บทความวิจัย

Computer Science Department
Faculty of Informatics, Mahasarakham University

ระบบค้นหาสุนัขด้วยการเรียนรู้เชิงลึก

Find My Dog application using Deep Learning

นัทชา หาญจิตต์,หรรษธร นรินทร์รัมย์,พรทิวา ปะวะระ
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยามหาสารคาม
natcha.hanjit@gmail.com,masterkiler7788@gmail.com, pornntiwa.p@msu.ac.th

บทคัดย่อ

ในปัจจุบันการคัดแยกประเภทหมวดหมู่มีบทบาทสำคัญในเรื่องการใช้จำแนกลักษณะเฉพาะของวัตถุ สิ่งมีชีวิตต่างๆ ดังนั้นการแยกประเภทหมวดสิ่งที่มีลักษณะคล้ายกันจึงเป็นเรื่องยากเมื่อสังเกตลักษณะทั่วไปโดยรวม จึงจำเป็นต้องมีการคัดแยกสิ่งที่มีลักษณะเฉพาะ ซึ่งมีเอกลักษณ์เด่นเฉพาะตัวยกตัวอย่างเช่น การคัดแยกสายพันธุ์สุนัข และจำแนกลักษณะเด่นเฉพาะที่มีเพียงสุนัขตัว

ดังนั้นโครงงานฉบับนี้จึงนำเสนอการสร้างโมเดลของการประมวลผลเพื่อตามหาสุนัขที่หายไป โดยในงานนี้จะอยู่ในลักษณะของการแยกประเภทสายพันธุ์สุนัข โดยอาศัยการจำแนกภาพ(Classification) และโครงข่ายประสาทเทียม(Convolutional Neural Network: CNN)

1. บทนำ

ในปัจจุบันนี้จำนวนสุนัขจรจัดเป็นปัญหามาอย่างยาวนาน เนื่องจากในแต่ละสถานที่ สามารถพบเห็นสุนัขแทบทุกพื้นที่ ไม่ว่าจะเป็นสุนัขที่มีเจ้าของ หรือสุนัขที่หลงทางจากบ้านของมันเอง ซึ่งเป็นอีกปัญหาที่เจ้าของต้องการความช่วยเหลือ จะไปแจ้งตำรวจอาจจะดำเนินการล่าช้า หรือโพสลงโซเชียล (Social) เพื่อต้องการตามหา หากแต่เป็นการค้นหาแบบวงกว้างเกินไปที่จะหาสุนัขพบ และสุนัขจรจัดที่เกิดจากการเพาะพันธุ์ขายและเกิดจากการผสมพันธุ์กันเองของสุนัข จนมีมาเกินความต้องการ และถูกปล่อยทิ้ง ดังนั้นสุนัขจรจัดเหล่านี้ส่วนมากแล้วจึงไม่เคยได้รับการทำหมัน ลูกสุนัขที่เกิดขึ้นในบ้านและเจ้าของไม่ต้องการเลี้ยง ก็ถูกนำเอามาปล่อยทิ้งแถวถนน ตลาด ตรอกข้างตึก หรือวัด ส่งผลกระทบด้านต่างๆมากมาย ตั้งแต่เห่าหอนเสียงดังสร้างความรำคาญ วิ่งตัดหน้ารถเสียหลักจนเกิดอุบัติเหตุบนท้องถนน ขับถ่ายไม่เป็นที่เป็นทาง เป็นต้น เนื่องจากสุนัขแต่ละสายพันธุ์จะมีเอกลักษณ์เฉพาะ แต่อาจมีบางสายพันธุ์ที่ลักษณะใกล้เคียงกันจนทำให้คนทั่วไปแยกไม่ออก จึงเป็นเรื่องยากในการคัดแยกด้วยการมองเห็นลักษณะทั่วไปโดยรวม นอกจากจะสังเกตคุณสมบัติเฉพาะของแต่ละสายพันธุ์ จึงจะสร้างโมเดลจำแนกประเภทสุนัข (Model dog classification) เพื่อช่วยในการจำแนกสุนัขด้วยอัลกอริทึมแบบโครงข่ายประสาทเทียม (Convolutional Neural Network: CNN) โดยมีการจำแนกประเภทสายพันธุ์สุนัข

