

การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก
ในการสำรวจแหล่งที่อยู่อาศัย
ด้วยภาพถ่ายจาก UAV
ตลอดแนวท่อส่งก๊าซธรรมชาติ

APPLIED DEEP LEARNING TECHNIQUE
FOR HOUSE DETECTION FROM UAV IMAGES IN GAS
TRANSMISSION PIPELINE AREA

พากย์สิริ ตรียุทธกุล (Parksiri Treeyuthakul)

อัจฉรา ภู่อ่าง (Ajchara Phu-ang)



การประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกในการสำรวจแหล่งที่อยู่อาศัย ด้วยภาพถ่ายจาก UAV ตลอดแนวท่อส่งก๊าซธรรมชาติ

Applied Deep Learning Technique For House Detection From UAV Images In Gas Transmission Pipeline Area

พาศย์สิริ ตริยยุทธกุล (Parksiri Treeyuthakul)

อจจรา ภู่อ่าง (Ajchara Phu-ang)

วิทยาลัยนวัตกรรมการ มหาวิทยาลัยธรรมศาสตร์

E-mail: au_wassarp@hotmail.com

บทคัดย่อ

งานวิจัยนี้จึงมีแนวคิดที่จะศึกษาการประยุกต์ใช้เทคโนโลยีอากาศยานไร้คนขับ (UAV) เทคโนโลยีการเรียนรู้เชิงลึก (Deep Learning) และเทคนิค Support Vector Machine ในการสำรวจชุมชนแนวท่อส่งก๊าซธรรมชาติ โดยใช้ข้อมูลรูปภาพจากเอกสารของบริษัท ปตท. จำกัด(มหาชน) ซึ่งนำมาผ่านการ Data Selection Data Augmentation ด้วยวิธีการติกรอบ (Crop) การพลิกรูปในแนวแกนนอน (Horizontal Flip) และการหมุนรูป (Rotation) จากนั้นนำข้อมูลรูปภาพทั้งหมดแบ่งเป็นข้อมูลสำหรับการฝึก (Training Dataset) จำนวน 80% และข้อมูลสำหรับการทดสอบ (Testing Dataset) จำนวน 20% นำเข้าประมวลผลกับโมเดล Convolutional Neural Network (CNN) และโมเดล Support Vector Machine (SVM) โดยข้อมูลที่เข้าประมวลผลในโมเดล SVM จะผ่านขั้นตอนการทำ Histogram of Oriented Gradients (HOG) เพื่อทำการสกัดคุณลักษณะ (Feature Extraction) จากนั้นทำการปรับค่าตัวแปรต่างๆ ของทั้ง 2 โมเดลเพื่อให้ได้โมเดลที่เหมาะสมที่สุด

Abstract

This independent study is research how to apply between the UAV technology, deep learning technology and support vector machine (SVM) technical to detect building in image. The image data source is come from PTT Public Company Limited. In data preparation process will have sub process as data selection process, data augmentation process by crop, horizontal flips, rotation. The total data were split to 80% for training and 20% of testing. For model SVM, the technic for feature extraction step is Histogram of Oriented Gradient (HOG). For both model will vary value of many parameters to find the best result of its.

1. บทนำ

ท่อส่งก๊าซธรรมชาติมีความสำคัญอย่างมากในด้านพลังงานของประเทศโดยเฉพาะในช่วงเวลาปัจจุบันที่มีการปรับเปลี่ยนการใช้พลังงานเป็นพลังงานสะอาดมากขึ้นจากกรณีสถานะโลกร้อนที่เกิดทุกแห่ง ทำให้ก๊าซธรรมชาติที่ถือว่าเป็นพลังงานสะอาดนั้นมีปริมาณและความต้องการการใช้เพิ่มขึ้นอย่างมากและมีแนวโน้มที่จะเพิ่มขึ้นอย่างต่อเนื่อง โดยการขนส่งก๊าซธรรมชาตินั้นจะการใช้การขนส่งทางท่อเป็นหลัก การออกแบบและบำรุงรักษาระบบท่อจึงเป็นสิ่งที่สำคัญมากและต้องเป็นไปตามมาตรฐานสากลซึ่งในมาตรฐาน ASME B31.8 นั้นตัวแปรที่สำคัญไม่ว่าจะเป็นเรื่องการออกแบบแรงดันท่อตันหรือขั้นตอนการทดสอบท่อจะขึ้นกับค่าตัวแปร Location Class โดยค่าตัวแปร Location Class จะกำหนดมาจากจำนวนอาคารที่อยู่อาศัยในระยะ 1 ไมล์นับจากแนวท่อส่งก๊าซธรรมชาติ ดังนั้นจึงจะต้องมีการสำรวจชุมชนตลอดแนวท่อเป็นประจำเพื่อกำหนดแรงดันสูงสุดของท่อส่งก๊าซ

ปัจจุบันเทคโนโลยีอากาศยานไร้คนขับ (UAV) ได้รับการพัฒนาอย่างต่อเนื่องให้มีการใช้งานได้อย่างยาวนานมีประสิทธิภาพการควบคุมที่ดี มีระยะควบคุมที่ไกลและสามารถติดตั้งกล้องความละเอียดสูงได้ และเทคโนโลยีการเรียนรู้เชิงลึก (Deep Learning) ได้ถูกพัฒนาและมีการนำมาใช้งานอย่างแพร่หลายโดยเฉพาะเทคนิค Convolution Neural Networks (CNN) และเทคนิค Support Vector Machine (SVM) ที่ผ่านการสกัดคุณลักษณะ (Feature Extraction) ด้วยเทคนิค Histogram of Oriented Gradient (HOG) เป็นเทคนิคที่ใช้กับการตรวจจับวัตถุผ่านรูปภาพหรือวิดีโอ งานวิจัยนี้จึงมีแนวคิดที่จะศึกษาการประยุกต์ใช้เทคโนโลยีทั้ง 3 อย่างที่ได้กล่าวมานั้นมาใช้ในการสำรวจชุมชนแนวท่อส่งก๊าซธรรมชาติเพื่อเพิ่มความรวดเร็ว ความแม่นยำและเพิ่มประสิทธิภาพในการบริหารบุคลากร

