาว	กระเจี๊ยบ		86 ข้าวสวย
87	ข้าวผัด	ลาบ	แกงเขียวหวาน	หมูทอด	เก็กฮวย		87 ข้าวสวย
88	ข้าวผัด	ลาบ	แกงเขียวหวาน	หมูทอด	กระเจี๊ยบ		88 ข้าวสวย
89	ข้าวผัด	ลาบ	ต้มจืด	ไข่ดาว	เก็กฮวย		89 แกงเขียวหวาน
90	ข้าวผัด	ลาบ	ต้มจืด	ไข่ดาว	กระเจี๊ยบ		90 แกงเขียวหวาน
91	เส้นก๋วยเตี๋ยว		พะแนง	ไข่มะตูม	น้ำใบเตย		91 พะแนง
92	เส้นก๋วยเตี๋ยว		พะแนง	ไข่มะตูม	น้ำจับเลี้ยง		92 พะแนง
93	เส้นก๋วยเตี๋ยว		พะแนง	ไข่มะตูม	เก็กฮวย		93 พะแนง
94	เส้นก๋วยเตี๋ยว		พะแนง	ไข่มะตูม	กระเจี๊ยบ		94 พะแนง
95	เส้นก๋วยเตี๋ยว		แกงเขียวหวาน	ไข่มะตูม	น้ำใบเตย		95 ข้าวสวย
96	เส้นก๋วยเตี๋ยว		แกงเขียวหวาน	ไข่มะตูม	น้ำจับเลี้ยง		96 ข้าวสวย

97	เส้นก๋วยเตี๋ยว		แกงเขียวหวาน	ไข่มะตูม	เก๊กฮวย		97 แกงเขียวหวาน
98	เส้นก๋วยเตี๋ยว		แกงเขียวหวาน	ไข่มะตูม	กระเจี๊ยบ		98 ข้าวสวย
99	ข้าวผัด	ผัดเต้าหู้	แกงเลียง	หมูทอด	น้ำใบเตย		99 ข้าวสวย
100	ข้าวผัด	ผัดเต้าหู้	แกงเลียง	หมูทอด	น้ำจับเลี้ยง		100 ข้าวสวย

ตารางที่ 4.1 แสดงค่าความแม่นยำในการแนะนำอาหาร

จากแสดงดังตารางที่ 4.1 จากการทำนายแนะนำอาหารจากชุดอาหารทั้งหมด 100 ชุด พบว่าเมื่อทำการสุ่มอาหารในชุดอาหาร รูปภาพอาหารที่คืนค่ากลับมาแล้วมีเหมาะสมกับเข้ากันกับชุดอาหารนั้น มีความถูกต้อง 38 ชุดอาหาร จาก 100 ชุดอาหาร คิดค่าเป็นมีความแม่นยำในการทำนายเท่ากับร้อยละ 38

4.3 วิเคราะห์ผลการทดลอง

ดังนั้นจากผลการทดลองการสร้างแบบจำลองระบบพบว่า ความแม่นยำในการจดจำและจำแนกรูปภาพอาหารอยู่ที่ 39.61% และความแม่นยำในการแนะนำอาหารอยู่ที่ 38% ซึ่งค่าความแม่นยำที่ได้ต่ำนี้เกิดจากฐานข้อมูลภาพอาหารที่มีอยู่อย่างไม่เพียงพอ ในการทดลองนี้ใช้ภาพถ่ายในการเทรนด์เพียง 2,500 ภาพ ซึ่งภาพอาหารไทยในแต่ละหมวดมีความคล้ายคลึงกันทำให้ เมื่อใช้ภาพถ่ายในปริมาณเพียงเท่านี้จึงไม่สามารถคืนค่าผลความแม่นยำที่เป็นที่พึงพอใจได้ โดยจากรูปภาพที่ 4.2 และตารางที่ 4.1 จะสังเกตได้ว่าระบบไม่สามารถแยกแยะภาพอาหารจำพวกผัดออกจากภาพอาหารจำพวกแกงได้เลย และไม่สามารถแยกภาพแกงที่มีสีใกล้เคียงกันได้เช่นกัน เช่น แกงเลียง กับแกงเขียวหวาน ดังแสดงในรูปภาพที่ 4.3