Computer Science Department
Faculty of Informatics, Mahasarakham University

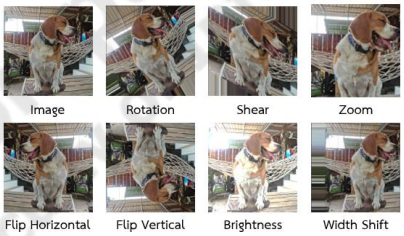
(Fine-grained classification) และคุณสมบัติเบื้องต้นของสุนัข(Multi-label classification) เพื่อทำการประมวลผลในการจัดกลุ่มสายพันธุ์

จากเหตุที่กล่าวมาข้างต้นจึงเกิดแนวคิดเพื่อประยุกต์ใช้ต้นแบบ Fine-grained classificationและMulti-label classification ในสร้างเป็นโมเดลจำแนกประเภทสุนัข(Model dog classification) ต้นแบบให้มีประสิทธิภาพและความแม่นยำสูง

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1. ทฤษฎีที่เกี่ยวข้อง

2.1.1. การเสริมข้อมูล (Data Augmentation) เป็นการเพิ่มข้อมูลรูปภาพหรือการแปลงภาพในข้อมูลการฝึกอบรม (Training Data) ให้เป็นภาพที่เปลี่ยนแปลง การซูม พลิก เลื่อน ตัด หมุน หรือเปลี่ยนความสว่างของภาพ จะทำให้ได้ภาพรูปแบบต่างๆไม่จำกัด เนื่องจากความแม่นยำของโมเดลขึ้นอยู่กับปริมาณข้อมูลเป็นปัจจัยสำคัญหลัก เพื่อเพิ่มความน่าเชื่อถือให้กับชุดข้อมูล



ภาพประกอบที่ 1 Data Augmentation

2.1.2. Convolutional Neural Network (CNN) หรือโครงข่ายประสาทเทียมจัดเป็นการเรียนรู้เชิงลึก (Deep Learning) มีการทำงานเริ่มจากการรับข้อมูลเข้ามาผ่านชั้นนำเข้าข้อมูล (Input layer) เพื่อส่งเข้าไปทำการประมวลผลโดยหน่วยย่อยของแต่ละชั้นที่ชั้นประมวลผล (Hidden Layers) เมื่อประมวลผลเสร็จชั้นสุดท้ายจะส่งไปยังชั้นแสดงผล (Output layer) ซึ่ง CNN สามารถจับคุณสมบัติที่เด่นของข้อมูล โดยใช้ตัวกรองที่เกี่ยวข้อง ในการลดจำนวนพารามิเตอร์ของข้อมูล โดยไม่ลดประสิทธิภาพของคุณสมบัติลง ขั้นตอนการหลักของ Convolutional Neural Network มีทั้งหมด 3 ชั้น คือ

(1) Convolution layer เป็นการค้นหาคุณลักษณะที่สำคัญของภาพ โดยขั้นตอนนี้จะใช้ตัวกรอง (Filter หรือ Kernel) เพื่อแยกองค์ประกอบต่างๆของภาพ เช่น ขอบ สี เป็นต้น โดยปกติภาพจะมีสีหลัก 3 สี (RGB) คือ สีแดง สีน้ำเงิน และสีเขียว แบ่งเป็น 3 Channel และภาพแสดงเป็นเมทริกซ์ 3 มิติ คือ ความกว้าง ความสูง และความลึก (Width, Height, Depth) โดยความลึกสอดคล้องกับช่องสี (RGB) ซึ่งแต่ละพิกเซลบอกความเข้มของสี ตั้งแต่ 0 ถึง 255 และทำการ Convolution เพื่อเก็บค่าไว้ในเมทริกซ์ชุดใหม่ที่เรียกว่า Feature Map

(2) Pooling layer เป็นการลดขนาดของ Feature Map ให้มีขนาดเล็กลงแต่ยังคงรักษาคุณสมบัติของข้อมูลที่สำคัญไว้ โดยมีการหาส่วนที่สำคัญที่สุดของข้อมูล และเพิ่มประสิทธิภาพการประมวลผลรวดเร็วมากขึ้นด้วยการหาค่าที่สูงที่สุด (Max Pooling), การหาค่าเฉลี่ย (Average Pooling) ของส่วนที่ครอบคลุมด้วยตัวกรองเก็บไว้