2. แนวคิด ทฤษฎี วรรณกรรม

2.1 ความรู้เกี่ยวกับมาตรฐาน ASME B 31.8

มาตรฐาน ASME B31.8 มีชื่อว่า Gas Transmission and Distribution Piping System เป็นมาตรฐานที่ระบุถึงระบบท่อส่งก๊าซธรรมชาติทั้งด้านการออกแบบและการบำรุงรักษา เช่น วัสดุที่ใช้ในระบบท่อส่งก๊าซ ขั้นตอนการออกแบบ การตรวจสอบท่อส่งก๊าซ การผลิตท่อส่งก๊าซ การใช้งานและการดูแลรักษาท่อส่งก๊าซ เป็นต้น ซึ่งหัวข้อที่เกี่ยวข้องกับงานวิจัยนี้จะมีหัวข้อการประเมิน Location Class จะแบ่งเป็น 4 ระดับหลักๆ คือ

1) Location Class 1 คือในระยะ 1 ไมล์ นับการแนวท่อส่งก๊าซจะมีอาคารที่มนุษย์อาศัยอยู่ไม่เกิน 10 อาคาร ซึ่งจะแบ่งย่อยออกเป็น 2 ประเภท ซึ่งส่งผลต่อค่า Design Factor และค่า Test Maximum Operating Pressure ที่ต่างกันพื้นที่ที่มาตรฐานยกตัวอย่างเช่น พื้นที่ภูเขา พื้นที่บริเวณเกษตรและพื้นที่ที่รกร้าง เป็นต้น

2) Location Class 2 คือในระยะ 1 ไมล์ นับการแนวท่อส่งก๊าซ จะมีอาคารที่มนุษย์อาศัยอยู่ตั้งแต่ 10 - 46 อาคารพื้นที่ที่มาตรฐานยกตัวอย่างเช่น พื้นที่โรงงานอุตสาหกรรมและพื้นที่ชานเมือง เป็นต้น

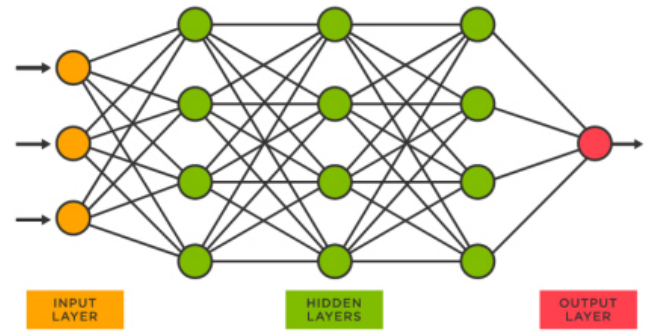
3) Location Class 3 คือในระยะ 1 ไมล์ นับการแนวท่อส่งก๊าซ จะมีอาคารที่มนุษย์อาศัยอยู่มากกว่า 46 อาคาร แต่ไม่ใช่พื้นที่ที่ถูกนิยามใน Location Class 4 พื้นที่ที่มาตรฐานยกตัวอย่างเช่น แหล่งที่อยู่อาศัย พื้นที่โรงงานอุตสาหกรรมและพื้นที่ที่มีห้างสรรพสินค้า เป็นต้น

4) Location Class 4 คือบริเวณที่มีอาคารสูงตั้งแต่ 4 ชั้นขึ้นไปจำนวนมาก มีการจราจรที่แออัดและมีสาธารณูปโภคที่ติดตั้งได้ดินหลากหลาย

2.2 ความรู้การเรียนรู้เชิงลึก (Deep Learning)

เป็นรูปแบบหนึ่งของ Machine Learning ซึ่งมีโครงสร้างลักษณะแบบโครงข่ายประสาทเทียม (Neural Network) ที่ประกอบไปด้วยชั้น Layer อย่างน้อยจำนวน 3 Layers คือชั้น Input Layer ทำหน้าที่รับข้อมูล Input เข้ามาและส่งข้อมูลต่อไปที่ Layer ชั้นถัดไป, Hidden Layer ทำหน้าที่รับข้อมูลจาก Input Layer มาประมวลผลและส่งค่าไปยัง Layer ชั้นถัดไปโดย Hidden Layer สามารถมีได้หลากหลายชั้น Layer แต่ละ Layer จะมีค่าที่ใช้ประมวลผลเช่น ค่าถ่วงน้ำหนัก (Weight) ค่าความเอนเอียง (Bias) และวิธีการประมวลผลทางคณิตศาสตร์ (Activation Function) ที่แตกต่างกันออกไปและเป็นอิสระต่อกันในชั้นต่อมาคือ Output Layer ทำหน้าที่รับค่าจาก Hidden Layer แล้วไปแสดงผลค่าผลลัพธ์ซึ่งโครงข่ายประสาทเทียมนี้เป็นการจำลองรูปแบบการทำงานของระบบประสาทมนุษย์

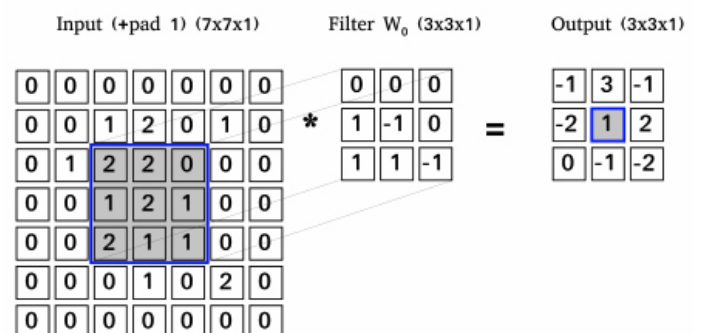
โดย Deep Learning จะมีการเรียนรู้และปรับเปลี่ยนค่าต่างๆ เพื่อให้ผลลัพธ์ที่ได้มีความถูกต้องมากขึ้น



