



การจำแนกภาพอวัยวะในระบบทางเดินอาหารจากกล้องแคปซูลไร้สาย
โดยใช้โครงข่ายประสาทเทียมคอนโวลูชัน

Digestive Organ Classification from Wireless Capsule Endoscopy
Images with Convolutional Neural Network

นางสาวสุกัศฉี ปลุกอร่าม
Supaksagee Plukaram

นายสิปปภาส ทรัพย์สนอง
Sippapart Subsanong

นางสาวสุประวีณ์ สร้อยทองเจริญ
Supawee Soythongcharoen

โครงการนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

มหาวิทยาลัยศรีนครินทรวิโรฒ ปีการศึกษา พ.ศ. 2563



คณะวิทยาศาสตร์มหาวิทยาลัยศรีนครินทรวิโรฒ

ชื่อโครงการ การจำแนกภาพอวัยวะในระบบทางเดินอาหารจากกล้องแคปซูลไร้สายโดยใช้โครงข่ายประสาทเทียมคอนโวลูชัน

Digestive Organ Classification from Wireless Capsule Endoscopy Images with Convolutional Neural Network

นิสิต	นางสาวสุภัคศรี	ปลุกอร่าม	60102010348
	นายสิปปภาส	ทรัพย์สนอง	60102010582
	นางสาวสุประวีณ์	สร้อยทองเจริญ	60102010584

ปริญญา วิทยาศาสตร์บัณฑิต (วท.บ)

ภาควิชา วิทยาการคอมพิวเตอร์

อาจารย์ที่ปรึกษาโครงการ ผศ.ดร.นุวิทย์ วิวัฒนวัฒนา

ลงชื่อ.....

(ผศ.ดร.นุวิทย์ วิวัฒนวัฒนา)

อาจารย์ที่ปรึกษาโครงการ

บทคัดย่อ

ตำแหน่งของบาดแผลเป็นข้อมูลสำคัญที่จะต้องรายงานโดยแพทย์ระบบทางเดินอาหารในการตีความรูปภาพจากกล้องแคปซูลไร้สาย ซึ่งในปัจจุบันนี้การศึกษา เกี่ยวกับการใช้ แบบจำลองการเรียนรู้เชิงลึกที่จะนำไปใช้กับการจำแนกตำแหน่งของระบบทางเดินอาหารมีจำนวนน้อย ดังนั้นกลุ่มผู้วิจัยจึงมีเป้าหมายในการสร้างแบบ จำลองการเรียนรู้เชิงลึกสำหรับการจำแนก คลาสของตำแหน่งภายในระบบทางเดินอาหาร ได้แก่ หลอดอาหาร กระเพาะอาหาร ลำไส้เล็ก และ ลำไส้ใหญ่ โดยใช้ชุดข้อมูล ภาพจากกล้องแคปซูลไร้สายจำนวน 723,681 ภาพ จากผู้ป่วย 174 คน ซึ่งถูกแบ่งในการ training เป็นจำนวน 40,000 ภาพ และ testing 683,681 ภาพ โดยภาพจะถูก label เป็น 4 ตำแหน่งอวัยวะ (ได้แก่ หลอดอาหาร กระเพาะอาหาร ลำไส้เล็ก และ ลำไส้ใหญ่) ทางกลุ่มผู้วิจัยใช้การเรียนรู้เชิงลึกสถาปัตยกรรม InceptionResnet V2 ในการ training ชุดข้อมูล จากนั้นจึง ตรวจสอบประสิทธิภาพของ Model ที่ถูก train โดยการใช้ชุดข้อมูลสำหรับ test ซึ่ง Confusion matrix ของแบบจำลองการเรียนรู้ เชิงลึกแสดงให้เห็นถึงประสิทธิภาพของอัลกอริทึมของเรา โดย Model ได้จำแนกรูปภาพจากกล้องแคปซูลไร้สายเป็น 4 คลาส ได้ค่า Accuracy 94.17%, Precision 97.12%, Recall (Sensitivity) 94.53% และ F1-score 95.49% มีค่าเฉลี่ยในการทำนาย รูปภาพอยู่ที่ 20 milliseconds ต่อ 1 รูปภาพ นอกจากนี้ ทางกลุ่มผู้วิจัยได้ทดลองทำ Ensemble โดยผลที่ได้นั้นในภาพรวมแล้ว ดีกว่าการใช้โมเดลตัวเดียว ได้ค่า Accuracy 94.21%, Precision 97.41%, Recall (Sensitivity) 94.22% และ F1-score 95.39% การสาธิตแบบจำลองการเรียนรู้เชิงลึก ดังกล่าวให้ผลลัพธ์ที่ดีซึ่งสามารถใช้ในการสร้างเป้าหมายในการสร้าง Model อัตโนมัติ อย่างเต็มรูปแบบสำหรับการตีความรูปภาพจากกล้องแคปซูลไร้สายได้

Abstract

Lesion location is a critical information that must be reported by gastroenterologists when interpreting capsule endoscopy images. There are few studies on how deep learning models can be applied to classify the locations of gastrointestinal tract. Therefore, we aim to create a deep learning model for classifying the locations of gastrointestinal tract (i.e., esophagus, stomach, small bowel, and colon) based on capsule endoscopy images. Dataset of capsule endoscopy images ($n = 723,681$ images from 174 patients) was divided into training ($n = 40,000$) and testing ($n = 683,674$). The images were labeled into 4 organ locations (esophagus, stomach, small bowel, and colon). We applied the deep learning architecture (InceptionResnet V2) to the training dataset. Then, the performance of the trained model was externally validated using the testing dataset. The confusion matrix of the deep learning model, visualizing the performance of our algorithm. The model classified capsule endoscopy images into 4 classes with accuracy of 97.38%, precision of 89.25%, recall (sensitivity) of 94.07%, and f1-score of 94.91%. The average prediction time was 20 milliseconds per image. In addition, we have used ensemble for comparison and the results were slightly better with accuracy of 94.21%, precision of 97.41%, recall (sensitivity) of 94.22%, and f1-score of 95.39%. The deep learning model demonstrated an excellent classification performance, which can be used as a building block for the ultimate goal of creating a fully automated model for interpreting capsule endoscopy images.

กิตติกรรมประกาศ

การดำเนินวิจัยในครั้งนี้ต้องขอขอบคุณอาจารย์ทางภาควิชาวิทยาการคอมพิวเตอร์ที่ได้ให้เซฟเวอร์การใช้งาน รวมถึงห้องวิจัยในการทำวิจัยของพวกเรา และต้องขอขอบคุณ อาจารย์ที่ปรึกษาวิจัยที่ได้ให้คำปรึกษา แนะนำ ช่วยเหลือ และ คอยสนับสนุนตลอดระยะเวลาที่ได้ทำงานจนกระทั่งโครงการชิ้นนี้ได้ออกมาสมบูรณ์และมีประสิทธิภาพ

ขอขอบคุณมหาวิทยาลัยศรีนครินทรวิโรฒแห่งนี้ที่ได้อำนวยความสะดวกเป็นแหล่ง ที่ศึกษาสำหรับการดำเนินการวิจัย

ขอขอบคุณพระคุณ ผศ.ดร.นุรีย์ วิวัฒน์วัฒนา ที่ได้กรุณาชี้แนะแนวทางและคอย สนับสนุนทุกอย่างพร้อมให้คำแนะนำ แนวความคิดและไตร่ตรองปัญหาต่าง ๆ ได้อย่างรอบ คอบมากยิ่งขึ้นจนทำให้โครงการนี้สำเร็จลงได้

ขอขอบพระคุณ คุณหม่อมพร อัครรุ่งเรืองกิจ และ ดร.กวี นำพาเจริญ ที่ได้ให้ความอนุเคราะห์ข้อมูลภาพถ่ายภายในระบบทางเดินอาหารจากกล้องแคปซูลไร้สายซึ่งสำคัญเป็นอย่างยิ่งในการทำโครงการชิ้นนี้

ขอขอบพระคุณคณะอาจารย์ทุกท่านที่ได้ให้ความรู้ คำแนะนำและคอยช่วยเหลือ พวกเราตั้งแต่ได้เข้ามาเรียนที่นี่จนมาถึงช่วงสุดท้าย และได้ทำวิจัยจบครั้งนี้สำเร็จไปด้วยดี

ขอขอบคุณ funding จาก Google Cloud Platform ซึ่งเป็นเครื่องมือหลักในการ ทำโครงการนี้

สุดท้ายนี้ทางกลุ่มผู้จัดทำนั้นต้องขอขอบคุณกัลยาณมิตรที่ได้ให้ความช่วยเหลือและ เป็นกำลังใจให้กลุ่มผู้จัดทำตลอดมา ขอขอบคุณพี่ๆ เพื่อนๆ น้องๆ ผู้ซึ่งคอยให้ความช่วย เหลือและคำแนะนำแก่กลุ่มผู้จัดทำด้วยความเต็มใจยิ่งเสมอมา

ดังนั้นคณะผู้จัดทำนั้นได้หวังว่าโครงการการจำแนกภาพอวัยวะในระบบทางเดินอาหารจากกล้องแคปซูลไร้สายโดยใช้โครงข่ายประสาทเทียมคอนโวลูชัน (Digestive Organ Classification from Wireless Capsule Endoscopy Images with Convolutional Neural Network) จะเป็นประโยชน์ต่อการศึกษานี้ในภายภาคหน้า ตลอดจนสามารถที่จะนำไปพัฒนา ต่อยอดได้ในอนาคต

สารบัญ

	หน้า
บทคัดย่อ	ข
Abstract	ค
กิตติกรรมประกาศ	ง
สารบัญ	ฉ
สารบัญรูปภาพ	ช
สารบัญตาราง	ฌ
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญ	1
1.2 วัตถุประสงค์ของโครงการ	4
1.3 ขอบเขตของโครงการ	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ	5
บทที่ 2 องค์ความรู้ที่เกี่ยวข้อง	6
2.1 องค์ความรู้ที่เกี่ยวข้อง	6
2.2 งานวิจัยที่เกี่ยวข้อง	14
บทที่ 3 วิธีการดำเนินโครงการ	19
3.1 วิธีการดำเนินงาน	19
3.2 อุปกรณ์และเครื่องมือที่ใช้	20
3.3 ชุดข้อมูล (Data set) ที่ใช้งาน	20
3.4 การออกแบบระบบ	21
3.5 ปัญหาและอุปสรรค	31
บทที่ 4 ผลการดำเนินโครงการ	32

บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ	50
5.1 สรุปผลการศึกษาค้นคว้า	50
5.2 อภิปรายผล	50
5.3 ข้อเสนอแนะ	51
บรรณานุกรม	52



สารบัญรูปภาพ

	หน้า
ภาพที่ 1 : ภาพอวัยวะเกี่ยวกับทางเดินอาหาร	2
ภาพที่ 2 : ตัวอย่างภาพที่ได้จากกล้องแคปซูลไร้สาย	2
ภาพที่ 3 : ตัวอย่างภาพเมื่อนำมาจำแนกเป็น 4 คลาส	3
ภาพที่ 4 : Capsule Endoscopy	6
ภาพที่ 5 : สถาปัตยกรรม ResNet50	9
ภาพที่ 6 : สถาปัตยกรรม Inception	10
ภาพที่ 7 : สถาปัตยกรรม InceptionResnetV2	11
ภาพที่ 8 : สูตรการหาค่า Precision	12
ภาพที่ 9 : สูตรการหาค่า Accuracy	12
ภาพที่ 10 : สูตรการหาค่า Specificity	12
ภาพที่ 11 : สูตรการหาค่า Sensitivity	13
ภาพที่ 12: สูตรการหาค่า F1 - Score	13
ภาพที่ 13 : แผนผังการดำเนินงาน	21
ภาพที่ 14 : โพลเดอร์ Images ภายใน Drop Box	22
ภาพที่ 15 : โพลเดอร์ Label จาก Dropbox	22
ภาพที่ 16 : โพลเดอร์ Time จาก Dropbox	23
ภาพที่ 17 : ตัวอย่างไฟล์ mapping และไฟล์ time	23
ภาพที่ 18 : ข้อมูลภาพของผู้ป่วย 1 คน	24
ภาพที่ 19 : โพลเดอร์ที่แบ่งตามจำนวนคลาสทั้ง 10 คลาส	24
ภาพที่ 20 : ตารางบ่งบอกขอบเขตของแต่ละคลาส	25
ภาพที่ 21 : ภาพไฟล์ mapping สำหรับการจำแนกภาพ	25
ภาพที่ 22 : ตัวอย่างการเลือกภาพในคลาส Other	26
ภาพที่ 23 : ภาพแรกของ Esophagus	26

ภาพที่ 24 : นำภาพใสในโพลเดอร์ Other	26
ภาพที่ 25 : นำภาพใสในโพลเดอร์ Esophagus	27
ภาพที่ 26 : นำภาพใสในโพลเดอร์ Esophagus	27
ภาพที่ 27 : โพลเดอร์ของแต่ละคนก็นำมารวม	28
ภาพที่ 28 : ภาพบางส่วนของโพลเดอร์ที่ทำการแบ่งกลุ่ม	29
ภาพที่ 29 : เลือกภาพจากหลายๆคนมาให้เหลือเพียง 4 class	30
ภาพที่ 30 : ภาพในส่วนของ Colon จากการสุ่มหลายๆคนมารวมกัน	30
ภาพที่ 31 : ค่า accuracy ของ InceptionResnetV2	40
ภาพที่ 32 : ค่า accuracy ของ InceptionV3	40
ภาพที่ 33 : ค่า accuracy ของ Resnet50	41
ภาพที่ 34 : Confusion matrix ของ model InceptionResnetV2	42
ภาพที่ 35 : Confusion matrix ของ Ensemble	43
ภาพที่ 36 : ตัวอย่างของ probability ของผลการทำนายจาก InceptionResnetV2	44
ภาพที่ 37 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะหลอดอาหาร ของ Ensemble	44
ภาพที่ 38 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะกระเพาะอาหาร ของ Ensemble	45
ภาพที่ 39 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะลำไส้เล็ก ของ Ensemble	45
ภาพที่ 40 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะลำไส้ใหญ่ ของ Ensemble	46
ภาพที่ 41 : การเทียบผลระหว่าง InceptionResnetV2 กับ Ensemble	47
ภาพที่ 42 : Graph classification reports ของ InceptionResnetV2 และ Ensemble	51

สารบัญตาราง

	หน้า
ตารางที่ 1 : ตัวอย่างจำนวนภาพที่ใช้สำหรับ train set	33 - 36
ตารางที่ 2 : ตัวอย่างจำนวนภาพที่ใช้สำหรับ valid set	36 - 38
ตารางที่ 3 : ตัวอย่างจำนวนภาพที่ใช้สำหรับ test set	38
ตารางที่ 4 : InceptionResnetV2 classification report	48
ตารางที่ 5 : Ensemble classification report	49



บทที่ 1 บทนำ

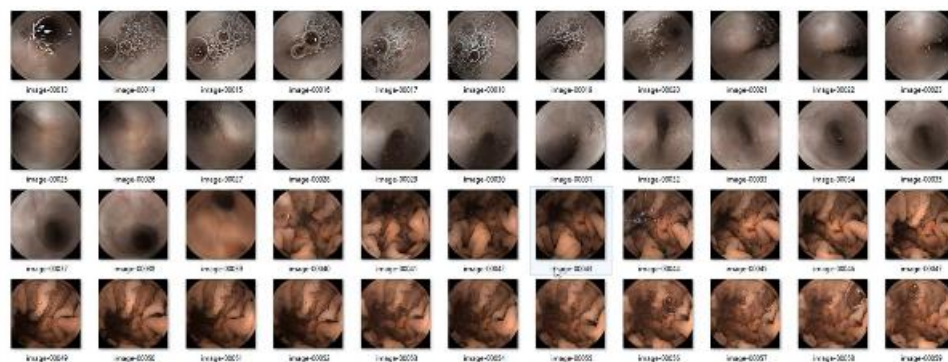
1.1 ที่มาและความสำคัญ

ในปัจจุบันนี้ เทคโนโลยีทางการแพทย์ที่ใช้ในการตรวจวินิจฉัยโรคในระบบทางเดินอาหารที่ดีที่สุดคือ กล้องแคปซูลไร้สาย (Wireless Capsule Endoscopy)[1] ซึ่งเป็นวัสดุขนาดเล็กและเบา ซึ่งมีลักษณะปลายมนโค้งเป็นพลาสติกใส มีตัวให้แสงสว่างและตัวบันทึกภาพ การใช้กล้องแคปซูลไร้สายในการตรวจจะทำให้สามารถเห็นภาพส่วนต่างๆของระบบทางเดินอาหาร (ดังภาพที่ 1) ตั้งแต่ต้นจนจบจากการ ถ่ายภาพนับหมื่นภาพ ทำให้แพทย์สามารถวินิจฉัย ได้อย่างแม่นยำ และมีประสิทธิภาพ