(3) Fully Connected Layer การเชื่อมต่อกันของแต่ละชั้นอย่างสมบูรณ์ จากกระบวน

Computer Science Department Faculty of Informatics, Mahasarakham University

การ 2 ขั้นตอน คือ Convolution layer ทำการสกัดคุณลักษณะที่สำคัญของภาพ และ Pooling layer ทำการลดมิติของ Feature Map ให้มีขนาดเล็กลงแต่ยังคงรักษาคุณสมบัติของข้อมูลที่สำคัญไว้ ซึ่งในขั้นตอนนี้จะเป็นการทำซ้ำตั้งแต่ Convolution layer ไปจนถึง Pooling layer จนกว่าจะเกิดการเชื่อมต่อกันอย่างสมบูรณ์ และทำกระบวนการ 2 ขั้นตอน ด้วย Activation Functions และ Loss Function เพื่อนำลักษณะเด่นที่สำคัญของภาพที่ได้มาทำการสร้างเป็น Neural Network สำหรับการเรียนรู้และทำนายประเภทของภาพ

2.1.3. สถาปัตยกรรม CNN (CNN Architecture)

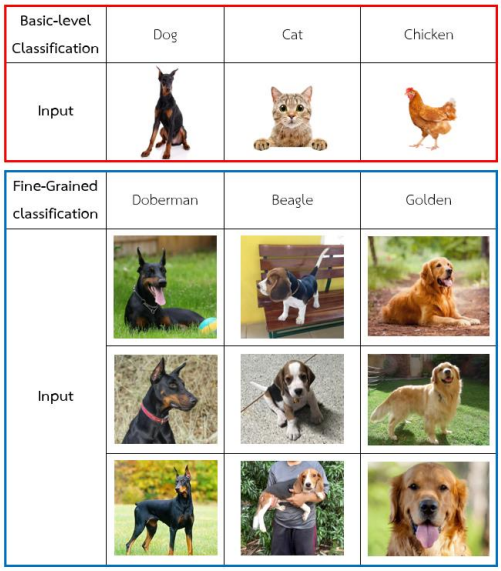
2.1.3.1. เป็น Image Classifier Neural Network คือ Convolutional Neural Network และรุ่นที่สามของ Inception Model ซึ่งพัฒนาขึ้นครั้งแรกโดยโมเดล GoogLeNet เป็นโครงข่ายประสาทเทียมขนาดใหญ่และมีระดับความแม่นยำที่ดีตามการจัดประเภทภาพ คุณลักษณะบางอย่างของโมเดลนี้ รวมถึงน้ำหนัก(weights) เป็นการจำแนกประเภทที่เป็นลักษณะเฉพาะที่สำคัญที่สุดบางประการสำหรับโมเดลนี้ Inception V3 เป็นโครงข่ายประสาทเทียมที่มีประสิทธิภาพเนื่องจากช่วยให้เรียนรู้ความซับซ้อนหลายอย่างและคุณลักษณะที่ยากของแบบจำลองหรือชุดข้อมูล มีขนาดตัวกรองหลายขนาด เช่น 1×1, 3×3, 5×5 ซึ่งสามารถใช้ได้ในระดับเดียวกัน

2.1.3.2. Residual Networks แบบ Deep Convolutional Networks มีแนวคิดพื้นฐานคือการเชื่อมต่อแบบ skipblock (หรือการเชื่อมต่อทางลัด) เพื่อให้พอดีกับ Input จากชั้นก่อนหน้าไปยังชั้นถัดไปโดยไม่ต้องปรับเปลี่ยนใดๆ ซึ่งในการทดลองของงานวิจัย จะทำการลบโมดูลทางลัดแบบลึกบางส่วนออกและเพิ่ม Classification layer ใหม่ที่ปรับให้เข้ากับชุดข้อมูลของงาน

2.1.3.3. การจัดประเภทภาพแบบ Fine-grained classification เป็นการแยกความแตกต่างระหว่างคลาสย่อย (Subclass) ของหมวดหมู่ (Class) เดียวกัน เพื่อลดความแปรปรวนภายในคลาสที่สูงและความแปรปรวนระหว่างคลาสที่ต่ำ เป็นการการออกแบบการแสดงคุณลักษณะและตัวแยกประเภทตามคำอธิบายระดับต่ำขั้นพื้นฐาน (Basic low-level descriptors) โดยตัวอย่างภาพอาจคล้ายกันมากและอาจแตกต่างกันในส่วนที่ละเอียดอ่อนบางคุณลักษณะที่แยกแยะได้