รูปภาพที่ 4.3 ภาพซ้ายบนแกงเลียง ภาพขวาบนแกงเขียวหวาน ภาพซ้ายล่างข้าว ภาพขวาล่างเส้นก๋วยเตี๋ยว

ซึ่งเป็นลักษณะเดียวกันกับการระบุจำแนกระหว่างข้าวสวยและเส้นก๋วยเตี๋ยว โดยความแม่นยำในการระบุเส้นก๋วยเตี๋ยวอยู่ที่ 42.4% ข้าวสวยอยู่ที่ 48.8% จะเห็นว่าเป็นค่าที่มีความคาบครึ่งกัน เนื่องจากอาหารที่มีสีขาวเป็นหลักจากข้อมูลภาพที่ป้อนเข้าไปมีเพียงข้าวสวยและเส้นก๋วยเตี๋ยว ซึ่งระบบสามารถจำแนกอาหารสองชนิดนี้ออกจากอาหารชนิดอื่นได้ แต่ไม่สามารถระบุแยกทั้งสองชนิดออกจากกันเองได้ ซึ่งจากการศึกษาคาดว่าหากต้องการแผ่นความสามารถในการจำแนก ต้องเพิ่มฐานข้อมูล เพื่อให้ระบบสามารถทำการสกัดแยกคุณลักษณะ

ในส่วนของพื้นผิว (Texture) และค่าสีออกมาได้ชัดเจนมากกว่าผลที่ได้ในการทดลองนี้ ในขณะเดียวกันภาพที่ไม่มีความซับซ้อนมาก เช่น ภาพน้ำเก๊กฮวย ระบบโครงข่ายประสาทสามารถทำการระบุได้อย่างค่อนข้างแม่นยำถูกต้องดังแสดงการอ่านค่าความแม่นยำในรูปที่ 4.2 โดยน้ำเก๊กฮวยมีค่าความแม่นยำ 99.6% และ 92.2% เมื่อทำการพิจารณาค่าการเรียนรู้จากภาพที่ 4.1 กราฟแสดงผลค่าความแม่นยำในการจำแนกจดจำภาพถ่าย จะเห็นว่าเมื่อทำการวนรอบ Epoch ระบบทำการละทิ้งตัวชี้วัด(Parameters) บางตัวในช่วงแรก แต่หลังจากนั้นการแปลงเป็นไปในระดับต่ำ เนื่องจากฐานข้อมูลที่ไม่เพียงพอ ทำให้การทำ Relu ในแต่ละขั้นของ Hidden layer ของโครงข่ายประสาทเป็นไปได้น้อย แลเกิดการเรียนรู้ที่ไม่เพียงพอ

บทที่ 5

สรุปผลการทดลอง

5.1 สรุปผลการทดลอง

รายงานฉบับนี้เป็นการทดลองสร้างแบบจำลองระบบแนะนำอาหารโดยใช้ deep learning และ recurrent neural network เพื่อเป็นการสร้างระบบแนะนำอาหารที่แม่นยำขึ้น