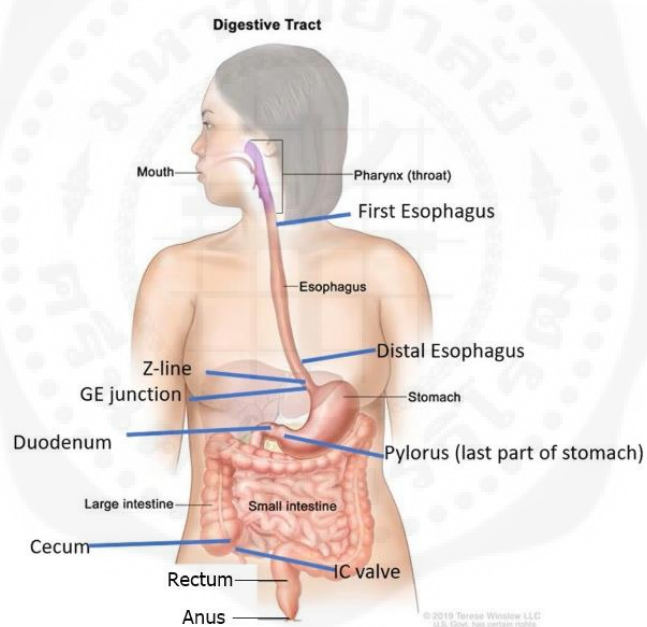
กล้องแคปซูลไร้สายทำให้สามารถเห็นความผิดปกติภายในระบบทางเดินอาหาร เช่น ก้อนเนื้อที่อาจตรวจพบภายหลังว่าเป็นมะเร็ง พยาธิ หรืออาจพบสภาวะเลือดออกในทางเดินอาหาร หรือ Gastrointestinal Bleeding (GI Bleeding)[2] เป็นสภาวะที่เกิดจากความผิดปกติของระบบทางเดินอาหารโดยผู้ป่วยจะมีอาการอาเจียน ปนเลือด หรือการมีเลือดปะปนอยู่ในอุจจาระซึ่งแพทย์ จะทำการวินิจฉัยอาการเลือดออกในทางเดินอาหารโดยการใช้กล้องแคปซูลไร้สาย ซึ่งทำให้สามารถเห็นจุดเลือดออกได้ โดยระหว่างการอ่านผลจากกล้องแคปซูลไร้สาย ทีมแพทย์จำเป็นต้องทราบภาพนั้น มาจากอวัยวะส่วนใดในระบบทางเดินอาหาร

ทางกลุ่มผู้วิจัยมีความสนใจที่จะนำภาพที่ได้จากกล้องแคปซูลไร้สาย (ดังภาพที่ 2) ผ่านกระบวนการ Deep convolutional neural networks (CNNs) ในการจำแนกภาพ (Classification) ภายในระบบทางเดินอาหารที่ต้องการจำแนกจำนวน 4 อวัยวะ ที่สนใจ ได้แก่ หลอดอาหาร (Esophagus) กระเพาะอาหาร (Stomach) ลำไส้เล็ก (Small Bowel) และ ลำไส้ใหญ่ (Colon) รวม 4 คลาส

(ดังภาพที่ 3) ซึ่งจะทำให้มีความสะดวกและความรวดเร็วมากขึ้นในการวิเคราะห์ อาการในขั้นต่อไป ทั้งนี้ในจำนวน 4 อวัยวะ ที่กล่าวมานั้นมีความน่าจะเป็น ที่จะสามารถแยกจุดที่น่าสนใจ (Landmark) ซึ่งเป็นจุดที่อยู่ระหว่างอวัยวะได้ด้วย

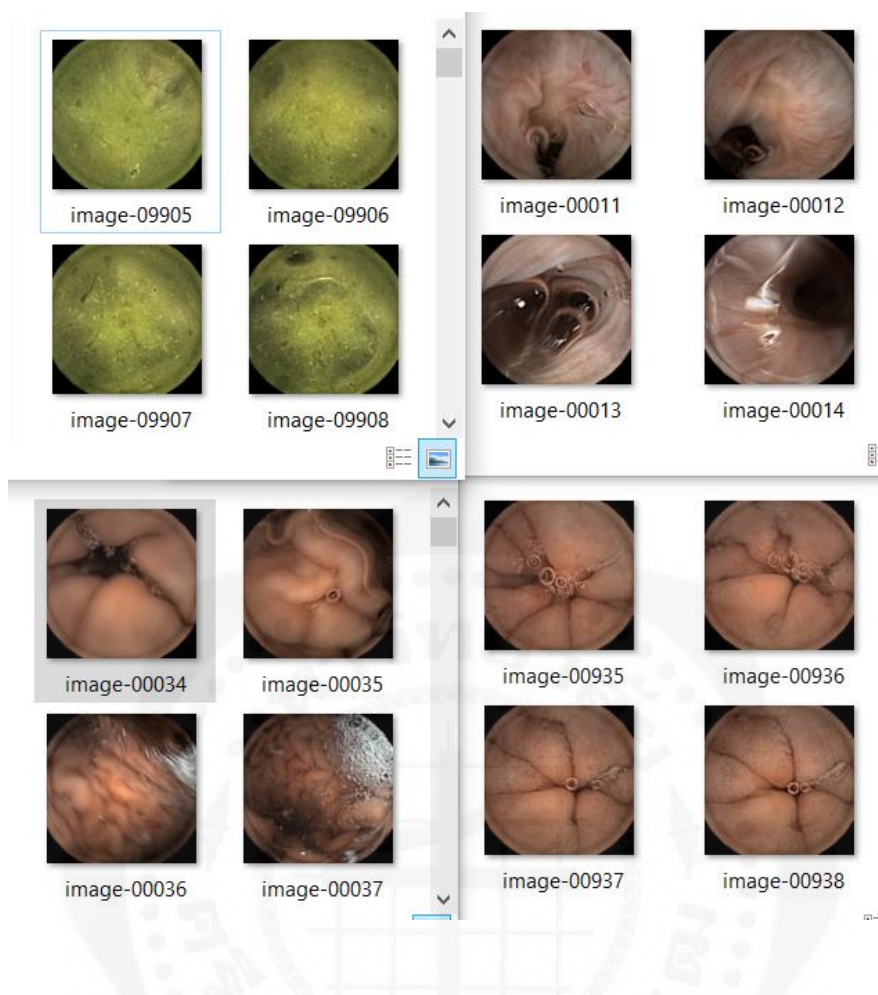


ภาพที่ 1 : ภาพอวัยวะเกี่ยวกับทางเดินอาหาร



ภาพที่ 2 : ตัวอย่างภาพที่ได้จากกล้องแคปซูลไร้สาย

(ที่มา : Brown University และ Rhode Island Hospital โดยความร่วมมือกับ Amporn Atsawarungruangkit, MD)



ภาพที่ 3 : ตัวอย่างภาพเมื่อนำมาจำแนกเป็น 4 คลาส

(ที่มา : Brown University และ Rhode Island Hospital โดยความร่วมมือกับ Amporn Atsawarungruangkit, MD)

1.2 วัตถุประสงค์ของโครงการ

- 1.2.1 เพื่อศึกษาการจำแนกภาพที่ได้จากกล้อง Wireless Capsule Endoscopy
- 1.2.2 เพื่อศึกษาการทำงานของ Convolutional neural networks (CNNs)
- 1.2.3 เพื่อสร้างแบบจำลองการจำแนกภาพอวัยวะ 4 อวัยวะ ได้แก่ หลอดอาหาร (Esophagus) กระเพาะอาหาร (Stomach) ลำไส้เล็ก (Small Bowel) และ ลำไส้ใหญ่ (Colon)
- 1.2.4 เพื่อประเมินประสิทธิภาพของแบบจำลองที่สร้าง

1.3 ขอบเขตของโครงการ

- 1.3.1 ศึกษาการจำแนกภาพที่ได้จากกล้องแคปซูลไร้สาย ภายในหลอดอาหาร กระเพาะอาหาร ลำไส้เล็ก ลำไส้ใหญ่ โดยใช้สถาปัตยกรรม CNN แบบ ResNet 50 , GoogLeNet/Inception และ InceptionResnetV2
- 1.3.2 สร้างแบบจำลองโดยใช้สถาปัตยกรรมที่ได้ศึกษา
- 1.3.3 ใช้ภาพจากกล้องแคปซูลไร้สายจาก Brown University และ Rhode Island Hospital โดยความร่วมมือกับ Amporn Atsawarungruangkit, MD ซึ่งได้ถูก label ไว้เรียบร้อยแล้ว
- 1.3.4 ทดสอบประสิทธิภาพของแบบจำลองโดยใช้เมตริก Accuracy Sensitivity Specificity Precision และ F1-score
- 1.3.5 วิเคราะห์ภาพที่น่าจะเป็นจุดที่น่าสนใจ (Landmark) จากผลการจำแนกภาพด้วย CNN

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1.4.1 เมื่อสามารถจำแนกภาพอวัยวะได้แล้วจะทำให้สามารถหาสิ่งผิดปกติ ภายในทางเดินอาหารได้ เช่น จุดเลือดออก
ก้อนเนื้อผิดปกติ พยาธิ เป็นต้น
- 1.4.2 แพทย์มีความสะดวกและรวดเร็วในการตรวจวินิจฉัยโรคมมากขึ้น
- 1.4.3 สามารถเพิ่มความเร็วและประสิทธิภาพในการจำแนกภาพมากกว่า บุคลากรทางการแพทย์ทั่วไป

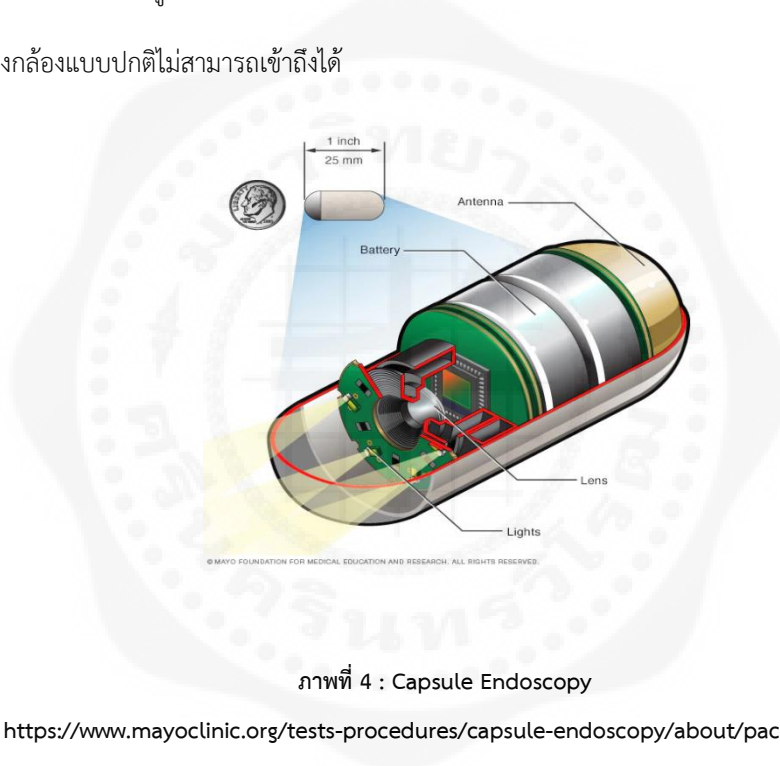


บทที่ 2 องค์ความรู้ที่เกี่ยวข้อง

2.1 องค์ความรู้ที่เกี่ยวข้อง

2.1.1 Capsule Endoscopy [3]

เป็นวิธีการที่ใช้กล้องไร้สายขนาดเล็กในการถ่ายภาพภายในทางเดินอาหาร โดยตัวกล้องจะอยู่ในแคปซูล ขนาดเท่าวิตามินที่ผู้รับการตรวจจะกลืนเข้าไป ขณะที่แคปซูลเคลื่อนตัวไปตามทางเดินอาหารเพื่อถ่ายภาพ ซึ่งภาพกว่าหมื่นภาพ ที่ได้จากกล้องจะถูกจัดเก็บในเครื่องบันทึก ซึ่งจะช่วยให้แพทย์สามารถมองเห็นภายในลำไส้เล็กหรือภายในพื้นที่ที่การส่องกล้องแบบปกติไม่สามารถเข้าถึงได้



(ที่มา : <https://www.mayoclinic.org/tests-procedures/capsule-endoscopy/about/pac-20393366>, 2019)

2.1.2 Classification [4]

เป็นกระบวนการสร้างโมเดลของการแบ่งกลุ่มข้อมูลให้อยู่ในรูปแบบที่กำหนดมาให้ เพื่อแสดงให้เห็นความแตกต่างระหว่างคลาส ซึ่งโมเดลที่ใช้จำแนกข้อมูลจะขึ้นอยู่กับการวิเคราะห์เซตของ training data ที่ใช้สอนให้ระบบเรียนรู้ว่า ข้อมูลใดควรอยู่ในคลาสเดียวกันบ้าง ซึ่งผลลัพธ์ที่ได้จากการเรียนรู้คือ classifier model

ข้อมูลส่วนที่เหลือจาก training data จะเป็นข้อมูลสำหรับทดสอบ testing data จะถูกนำมาเปรียบเทียบกับกลุ่มที่หามาได้จากโมเดลเพื่อตรวจสอบความถูกต้อง โดยเราจะปรับปรุงโมเดลจนกว่าจะได้ค่า accuracy ในระดับที่พอใจ หลังจากนั้นเมื่อมีข้อมูลใหม่เข้ามา เราจะนำข้อมูลผ่านโมเดล

โดยโมเดลจะสามารถทำนายกลุ่มของข้อมูลนี้ได้

การ Classification มี 2 ขั้นตอนคือ

1. Learning เป็นขั้นตอนที่จะวิเคราะห์ training data โดย algorithm ของ classification และ learning model ถูกแทนในรูปของ classification rules
2. Classification เป็นขั้นตอนที่เมื่อได้ classification rules จะมีการตรวจสอบว่ากฎที่ได้สามารถทำนายได้อย่างถูกต้องหรือไม่ โดยการนำ test data ที่เราทราบว่าอยู่ในกลุ่มใดไปเปรียบเทียบกับ learning model

2.1.3 Convolutional Neural Network (CNN) [5]

Convolutional Neural Network (ConvNet/CNN) คืออัลกอริทึมการเรียนรู้เชิงลึกที่สามารถใช้ในภาพอินพุตกำหนดความสำคัญ (น้ำหนักและความเอนเอียง ที่เรียนรู้ได้ ไปยังแง่มุม) วัตถุต่างๆในภาพและสามารถแยกความแตกต่างจากอีกด้านหนึ่งได้ การประมวลผลล่วงหน้าที่เป็นใน ConvNet นั้นต่ำกว่ามากเมื่อ เทียบกับอัลกอริทึมการจำแนกประเภทอื่น ๆ แม้ว่าตัวกรองในวิธีการดั้งเดิมจะได้ รับการออกแบบด้วยมือ ด้วยการ training ข้อมูลที่เพียงพอ ConvNets จึงจะสามารถเรียนรู้ตัวกรองคุณสมบัติเหล่านี้ได้ สถาปัตยกรรมของ ConvNet นั้นคล้ายคลึงกับรูปแบบการเชื่อมต่อของเซลล์ประสาทในสมองมนุษย์และได้รับแรง บันดาลใจจากองค์กรของ Visual Cortex เซลล์ประสาทแต่ละเซลล์ตอบสนองต่อสิ่งเร้าเฉพาะในขอบเขตที่จำกัดของมุมมองที่มองเห็นเรียกว่า Receptive field โดยการรวมกลุ่มหลายๆตัวของเขตข้อมูลดังกล่าวทับซ้อนกันเพื่อให้ครอบคลุมพื้นที่ภาพทั้งหมด โดย CNN เป็น neural network ที่มีหลายเลเยอร์ที่จะวิเคราะห์ input ที่เป็นรูปภาพ และปฏิบัติงาน เช่น การจดจำภาพ การจำแนกรูปภาพ และ การ detection

CNN จะประกอบด้วยเลเยอร์หลายชนิด คือ

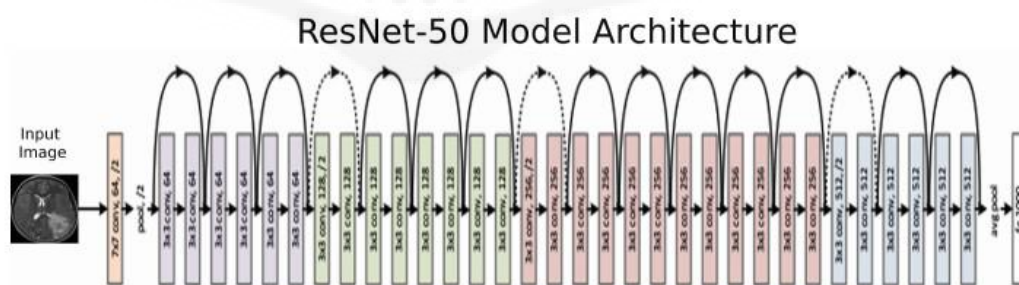
- **Convolutional layer** : เป็นการสร้าง feature map เพื่อทำนายความน่าจะเป็นของคลาสสำหรับแต่ละ feature โดยใช้ filter ในการพิจารณาทั้งภาพในแต่ละพิกเซลจุดประสงค์ของเลเยอร์นี้ คือการค้นหา คุณสมบัติที่มีอยู่ในรูปภาพตัวอย่าง เช่น ขอบแนวตั้ง / แนวนอนการไล่ระดับสี ฯลฯ เพื่อให้มีการตรวจสอบคุณสมบัติ หลายอย่าง และ จะมีตัวกรอง (filters) ที่แตกต่างกันหลายตัว เมื่อนำทุกๆตัวมารวมกันก็จะสร้างเป็นผลลัพธ์ของ Neurons หลายๆตัวภายในพื้นที่ส่วนๆ นั้นที่ทำการอินพุตเข้าไปหลังจากเลเยอร์นี้เป็นคุณสมบัติที่ดึงออกมาจากอินพุตของพื้นที่ในรูปภาพ เพื่อให้ได้ผลลัพธ์เราจะดำเนินการ dot-product ระหว่าง Conv-Layer และชั้นอินพุต
- **Pooling layer (downsampling)** : เป็นเรื่องปกติที่จะแทรก Pooling Layer เป็นระยะๆ ระหว่าง Conv-Layers ที่ต่อเนื่องกันในสถาปัตยกรรม CNN หน้าที่ของมัน คือการลดขนาดของการแสดงผล (เมทริกซ์) เพื่อลดจำนวนพารามิเตอร์ (เซลล์ประสาทแต่ละตัวถือได้ว่าเป็นพารามิเตอร์) และการ