ความยากในการจำแนกประเภทหลายการแสดงผล คือปัญหาของการจำแนกประเภทคลาสที่มีลาเบลหลายรายการให้กับแต่ละข้อมูลเป็นการสร้างแบบจำลองเชิงคาดการณ์ที่เกี่ยวข้องกับการแสดงลาเบลหลายลาเบลของคลาสที่ได้รับข้อมูลเข้ามาเพื่อจดจำคุณสมบัติเฉพาะของข้อมูล

Computer Science Department
Faculty of Informatics, Mahasarakham University



ภาพประกอบที่ 2 Fine-grained classification

2.1.3.4. การจำแนกประเภทหลายลาเบล (Multi-label classification) เมื่อรับข้อมูลเข้ามา สามารถอยู่ในลาเบลได้มากกว่าหนึ่ง ซึ่งเป็นตัวแปรของสองคลาสขึ้นไปพร้อมกันได้ แต่ละข้อมูลสามารถเป็นได้มากกว่าหนึ่งคลาสในเวลาเดียวกัน ตัวอย่างเช่น สายพันธุ์สามารถอยู่ในประเภทใดประเภทหนึ่งเหล่านี้ ปั๊ก, ไซบีเรียน, บีเกิ้ล, คอร์กี้ หรืออาจเป็นของทั้งสี่ประเภทพร้อมกัน นอกจากนี้ในปัญหาการจำแนกหลายลาเบลคือไม่จำกัดว่าสามารถกำหนดข้อมูลที่นำเข้ามา ว่าแต่ละข้อมูลจะเป็นได้กี่คลาส

2.1.6 การประเมินประสิทธิภาพ (Evaluation)

(1) Confusion Matrix

ในการประเมินประสิทธิภาพการทำนาย หรือการทำนาย (Prediction) กับโมเดลที่สร้างขึ้นในโปรแกรมแอปพลิเคชั่น เปรียบเทียบกับผลลัพธ์จริงที่ถูกต้อง โดยมีความหมายแต่ละตัวดังนี้

True Positive (TP) คือ ค่าทำนายว่าจริงและผลการทำนายบอกว่าจริง

False Positive (FP) คือ ค่าทำนายว่าจริงและผลการทำนายบอกว่าไม่จริง

True Negative (TN) คือ ค่าทำนายว่าไม่จริงและผลการทำนายบอกว่าไม่จริง

False Negative (FN) คือ ค่าทำนายว่าไม่จริงและผลการทำนายบอกว่าจริง

(2) ค่าระลึก (Recall)

ค่าที่โมเดลทำนายเป็นคลาสที่กำลังพิจารณาถูก (TP) หารค่าเหตุการณ์จริงเป็นคลาสที่กำลังพิจารณาทั้งถูกและผิด (TP+FN) คำนวณโดยสมการดังนี้

$$Recall = \frac{TP}{TP + FN} \qquad (1)$$

(3) ค่าความแม่นยำ (Precision)

ค่าที่โมเดลทำนายเป็นคลาสที่กำลังพิจารณาถูก (TP) หารค่าที่โมเดลทำนายว่าเป็นคลาสที่กำลังพิจารณาทั้งถูกและผิด (TP+FP) คำนวณโดยสมการดังนี้

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

(4) ค่าความถูกต้อง (Accuracy)

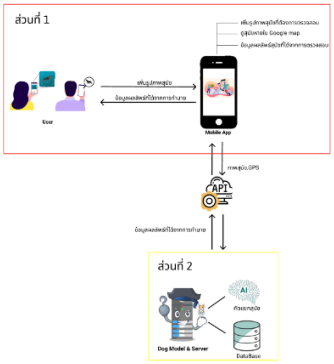
การคำนวณความถูกต้องเพื่อเปรียบเทียบประสิทธิภาพการทำนาย จากค่าที่โมเดลทำนายถูกทั้งหมด (TP+TN) หารค่าทั้งหมด (TP+TN+FP+FN) คำนวณโดยสมการดังนี้