ผลการทดลองพบว่า ค่า accuracy ของ neaural network อยู่ที่ 39.61% ซึ่งเป็นค่าที่ไม่สูงมากพอ ทำให้เกิดความผิดพลาดในการทำนายภาพของ Neural network ซึ่งเมื่อมีการทำนายภาพผิดก็จะส่งผลกระทบไปสู่ขั้นตอนการแนะนำอาหารซึ่งในส่วนของการแนะนำอาหาร ก็มีความแม่นยำในการทำนายอาหารอยู่ที่ประมาณ 38% ซึ่งยังเป็นค่าที่ไม่สูงนัก ทั้งนี้ปัจจัยที่มีผลต่อความแม่นยำในการทำนายภาพมีดังนี้ 1. จำนวนรูปภาพที่ใช้ 2. จำนวนชนิดของอาหาร 3. ชนิดของ Layer ที่ใช้ในการสกัดหา Feture 4. ความหลากหลายของข้อมูลที่ใช้ โดยในส่วนของ RNN พบว่าสามารถนำมาใช้งานในการแนะนำอาหารได้หากมีฐานข้อมูลที่สมบูรณ์พร้อมมากพอ เช่น การใช้ฐานข้อมูลรูปภาพจาก Food-101 ซึ่งมีการคลังรูปภาพราว 101,000 รูป ดังนั้นจึงสรุปผลจากการสร้างระบบจำลองการแนะนำอาหารได้ว่าการใช้โครงข่ายประสาทเทียมในการจดจำและจำแนกรูปภาพสามารถทำได้ และ RNN มีแนวโน้มที่จะสามารถสร้างระบบแนะนำได้ แต่ต้องใช้ฐานข้อมูลในการเทรนด์ระบบที่มีจำนวนและความหลากหลายมากกว่านี้ และจากการทดลองในครั้งนี้ใช้ภาพอาหารที่เห็นแต่เพียงภาพอาหารเพียงอย่างเดียว แนวทางในการพัฒนาในอนาคตคาดว่าควรทดลองใช้ภาพที่ประกอบด้วยองค์ประกอบหลายอย่าง และใช้การทำ ROI หรือการใช้ฟังก์ชั่นสำเร็จรูป Bounding box ในโปรแกรม MATLAB ทำการระบุพื้นที่ที่สนใจ เพื่อเพิ่มองค์ประกอบในการสกัดแยกคุณลักษณะออกมา ทั้งนี้หากฐานข้อมูลมีมากเพียงพอจะสามารถทำ Hidden layer ของโครงข่ายประสาทเทียมที่ซับซ้อนขึ้นได้ และทำการเวท (Weight) น้ำหนักข้อมูล เพื่อเพิ่มประสิทธิภาพการเรียนรู้ของโครงข่ายประสาทได้เช่นกัน

บรรณานุกรม

[1] Food object recognition using a mobile device: Evaluation of currently implemented systems, Simon Knez, and Luka Šajn, Elsevier; Trends in Food Science & Technology (2020)

[2] Human activity classification based on sound recognition and residual convolutional neural network, Minhyuk Junga and Seokho Chib, Elsevier; Automation in Construction (2020)

[3] Kurunakaran, D., 2018 “Deep learning series 1: Intro to deep learning,” Available at: https://medium.com/intro-to-artificial-intelligence/deep-learning-series-1-intro-to-deep-learning-abb1780ee20 (Accessed: 9 January 2020)

[4] “Pretrained Deep Neural Networks”, 2019 Available at: https://www.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html (Accessed: 9 January 2020)

[5] JahKnows, Moller, D., and Pinetz, T., 2017, “Merge two different deep learning models in Keras” Available at: https://stackoverflow.com/questions/47035367/merge-two-different-deep-learning-models-in-keras (Accessed: 9 January 2020)

[6] Prabhu, 2018, “Understanding of Convolutional Neural Network (CNN) — Deep Learning,” Available at: https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148 (Accessed: 9 January 2020)

[7] Aizawa, K., Maruyama, Y., Li, H., & Morikawa, C. (2013). Food balance estimation by using personal dietary tendencies in a multimedia food log. IEEE Transactions on Multimedia, 15, 2176–2185

[8] Naïve Bayesian Classification, Available at: wipawanblog.files.wordpress.com (Accessed: 9 April 2020)

[9] Oliveira, L., Costa, V., Neves, G., Oliveira, T., Jorge, E., & Lizarraga, M. (2014). A mobile, lightweight, poll-based food identification system. Pattern Recognition, 47, 1941–1952. S.-H. Park, J. Fü

[10] Image processing, Available at: https://www.cs.auckland.ac.nz (Accessed: 9 April 2020)

[11] แนะนำสู่การประมวลภาพดิจิตอล, Available at: http://fivedots.coe.psu.ac.th/~montri/Teaching/image/chap1.htm (Accessed: 9 April 2020)

[12] G. Waltner, M. Schwarz, S. Ladstätter, A. Weber, P. Luley, H. Bischof, M. Lindschinger, I. Schmid, L. Paletta, Mango-mobile augmented reality with functional eating guidance and food awareness, in: International conference on image analysis and processing, Springer, pp. 425–432.

[13] Bounding Box, Available at: https://www.mathworks.com/matlabcentral/answers/87597-rectangle-around-the-object-bounding-box (Accessed: 9 April 2020)

[14] Mezgec, S., & Korousˇic Seljak, B. (2017). Nutrinet: A deep learning food and drink image recognition system for dietary assessment. Nutrients, 9, 657.