คำนวณในแต่ละเครือข่ายของมัน โดยในการคำนวณสามารถใช้ค่าต่ำสุด (Min Pooling) ค่าสูงสุด (Max Pooling) ผลรวม (Sum Pooling) และค่าเฉลี่ย (Average Pooling) ในการคำนวณ Feature Map จะถูก แบ่งออกเป็นบล็อก ขนาดต่างๆ ซึ่งหากใช้วิธี Max Pooling ในการคำนวณ ค่าที่ได้ ก็คือค่าสูงสุด (Max Value) ของแต่ละบล็อก Conv-Layer และ Pooling Layer จะทำซ้ำหลายๆ ครั้งเพื่อดึงคุณสมบัติที่ลึกลงไปในรูปแบบจากนั้นมันจะแบนเป็นเวกเตอร์คอลัมน์ซึ่งจะเชื่อมต่อกับเลเยอร์อื่นอย่างสมบูรณ์ (Fully connected layer)

- **Fully connected layer** : เป็น Hidden Layer และ Output Layer ของโครงข่ายประสาทเทียม ดังนั้น Fully-Connected Layer จึงทำหน้าที่ในการเรียนรู้ (Training) และการจำแนก ประเภทของวัตถุโดยผลลัพธ์ที่ได้ก็คือจำนวนของ Class ที่ต้องการ จำแนก

สถาปัตยกรรม CNN ที่ใช้

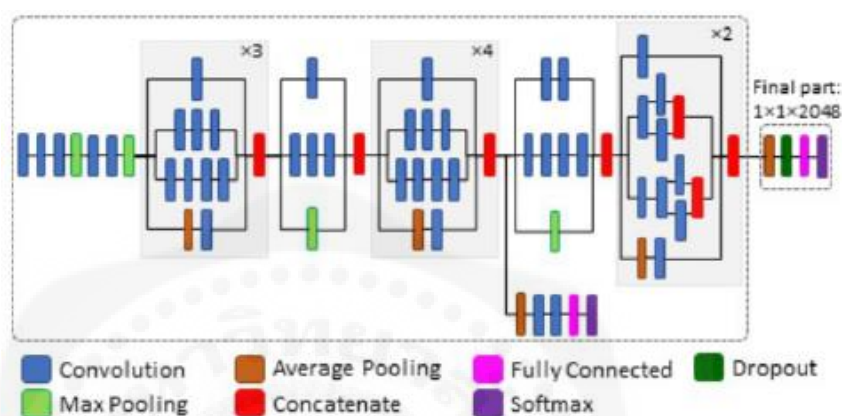
- **ResNet50 [6]** : ResNet50 ย่อมาจาก Deep Residual Network ที่มีจำนวนเลเยอร์ในบล็อกใหญ่ 48 ชั้น + ชั้นคอนโวลูชันที่ติดกับชั้น input + ชั้น Dense ที่ติดชั้น output เป็นการแก้ปัญหา Vanishing gradient โดยการใส่ทางลัด (Shortcut) ในโครงข่าย ดังในภาพที่ 5



ภาพที่ 5 : สถาปัตยกรรม ResNet 50

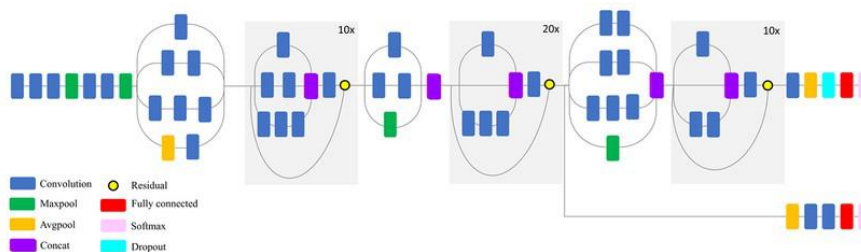
(ที่มา : <https://www.aimspress.com/aimspress-data/mbe/2020/5/PIC/mbe-17-05-328-g007.jpg>)

- **GoogLeNet/Inception [7]** : สถาปัตยกรรมที่มี 22 เลเยอร์ ซึ่งนักวิจัยได้สร้างวิธีการใหม่เรียกว่า Inception Module เป็นการเปลี่ยนแปลงจากสถาปัตยกรรมที่เป็นลำดับชั้น ในหนึ่งเลเยอร์จะมี feature extractors หลายชนิดจะทำให้ network ปฏิบัติการได้ดีขึ้นในทางอ้อม ซึ่งภายในสถาปัตยกรรม จะมีหลาย Inception Module ที่ซ้อนกัน ดังในภาพที่ 6



ภาพที่ 6 : สถาปัตยกรรม InceptionV3

- **InceptionResnet [8]** : สร้างขึ้นโดยการรวม ResNet และ Inception โดยที่ใน batch-normalization จะใช้เฉพาะกับเลเยอร์ ดั้งเดิมเท่านั้นแทนที่จะเป็นต้นบนของการสรุปโดยเฉพาะอย่างยิ่งโมดูลที่เหลือ ถูกนำมาใช้เพื่อให้สามารถเพิ่มจำนวนบล็อก Inception และทำให้ความลึกของเครือข่ายเพิ่มขึ้น ดังที่ได้กล่าวไว้ก่อนหน้านี้ปัญหาคือที่เด่นชัดที่สุดที่เกี่ยวข้องกับเครือข่ายที่ลึกมากคือขั้นตอนการ training ซึ่งสามารถแก้ไขได้โดยใช้การเชื่อมต่อที่ ข้ามชั้นในแบบของ ResNet เครือข่ายจะลดขนาดส่วนที่เหลือลง เป็นแนวทางที่มีประสิทธิภาพเพื่อแก้ไขปัญหาคือการ training ที่นานเมื่อ มีการใช้ตัวกรองจำนวนมาก (มีมากกว่า 1,000 ตัวกรอง) ที่ถูกใช้ในเครือข่าย ดังในภาพที่ 7 โดยความแตกต่างของ InceptionResnetV1 และ InceptionResnetV2 ก็คือ ตัวของ InceptionResnetV1 นั้นมีการคำนวณที่ใกล้เคียงกับ InceptionV3 ส่วน InceptionResnetV2 นั้นมีการคำนวณที่ใกล้เคียงกับ InceptionV4



ภาพที่ 7 : สถาปัตยกรรม InceptionResnetV2

(ที่มา : [https://www.researchgate.net/profile/Masoud-](https://www.researchgate.net/profile/Masoud-Mahdianpari/publication/326421398/figure/fig9/AS:649354851385344@1531829669740/Schematic-diagram-of-InceptionResNetV2-model-compressed-view.png)

[Mahdianpari/publication/326421398/figure/fig9/AS:649354851385344@1531829669740/Schematic-diagram-of-InceptionResNetV2-model-compressed-view.png](https://www.researchgate.net/profile/Masoud-Mahdianpari/publication/326421398/figure/fig9/AS:649354851385344@1531829669740/Schematic-diagram-of-InceptionResNetV2-model-compressed-view.png))

2.1.4 Ensemble [9]

เป็นเทคนิคที่มีการนำมาใช้มากในการพัฒนา learning model ใน machine learning โดยเฉพาะเป้าหมายการเพิ่มประสิทธิภาพการทำงานของโมเดลโดยเทคนิคของ Ensemble จะมีอยู่ 3 แบบหลักๆ

- Bagging (Decision Trees, Random Forest และ Extra Trees)
- Boosting (AdaBoost และ Stochastic Gradient Boosting)
- Voting (Decision Tree, SVM , K-Nearest Neighbors)

2.1.5 Evaluation Metric[10]

เปรียบเป็นคะแนนสำหรับบ่งชี้ความถูกต้องของค่าที่โมเดลทำนายหากค่าคะแนนสูงก็จะบ่งบอกว่าโมเดลนั้นมีประสิทธิภาพมาก ตัวอย่างของ Evaluation Metric เช่น Confusion Matrix ซึ่งเป็นเทคนิคในการสรุปผลประสิทธิภาพของ Classification Algorithm

2.1.6 สูตรในการคำนวณค่าที่ใช้ในการประเมินประสิทธิภาพของโมเดล[11]

Precision: เป็นความน่าจะเป็นของจำนวนครั้งที่โมเดลจะทำนาย ถูกต้องจากการทำนายทั้งหมด โดยมีสูตรในการทำนายดังภาพที่ 8

$$\text{Precision} = \frac{TP}{TP + FP}$$

ภาพที่ 8 : สูตรการหาค่า Precision

(ที่มา: <https://developers.google.com/machine-learning/crash-course/classification/precision-and-recall>)

Accuracy: เป็นค่าที่บ่งบอกว่าโมเดลของเราทำนายถูกทั้งหมดกี่เปอร์เซ็นต์ โดยจะมีค่าอยู่ระหว่าง 0-1 ซึ่งค่าถ้าที่ได้เข้าใกล้ 1 มาก หมายถึงมีความแม่นยำมาก มีสูตรในการทำนายดังภาพที่ 9

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

ภาพที่ 9 : สูตรการหาค่า Accuracy

(ที่มา: <https://developers.google.com/machine-learning/crash-course/classification/accuracy>)

Specificity: เป็นการประเมินค่า True Negative ที่ได้จากการทำนายของโมเดล โดยมีสูตรในการทำนายดังภาพที่ 10

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

ภาพที่ 10 : สูตรการหาค่า Specificity

(ที่มา: <https://medium.com/fintechexplained/sensitivity-vs-specificity-in-data-science-2f673039dbd1>)

Sensitivity: Sensitivity หรือ Recall เป็นสัดส่วนของค่า True Positive ที่ได้จากการทำนายของโมเดล โดยมีสูตรในการทำนัยดังภาพที่ 11

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

ภาพที่ 11 : สูตรในการหาค่า Sensitivity

(ที่มา: <https://medium.com/fintechexplained/sensitivity-vs-specificity-in-data-science-2f673039dbd1>)

F1-Score: เป็นค่าเฉลี่ยแบบ Harmonic mean ระหว่าง Precision และ Recall ใช้วัดความสามารถของโมเดล โดยมีสูตรในการทำนัยดังภาพที่ 12

$$F1 = 2 * \left(\frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \right)$$

ภาพที่ 12 : สูตรในการหาค่า F1 - Score

(ที่มา: <https://medium.com/@615162020027/metrics-พื้นฐานสำหรับวัดประสิทธิภาพของโมเดล-machine-learning-c00fcc32fa30>)

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 Automatic Classification of Digestive Organs in Wireless Capsule Endoscopy Videos [12]

บทความนี้เป็นการนำวิดีโอจากกล้องแคปซูลไร้สายมาแบ่งเป็นภาพ และทำการหา feature ที่สำคัญที่สุด ซึ่งเป็น pattern ของสีรวมไปถึงการเคลื่อนไหวซึ่งจะเป็นลักษณะใกล้เคียงกันภายในอวัยวะเดียวกัน ทำให้สามารถจำแนกอวัยวะภายในระบบทางเดินอาหารได้

2.2.2 Deep Convolutional Neural Network for Ulcer Recognition in Wireless Capsule Endoscopy: Experimental Feasibility and Optimization [13]

บทความนี้เป็นการสำรวจความเป็นไปได้ของการเรียนรู้เชิงลึกสำหรับการจดจำแผลในกระเพาะอาหารและปรับปรุงการจดจำแผลในกระเพาะอาหารโดยใช้ CNN-based สำหรับภาพจากกล้องแคปซูลไร้สาย โดยการวิเคราะห์งานการจำแนกแผลและลักษณะของการเรียนรู้เชิงลึกแบบคลาสสิก จึงได้เสนอสถาปัตยกรรม HAnet ที่ใช้ ResNet-34 เป็นเครือข่ายพื้นฐานและรวมคุณสมบัติไฮเปอร์จากเลเยอร์บนด้วยคุณสมบัติที่ลึกลงไปในเลเยอร์ที่ลึกกว่าเพื่อใช้ในการตัดสินใจวินิจฉัยขั้นสุดท้าย

2.2.3 An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features [14]

บทความนี้เกี่ยวกับระบบอัตโนมัติเพื่อสนับสนุนการตรวจจับ ตั้งเนื้อในทางเดินอาหาร ระบบนี้จับข้อมูลวิดีโอจากวิดีโอส่องกล้อง (Endoscopic video) และในผลลัพธ์จะแสดงติ่ง (polyp) ที่จะระบุ โดยได้ใช้ฟิเจอร์ Color wavelet (CW) และ Convolutional neural network (CNN) ของเฟรมวิดีโอถูกแยกและรวมเข้าด้วยกัน ซึ่งใช้ในการเทรน support vector machine (SVM) การประเมินฐานข้อมูลมาตรฐานของสาธารณะแสดงให้เห็นว่าระบบที่เสนอมีประสิทธิภาพเหนือกว่าวิธีการที่ทันสมัย

2.2.4 Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network [15]

Yuxian Zou และคณะ ได้เสนอการใช้ deep convolutional neural network (DCNN) framework เพื่อที่จะทำการจำแนกอวัยวะย่อยอาหารในภาพ Wireless Capsule Endoscopy (WCE) เพื่อการจดจำคุณสมบัติทางความหมายของภาพ ที่สูงขึ้นด้วย เพื่อให้บรรลุเป้าหมายดังกล่าวจึงมีการสร้างระบบการจำแนกประเภท WCE โดยใช้ CNN-based ให้ได้อย่างมีประสิทธิภาพ (DCNN-WCE-CS) ด้วยภาพ WCE จริงจากประมาณ 1 ล้านภาพ การทดลองอย่างตั้งใจที่จะดำเนินการเพื่อประเมินประสิทธิภาพโดยการตั้งค่าพารามิเตอร์ที่แตกต่างกัน ผลลัพธ์แสดงให้เห็นถึงประสิทธิภาพที่เหนือกว่าเมื่อเปรียบเทียบกับวิธีการจำแนกแบบดั้งเดิมโดยที่ความแม่นยำในการจำแนกประมาณ 95% สามารถทำได้โดยการเฉลี่ย ยิ่งไปกว่านั้น DCNN-WCE-CS นั้นมีความทนทานต่อการเปลี่ยนแปลงของภาพ WCE อันเนื่องด้วยบุคคลและสถานการณ์ระบบทางเดินอาหารที่ซับซ้อนรวมถึงการหมุนการเปลี่ยนแปลงความสว่างของภาพ WCE

2.2.5 A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images [16]

Xiao Jia และ Max Q.-H. Meng ได้นำเสนอกลยุทธ์การตรวจจับภาวะเลือดออกอัตโนมัติใหม่ โดยใช้โครงข่ายประสาทเทียมเชิงลึก (CNN) และใช้วิธีของพวกเขา ในการประเมิน 10,000 ภาพ จากกล้องแคปซูลไร้สาย (WCE) ผลการทดลองที่เพิ่มขึ้นประมาณ 2% ในคะแนน F แสดงให้เห็นว่าวิธีการของเรามีประสิทธิภาพเหนือกว่าวิธีการที่ทันสมัย ในการตรวจจับการตกเลือด WCE คะแนน F ที่ทำได้คือ 0.9955

2.2.6 Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module [17]

Yun Jiang และคณะ ได้ออกแบบโครงข่ายประสาทเทียมแบบใหม่ซึ่งประกอบด้วย เลเยอร์ convolutional โมดูล SE-ResNet ขนาดเล็กและเลเยอร์ที่เชื่อมต่ออย่างสมบูรณ์ พวกเขาได้เสนอโมดูล SE-ResNet ขนาดเล็กซึ่งเป็นการ

ปรับปรุงการรวมกันของ Module ที่เหลือ และ Squeeze-and-Excitation block และบรรลุประสิทธิภาพที่คล้ายกัน โดยมีพารามิเตอร์น้อยลง

นอกจากนี้เราเสนอตัวจัดตารางอัตราการเรียนรู้ใหม่ซึ่งสามารถรับประสิทธิภาพที่ยอดเยี่ยมโดยไม่ต้องปรับอัตราการเรียนรู้อย่างซับซ้อนทางคณะใช้แบบจำลองสำหรับการจำแนกภาพเนื้อเยื่อมะเร็งเต้านมโดยอัตโนมัติ (ชุดข้อมูล BreakHis) ออกเป็น 8 ชนิดย่อย ผลการวิจัยพบว่าแบบจำลองของเรามีความถูกต้องระหว่าง 98.87% และ 99.34% สำหรับการจำแนกเลขฐานสองและบรรลุความแม่นยำระหว่าง 90.66% และ 93.81% สำหรับการจำแนกประเภทหลายชั้น

2.2.7 Hookworm Detection in Wireless Capsule Endoscopy Images With Deep Learning [18]

พยาธิปากขอเป็นสาเหตุสำคัญอันดับต้นๆของการเจ็บป่วยของมารดาและเด็ก ซึ่งเป็นภัยคุกคามต่อสุขภาพของมนุษย์เมื่อเร็วๆ นี้การส่องกล้องไร้สายแคปซูล (WCE) ได้ถูกนำไปใช้กับการตรวจหาพยาธิปากขอ โดยอัตโนมัติ นำเสียดายที่มันยังคงเป็นงาน ที่ท้าทายเกินไป

ในช่วงไม่กี่ปีที่ผ่านมาเครือข่ายประสาทเทียมเชิงลึก (CNN) ได้แสดงให้เห็นถึงประสิทธิภาพที่น่าประทับใจในงานการวิเคราะห์ภาพ และวิดีโอในงานวิจัยนี้เสนอกรอบการตรวจหาพยาธิปากขออย่างล้ำลึกที่แปลกใหม่สำหรับภาพ WCE ซึ่งจำลองรูปแบบการมองเห็นและรูปแบบท่อของพยาธิปากขอพร้อมกัน

การทดลองดำเนินการกับหนึ่งในชุดข้อมูล WCE ที่ใหญ่ที่สุดที่มีรูปภาพประมาณ 440,000 WCE ซึ่งแสดงให้เห็นถึงประสิทธิภาพของกรอบการตรวจหาพยาธิปากขอที่เสนอว่ามีประสิทธิภาพสูงกว่าวิธีการที่ทันสมัยความไวและความแม่นยำสูงของวิธีการ ที่นำเสนอการตรวจหาพยาธิปากขอแสดงให้เห็นถึงศักยภาพในการใช้ทางคลินิก