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \qquad (3)$$

Computer Science Department
Faculty of Informatics, Mahasarakham University

3. วิธีดำเนินงานวิจัย

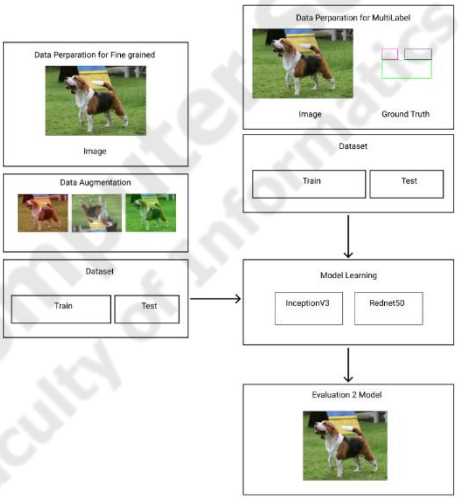
3.1. กรอบการดำเนินงาน



ภาพประกอบที่ 3 การดำเนินงานของระบบ

3.2. ขั้นตอนการดำเนินงานระบบค้นหาสุนัขด้วยการเรียนรู้เชิงลึก

การสร้างแบบจำลองโมเดลแบบ Classification โดยใช้ CNN ในงานวิจัยนี้ใช้สถาปัตยกรรม CNN แบบ Inception V3 และ ResNet50



ภาพประกอบที่ 4 กรอบการดำเนินงาน

3.3. การรวบรวมชุดข้อมูล (Data Preparation)

ในการเก็บข้อมูลในการวิจัยนี้ ได้ทำการเก็บชุดข้อมูลสุนัข 11 สายพันธุ์จากเว็บไซต์ต่างๆ จาก Google Search, Facebook, Stanford Dogs จาก ImageNet และจากการภาพถ่ายทางมือถือ โดยจะแบ่งชุดข้อมูลสำหรับ 2 โมเดล คือ

3.3.1. ชุดข้อมูลที่ใช้ในการสร้างโมเดลมีทั้งหมด 16,243 ภาพ จะแบ่งเป็นชุดข้อมูลสำหรับการสอน 12,994 ภาพ การทดสอบ 3,249 ภาพ

ภาพประกอบที่ 5 สุนัขแต่ละสายพันธุ์

3.3.2. ชุดข้อมูลสำหรับ Multi-label Modelชุดข้อมูลที่ใช้ในการสร้างโมเดลมีทั้งหมด 13,872 ภาพ ทำการปรับแต่ละภาพเป็นขนาด 227 × 227 และนำมาทำ Ground Truth จากนั้นจะแบ่งเป็นชุดข้อมูลสำหรับการสอน 80% และการทดสอบ 20% ด้วยมือ โดยจะกำหนดให้แต่ละ Label ของ Feature หู หาง ขน ปลอกคอ มีอยู่ 2 กรณี คือ 0 หรือ 1 เช่น หูตั้ง หูตก หางสั้น หาง ยาว เป็นต้น

Computer Science Department
Faculty of Informatics, Mahasarakham University

4. ผลการทดลอง

4.1. วิธีการทดลอง

จากชุดข้อมูลสำหรับ Fine-grained มีทั้งหมด 16,967 ภาพ และชุดข้อมูลสำหรับ Multi-label มีทั้งหมด 13,864 ภาพ โดยแต่ละโมเดลสามารถแบ่งชุดข้อมูลออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลการเรียนรู้ 80 เปอร์เซ็นต์ และชุดข้อมูลการตรวจสอบ 20 เปอร์เซ็นต์

ตารางที่ 1 การแบ่งข้อมูล

Classification	จำนวนข้อมูล		
	การเรียนรู้ (80%)	การตรวจสอบ (20%)	การทดสอบ
Fine-grained	12,994	3,973	2,558
Multi-label	11,092	2,772	

การตั้งค่าการฝึกการเรียนรู้

ตารางที่ 2 การตั้งค่าการฝึกการเรียนรู้

Function	Description
Threshold ค่ากลางในการแปลงค่าทำนาย	0.5
Epoch กำหนดรอบการเรียนรู้	15
Batch size ขนาดของชุดข้อมูล	32
Train base CNN ฝึกการเรียนรู้โมเดลส่วนที่หาคุณลักษณะเด่นของรูปภาพ (Feature Extraction)	Inception V3, ResNet50
Image size กำหนดขนาดภาพที่จะทำการฝึกทั้งหมด	277
Activation for Multi class แปลงผลรวมเพื่อให้ได้ผลลัพธ์สำหรับการจัดหมวดหมู่สายพันธุ์ หู หาง ขน ปลอกคอ	SoftMax
Activation for Multi-label แปลงผลรวมเพื่อให้ได้ผลลัพธ์สำหรับการจัดหมวดหมู่สีตัว	Sigmoid
Optimizer	Adam, Learning rate 0.00001