ภาพที่ 1 แบบจำลอง Model Neural Network

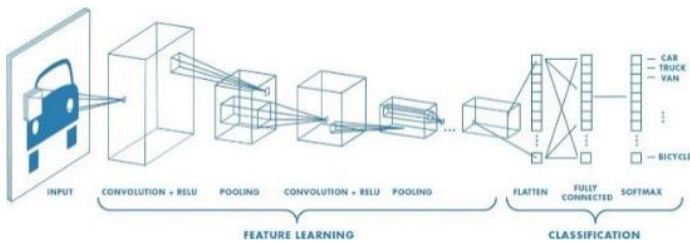
2.2.2 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network, CNN)

เป็นเทคนิคการเรียนรู้เชิงลึกประเภทหนึ่ง ที่มีชั้น Convolution Layer เพื่อสกัดคุณลักษณะ (Feature Extraction) จากรูปภาพข้อมูล Input โดยใช้ Filter ในการใช้สกัดคุณลักษณะ โดย Filter คือขนาดพื้นที่ย่อยๆ ที่กำหนดขึ้นจากนั้นนำ Filter เคลื่อนที่เรียงตามข้อมูลรูปภาพเพื่อทำการ Convolution โดยขนาดหรือระยะการเคลื่อนที่เรียกว่า Stride ผลที่ได้จาก Convolution คือชุดข้อมูลที่เรียกว่า Feature Maps ซึ่งจะมีขนาดน้อยกว่าขนาดของภาพ Input ยิ่ง Stride มีค่ามากขึ้นจะทำให้ Feature Map ยังมีขนาดเล็กลงแต่ทำให้การคำนวณนั้นมีพื้นที่ทับซ้อนกันน้อยลง หากต้องการให้ขนาดของ Feature Map ไม่ลดลงจะมีเทคนิคที่ใช้เพิ่มเติมเรียกว่าการ Zero Padding และจะมีเทคนิคการนำฟังก์ชันกระตุ้น (Activation Function) ที่เรียกว่า ReLu มาใช้เพื่อปรับค่าให้กับ Feature Map มีความซับซ้อนน้อยลงคือให้คงเหลือแต่ค่าที่มากกว่า 0 ส่วนค่าที่น้อยกว่า 0 นั้นจะถูกตัดทิ้งจะทำให้ได้ Feature Map ที่นำไปประมวลผลโมเดลได้รวดเร็วยิ่งขึ้น



ภาพที่ 2 ตัวอย่างวิธีการทำ Padding และการคำนวณ Feature Map

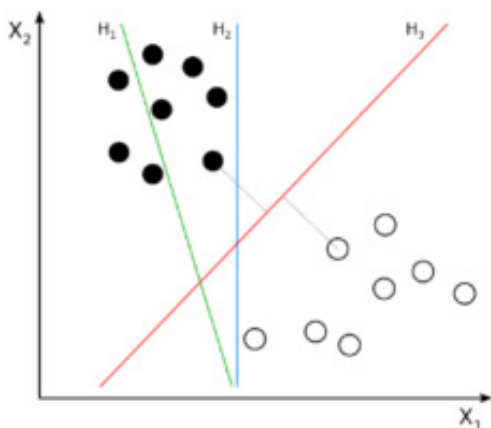
CNN จะมีการใช้ชั้น Pooling Layer มาวางต่อจากชั้น Convolution Layer เพื่อลดขนาดของ Feature Map ลงโดยยังคุณลักษณะเด่นของ Feature Map นั้นๆ ไว้เพื่อส่งต่อไปให้กับ Layer ในชั้นถัดไปนำไปวิเคราะห์ได้ง่ายขึ้น การ Pooling นั้นจะต้องกำหนดขนาดของ Pooling และค่า Stride ที่ใช้เลื่อน Pooling การ Pooling มีวิธีการอยู่หลักๆ อยู่ 2 วิธีคือ Max pooling คือการนำค่าที่สูงที่สุดที่ปรากฏใน Pool ออกมาเป็นค่าตัวแทนและ Average Pooling คือการนำค่าใน Pool มาหาค่าเฉลี่ยเพื่อเป็นตัวแทนของ Pool ชั้นสุดท้าย Fully Connected Layer เป็นช่วงที่รับข้อมูลหลังจากทำการสกัดคุณลักษณะแล้วเสร็จจะเป็นช่วงที่ทำการจำแนกประเภท (Classification) ซึ่งจะประกอบไปด้วยชั้น Hidden Layer และชั้น Output Layer โดยจำนวน Node ของ Output Layer จะให้เท่ากับจำนวน Label ที่ต้องการจะตรวจจับ



ภาพที่ 3 แบบจำลอง Model CNN

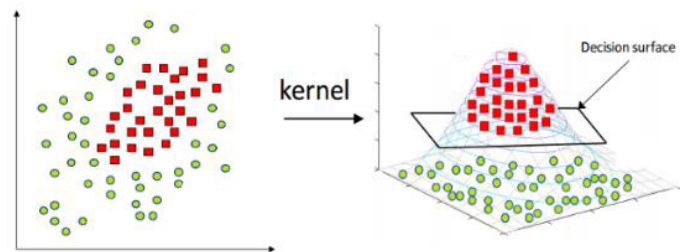
2.2.3 Support Vector Machines (SVM)

เป็น Machine Learning แบบ Supervised Learning กล่าวคือเป็นโมเดลที่จะต้องมีการฝึกฝนก่อนนำไปใช้งาน โดยการทำงานของโมเดลคือการหาเส้น Hyperplane เพื่อแบ่งข้อมูลแต่ละประเภท (Class) ออกจากกัน โดย Model ที่ดีจะสร้างเส้น Hyperplane ที่แบ่งประเภทข้อมูลออกจากกันแล้วผลรวมระยะห่างของเส้นแบ่งกับข้อมูลที่ใกล้เส้นที่สุดจะมีค่ามากที่สุด (Maximum Margin) เนื่องจากหากมีข้อมูลอื่นๆ มาผ่านประมวลผลในโมเดลจะทำให้โอกาสที่จะเกิดข้อผิดพลาดนั้นน้อยลง คือลดการ Overfitting