2.2.8 A hybrid convolutional neural networks with extreme learning machine for WCE image classification [19]

Jia-sheng Yu และคณะได้ศึกษาปัญหาการจำแนกภาพอวัยวะที่ได้จากกล้องแคปซูลไร้สาย ซึ่งก่อนหน้านี้ทางคณะผู้วิจัยได้พิสูจน์ว่าระบบการจำแนก Convolutional Neural Networks (CNN)-based WCE มีความถูกต้อง 95% โดยเฉลี่ย ซึ่งการวิจัยนี้จะแสดงให้เห็นว่ามีสองวิธีที่เป็นไปได้ในการปรับปรุงประสิทธิภาพในการจำแนก ได้แก่ การจำแนก feature ของภาพให้มากขึ้น และการใช้ตัวแบ่งคลาส (Classifier) ที่มีประสิทธิภาพมากขึ้น

ในการวิจัยนี้คณะผู้วิจัยได้เสนอการออกแบบ WCE classification โดย hybrid CNN ด้วย Extreme Learning Machine (ELM) ซึ่งได้สร้าง CNN เป็นตัวจำแนก features ที่ขับเคลื่อนด้วยข้อมูลและให้ ELM เป็น classifier แทน full-connection classifier ที่ใช้กันปกติในระบบการจำแนกคลาสของ deep CNN

ระบบจำแนกภาพด้วยกล้องแคปซูลไร้สายที่ทางผู้วิจัยได้พัฒนาขึ้นชื่อว่า HCNN-NELM ซึ่งผลลัพธ์แสดงให้เห็นว่ามีความถูกต้องอยู่ที่ 97.25% โดยเฉลี่ย ซึ่งมีประสิทธิภาพดีกว่าวิธีการจำแนกคลาสแบบทั่วไป และวิธี CNN-base แบบปกติ

2.2.9 Convolutional Neural Network Technology in Endoscopic Imaging: Artificial Intelligence for Endoscopy [20]

เมื่อเร็วๆ นี้มีการปรับปรุงที่สำคัญในปัญญาประดิษฐ์โครงข่ายประสาทเทียมถูกนำมาใช้ในปี 1950 อย่างไรก็ตามเนื่องจากในเวลานั้นกำลังในการประมวลผลต่ำ และชุดข้อมูลที่ไม่เพียงพอเครือข่ายประสาทเทียมที่ได้รับความนิยมจากปัญหาการไล่ระดับสีที่มากเกินไปและการหายไปของการไล่ระดับสีสำหรับการฝึกอบรมเครือข่ายเชิงลึก (deep network)

แนวคิดนี้มีแนวโน้มมากขึ้นเนื่องจากความสามารถในการประมวลผลข้อมูลขนาดใหญ่ที่เพิ่มขึ้นการปรับปรุงกำลังการประมวลผลด้วยหน่วยประมวลผลแบบขนานและอัลกอริธึมใหม่สำหรับเครือข่ายนิวรัลลึก (deep neural network) ซึ่งประสบความสำเร็จมากขึ้นและดึงดูดความสนใจในหลายโดเมนการศึกษาล่าสุดในเทคโนโลยีนี้เป็นประโยชน์อย่างยิ่งสำหรับการใช้งานทางการแพทย์และการดูแลสุขภาพโดยเฉพาะอย่างยิ่งในการถ่ายภาพการส่องกล้อง

2.2.10 Uncertainty-Aware Organ Classification for SurgicalData Science Applications in Laparoscopy [21]

วิทยาศาสตร์ข้อมูลการผ่าตัดได้รับการพัฒนาเป็นชุดข้อมูลการค้นหาซ้ำที่มีวัตถุประสงค์เพื่อสังเกตทุกสิ่งที่เกิดขึ้นภายในและรอบกระบวนการบำบัดเพื่อให้ความช่วยเหลือโดยใช้ข้อมูลเป็นตัวขับเคลื่อนสถานการณ์ ในบริบทของการวิเคราะห์วิดีโอส่องกล้องการจำแนกอวัยวะที่แม่นยำในมุมมองของ กล้องถ่ายภาพช่วยให้ความท้าทายทางเทคนิคในที่นี้ทางคณะขอเสนอวิธีการใหม่ ในการจำแนกโครงสร้างทางกายวิภาคและการติดตามภาพซึ่งมีการวัดความมั่นใจภายในเพื่อประเมินประสิทธิภาพของตัวเองด้วยความน่าเชื่อถือสูงและสามารถนำไปใช้กับทั้งข้อมูล RGB และ multispectral (MI)

วิธีการ : การรู้จำอวัยวะจะดำเนินการโดยใช้การจัดหมวดหมู่แบบ Superpixel ตามข้อมูลของรูปแบบและการสะท้อนกลับ การจำแนกความเชื่อมั่นถูกประเมินโดย การวิเคราะห์การกระจายตัวของความน่าจะเป็นในระดับคลาส การประเมินผลของเทคโนโลยีที่น่าเสนอจะดำเนินการผ่านภายในร่างกายของหมูทั้ง 7 ตัว

ผลลัพธ์ : เมื่อนำไปใช้กับการถ่ายภาพความแม่นยำในการทดลองของกลุ่มคณะ เพิ่มขึ้นจาก 65% (RGB) และ 80% (MI) เป็น 90% (RGB) และ 96% (MI) ด้วยการวัด ความเชื่อมั่น

บทที่ 3 วิธีการดำเนินโครงการ

3.1 วิธีการดำเนินงาน

3.1.1 เริ่มต้นและวางแผนโครงการ

3.1.1.1 เลือกหัวข้อโครงการ

3.1.1.2 หาข้อมูลศึกษาแนวคิดและทฤษฎี

3.1.1.3 เลือกเครื่องมือ

3.1.1.4 กำหนดรายละเอียดของโครงการ

3.1.2 การวิเคราะห์ระบบ

3.1.2.1 ศึกษาหลักการของ CNN

3.1.2.2 ศึกษาสถาปัตยกรรม ResNet 50 และ GoogLeNet/Inception และ InceptionResnetV2

3.1.2.3 ศึกษาหลักการของ Ensemble

3.1.3 การออกแบบระบบ

3.1.3.1 ทดสอบใช้ Algorithm ที่ได้เลือกไว้

3.1.3.2 เตรียมข้อมูลภาพจากกล้องแคปซูลไร้สาย

3.1.3.3 สร้างโมเดลแบบจำลองของ CNN

3.1.3.4 ปรับปรุงและแก้ไขค่าที่เหมาะสม

3.1.3.5 นำข้อมูลทั้งหมดมาแสดงผล

3.1.3.6 สรุปผลงานวิจัยและเผยแพร่

3.2 อุปกรณ์และเครื่องมือที่ใช้

3.2.1 ฮาร์ดแวร์ (Hardware):

3.3.1.1 คอมพิวเตอร์

3.2.2 ซอฟต์แวร์และไลบรารีที่ใช้ (Software and Library):

3.2.2.1 Jupyter

3.2.2.2 Pandas

3.2.2.3 Keras

3.2.3 ภาษาที่ใช้ (Language):

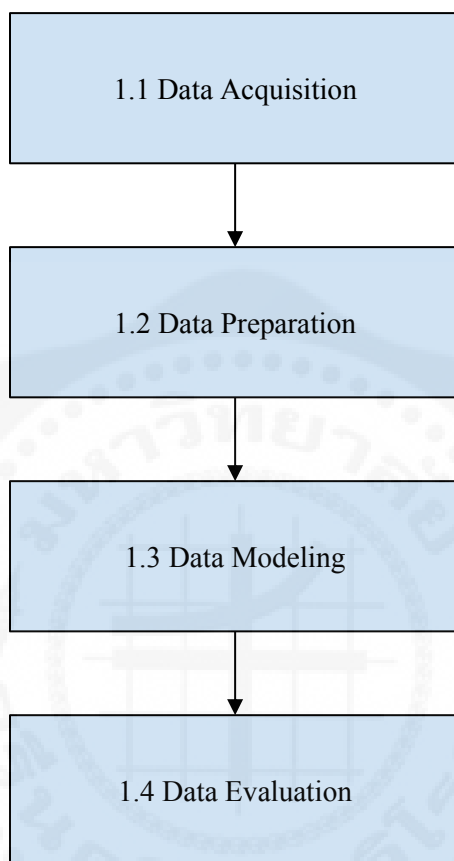
3.2.3.1 Python

3.3 ชุดข้อมูล (Data set) ที่ใช้งาน

ใช้ภาพจากกล้องแคปซูลไร้สายจาก Brown University และ Rhode Island Hospital โดยความร่วมมือกับ Amporn Atsawarungruangkit, MD และ Dr.Kawee Numpacharoen ซึ่งได้ถูก label ไว้เรียบร้อยแล้ว เป็นจำนวนทั้งหมด 174 คน แบ่งได้เป็น training set มีทั้งหมด 123 คนมีจำนวนภาพทั้งหมด 40,000 ภาพ โดยใช้ Esophagus 3,000 รูป Stomach 12,250 รูป Small Bowel 15,000 รูป และ Colon 9,750 รูป และ testing set มีทั้งหมด 51 คนมีจำนวนภาพทั้งหมด 683,681 ภาพ โดยใช้ Esophagus 918 รูป Stomach 71,912 รูป Small Bowel 558,791 รูป และ Colon 52,060 รูป

3.4 การออกแบบระบบ

แผนผังการดำเนินงาน



ภาพที่ 13 : แผนผังการดำเนินงาน

โดยรายละเอียดของการดำเนินงานจะสามารถอธิบายอย่างละเอียดได้ดังนี้

Data Acquisition

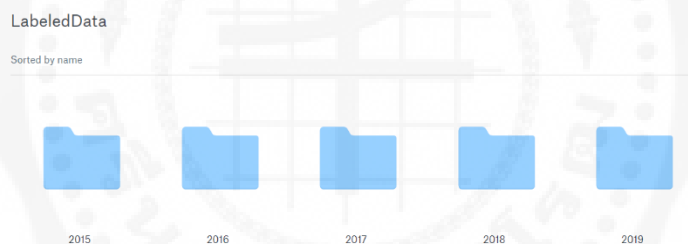
ขั้นตอนการจำแนกภาพด้วยมือสำหรับใช้ train model

1. ก่อนเริ่มการจำแนกภาพทำการโหลดไฟล์ภาพและไฟล์ Excel ที่ใช้สำหรับการจำแนกภาพจาก Dropbox ซึ่งจะแบ่งเป็นโฟลเดอร์ของข้อมูลตั้งแต่ปี 2015-2019 ซึ่งภายในโฟลเดอร์ของแต่ละปีจะประกอบด้วย โฟลเดอร์ย่อย ได้แก่

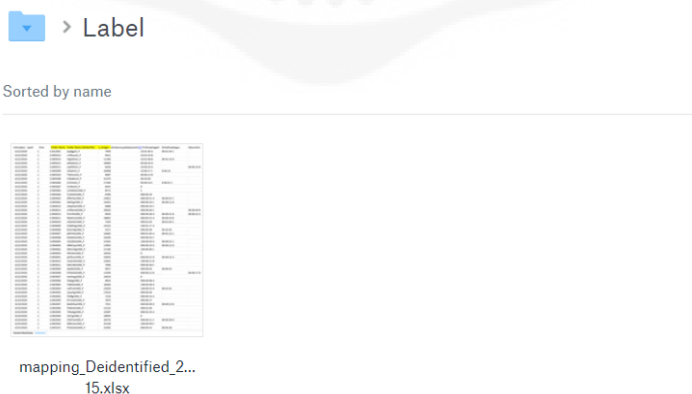
1.1 Images ซึ่งจะเก็บ zip ข้อมูลภาพของผู้ป่วยแต่ละคน ดังภาพที่14

1.2 Label เก็บไฟล์ mapping ซึ่งเป็นข้อมูลที่จะบอกว่า คลาสหรืออวัยวะใด เริ่มต้นและสิ้นสุดที่เวลาใด ดังภาพที่15

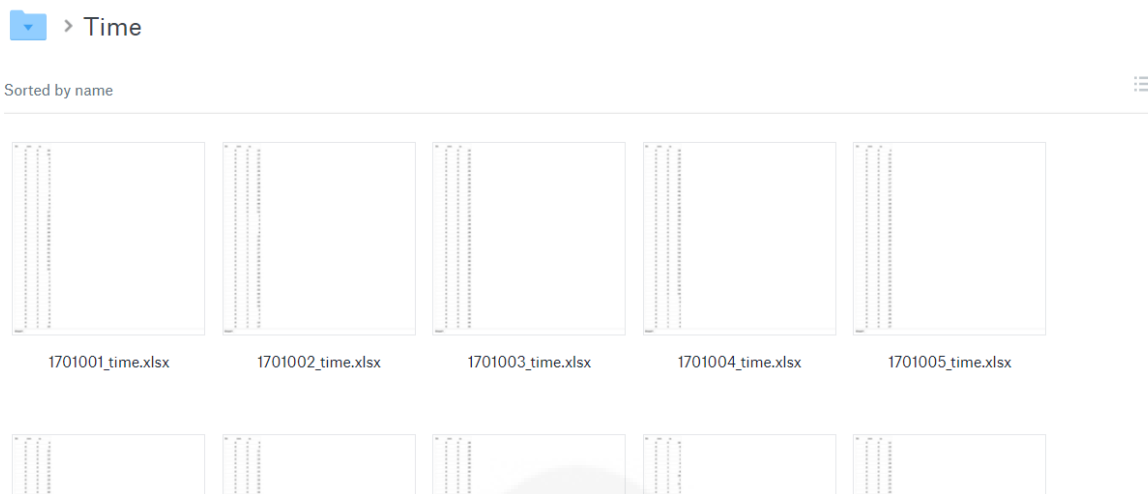
1.3 Time ซึ่งจะรวบรวมไฟล์สำหรับบ่งบอกว่าลำดับภาพจากโฟลเดอร์ Images ของผู้ป่วยแต่ละคนตรงกับเวลาใด ดังภาพที่ 16



ภาพที่ 14 : โฟลเดอร์ Images ภายใน Drop Box

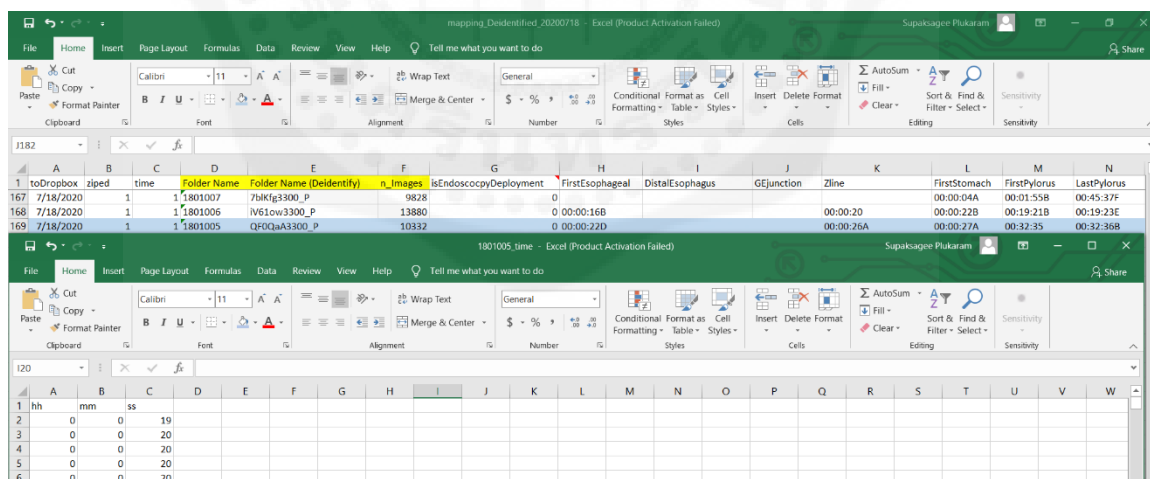


ภาพที่ 15 : โฟลเดอร์ Label จาก Dropbox



ภาพที่ 16 : โฟลเดอร์ Time จาก Dropbox

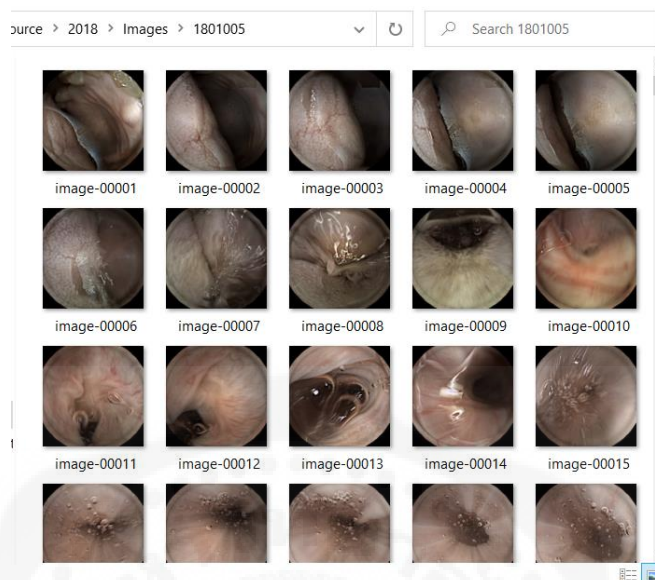
- เปิดไฟล์ mapping ที่อยู่ในโฟลเดอร์ Label และหา row ที่มี column Folder Name ตรงกับโฟลเดอร์ที่ต้องการ จะทำการจำแนกภาพ เช่น 1801005
- เปิดโฟลเดอร์ Time และ เลือกไฟล์ของผู้ป่วยที่จะทำการจำแนกภาพ เช่น 1801005 ดังภาพที่ 17