ตารางที่ 3 การตั้งค่าการฝึกการเรียนรู้ (ต่อ)

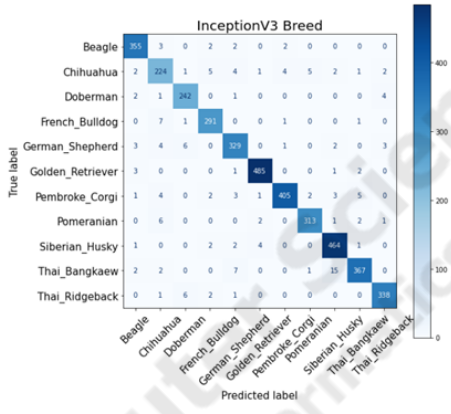
4.2. การเปรียบเทียบประสิทธิภาพการรู้จำจากสถาปัตยกรรม CNN

ในการจำแนกสายพันธุ์ได้ประยุกต์ใช้ร่วมกับสถาปัตยกรรม CNN โดยทำการเลือกใช้สถาปัตยกรรม Inception V3 และ ResNet50 ตามลำดับเพื่อหาผลการเรียนรู้ที่มีประสิทธิภาพที่ดีที่สุด โดยทำการปะเมินประสิทธิภาพโมเดลด้วย confusion matrix ในการทดลองนี้ ได้ใช้ข้อมูลในการเรียนรู้ในการทำ Fine-grained

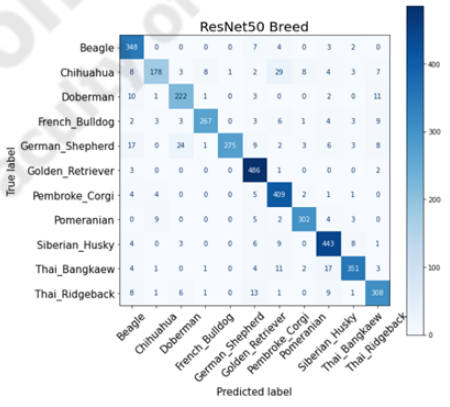
Computer Science Department Faculty of Informatics, Mahasarakham University

classification สำหรับสายพันธุ์ การทำ Multi-label classification สำหรับหู หาง สีตัว ขน ปลอกคอ

การประเมินผลประสิทธิภาพด้วย confusion matrix กับชุดข้อมูลตรวจสอบ 20 เปอร์เซ็นต์จากข้อมูลทั้งหมด จากนั้นนำข้อมูลสำหรับประเมินประสิทธิภาพแต่ละหมวดหมู่หาค่าความแม่นยำ (Precision), ค่าความระลึก (Recall) ด้วย Confusion Matrix ที่มีค่ากลางเป็น 0.5 เมื่อได้ผลลัพธ์แล้วนำค่าทั้งสองหาค่าความถูกต้อง (Accuracy) สามารถแสดงตัวอย่างการทำ confusion matrix ได้ดังนี้



ภาพประกอบที่ 6 InceptionV3 Confusion matrix



ภาพประกอบที่ 7 ResNet50 Confusion matrix

จากผลการประเมินประสิทธิภาพพบว่าค่าความถูกต้องของสถาปัตยกรรม Inception V3 ดีกว่าหรือใกล้เคียงอยู่ที่ประมาณ 2-3 เปอร์เซ็นต์เมื่อเทียบกับ ResNet50 และได้ทำการทดสอบกับภาพ 100 ภาพ พบว่าโมเดลที่ประยุกต์ใช้กับสถาปัตยกรรม Inception V3 ได้ผลลัพธ์ที่ถูกต้องมากกว่า ดังนั้นเราจึงเลือกใช้สถาปัตยกรรม Inception V3 ประยุกต์ร่วมกับโมเดล โดยการทดลองดังกล่าว สามารถแสดงภาพตัวอย่างที่ได้จากการทำนายผลลัพธ์โดยยกตัวอย่าง ได้ดังต่อไปนี้