ภาพที่ 4 ตัวอย่างเส้น Hyperplane จาก Model ทั้ง 3 Model

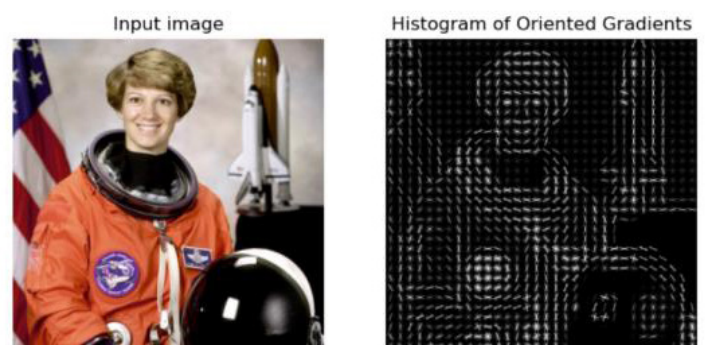
ซึ่งโดยปกติเส้น Hyperplane จะเป็นเส้นตรง (Linear) ทำให้มีข้อจำกัดในการสร้างโมเดลในกรณีที่มีข้อมูลมีความซับซ้อนมาก จึงได้นำ Kernel Function มาใช้ในการปรับเปลี่ยนข้อมูลให้มีมิติที่เพิ่มมากขึ้นเพื่อจะสามารถสร้างเส้น Hyperplane ที่เหมาะสมในการแบ่งประเภทกับข้อมูลที่มีความซับซ้อนได้



ภาพที่ 5 ตัวอย่างการทำงานของ Kernel Function

2.2.4 Histogram of Oriented Gradients (HOG)

เป็นเทคนิคที่ใช้ในการทำ Image Processing โดยมีจุดประสงค์เพื่อสกัดคุณลักษณะ (Feature Extraction) ซึ่งนำมาใช้ในการตรวจจับวัตถุ หลักการทำงานคือการหารูปทรงของวัตถุในรูปจากความหนาแน่นของ Gradients โดยการแบ่งรูปออกเป็น Cells ย่อยๆ ในแต่ละ Cell นั้นจะมีจำนวน Pixels ตามที่กำหนด จากนั้นคำนวณหา Gradient ในแต่ละ Cell หลังจากได้ค่า Gradient มาแล้วจะนำไปคำนวณค่า Magnitude และ ค่า มุมองศา (Angle) จากนั้นจะจัด Block โดยแต่ละ Block จะประกอบไปด้วยจำนวน Cell ที่กำหนดเพื่อนำค่าในแต่ละ Block มาคำนวณ Histogram แบ่งตามความถี่ของจำนวน Bins ต่างๆ ตามที่กำหนดสิ่งที่ได้จะเป็น Feature Vector



ภาพที่ 6 ตัวอย่างการทำ HOG

2.2.5 กระบวนการวิเคราะห์ข้อมูลวิธี CRISP-DM (Cross-Industry Standard For Data Mining)

กระบวนการ CRISP-DM เป็นวิธีการทำเหมืองข้อมูลเพื่อใช้ในการวิเคราะห์ข้อมูลขนาดใหญ่เพื่อแก้ไขปัญหาหรือสร้างโอกาสทางธุรกิจใหม่ๆ ในการตอบโต้ธุรกิจนั้นๆ โดยจะประกอบไปด้วยขั้นตอนทั้งหมด 6 ขั้นตอนดังนี้

1) Business Understanding เป็นขั้นตอนแรก ซึ่งจะเน้นไปทางการเข้าใจปัญหาของธุรกิจหรือโอกาสทางธุรกิจเพื่อจะได้กำหนดวัตถุประสงค์ของการจัดทำเหมืองข้อมูล การวางแผนและกรอบการทำงานในแต่ละขั้นตอนต่อไป

2) Data Understanding เป็นขั้นตอนการรวบรวมข้อมูลที่เกี่ยวข้องในการจัดทำเหมืองข้อมูลตามวัตถุประสงค์จากขั้นตอน Business Understanding และดำเนินการพิจารณาข้อมูลว่ามีความน่าเชื่อถือ เหมาะสม และเพียงพอต่อการนำไปวิเคราะห์ในขั้นตอนต่อไปหรือไม่

3) Data Preparation เป็นขั้นตอนการแปลงข้อมูลที่ได้รวบรวมมาจากกระบวนการก่อนหน้านี้ซึ่งอยู่ในรูปแบบของข้อมูลดิบ (Raw Data) มาดำเนินการแปลงให้เป็นข้อมูลที่เหมาะสมในการวิเคราะห์

4) Modeling เป็นการกำหนดเทคนิคและดำเนินการทำเหมืองข้อมูลจากข้อมูลที่ได้รับจากขั้นตอนก่อนหน้านี้เพื่อตอบโจทย์วัตถุประสงค์ โดยจะมีการปรับ Parameter ต่างๆ ให้เหมาะสมกับข้อมูลที่ใช้และสามารถใช้หลากหลายเทคนิคในการทำเหมืองข้อมูลได้

5) Evaluation ขั้นตอนการประเมินหรือวัดประสิทธิภาพของโมเดลที่ใช้วิเคราะห์ข้อมูลซึ่งจะส่งผลถึงการปรับ Parameter ต่างๆ ในช่วง Modeling และช่วงการทำ Data Preparation โดยในการประเมินการทำ Classification ก็จะมีค่าที่นิยมนำมาใช้ดังนี้

- ค่าความถูกต้อง (Accuracy) เป็นค่าความถูกต้องแม่นยำของโมเดลโดยนิยามแสดงในรูปของ Percentage

- ค่า Precision เป็นการประเมินทางด้านฝั่ง Positive ว่ามีความแม่นยำมากน้อยเท่าไรโดยยังมีค่าที่สูงยิ่งดี ซึ่งค่าที่มากที่สุดคือ 1 โดยคำนวณจากสมการดังนี้

- ค่า Recall เป็นการประเมินความไวของโมเดลยังมีค่าที่สูงยิ่งดีโดยค่ามากที่สุดคือ 1