ภาพที่ 17 : ตัวอย่างไฟล์ mapping และไฟล์ Time

- ลบ row แรกของไฟล์ในโฟลเดอร์ Time ที่มีข้อมูลเป็น hh, mm, ss เพื่อให้เลขของเวลาในภาพตรงกับลำดับของภาพ ในโฟลเดอร์ Images ซึ่งจะช่วยลดความสับสนในการทำงาน

5. เปิดโฟลเดอร์ข้อมูลภาพของผู้ป่วยที่ต้องการ เช่น 1801005 เพื่อเตรียมทำการจำแนกภาพ ดังภาพที่ 18



ภาพที่ 18 : ข้อมูลภาพของผู้ป่วย 1 คน

6. สร้างโฟลเดอร์เปล่าสำหรับเตรียมใส่ข้อมูลที่จะทำการจำแนกภาพ

ซึ่งภายในจะแบ่งเป็น

โฟลเดอร์ย่อยตามคลาสที่เราต้องการจำนวน 10 คลาส ดังภาพที่ 19

ata > Labeled > 2018 > 1801005					Search 1801005
Name	Date	Type	Size	Tags	
Colon	11/22/2020 7:50 PM	File folder			
Distal Esophagus	11/22/2020 7:50 PM	File folder			
Esophagus	11/22/2020 7:49 PM	File folder			
GE Junction	11/22/2020 7:50 PM	File folder			
IC-Valve	11/22/2020 7:51 PM	File folder			
Other	11/22/2020 7:50 PM	File folder			
Pylorus	11/22/2020 7:51 PM	File folder			
Small Bowel	11/22/2020 7:50 PM	File folder			
Stomach	11/22/2020 7:50 PM	File folder			
Z-Line	11/22/2020 7:50 PM	File folder			

ภาพที่ 19 : โฟลเดอร์ที่แบ่งตามจำนวนคลาสทั้ง 10 คลาส

7. ในไฟล์ mapping จะมี Sheet ที่ 2 สำหรับบอกขอบเขตว่าคลาสใดจะเริ่มต้นและสิ้นสุดที่ตำแหน่งใดดังภาพที่ 20

IsEndoscopyDeployment		1 if the capsule is deployed using upper endoscopy, 0 if the capsule is swallowed (typical method)	
If LastImage = "colon"			
Variable	Description	main organ / landmark	description
I1	FirstEsophageal	main	Esophagus
I2	DistalEsophagus	landmark	Distal esophagus
I3	GEjunction	landmark	GE junction
I4	Zline	landmark	Z-line
I5	FirstStomach	main	Stomach
I6	FirstPylorus	landmark	Pylorus
I7	LastPylorus	main	small bowel
I8	FirstDuodenum	landmark	IC valve
I9	First_IC_valve	main	colon
I10	FirstCecum		
LastImage	last image of organ		
If LastImage = "small bowel"			
main organ / landmark	description	Range of Images	
main	Esophagus	[1-15]	
landmark	Distal esophagus	[12-15]	
landmark	GE junction	[13-15]	
landmark	Z-line	[14-15]	
main	Stomach	[15-18]	
landmark	Pylorus	[16-17]	
main	small bowel	[18 - LastImage]	
landmark	IC valve	Not applicable	
main	colon	Not applicable	
If LastImage = "stomach"			
main organ / landmark	description	Range of Images	
main	Esophagus	[1-15]	
landmark	Distal esophagus	[12-15]	
landmark	GE junction	[13-15]	
landmark	Z-line	[14-15]	
main	Stomach	[15 - LastImage]	
landmark	Pylorus	Not applicable	
main	small bowel	Not applicable	
landmark	IC valve	Not applicable	
main	colon	Not applicable	

ภาพที่ 20 : ตารางบ่งบอกขอบเขตของแต่ละคลาส

8. เริ่มทำการแบ่งภาพ โดยไปที่ไฟล์ mapping และดูที่ column FirstEsophagus ที่ตรงกับ row ของผู้ป่วยที่เราต้องการจำแนก ซึ่งจากภาพที่ 21 จะเห็นว่า ข้อมูลใน cell คือ 00:00:22D

	A	B	C	D	E	F	G	H	I
167	7/18/2020	1	1	Folder Name	Folder Name (Deidentify)	n_Images	isEndoscopyDeployment	FirstEsophageal	Distance
168	7/18/2020	1	1	1801007	7blKfg3300_P	9828	0		
169	7/18/2020	1	1	1801006	iV61ow3300_P	13880	0	00:00:16B	
170	7/18/2020	1	1	1801005	QF0QaA3300_P	10332	0	00:00:22D	
171	7/18/2020	1	1	1801004	za7jew3300_P	10427	0	00:00:28D	
172	7/18/2020	1	1	1801003	YKn3Gw3300_P	10599	0		
173	7/18/2020	1	1	1801002	RPI88A3300_P	8804	0		

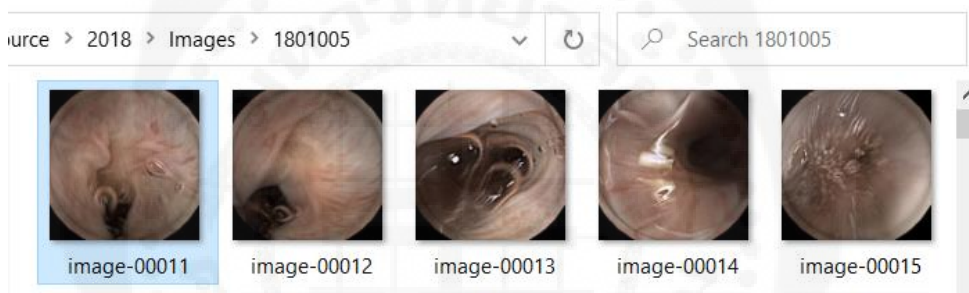
ภาพที่ 21 : ภาพไฟล์ mapping สำหรับการจำแนกภาพ

9. ให้ไปที่ไฟล์จากโฟลเดอร์ Time เช่น 1801005_time แล้วดูเวลาที่เรได้จากไฟล์ mapping คือ 00:00:22D จากนั้นนำไปเทียบกับ column แรกที่เป็น 0 column ที่สองที่เป็น 0 และ column ที่สามที่เป็น 22D (22อันที่4) จะเห็นว่าตรงกับ row ที่ 11 ซึ่งหมายถึง ภาพเริ่มต้นของ Esophagus เป็นภาพที่ 11 ในโฟลเดอร์ภาพของผู้ป่วย ซึ่งดังนั้น

ภาพที่ 0-10 จะอยู่ในโฟลเดอร์ที่ชื่อว่า Other ซึ่งจะเก็บภาพตั้งแต่เริ่มกลืนกลืนแคลปูลไว้สายจนถึงก่อน Esophagus
ดังภาพที่ 22 และ 23

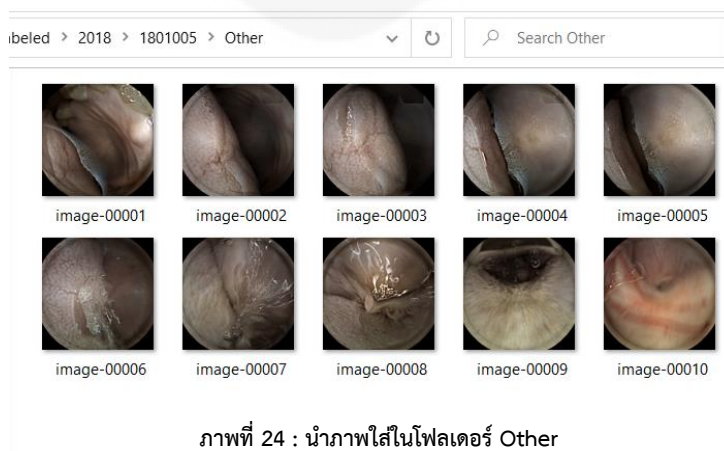
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	0	0	19									In Folder		
2	0	0	20									Other		
3	0	0	20									Esophagus		
4	0	0	20									Stomach		
5	0	0	20									Small Bowel		
6	0	0	21									Colon		
7	0	0	21											
8	0	0	22									Distal Esophagus		
9	0	0	22									GE Junction		
10	0	0	22									Z-Line		
11	0	0	22									Pylorus		
12	0	0	22									IC Valve		
13	0	0	22											
14	0	0	23											
15	0	0	23											

ภาพที่ 22 : ตัวอย่างการเลือกภาพในคลาส Other



ภาพที่ 23 : ภาพแรกของ Esophagus

10. นำรูปภาพที่ 1-10 จากภาพที่ 23 ใส่ในโฟลเดอร์ Other ดังภาพที่ 24

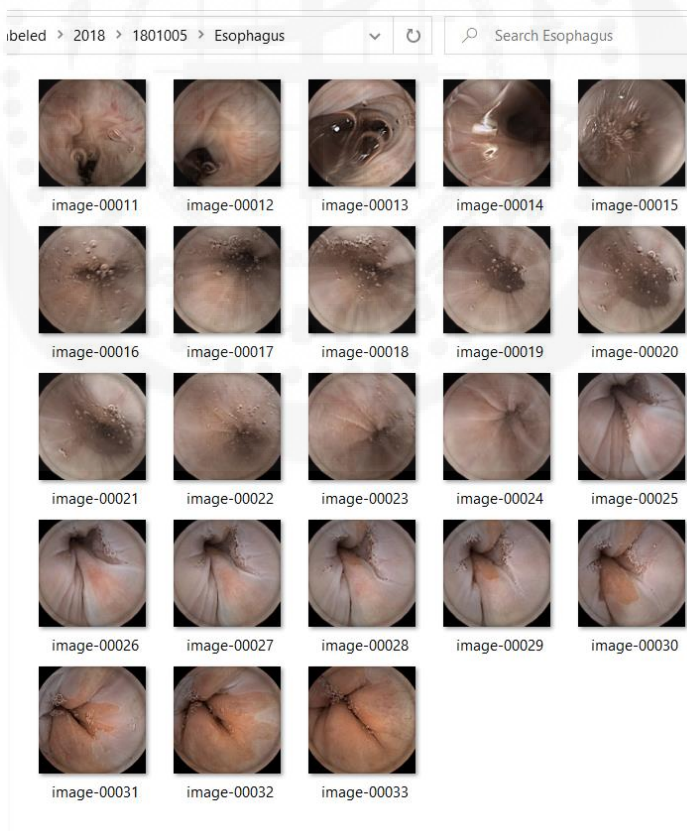


ภาพที่ 24 : นำภาพใส่ในโฟลเดอร์ Other

11. จุดเริ่มต้นของ Stomach จะเริ่มที่ column FirstStomach ในไฟล์ mapping ซึ่งข้อมูลใน cell คือ 00:00:27A เมื่อนำมาดูในไฟล์ 1801005_time จะตรงกับ row ที่ 34 ดังนั้นขอบเขตของ Esophagus จะอยู่ระหว่างภาพที่ 11-33 จากนั้นเราจึงนำภาพที่ 11-33 ไว้ที่โฟลเดอร์ชื่อ Esophagus ดังภาพที่ 24 และ 25

Row	Column	Value
167	FirstEsophagus	00:00:16B
168	FirstEsophagus	00:00:22D
169	FirstEsophagus	00:00:28D
170	FirstEsophagus	00:00:28D
171	FirstEsophagus	00:00:28D
172	FirstEsophagus	00:00:28D
173	FirstEsophagus	00:00:28D
174	FirstEsophagus	00:00:28D
175	FirstEsophagus	00:00:28D
176	FirstEsophagus	00:00:28D
177	FirstEsophagus	00:00:28D
178	FirstEsophagus	00:00:28D
179	FirstEsophagus	00:00:28D
180	FirstEsophagus	00:00:28D
181	FirstEsophagus	00:00:28D
182	FirstEsophagus	00:00:28D
183	FirstEsophagus	00:00:28D
184	FirstEsophagus	00:00:28D
185	FirstEsophagus	00:00:28D
186	FirstEsophagus	00:00:28D
187	FirstEsophagus	00:00:28D
188	FirstEsophagus	00:00:28D
189	FirstEsophagus	00:00:28D
190	FirstEsophagus	00:00:28D
191	FirstEsophagus	00:00:28D
192	FirstEsophagus	00:00:28D
193	FirstEsophagus	00:00:28D
194	FirstEsophagus	00:00:28D
195	FirstEsophagus	00:00:28D
196	FirstEsophagus	00:00:28D
197	FirstEsophagus	00:00:28D
198	FirstEsophagus	00:00:28D
199	FirstEsophagus	00:00:28D
200	FirstEsophagus	00:00:28D

ภาพที่ 25 : นำภาพใส่ในโฟลเดอร์ Esophagus

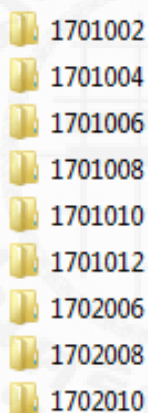


ภาพที่ 26 : นำภาพใส่ในโฟลเดอร์ Esophagus

12. ทำตั้งแต่ขั้นตอนที่ 8-11 ไปจนครบทุกอวัยวะ โดยดูขอบเขตของแต่ละอวัยวะ ดังภาพที่ 20
13. ภาพที่ใส่ในโพลเดอร์ที่เป็น Landmark ได้แก่ Distal Esophagus, GE Junction, Z-Line, Pylorus, IC-Valve
จะเป็นภาพที่มีอยู่ใน 4 คลาสหลัก ได้แก่ Esophagus, Stomach, Small Bowel, Colon เช่นกัน

Data Preparation

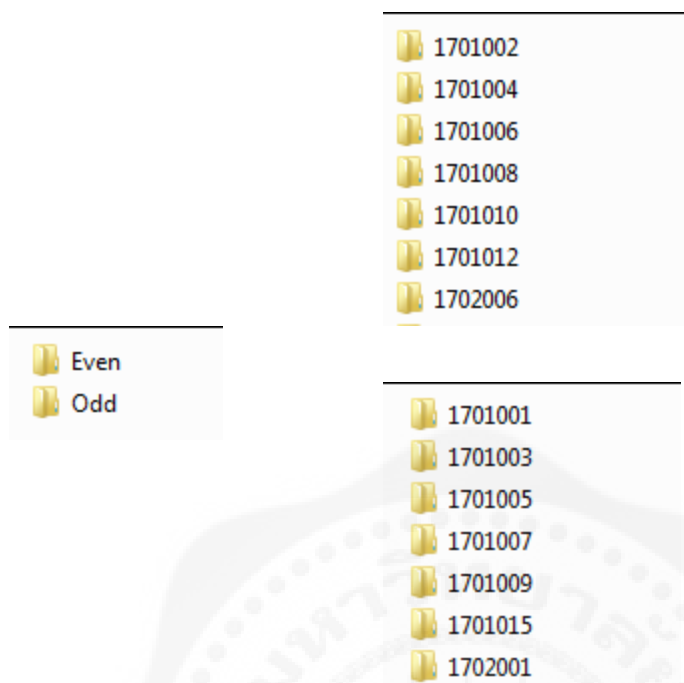
1. นำภาพที่ได้จากการจำแนกมาแบ่งใส่โพลเดอร์ไว้แยกเป็นแต่ละคนตามเดิม โดยที่อวัยวะของแต่ละคนยังคงแยกกัน
ไม่นำมารวมเป็นโพลเดอร์หลักโพลเดอร์เดียว ดังภาพที่ 27



ภาพที่ 27 : โพลเดอร์ของแต่ละคนที่นำมารวมกัน

โดยที่ภายในโพลเดอร์ยังคงมี 10 คลาสอยู่ ตามภาพที่ 19

2. ทำการแบ่งกลุ่มของภาพที่จะนำมาใช้ คือ เลขคู่ จะเป็นไฟล์ภาพสำหรับ train เลขคี่ จะเป็นไฟล์ภาพสำหรับ valid
ดังภาพที่ 28 เหตุผลที่ต้องแยกกันเพื่อให้่ายต่อการเลือกภาพ และ ที่ไม่นำภาพมารวมกันเป็นโพลเดอร์เดียวกันเพราะเนื่องจาก
โพลเดอร์จะมีขนาดใหญ่มาก อีกทั้งเพื่อให้่ายต่อการหาว่าบุคคลใดบ้างที่ใช้สำหรับ train หรือ valid



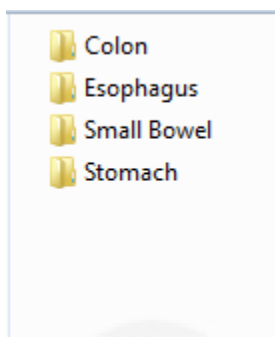
ภาพที่ 28 : ภาพบางส่วนของโฟลเดอร์ที่ทำการแบ่งกลุ่ม

3. เมื่อทำการแบ่งโฟลเดอร์แล้ว ให้ทำการ resize ขนาดของรูปภาพ โดยที่จะแบ่งโฟลเดอร์เป็น 2 ส่วนคือ โฟลเดอร์ภาพตั้งต้น และ โฟลเดอร์สำหรับเก็บภาพที่ทำการ resize แล้ว โดยในการ resize นี้ทำในเครื่องตนเองจะเร็วกว่าทำบน platform บนเว็บต่างๆ (เช่น google colab) เพราะ ถ้าต้องการเคลื่อนย้ายหรือเปลี่ยนตำแหน่งไฟล์ จะเสียเวลามากกว่าทำบนเครื่องของตนเอง