ตารางที่ 4 ผลลัพธ์จากการทดสอบของสถาปัตยกรรมที่ประยุกต์ใช้

ภาพที่ 5			
Model	ผลเฉลย	ผลทำนายจากสถาปัตยกรรม	
		Inception V3	ResNet50
สายพันธุ์	หลังอาน	หลังอาน	บีเกิ้ล
หู	หูตั้ง	หูตั้ง	หูตก
หาง	หางยาว	หางยาว	หางสั้น
สีตัว	สีน้ำตาลเข้ม	สีน้ำตาลเข้ม	สีขาว
ขน	ขนสั้น	ขนสั้น	ขนสั้น
ปลอกคอ	ไม่มีปลอกคอ	ไม่มีปลอกคอ	ไม่มีปลอกคอ

ตารางที่ 5 สรุปผลการประเมินประสิทธิภาพ

แบบเดิม	Accuracy (%)	เลือกใช้
สายพันธุ์	96	✓
หู	96	✓
หาง	76	✓
สีตัว	68	✓
ขน	45	×
ปลอกคอ	86	×

Computer Science Department Faculty of Informatics, Mahasarakham University

จากผลการเรียนรู้ของ 2 ปัตยกรรม ทั้ง 6 โมเดล โดยการทดลองดังกล่าว โดยที่โมเดลจากที่ได้เลือกเรียงจากค่า accuracy สูงไปต่ำ พบว่าโมเดลขนและปลอกคอมีค่า Precision และค่า Recall ต่ำ และเมื่อนำโมเดลทั้ง 2 ไปทำนายผลกับรูปภาพสุนัข 100 รูป มีค่าเฉลี่ยการทำนายถูกต้องที่ต่ำ จึงเลือกมาโมเดล 4 โมเดลจากทั้ง 6 โมเดล คือ โมเดลสายพันธุ์ หู หาง สีตัว

5. สรุปและอภิปรายผลการทดลอง

5.1. สรุปผลและอภิปรายผล

โครงงานปริญญานิพนธ์ฉบับนี้นำเสนอระบบค้นหาสุนัขด้วยการเรียนรู้เชิงลึก (Find My Dog application using Deep Learning) โดยใช้สถาปัตยกรรม Inception V3 และ ResNet50
มาประยุกต์ใช้ในงานนี้ ร่วมกับการเรียนเรียนรู้เชิงลึก ซึ่งสามารถใช้งานการประมวลผลกับโมเดลที่จำเป็นต่อการนำไปประยุกต์ใช้บน Application

จากการทดลองประเมินประสิทธิภาพด้วย Confusion Matrix การรู้จำภาพสุนัขโดยใช้ชุดข้อมูลรูปภาพสายพันธุ์สุนัข 3,973 ภาพ และภาพคุณลักษณะของสุนัข 2,772 ภาพ ด้วยการประยุกต์ใช้ร่วมกับสถาปัตยกรรม Inception V3 มีประสิทธิภาพในการทำนายสูงสุด และได้ทำการเลือกโมเดลที่จำเป็นเพื่อนำมาประมวลผลบน Application ได้แก่โมเดลสายพันธุ์ หู หาง สีตัว เนื่องจากมีค่าความถูกต้องสูงและผลการทดสอบกับภาพสุนัข 100 ภาพมีผลการทำนายที่ถูกต้องมากกว่าเมื่อเทียบเท่ากับโมเดลขนและปลอกคอ

โดยจะพบปัญหากับภาพสำหรับการเรียนรู้ Fine-grained classification มีภาพสุนัขสายพันธุ์ต่างกันหลายตัวในหนึ่งภาพ อีกทั้งมีบางลักษณะคล้ายกับอีกสายพันธุ์ และภาพที่ใช้ในการเรียนรู้เพื่อทำ Ground truth สำหรับ Multi-label classification ส่วนใหญ่เป็นภาพทางด้านหน้าสุนัขและไม่เห็นคุณลักษณะอื่นที่ชัดเจน รวมทั้งคุณลักษณะที่สัดส่วนข้อมูลไม่เท่ากัน ทำให้เรียนรู้ข้อมูลและรู้จำข้อมูลที่มีจำนวนมากกว่า ทำให้การประมวลผลมีการทำนายออกมาผิดพลาดและค่าความถูกต้องของโมเดลต่ำลง