- ค่า F1 Score เป็นการเฉลี่ยแบบ Harmonic Mean ระหว่างค่า Precision กับค่า Recall

โดยตัวแปรทั้งหมดที่กล่าวมานั้นมาจากค่าใน Confusion Matrix ดังตารางที่ 1

ตารางที่ 1 Confusion Matrix

		Predicted Values	
		Positive	Negative
Actual Values	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

6) Deployment ขั้นตอนการนำโมเดลที่มีประสิทธิภาพดีที่สุดไปใช้เพื่อตอบโจทย์วัตถุประสงค์



ภาพที่ 7 กระบวนการวิเคราะห์ข้อมูล CRISP-DM

3. วิธีวิจัย

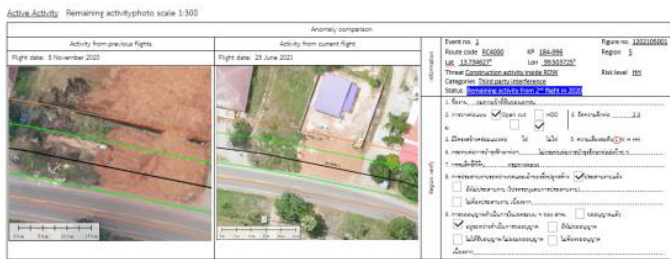
3.1 Business Understanding

ในการสำรวจชุมชนบริเวณท้องฟ้าชั้นปัจจุบันจะใช้วิธีขั้วรถตามแนวท้องฟ้าทำให้เกิดปัญหาเรื่องบุคลากรและเวลาจำนวนมากอีกทั้งหลายพื้นที่ที่ไม่สามารถนำพาหะเข้าถึงได้จะต้องใช้การเดินสำรวจแทนจึงทำให้อาจจะมีการสำรวจที่ไม่ทั่วถึงทั้งหมดประกอบกับการขยายตัวของชุมชนที่รวดเร็วขึ้นในปัจจุบันทำให้การสำรวจจึงมีความสำคัญมากขึ้น จึงได้นำเอาประเด็นความสำคัญและปัญหาดังกล่าวมาทำการวิจัยโดยจะนำเอาข้อมูลภาพถ่ายจากเทคโนโลยีโดรนมาทำการประยุกต์ใช้ในการประเมิน Location Class โดยประยุกต์ร่วมกับการเรียนรู้เชิงลึก (Deep Learning) เพื่อตรวจจับที่ผิดปกติจากข้อมูลภาพถ่าย

3.2 Data Understanding

ในขั้นตอนนี้ข้อมูลรูปภาพจะมาจากเอกสารรายงานผลการปฏิบัติงาน การจัดทำภาพถ่ายทางอากาศเพื่อประเมินการเปลี่ยนแปลงของพื้นที่แนวท้องฟ้าอากาศยานไร้คนขับ (UAV) ซึ่งมีทั้งหมด 10 รายงานแบ่งเป็นตามพื้นที่ศูนย์ปฏิบัติการของระบบท้องฟ้าธรรมชาติของ บริษัท ปตท. จำกัด (มหาชน) ทั้งหมด 10 เขต ทั้งหมดจะครอบคลุม 18 จังหวัดประกอบไปด้วย ชลบุรี, ระยอง, อุดรธานี, สระบุรี, ระยอง, ราชบุรี, นครปฐม, กทม., นนทบุรี, กาญจนบุรี, ปทุมธานี, นครนายก, ปราจีนบุรี, อ่างทอง, สิงห์บุรี, ชัยนาท, นครสวรรค์ และ นครราชสีมา โดยรูปภาพที่ได้จากการบิน UAV จะนำมาผ่านกระบวนการ Data Preparation ให้เป็นภาพถ่ายแบบออร์โธ (ภาพถ่ายทางอากาศซึ่งผ่านกระบวนการปรับแก้ความผิดเพี้ยนเนื่องจาก

เรขาคณิตของการถ่ายภาพ และความสูงต่างที่กันของภูมิประเทศ (Relief Displacement)) และนำภาพมารวมกันออกเป็นภาพ Orthomosaic ก่อนจัดทำออกมาในรูปแบบเอกสารรายงาน โดยการตีกรอบบริเวณที่สนใจมาเสนอในเอกสารรายงาน ทำให้ขนาดรูปภาพที่ได้มาจะมีขนาดที่ไม่เท่ากันในแต่ละรูป ซึ่งรูปภาพข้อมูลที่จะนำมาจัดทำต่อในงานวิจัยนี้จะทำการ Data Selection โดยการ Save มาจากรายงานแต่ละฉบับที่มีเฉพาะรูปภาพที่ปักอาศัยอยู่ในรูปภาพของรายงานปี 2564 จะได้รูปทั้งหมด 119 รูป โดยรูปทั้งหมดเป็นไฟล์นามสกุล JPG



ภาพที่ 8 ตัวอย่างรายละเอียดในเอกสารรายงาน

3.3 Data Preparation

ในขั้นตอนการเตรียมข้อมูลนี้ จากรูปทั้งหมด 119 รูป จะทำการตีกรอบ (Crop) บริเวณที่สนใจคือบริเวณที่เป็นที่ปักอาศัยออกมาเป็นรูปย่อยๆของแต่ละอาคารเพื่อเพิ่มปริมาณข้อมูลและตัดข้อมูลที่ไม่จำเป็นออกจะทำให้ข้อมูลมีรายละเอียดเฉพาะในส่วนที่ต้องการใช้สำหรับการฝึกฝนและทดสอบโมเดลเพื่อตรวจจับเท่านั้น โดยชุดข้อมูลรูปภาพจะมีรายละเอียดของที่อยู่อาศัยของแต่ละรูปเป็น 1 อาคารที่ปักอาศัยต่อ 1 รูป