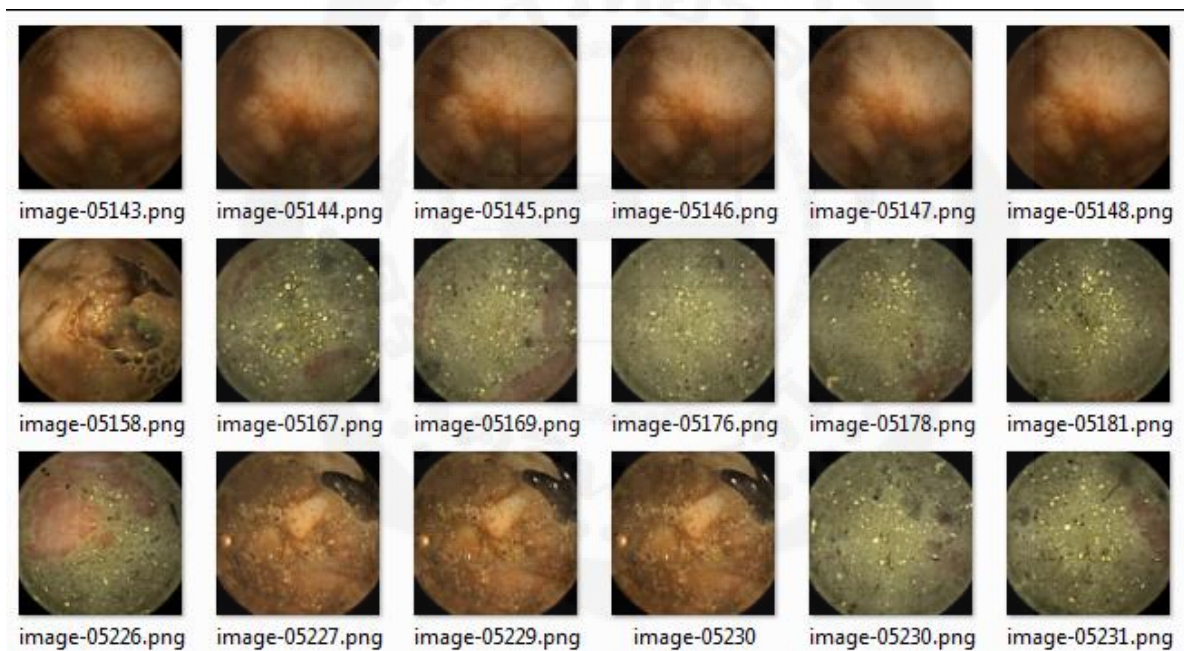
หมายเหตุ : ขั้นตอนที่ 3 นี้เป็น optional ไม่จำเป็นต้องทำได้

4. เมื่อทำการเตรียมภาพเรียบร้อยแล้ว ทำการเลือกเฉพาะ 4 class คือ Esophagus , Stomach , Small Bowel และ Colon ดังภาพที่ 29 และทำการสุ่มภาพจากกลุ่มโฟลเดอร์เลขคู่ เพื่อไปใส่ในโฟลเดอร์สำหรับ train โดยการสุ่มภาพนั้น ทำการสุ่มโดยเลือกเอง โดยยกตัวอย่างในภาพที่ 30 เนื่องจากการสุ่มภาพด้วยโค้ด มีโอกาสเกิดความผิดพลาดคือภาพมีความไม่ใกล้เคียงกัน เพราะในอวัยวะเดียวกันนั้น จุดเริ่มของอวัยวะจะมีความแตกต่างกับ จุดสุดท้ายของอวัยวะค่อนข้างมาก อีกทั้ง

บางภาพที่เจอไม่ใช่ความปกติของอวัยวะ นั้นด้วย เช่น มีแผล หรือมีเลือดออกเกิดขึ้นในตำแหน่งนั้นๆ หรือมีการ label ภาพผิด เป็นต้น ในส่วนของกลุ่มโพลเดอร์ เลขคี่ นำไปใส่โพลเดอร์ valid



ภาพที่ 29 : เลือกภาพจากหลายๆคนมาให้เหลือเพียง 4 class



ภาพที่ 30 : ภาพในส่วนของ Colon จากการสุ่มหลายๆคนมารวมกัน

5. สำหรับกลุ่มที่จะนำมาใช้ test จะใช้ภาพของกลุ่มบุคคลในปี 2019 โดยไม่มีการแยกกลุ่มคน จะทำการเลือกสุ่มว่าจะใช้ภาพจากบุคคลไหน

3.5 ปัญหาและอุปสรรค

1. Dataset ที่เป็นรูปภาพที่ได้มามีจำนวนเยอะมาก ทำให้ใช้เวลาในการจัดเตรียมข้อมูลที่จะใช้ในการ training นานเนื่องจากต้องคัดเลือกภาพที่เหมาะสม
2. ในช่วงแรกทางกลุ่มนิสิตยังไม่มีความชำนาญในการใช้ CNN ทำให้ต้องใช้เวลาในการศึกษาเพิ่มเติมและแก้ไขปัญหาที่ไม่คุ้นเคย
3. เนื่องจากใช้เวลาในการ train ข้อมูลนาน การปรับค่า parameter ต่างๆ จึงทำให้ปรับได้ยากตามไปด้วย
4. ภาพบางภาพมีเลือดและสิ่งแปลกปลอม ทำให้ไม่สามารถใช้ภาพเหล่านั้นได้ และต้องทำการตัดภาพเหล่านั้นออก
5. ภาพบางส่วนมีการ label ผิด ทำให้มีอวัยวะที่ไม่เกี่ยวข้องเข้ามาปนอยู่ในอวัยวะของส่วนนั้นๆ

บทที่ 4 ผลการดำเนินโครงการ

Data Modeling

ในงานนี้ได้ทำการนำเทคนิคในการวิเคราะห์รูปภาพเพื่อใช้ในการทำนายหา ความเป็นไปได้ของภาพว่าเหมือนกับอวัยวะ ส่วนใดในภายในระบบทางเดินอาหาร โดยใช้หลักการของโครงข่ายประสาทเทียมเข้ามาทำงาน ซึ่งชุดข้อมูลได้มาจาก Brown University และ Rhode Island Hospital โดยความร่วมมือกับ Amporn Atsawarungruangkit, MD ทางผู้วิจัยได้ดำเนินการ ตามกระบวนการและขั้นตอนต่างๆ ตามที่ได้เสนอไปในบทที่ 3 และเพื่อวัดประสิทธิภาพความถูกต้อง ของข้อมูล จึงได้มีการ ทำข้อมูลภาพในส่วนของ train , valid และ test เพื่อดูผลความแม่นยำที่เกิดขึ้น ผู้จัดทำได้เสนอการดำเนิน โครงการ โดยได้ผลลัพธ์ดังต่อไปนี้

ใช้สถาปัตยกรรม CNN : InceptionRensetV2 , Resnet50 , InceptionV3 และการทำ Ensemble โดยข้อมูลที่ใช้เป็นข้อมูลชุดเดียวกันดังนี้

train set จำนวน 30,000 Images

valid set จำนวน 10,000 Images

test set จำนวน 683,674 Images

โดยรูปภาพที่ใช้นั้นจะมีขนาด 256x256 pixel และ มีทั้งหมด 4 Classes ประกอบไปด้วย Colon , Esophagus , Small Bowel , Stomach

โดยแต่ละ Class นั้นก็จะเก็บภาพภายในระบบทางเดินอาหารที่เป็นในส่วนๆ นั้น และจำนวนภาพทั้งหมดที่ใช้ทั้ง 3 โมเดล และการทำ Ensemble มีจำนวนภาพเท่ากัน

ปี 2017	Colon	Esophagus	Small Bowel	Stomach	Sum
1701002	100	9	300	150	559
1701004	250	-	450	100	800
1701006	-	2	400	100	502
1701008	150	100	400	100	750
1701010	150	-	300	100	550
1701012	150	-	400	100	650
1702006	250	5	400	-	655
1702008	150	8	350	100	608
1702010	-	1	100	250	351
1702012	250	1	300	200	751
1702016	250	11	300	200	761
1703002	100	18	200	-	318
1703004	150	1	300	100	551
1703006	50	-	300	-	350
1703008	100	-	300	-	400
1703010	100	34	-	150	284
1703012	150	20	400	150	720

1703014	-	-	-	200	200
1704002	350	100	150	50	650
1704004	-	-	150		150
1704006	-	30	-	200	230
1704008	200	-	100	-	300
1704010	150	20	-	-	170
1704012	100	-	-	-	100
1704014	-	16	-	-	16
1704018	150	-	150	150	450
1704020	150	-	200	150	500
1705002	-	100	-	-	100
1705004	-	1	-	100	101
1705008	250	9	200		459
1705010	-	20	100	20	140
1705012	100	-	100	100	300
1705014	-	17	100	-	117
1705016	-	10	50	-	60
1705018	100	-	-	100	200

1705020	-	15	-	100	115
1705022		11	-	100	111
1706004	100	-	-	150	250
1706006	100	-	-	-	100
1706010		18	-	200	218
1706012	100	18	-	150	268
1706014	-	3	-	100	103
1706016	-	11	-	-	11
1706018	100	127	-	150	377
1706020	100	-	-	80	180
1706022	100	-	-	-	100
1706024	-	1	-	-	1
1707002	-	-	-	150	150
1707004	100	31	-	-	131
1707006	100	-	-	100	200
1707008	-	-	-	200	200
1707010	100	62	-	100	262
1707012	100	3	-	100	203

1707014	100	-	-	-	100
1708006	-	24	-	100	124
1708012	-	10	-	100	110
1709002	-	23	-	150	173
1709008	-	10	-	-	10
1709010	-	-	-	100	100
	...	...	...	...	...
Total Image	7500	2500	10000	10000	30000

ตารางที่ 1 : ตัวอย่างจำนวนรูปที่ใช้สำหรับ train set

จากตารางที่ 1 คือจำนวนภาพในส่วนของ Colon , Esophagus , Small Bowel และ Stomach ตามลำดับ โดยภาพทั้งหมดของ train set จะมีการสุ่มเลือกภาพมาทั้งหมด 91 คน มีทั้งหมด 30,000 ภาพ

ปี 2017	Colon	Esophagus	Small Bowel	Stomach	Sum
1701001	100	1	200	50	351
1701003	100	10	200	150	460
1701005	100	10	200	100	410
1701007	-	1	150	50	201

1701009	-	3	-	-	3
1701015	100	-	-	-	100
1702001		2	-	100	102
1702003	-	-	200	50	250
1702005	50	20	200	-	270
1702009	50	19	100	50	219
1702011	-	-	50	-	50
1703001	-	-	100	100	200
1703003	100	13	100	50	263
1703005	100	7	100	50	257
1703007	100	14	200	50	364
1703009	50		150	50	250
1703011	100	15	100	-	215
1703013	-	5	-	-	5
1704001	50	13	100	20	183
1704003	50	-	100	100	250
1704005	100	71	50	50	271
1704009	50	16	100	100	266

1704011	100	-	100	100	300
	...	...	...	...	...
Total image	2250	500	5000	2250	10000

ตารางที่ 2 : ตัวอย่างจำนวนภาพที่ใช้สำหรับ valid set

จากตารางที่ 2 คือจำนวนภาพในส่วนของ Colon , Esophagus , Small Bowel และ Stomach ตามลำดับ โดยภาพทั้งหมดของ valid set จะมีการสุ่มเลือกภาพมาทั้งหมด 32 คน มีทั้งหมด 10,000 ภาพ

โดยจากในตารางที่ 1 และ ตารางที่ 2 จะทำการรวมเป็น training set และจะได้ภาพทั้งหมด 40,000 ภาพ สำหรับ training set

ปี 2019	Colon	Esophagus	Small Bowel	Stomach	Sum
1901001	682	21	24706	34	25443
1901002	1141	16	7770	1250	10177
1901004	2123	9	4821	559	7512
1901005	1006	6	6011	16	7039
	...	...	...	...	...
Total image	52060	925	558784	71912	683681

ตารางที่ 3 : ตัวอย่างจำนวนภาพที่ใช้สำหรับ test set

จากตารางที่ 3 คือคือจำนวนภาพในส่วนของ Colon, Esophagus, Small Bowel และ Stomach ตามลำดับ จะนำภาพทั้งหมดของ 51 คน มาทำงานทดลอง มีทั้งหมด 683,681 ภาพ

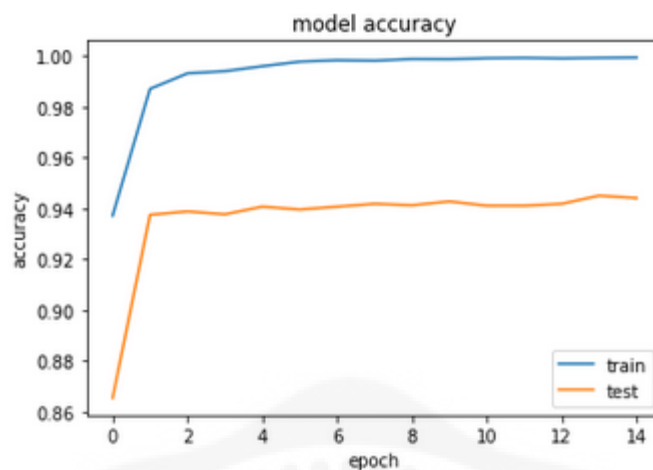
ในงานนี้ได้ทดลองใช้ Architecture ของ Resnet50 (ภาพที่ 5), InceptionV3 (ภาพที่ 6) และ InceptionResnetV2 (ภาพที่ 7) โดยได้กำหนด layer สุดท้ายเพิ่มเติมคือ AveragePooling2D เพราะในชั้นสุดท้ายของทั้ง 3 โมเดลนั้น ขนาดของภาพที่ได้ยังไม่เป็นในภาพแบบของ vector

พารามิเตอร์ที่กำหนดต่างๆในทุกๆ model คือ

- กำหนดจำนวน n_class (จำนวน class ที่จะรับข้อมูล)
- กำหนดค่า learning rate และค่า momentum ของ SGD optimizer ในการทำครั้งนี้ ให้ learning rate = 0.01 และ momentum = 0.9
- กำหนดจำนวนรอบในการ train (epochs)

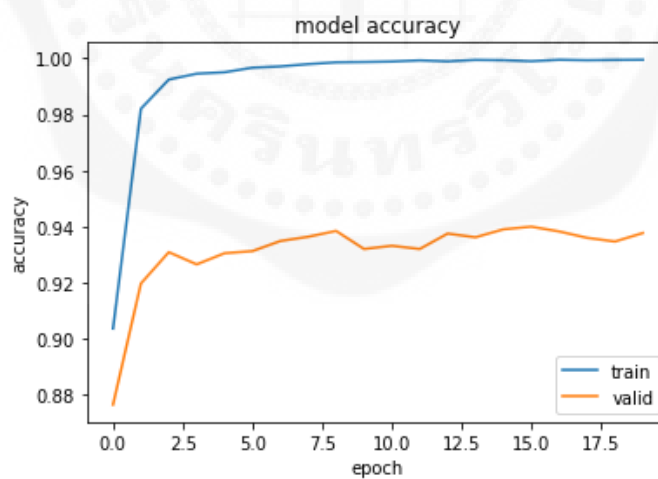
เวลาในการ train แต่ละครั้งประมาณ 1 - 4 ชั่วโมง เวลาในการ train จะขึ้นอยู่กับ input ของภาพที่เพิ่มขึ้น

Data Evaluation



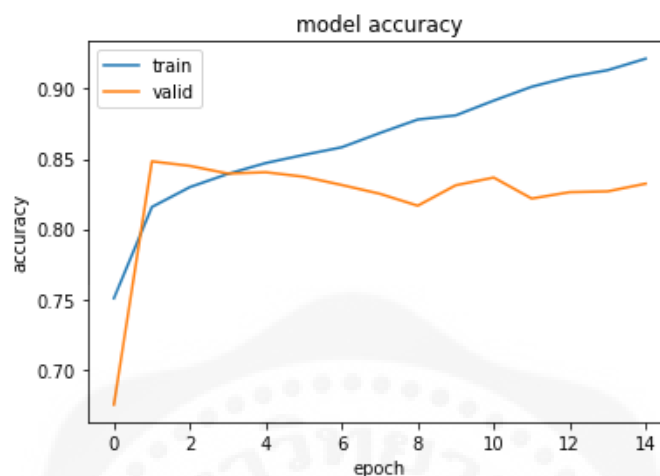
ภาพที่ 31 : ค่า accuracy ของ InceptionResnetV2

จากภาพที่ 31 ในส่วนของค่า accuracy ที่ได้ในระหว่างการทำโมเดล สำหรับ train model เทียบกับ valid ของโมเดล InceptionResnetV2 ได้ผล accuracy 94.41% คือค่าสูงสุด



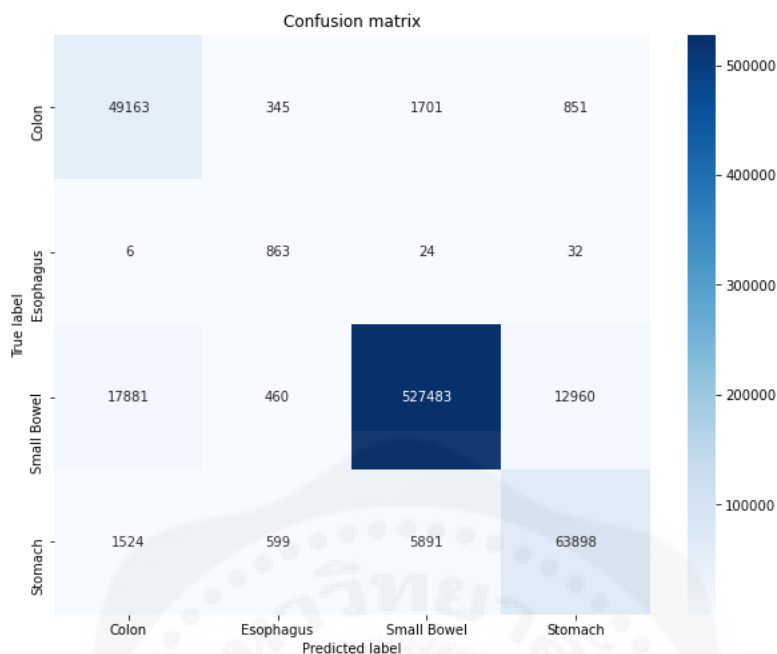
ภาพที่ 32 : ค่า accuracy ของ InceptionV3