5.2. ปัญหาและอุปสรรคในการดำเนินงาน

5.2.1. ชุดข้อมูลในการเรียนรู้มีสัดส่วนของแต่ละลักษณะต่างกันมาก(Imbalanced dataset) ทำให้การรู้จำข้อมูลไม่เท่ากัน ข้อมูลของลักษณะหูตั้งมีมากกว่าหูตก หางยาวมากกว่าหางสั้น ไม่มีปลอกคอมากกว่าสีปลอกคออื่นๆ ทำให้โมเดลรู้จำข้อมูลที่มีจำนวนมากกว่าและทำนายผลลัพธ์ได้ไม่ดีพอ

5.2.2. ค่าคะแนน และ threshold อาจยังไม่เหมาะสม ทำให้ข้อมูลที่ต้องการไม่แสดงผล

5.2.3. ทรัพยากรบนเซิร์ฟเวอร์มีจำกัด ทำให้ต้องใช้เซิร์ฟเวอร์ 3 เซิร์ฟในการทำงานด้านการประมวลผลภาพ อัพโหลดและรับส่งข้อมูลภาพรวมทั้งตำแหน่ง GPS ออกมาใช้งานและจัดเก็บข้อมูลผลการทำนายภาพ

Computer Science Department
Faculty of Informatics, Mahasarakham University

5.2.4. เนื่องจากภาพถ่ายที่ถ่ายจากมือถือมีขนาดใหญ่ จำเป็นต้องใช้ทรัพยากรในการประมวลผลบนเซิร์ฟเวอร์อย่างมาก จึงมีปัญหาในการส่งภาพไปทำนายบนเซิร์ฟเวอร์

5.3. ข้อเสนอแนะ

5.3.1. ควรเพิ่มข้อมูลของลักษณะหูหาง และปลอกคอ ในชุดข้อมูลสำหรับการเรียนรู้ให้มีสัดส่วนใกล้เคียงกันมากขึ้น

5.3.2. ปรับค่า threshold ความถูกต้อง จากช่วงของค่า threshold เพื่อปรับเกณฑ์การให้คะแนนแต่ละรายการข้อมูล

5.3.3. ควรพัฒนาส่วน Application ในการปรับลดขนาดภาพก่อนส่งไปประมวลผลผ่านเซิร์ฟเวอร์

เอกสารอ้างอิง

1. พิมพา ชีวาประกอบกิจ. (2019). การปรับปรุงประสิทธิภาพในการจำแนกภาพด้วยโครงข่ายประสาท แบบคอนโวลูชันโดยใช้เทคนิคการเพิ่มภาพ. Retrieved 7 July 2021 from 2019 TNI Journal ENG.indd - ThaiJO
2. Machine Learning Tutorial. (2020). Convolutional Neural Networks. Retrieved 7 August 2021 from https://sci2lab.github.io/ml_tutorial/cnn/

5 International Conference on Intelligent Technologies (CONIT). (2021). Dog Breed Classification Using Deep Learning. Retrieved 1 July 2022 from https://www.researchgate.net/publication/353693466_Dog_Breed_Classification_Using_Deep_Learning

6 creativecommons. (2019). Optimized Deep Convolutional Neural Networks for Identification of Macular Diseases from Optical Coherence Tomography Images. Retrieved 1 July 2022 from https://www.researchgate.net/publication/331364877_Optimized_Deep_Convolutional_Neural_Networks_for_Identification_of_Macular_Diseases_from_Optical_Coherence_Tomography_Images

7. Tiago Carvalho. (2019). Deep Feature-Based Classifiers for Fruit Fly Identification (Diptera: Tephritidae). Retrieved 1 July 2022 from https://www.researchgate.net/publication/330478807_Deep_Feature-Based_Classifiers_for_Fruit_Fly_Identification _Diptera_Tephritidae
8. Saeed Anwar, et al. (2020). A Systematic Evaluation: Fine Grained CNN vs. Traditional CNN Classifiers. Retrieved 1 July 2022 from https://deepai.org/publication/a-systematic-evaluation-finegrained-cnn-vs-traditional-cnn-classifiers

Computer Science Department Faculty of Informatics, Mahasarakham University

12. Saugata Paul. (2019). A detailed case study on Multi-Label Classification with Machine Learning algorithms and predicting movie tags based on plot summaries!. Retrieved 7 August 2021fromhttps://medium.com/@saugata.paul1010/a-detailed-case-study-on-multi-label-classification-with-machine-learning-algorithms-and-72031742c9aa

Computer Science Department
Faculty of Informatics, Mahasarakham University