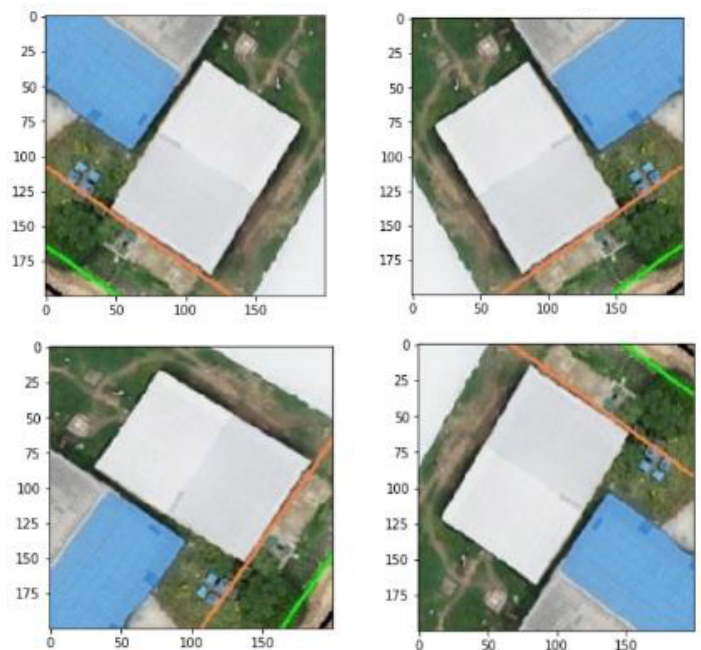
ภาพที่ 9 ตัวอย่างข้อมูลรูปภาพที่ทำการตีกรอบบริเวณที่อยู่อาศัย

ส่งผลให้ได้จะได้รูปทั้งหมด 332 รูปสำหรับรูปที่ปักอาศัยและทำการตีกรอบรูปภาพอื่นๆ ที่ไม่ใช่รูปที่อยู่อาศัยอีก 200 รูปเพื่อเป็นอีก 1 Label รวมเป็นทั้งหมด 532 รูป ซึ่งข้อมูลแต่ละรูปจะมีขนาด (Dimension) ที่แตกต่างกันจึงได้ทำการปรับให้ขนาดของภาพเท่ากันทั้งหมดเพื่อให้สามารถนำเข้าประมวลผลในโมเดลได้และเมื่อทำการตรวจสอบค่าสีพบว่าข้อมูลบางรูปมีค่าสีเป็น RGBA ปนอยู่กับรูปที่มีค่าสีเป็น RGB จึงได้ทำการเปลี่ยนให้เป็น RGB ทั้งหมด

(200, 200, 3)
(200, 200, 3)
(200, 200, 3)
(200, 200, 3)
(200, 200, 3)
(200, 200, 3)
(200, 200, 4)
(200, 200, 4)
(200, 200, 4)
(200, 200, 4)
(200, 200, 4)

ภาพที่ 10 ตัวอย่างการตรวจสอบ Shape ของข้อมูล

เนื่องจากข้อมูลรูปยังมีจำนวนที่น้อยอยู่อาจจะมีผลต่อการฝึกและทดสอบโมเดลได้ซึ่งมีโอกาสที่จะทำให้อาจได้ค่าความแม่นยำ (Accuracy) ที่ต่ำหรือเกิดการ Over Fit ขึ้น จึงดำเนินการจัดทำ Data Augmentation เพื่อเพิ่มจำนวนข้อมูลรูปโดยการพลิก (Flip) ในแนวนอน (Horizontal) ทุกรูปและนำแต่ละรูปมาทำการหมุน (Rotation) ด้วยมุมมอง 90, 180, 270 ทุกรูปทำให้จะได้ข้อมูลรูปทั้งหมด 2,660 รูป



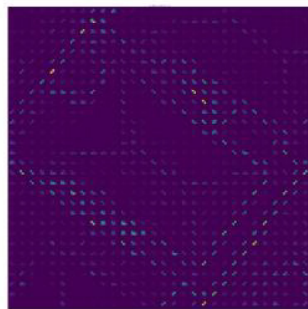
ภาพที่ 11 ตัวอย่างการทำ Data Augmentation

เมื่อทำการตรวจสอบข้อมูลรูปภาพนามสกุล JPG ซึ่งโดยปกติข้อมูลรูปภาพ RGB จะประกอบไปด้วยชั้น สีแดง (Red) สีเขียว (Green) และสีน้ำเงิน (Blue) อย่างละ 1 Matrix และค่าที่แสดงจะเป็นค่าตัวเลขสีจาก 0-255 แต่เมื่อผ่าน Function ที่ทำการปรับขนาด (Resize) จะส่งผลให้ข้อมูลจะถูกแปลงจาก ชนิดข้อมูล uint8 เป็นชนิด Float โดยอัตโนมัติ โดยส่งผลให้ ค่าตัวเลขจะถูกแปลงจากค่า 0-255 เป็นค่า 0-1

uint8
float64
float64

ภาพที่ 12 ตัวอย่าง Data Type ที่เปลี่ยนไป

ในขั้นตอนถัดมาจะทำการกำหนด Label ของข้อมูล รูปภาพอาคารที่พักอาศัยจาก House เป็นเลข 0 และข้อมูลรูปภาพ อื่นๆ จาก Other เป็น 1 เพื่อใช้ในการฝึกและทดสอบโมเดล ทั้งนี้ข้อมูลรูปภาพทั้งหมดจะแบ่งข้อมูลออกเป็น 80% สำหรับการฝึกและ 20% สำหรับการทดสอบในส่วนด้านการทำ Data Preprocessing ของเทคนิค SVM จะมีขั้นตอนเพิ่มขึ้นมาจาก ที่กล่าวมาในข้างต้นคือการทำ Histogram of Oriented Gradients (HOG) โดยจะมีการปรับตัวแปร : Pixel/Cell, Cells/Block และ จำนวน Bin Orientation เพื่อจะทำการสกัดคุณลักษณะ (Feature Extraction) ออกมาให้ได้ผลลัพธ์ที่เหมาะสมสำหรับ ใช้ในการทำโมเดลขั้นตอนถัดไป และเนื่องจากเมื่อใช้เทคนิค การทำ HOG ผลลัพธ์ที่ได้จะมีลักษณะที่เป็นรูปทรงเพื่อให้่าย และรวดเร็วในการทำจึงได้มีการเปลี่ยนรูปจากค่าสี RGB เป็นค่าสี Grayscale เพื่อลดความซับซ้อนของข้อมูลรูป