จากภาพที่ 32 ในส่วนของค่า accuracy ที่ได้ในระหว่างการทำโมเดล สำหรับ train model เทียบกับ valid ของโมเดล InceptionV3 ได้ผล accuracy 93.92% คือค่าสูงสุด



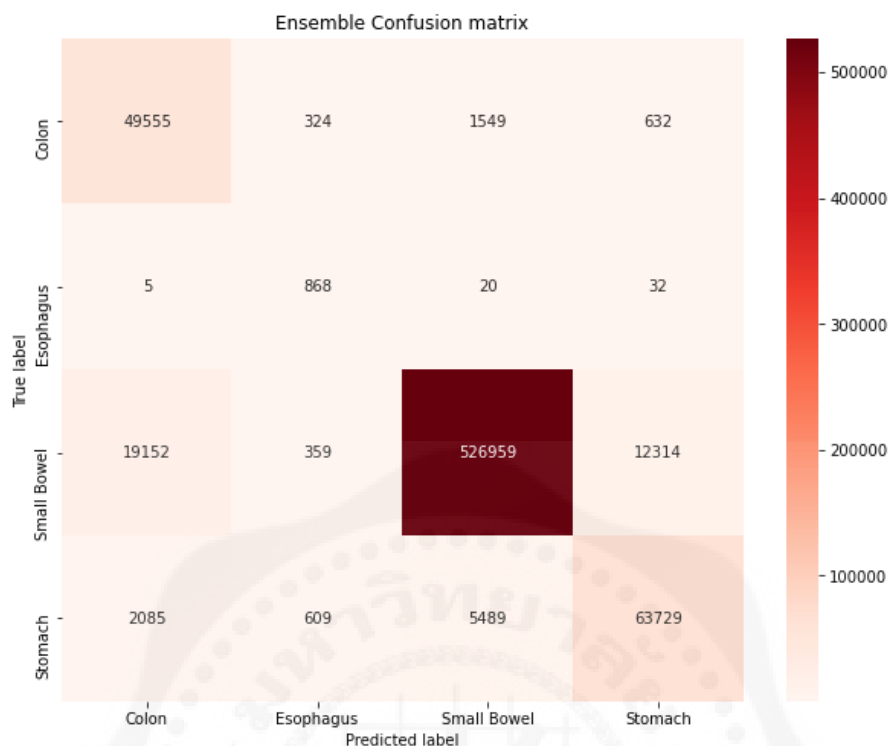
ภาพที่ 33 : ค่า accuracy ของ Resnet50

จากภาพที่ 33 ในส่วนของค่า accuracy ที่ได้ในระหว่างการทำโมเดล สำหรับ train model เทียบกับ valid ของโมเดล Resnet50 ได้ผล accuracy 85.02% คือค่าสูงสุด



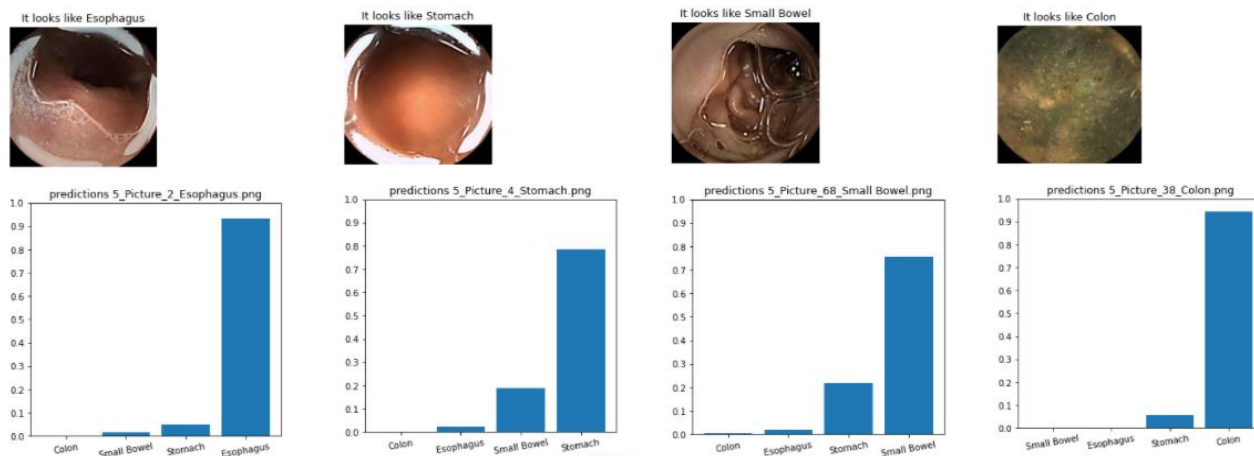
ภาพที่ 34 : Confusion matrix ของ model InceptionResnetV2

Confusion matrix จากภาพที่ 34 ได้อธิบายถึงว่า ในแนวดิ่งเป็น Actual label หรือ label ที่เป็นข้อมูลของภาพที่ถูกต้อง ส่วนในแนวนอนเป็น Predicted label สำหรับภาพนี้เป็นการอธิบายว่ามีการทำนายภาพถูกต้องมากแค่ไหนเมื่อทดสอบรูปภาพในส่วนของลำไส้ใหญ่ (Colon) มีการทำนาย 52,060 ภาพ ทำนายได้ถูกต้อง 49,163 ภาพ และมี 2,897 ภาพที่ทำนายผิดพลาดเป็น class ต่างๆตามภาพที่ 34 ในส่วนของหลอดอาหาร (Esophagus) มีการทำนาย 925 ภาพ ทำนายได้ถูกต้อง 863 ภาพ ในส่วนของลำไส้เล็ก (Small Bowel) มีการทำนาย 558,784 ภาพ ทำนายได้ถูกต้อง 527,483 ภาพ ในส่วนของกระเพาะอาหาร (Stomach) มีการทำนาย 71,912 ภาพ ทำนายได้ถูกต้อง 63,898 ภาพ

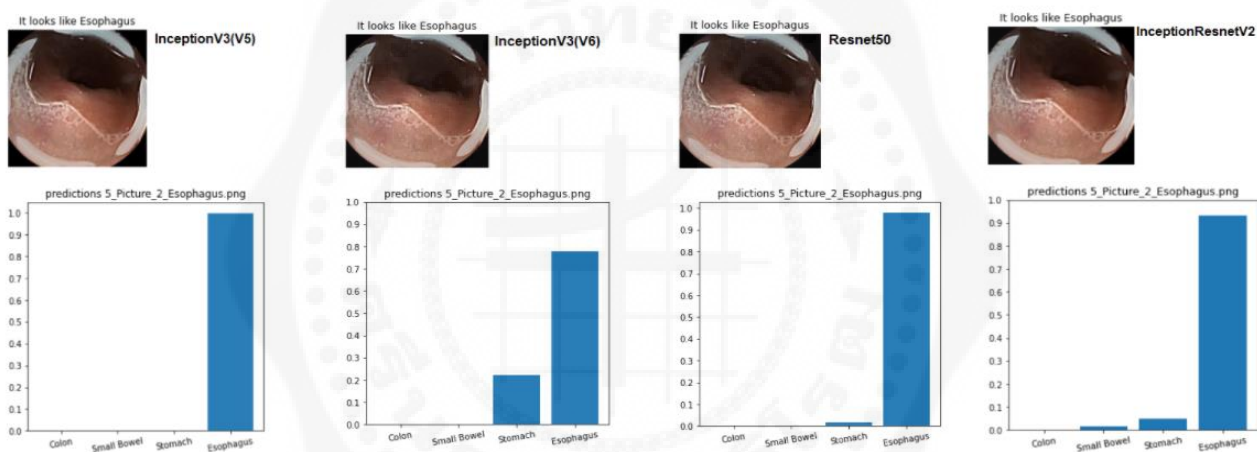


ภาพที่ 35 : Confusion matrix ของ Ensemble

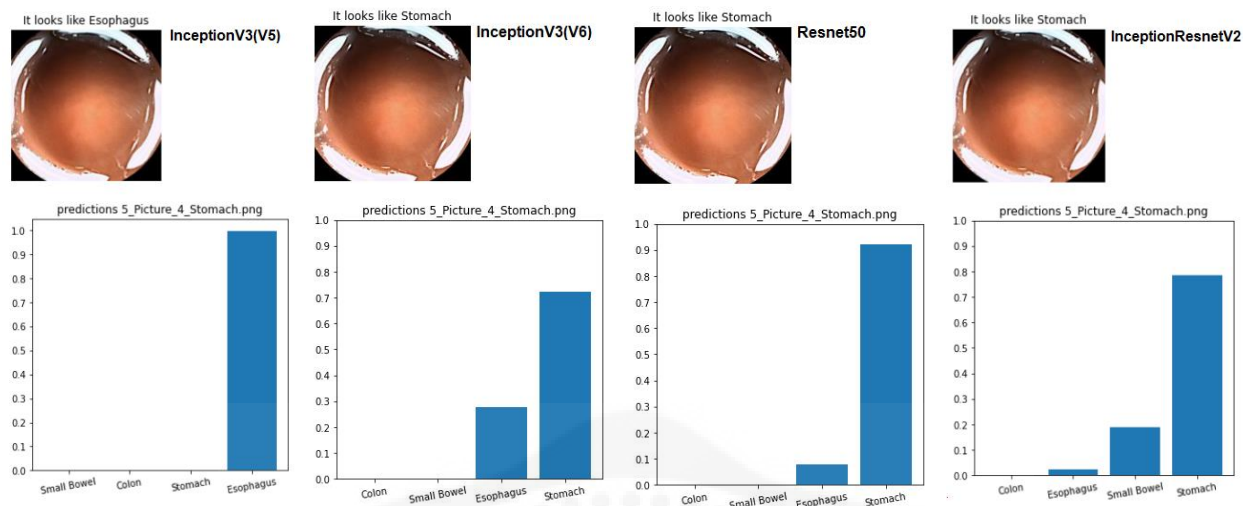
Confusion matrix จากภาพที่ 35 ได้อธิบายถึงว่า ในแนวนอนเป็น Actual label หรือ label ที่เป็นข้อมูลของมันที่ถูกต้อง ส่วนในแนวนอนเป็น Predicted label สำหรับภาพนี้เป็นการอธิบายว่ามีการทำนายภาพถูกต้องมากแค่ไหน เมื่อทดสอบรูปภาพในส่วนของลำไส้ใหญ่ (Colon) มีการทำนาย 52,060 ภาพ ทำนายได้ถูกต้อง 49,555 ภาพ และมี 2,505 ภาพที่ทำนายผิดพลาดเป็น class ต่างๆตามภาพที่ 35 ในส่วนของหลอดอาหาร (Esophagus) มีการทำนาย 925 ภาพ ทำนายได้ถูกต้อง 868 ภาพ ในส่วนของลำไส้เล็ก (Small Bowel) มีการทำนาย 558,784 ภาพ ทำนายได้ถูกต้อง 526,959 ภาพ ในส่วนของกระเพาะอาหาร (Stomach) มีการทำนาย 71,912 ภาพทำนายได้ถูกต้อง 63,729 ภาพ



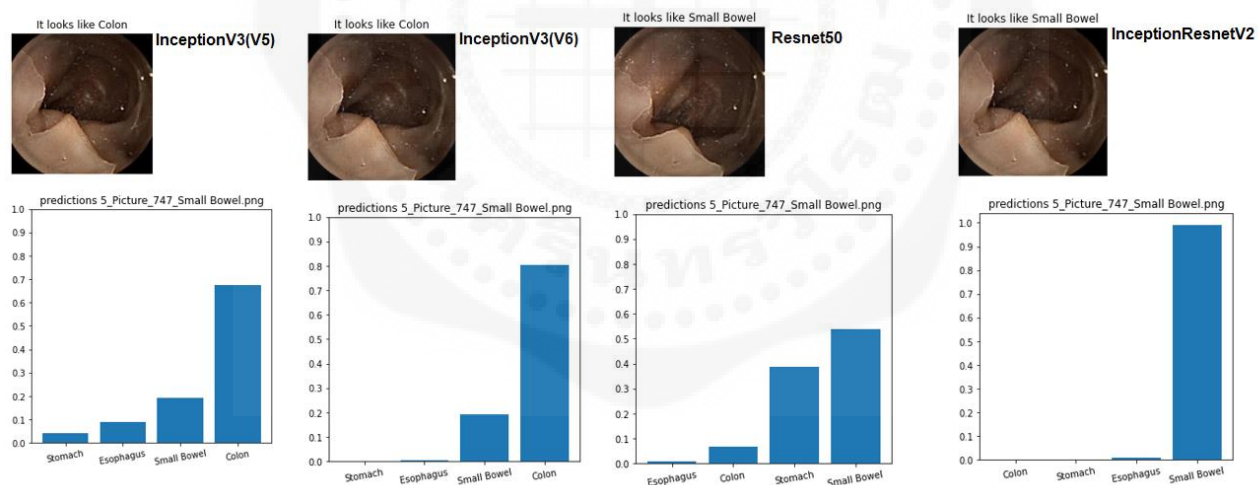
ภาพที่ 36 : ตัวอย่างของ probability ของผลการทำนายจาก InceptionResnetV2



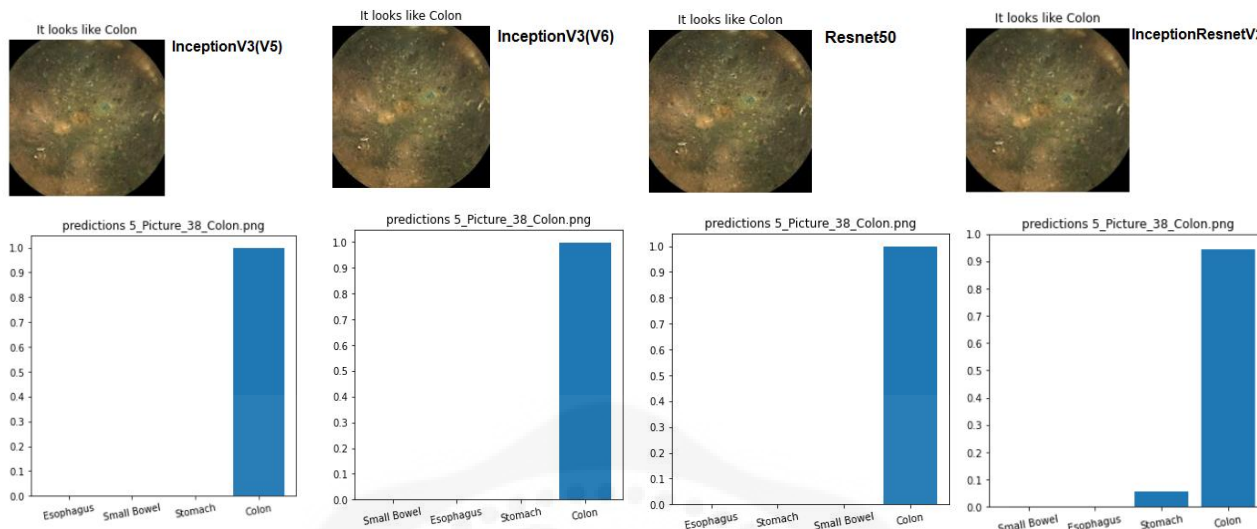
ภาพที่ 37 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะหลอดอาหาร ของ Ensemble



ภาพที่ 38 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะกระเพาะอาหาร ของ Ensemble



ภาพที่ 39 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะลำไส้เล็ก ของ Ensemble



ภาพที่ 40 : ตัวอย่าง ผลสำหรับการ Voting อวัยวะลำไส้ใหญ่ ของ Ensemble

จากภาพที่ 36 เป็นรูปส่วนหนึ่งของการทำนายรูปภาพจากกลุ่ม test set ภาพแรกนับจากทางซ้ายได้ผลการทำนายคือ ภาพของหลอดอาหาร (Esophagus) ด้วยความน่าจะเป็นคือประมาณ 0.91 ภาพที่สองนับจากทางซ้ายได้ผลการทำนายคือภาพของกระเพาะอาหาร (Stomach) ด้วยความน่าจะเป็นคือประมาณ 0.79 ภาพที่สามนับจากทางซ้ายได้ผลการทำนายคือภาพของลำไส้เล็ก (Small Bowel) ด้วยความน่าจะเป็นคือประมาณ 0.77 ภาพที่สี่นับจากทางซ้ายได้ผลการทำนายคือภาพของลำไส้ใหญ่ (Colon) ด้วยความน่าจะเป็นคือประมาณ 0.95

จากภาพที่ 37, 38, 39 และ 40 ภาพเริ่มจากภาพทางซ้ายเป็นภาพของ InceptionV3(V5), InceptionV3(V6), Resnet50 และ InceptionResnetV2 ตามลำดับ ในการทำ Ensemble นั้น เป็นการดูผลการทำนายที่แต่ละโมเดลทำนายมาเป็นอวัยวะใดบ้าง และจึงนำผล ที่เป็น majority มาทำการ voting เพื่อให้ได้ผลสุดท้ายออกมา

จากภาพที่ 37 นั้น ผลที่ได้จากทั้ง 4 โมเดลที่ทำนายได้คือหลอดอาหาร (Esophagus) ดังนั้นผล vote ของ Ensemble จึงเป็น หลอดอาหาร (Esophagus)

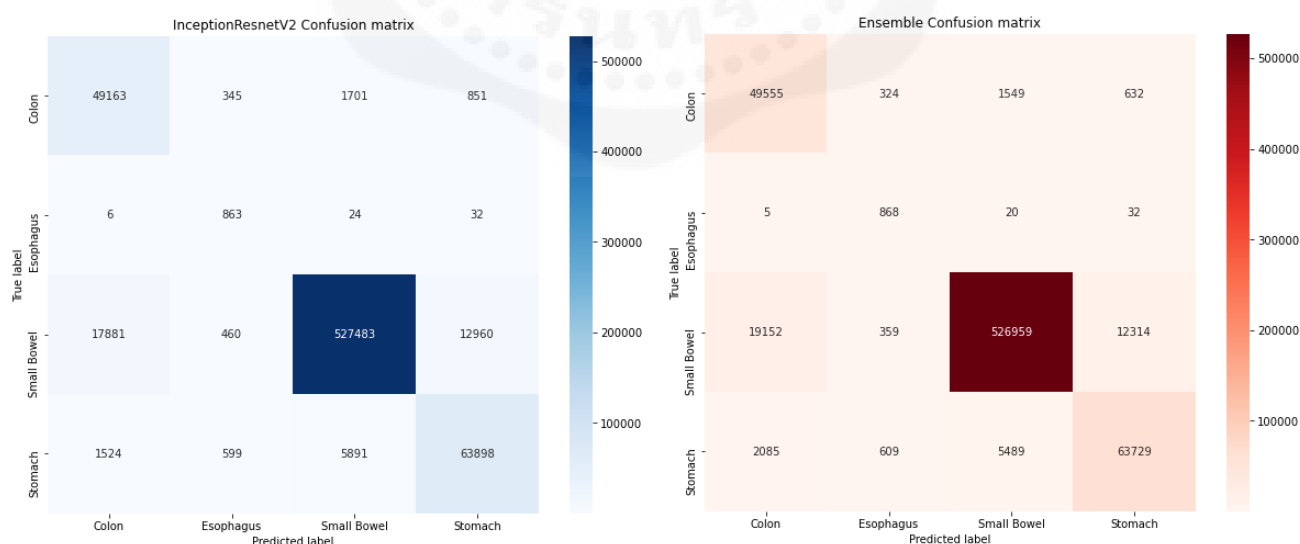
จากภาพที่ 38 นั้น ผลที่ได้จากโมเดล InceptionV3(V5) ได้เป็น หลอดอาหาร (Esophagus) อีก 3 โมเดลที่เหลือได้ผลเป็น กระเพาะอาหาร (Stomach) ดังนั้นผล vote ของ Ensemble จึงเป็น กระเพาะอาหาร (Stomach)

จากภาพที่ 39 นั้น ผลที่ได้จากโมเดล InceptionV3(V5) และ InceptionV3(V6) ได้เป็น ลำไส้ใหญ่ (Colon) ส่วนโมเดล InceptionResnetV2 และ Resnet50 ได้ผลเป็น ลำไส้เล็ก (Small Bowel) ในกรณีนี้โดยผล vote ของ Ensemble นั้น ได้เป็นจำนวนครั้งหนึ่งสำหรับโมเดลที่ใช้เป็นเลขคู่ ซึ่งในที่นี้คือ 4 โมเดล ได้แก่ InceptionV3(V5) , InceptionV3(V6) , Resnet50 และ InceptionResnetV2 ดังนั้นจะทำการยึดผลที่ได้จากการทำนายของ InceptionResnetV2 ซึ่งเป็นโมเดลหลักสำหรับการทำ Ensemble นี้เนื่องจากได้ผลภาพรวมที่ดีที่สุด ดังนั้นผล vote ของ Ensemble จึงเป็น ลำไส้เล็ก (Small Bowel)

จากภาพที่ 40 นั้น ผลที่ได้จากทั้ง 4 ที่ทำนายได้คือลำไส้ใหญ่ (Colon) ดังนั้นผล vote ของ Ensemble จึงเป็น ลำไส้ใหญ่ (Colon)

ทางกลุ่มนิสิตได้ทำการเปรียบเทียบและพัฒนาผลเพิ่มเติมโดยใช้ Ensemble ในการช่วยเพิ่มประสิทธิภาพของโมเดลที่ได้ ออกแบบมา โดยยกมาใช้งานทั้งหมด 4 โมเดล ได้แก่ InceptionV3(V5) , InceptionV3(V6) , InceptionResnetV2 และ Resnet50 โดย Ensemble นั้นจะมีการใช้ majority Vote เพื่อดูว่าในแต่ละโมเดลนั้นมีการทำนายเป็นอวัยวะใดบ้าง และหากอวัยวะใดเยอะสุดก็จะทำการแสดงผลออกมาเป็นอวัยวะนั้นๆ และถ้าหาก Majority Vote มีค่าเท่ากับ 2 จะให้ทำการเลือกผลการทำนายที่ได้จากโมเดล InceptionResnetV2 เพราะได้ทำการเลือกโมเดลนี้เป็นโมเดลหลักสำหรับการทำ Ensemble ในครั้งนี้และโมเดลนี้ให้ผลที่ดีที่สุดเมื่อเทียบกับโมเดลอื่น

ผลการทดลองที่ได้จากการใช้ Ensemble ด้วย majority vote โดยเทียบกับ InceptionResnetV2 สรุปได้ดังภาพที่ 41 และตารางที่ 4 และตารางที่ 5



ภาพที่ 41 : การเทียบผลระหว่าง InceptionResnetV2 กับ Ensemble

	Precision	Recall	F1-Score	Support
Colon	0.6592	0.8707	0.7192	52060
Esophagus	0.5358	0.9043	0.6100	925
Small Bowel	0.9666	0.9312	0.9474	558784
Stomach	0.6862	0.7454	0.6892	71912

Macro Avg	0.6988	0.8493	0.7293	683681
Weight Avg	0.9712	0.9453	0.9549	683681

Specificity	0.9744
Accuracy	0.9417

ตารางที่ 4 : InceptionResnetV2 classification report

	Precision	Recall	F1-Score	Support
Colon	0.6322	0.8781	0.7300	52060
Esophagus	0.5682	0.9100	0.6370	925
Small Bowel	0.9678	0.9279	0.9463	558784
Stomach	0.6624	0.7425	0.6704	71912

Macro Avg	0.7609	0.8650	0.7380	683681
Weight Avg	0.9741	0.9422	0.9539	683681

Specificity	0.9748
Accuracy	0.9421

ตารางที่ 5 : Ensemble classification report

บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ

5.1 สรุปผลการศึกษาค้นคว้า

จากการดำเนินโครงการผลการค้นคว้าแสดงได้ตามตารางที่ 4 และ 5 เป็นตารางผลต่างๆของ โมเดล InceptionResnetV2 และ การทำ Ensemble ตามลำดับ โดยมีค่าเฉลี่ยในการทำนายรูปภาพอยู่ที่ 20 milliseconds ต่อ 1 รูปภาพ ซึ่งภาพที่นำมาใช้ได้ถูกแบ่งเป็นภาพสำหรับ training 40,000 ภาพ และ testing 683,674 ภาพ โดยมีรายละเอียดดังตารางที่ 4 ที่แสดงให้เห็นถึง Classification report ของ InceptionResnetV2 และ ตารางที่ 5 ที่แสดงให้เห็นถึง Classification report ของการทำ Ensemble

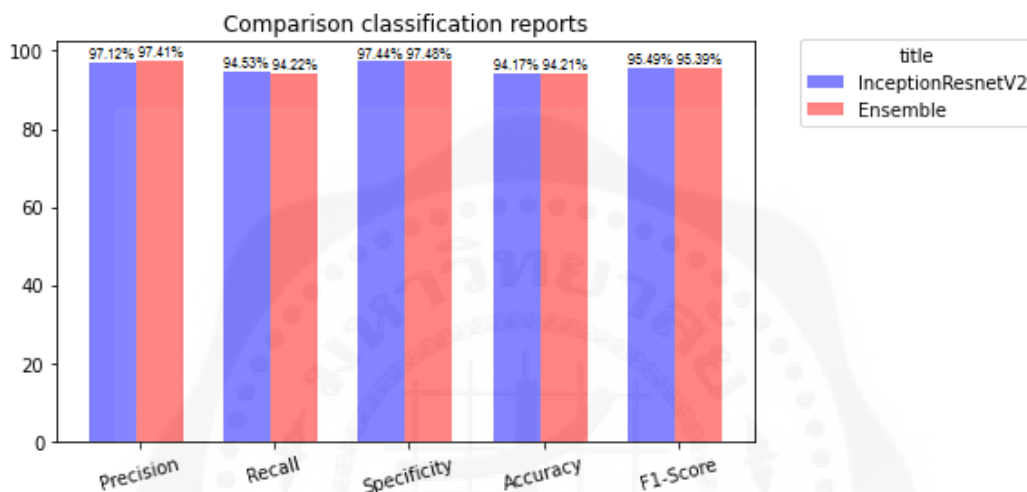
จากการสรุปผลเมื่อเทียบกันทั้ง 2 ตารางข้างต้นนี้ ผลที่ได้จาก Ensemble จะมีทั้งดีกว่าและแย่กว่า จากโมเดล InceptionResnetV2 ดีกว่าก็คือผล ลำไส้ใหญ่ (Colon) และ หลอดอาหาร (Esophagus) ส่วน ลำไส้เล็ก (Small Bowel) และ กระเพาะอาหาร (Stomach) ได้ผลที่ต่ำกว่าเพียงเล็กน้อย แต่ภาพรวม การทำ Ensemble จะทำให้ผลการทำดีขึ้น

5.2 อภิปรายผล

จากการดำเนินโครงการ ในส่วนของรูปภาพนั้น รูปภาพมีปริมาณมากทำให้ยากต่อการเลือกรูปภาพ และไม่สามารถใช้คำสั่งสุ่มรูปภาพได้เนื่องจากมีภาพที่ไม่สามารถใช้ได้ปะปนอยู่ เช่น ภาพอวัยวะที่มีเลือดหรือส่วนแปลกปลอม และในส่วนของการใช้โมเดลนั้น ได้มีการใช้โมเดลของ InceptionResnetV2 ได้ผลออกมาดีกว่าโมเดลอื่นที่ใช้ในการทดลองนี้ การปรับแก้โค้ดหรือ parameter ต่างๆ เป็นไปได้ลำบากเล็กน้อย เพราะเนื่องจากใช้เวลาในการทำงานในแต่ละรอบนาน จึงส่งผลให้งานเกิดความล่าช้าได้ และในส่วนของการเพิ่มประสิทธิภาพนั้นได้มีการทำ Ensemble ขึ้นมา โดยเทียบจาก majority vote ของโมเดลต่างๆที่ได้มีการทดลองมา ได้แก่ InceptionV3(V5) , InceptionV3(V6) , Resnet50 และ InceptionResnetV2 ผลที่ได้จากจำแนกรูปภาพจากกล้องแคปซูลไร้สาย เป็น 4 คลาส ของโมเดลที่ดีที่สุดคือ InceptionResnetV2 และได้ค่า Accuracy 94.17%, Precision 97.12%, Recall (Sensitivity) 94.53% และ F1-score 95.49%

มีค่าเฉลี่ยในการทำนายรูปภาพอยู่ที่ 20 milliseconds ต่อ 1 รูปภาพ นอกจากนี้ ทางกลุ่มผู้วิจัยได้ทดลองทำ Ensemble โดยใช้ทั้ง 4 โมเดล ได้ค่า Accuracy 94.21%, Precision 97.41%, Recall (Sensitivity) 94.22% และ F1-score 95.39% ซึ่งจะเห็นได้ว่าในภาพรวมนั้นการทำ Ensemble ดีขึ้นเล็กน้อย

Classification report จากตารางที่ 4 และ 5 สามารถสรุปผลลงในกราฟ ได้ดังภาพที่ 42



ภาพที่ 42 : Graph classification reports ของ InceptionResnetV2 และ Ensemble

โดยรูปทางซ้ายจะเป็น classification report ของ InceptionResnetV2 และทางขวาจะเป็น classification report ของการทำ Ensemble ทั้ง 4 โมเดล

5.3 ข้อเสนอแนะ

ทำโค้ดเพิ่มเติมเพื่อแยกภาพผิดปกติออกจากภาพปกติโดยที่ไม่จำเป็นต้องทำการแยกด้วยมือเอง, ในส่วนของ Ensemble อาจจะมีการใช้โมเดลมาทดสอบเพิ่มเติม, ทำ Ensemble เพิ่มเติมโดยใช้ค่าจาก probability ของรูปภาพที่ทำนายได้

บรรณานุกรม

- [1] **ตรวจลำไส้เล็กด้วยการกลืนแคปซูล “แคปซูล เอ็นโดสโคป”**. สืบค้นเมื่อ 7 มิถุนายน 2563. จาก <https://www.phukethospital.com/th/center/gastrointestinal-liver-center/capsule-endoscopy/>
- [2] **GI Bleeding (Gastrointestinal Bleeding)**. สืบค้นเมื่อ 7 มิถุนายน 2563. จาก <https://www.pobpad.com/เลือดออกในทางเดินอาหาร>
- [3] **Capsule endoscopy**. สืบค้นเมื่อ 7 มิถุนายน 2563. จาก <https://www.mayoclinic.org/tests-procedures/capsule-endoscopy/about/pac-20393366>
- [4] Asiri, S. (2018, Jun 11). **Machine Learning Classifiers**. Retrieved June 8, 2020, from <https://towardsdatascience.com/machine-learning-classifiers-a5cc4e1b0623>
- [5] **Convolutional Neural Network Architecture: Forging Pathways to the Future**. Retrieved June 17, 2020, from <https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future/?fbclid=IwAR2fTx7L7RDmXgVwyxqF4QaCALucvIN3Mw7OopnRt4jBS5Rv6UN6RokjYOY>
- [6] Phongchit, N. (2020, Mar 18). **มาทำความรู้จัก ResNet กันดีกว่า**. Retrieved June 19, 2020, from <https://medium.com/@natthawatphongchit/มาทำความรู้จัก-resnet-กันดีกว่า-aec3a8c10793>
- [7] Shaikh, F. (2017, Aug 9). **10 Advanced Deep Learning Architectures Data Scientists Should Know!**. Retrieved June 20, 2020, from <https://www.analyticsvidhya.com/blog/2017/08/10-advanced-deep-learning-architectures-data-scientists/?fbclid=IwAR00Kr9syEmbBg-mDw4uBsPFcY1-jZAlNoUAsYponl9Gg3bSZ4w2J66hA-E>
- [8] Elhamraoui, Z. (2020, May 16). **InceptionResNetV2 Simple Introduction**. Retrieved March 1, 2021. from <https://medium.com/@zahraelhamraoui1997/inceptionresnetv2-simple-introduction-9a2000edc6b6>
- [9] Cway Investment. (2018, Oct 6). **Ensemble Learning Method**. Retrieved April 18, 2021. from <https://medium.com/cw-quantlab/ensemble-learning-method-98359636adf9>

- [10] Visitsora-at,P. (2019, Oct 29). **Metrics พื้นฐานสำหรับวัดประสิทธิภาพของโมเดล Machine Learning**. Retrieved April 23, 2021. from <https://medium.com/@615162020027/metrics-พื้นฐานสำหรับวัดประสิทธิภาพของโมเดล-machine-learning-c00fcc32fa30>
- [11] **A Tour of Evaluation Metrics for Machine Learning**. Retrieved April 23,2021. from <https://www.analyticsvidhya.com/blog/2020/11/a-tour-of-evaluation-metrics-for-machine-learning/>
- [12] Lee, J., Oh, J., Shah, S. K., Yuan, X., & Tang, S. J. (2007). **“Automatic classification of digestive organs in wireless capsule endoscopy videos”**. *Proceedings of the 2007 ACM Symposium on Applied Computing - SAC 07*. doi:10.1145/1244002.1244230
- [13] Wang, S., Xing, Y., Zhang, L., Gao, H., & Zhang, H. (2019). **“Deep Convolutional Neural Network for Ulcer Recognition in Wireless Capsule Endoscopy: Experimental Feasibility and Optimization”**. *Computational and Mathematical Methods in Medicine*, 2019, 1-14. doi:10.1155/2019/7546215
- [14] Billah, M., Waheed, S., & Rahman, M. M. (2017). **“An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features”**. *International Journal of Biomedical Imaging*, 2017, 1-9. doi:10.1155/2017/9545920
- [15] Zou, Y., Li, L., Wang, Y., Yu, J., Li, Y., & Deng, W. J. (2015). **“Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network”**. 2015 IEEE International Conference on Digital Signal Processing (DSP). doi:10.1109/icdsp.2015.7252086
- [16] Jia, X., & Meng, M. Q. (2016). **“A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images”**. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). doi:10.1109/embc.2016.7590783

- [17] Jiang, Y., Chen, L., Zhang, H., & Xiao, X. (2019). **“Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module”**. Plos One, 14(3).
doi:10.1371/journal.pone.0214587
- [18] He, J., Wu, X., Jiang, Y., Peng, Q., & Jain, R. (2018). **“Hookworm Detection in Wireless Capsule Endoscopy Images With Deep Learning”**. IEEE Transactions on Image Processing, 27(5), 2379-2392.
doi:10.1109/tip.2018.2801119
- [19] Yu, J., Chen, J., Xiang, Z., & Zou, Y. (2015). **“A hybrid convolutional neural networks with extreme learning machine for WCE image classification”**. 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO). doi:10.1109/robio.2015.7419037
- [20] Choi, J., Shin, K., Jung, J., Bae, H., Kim, D. H., Byeon, J., & Kim, N. (2020). **“Convolutional Neural Network Technology in Endoscopic Imaging: Artificial Intelligence for Endoscopy”**. *Clinical Endoscopy*, 53(2), 117-126. doi:10.5946/ce.2020.054
- [21] Moccia, S., Wirkert, S. J., Kenngott, H., Vemuri, A. S., Apitz, M., Mayer, B., . . . Maier-Hein, L. (2018). **“Uncertainty-Aware Organ Classification for Surgical Data Science Applications in Laparoscopy”**. *IEEE Transactions on Biomedical Engineering*, 65(11), 2649-2659. doi:10.1109/tbme.2018.2813015