ภาพที่ 13 ตัวอย่างการทำ HOG กับ Data

3.4 Modeling

3.4.1 Deep Learning Model

เทคนิคที่จะนำมาประยุกต์ใช้ในการวิจัยนี้คือ การเรียนรู้เชิงลึก (Deep Learning) โดยจะใช้เทคนิคที่เรียกว่า โครงข่ายประสาทเทียมแบบ Convolutional Neural Network (CNN) ซึ่งขั้นตอนการสร้างโมเดลจะแบ่งเป็นออกเป็น 2 ช่วงดังนี้

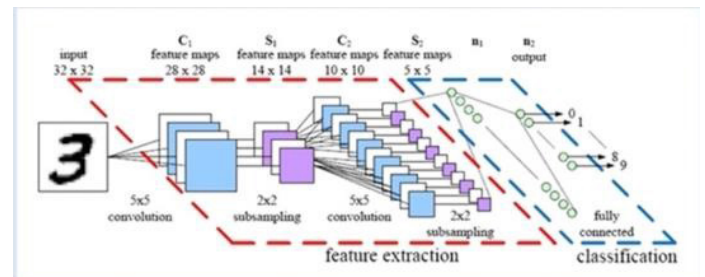
3.4.1.1 ช่วงการสกัดคุณลักษณะ (Feature Extraction)

จะประกอบไปด้วย

1. ชั้น Convolution Layer โดยขั้นตอนนี้จะมีการทดลองปรับเปลี่ยนจำนวนชั้น, ขนาดของ Kernel, ขนาดของ Stride และ Activation Function Relu เป็นต้น เพื่อให้โมเดลได้ค่าแม่นยำที่สุด
2. ชั้น Pooling Layer โดยขั้นตอนนี้จะมีการทดลองปรับการใช้ค่าสูงสุดหรือค่าเฉลี่ยในการคำนวณและค่า Pool size รวมถึงลำดับชั้นและจำนวนชั้นของ Pooling เป็นต้น เพื่อให้ได้ค่าความแม่นยำสูงสุด
3. ชั้น Flatten Layer จะมีชั้นเดียวเป็นชั้นสุดท้ายของช่วงการสกัดคุณลักษณะเพื่อเตรียมข้อมูลไว้ส่งต่อไปยังช่วงการจำแนกประเภทต่อไป

3.4.1.2 ช่วงการจำแนกประเภท (Classification) (Fully Connected) จะประกอบไปด้วย

1. Hidden Layer โดยขั้นตอนนี้จะมีการทดลองปรับจำนวนของ Node และจำนวนชั้นของ Hidden Layer เป็นต้น เพื่อให้ได้ค่าความแม่นยำที่สุด
2. Output Layer โดยในชั้นนี้จะมีจำนวน Node 2 Node เนื่องจาก Label มี 2 Label



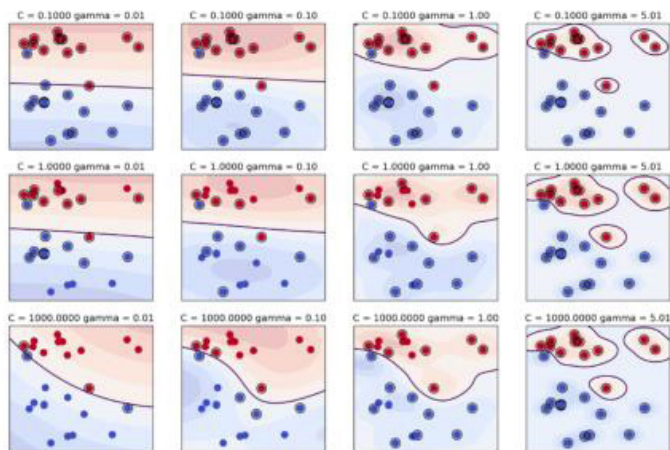
ภาพที่ 14 ภาพตัวอย่างการทำ Architecture Model CNN

จะมีการใช้ Optimizer ที่ชื่อว่า Adaptive Moment Estimation (Adam) ของ Tensorflow ในการปรับค่า Loss และ Bias ซึ่ง Loss Function จะมีการทดลองใช้ 2 Function ที่นิยมในงานการจำแนกประเภท (Classification) คือ 1. Binary Cross Entropy 2. Categorical Cross Entropy เพื่อเปรียบเทียบหาผลลัพธ์ที่ดีที่สุด นอกจากนี้จะมีการปรับค่า Learning Rate เพื่อหาค่าที่เหมาะสมที่สุดรวมถึงจะมีการปรับรอบการฝึกฝนโมเดลเพื่อหาจำนวนรอบที่เหมาะสมโดยจะแบ่งชุดข้อมูลสำหรับการฝึกออกอีก 20 % ให้เป็นข้อมูลที่ใช้ทดสอบในแต่ละรอบ (Validation Data) เพื่อใช้ในการประเมินโมเดลในแต่ละรอบ



3.4.2 SVM Model

เทคนิคที่ 2 ที่จะนำมาประยุกต์ใช้ในการวิจัยนี้คือ Support Vector Machine (SVM) โดยจะทำการเขียน Code ด้วยภาษา Python จะใช้ Kernel Function 'rbf' (Radial Basis Function) ที่เป็น Default ของ Library Skimage เนื่องจากข้อมูลเราเป็นข้อมูลรูปภาพมีความซับซ้อนและยังไม่มีคุณลักษณะของข้อมูลที่ชัดเจนจากนั้นจะทำการปรับค่าตัวแปร (Parameter) ต่างๆ ของโมเดล SVM เช่น ค่า C และค่า Gamma โดยที่ค่า C เมื่อมีค่าที่มากขึ้นจะส่งผลให้โมเดลนั้นมีค่าโอกาสเกิดข้อผิดพลาดที่น้อยลงแต่อาจจะส่งผลให้เกิดการ Overfit ขึ้นได้ หรือทำให้รับชุดข้อมูลในอนาคตที่มีความซับซ้อนเพิ่มขึ้นไม่ได้ ส่วนค่า Gamma เมื่อมีค่าสูงขึ้นจะทำให้ระยะที่วาดออกจากจุดข้อมูลแต่ละจุดมีขนาดเล็กลงซึ่งส่งผลให้มีโอกาสเกิดข้อผิดพลาดได้น้อยแต่จะทำให้เกิดการ Overfit ขึ้นได้ หรือทำให้รับชุดข้อมูลในอนาคตที่มีความซับซ้อนเพิ่มขึ้นไม่ได้ รวมถึงการปรับตัวแปรในช่วงของการสกัดคุณลักษณะ (Feature Extraction) ด้วยเทคนิค HOG ด้วย เพื่อหาโมเดลที่เหมาะสมกับข้อมูลมากที่สุด



ภาพที่ 15 ตัวอย่างความสัมพันธ์ระหว่าง Hyperplane กับค่าตัวแปร C และ Gamma

3.5 Evaluation

การประเมินผล Model จะใช้ 3 วิธีในการประเมินผล

3.5.1 ค่าความแม่นยำ (Accuracy)

ในช่วงการฝึกโมเดลโดยทดสอบจาก Validation Data และเปรียบเทียบกับค่าความแม่นยำจากช่วงทดสอบ โมเดลจาก Test Dataset ประเมินควบคู่กัน

3.5.2 ค่า Precision, Recall และค่า F1 Score

เป็นการคำนวณค่าเฉลี่ยของค่า Precision และค่า Recall ของผลลัพธ์ในการทดสอบโมเดลด้วย Test Dataset เพื่อดูค่า False Negative และ False Positive ว่ามีนัยยะสำคัญในด้านไหนหรือไม่เพื่อใช้ในการปรับปรุงโมเดล

3.5.3 ค่า Loss Validation

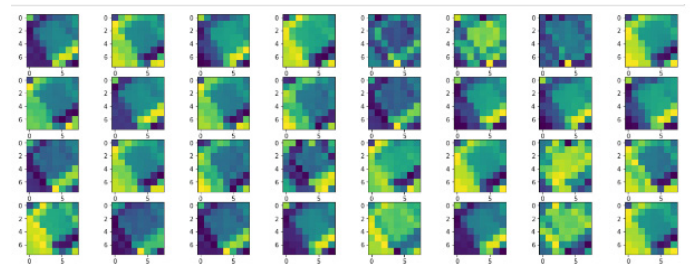
เพื่อประเมินว่าโมเดลยังสามารถเรียนรู้ได้มากกว่านี้หรือไม่ และสามารถสะท้อนถึงการเกิด Overfit ของโมเดลได้ โดยจะการประเมินผลจะใช้ทั้ง 3 วิธีควบคู่กันในการวิเคราะห์ประสิทธิภาพของโมเดล

4. ผลการวิจัยและอภิปรายผล

4.1 ผลการสกัดคุณลักษณะ

4.1.1 Model CNN

จะมีการปรับค่าจำนวน Feature Map แต่ละชั้นของ Convolution Layer, ขนาดของ Kernel Size และขนาดของ Stride โดยจะใช้การประเมินจากการตรวจสอบด้วยสายตา (Visual Check) ในการพิจารณาความชัดเจนของ Feature Map ควบคู่กับค่าความแม่นยำ (Accuracy) และค่า Loss เพื่อทำการปรับโมเดลแบบหยาบก่อนจะไปพิจารณาปรับแบบละเอียด โดยการดูค่า Precision, Recall และค่า F1 Score อีกครั้ง

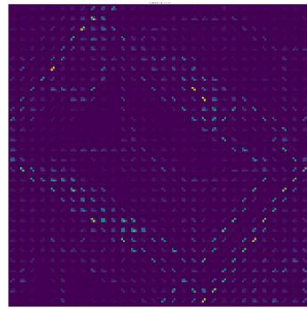


ภาพที่ 16 ตัวอย่าง Feature Map ชั้นที่ 4 จำนวน Feature= 32, Kernel_size = (4,4), Stride = (2,2)

ผลจาก Visual Check จะพบว่า Feature Map มีลักษณะที่ดูซ้ำกันอยู่จำนวนหลาย Feature Map จึงพิจารณาลดจำนวน Feature Map ลงเพื่อไม่ให้โมเดลมีความซับซ้อนจนเกินไปและเมื่อพิจารณาค่า Loss จากผลลัพธ์การประเมินจาก Validation Data โดยทำการฝึกทั้งหมด 20 รอบพบว่าค่า Loss สูงถึงหลักล้าน ซึ่งเป็นค่าที่สูงมากจึงเป็นจุดบ่งชี้ว่าโมเดลมีความซับซ้อนเกินไปเมื่อวิเคราะห์เพิ่มเติมพบว่ากับ Feature Map ที่สกัดออกมานั้นดูไม่มีความชัดเจนจึงพิจารณาลดจำนวนชั้น CNN ลงเพื่อให้ Feature Map มีรูปร่างที่ชัดเจนมากขึ้นและทำให้ลดความซับซ้อนของโมเดลลงด้วย พบว่าเมื่อลดชั้นของ CNN ลงผลลัพธ์ทำให้ Feature Map ที่สกัดออกมาดูไม่ซ้ำซ้อนกันจนเกินไปและตัว Feature Map มีความชัดเจนซึ่งคาดว่าจะทำให้ใช้ในการตรวจจับได้ดีขึ้นและเมื่อพิจารณาค่า Loss จะพบว่าค่าลดลงเหลือหลักพันเท่านั้นเมื่อเริ่มได้ค่าตั้งต้นแล้วจะนำไปปรับตัวแปรอย่างอื่นเพิ่มโดยดูการประเมินจากค่าอื่นๆ ควบคู่กันไป

4.1.2 Model SVM

จะมีการปรับค่าจำนวน Orientation ซึ่งจะเป็นจำนวน Bin ของมุมมองสาโดยจะเริ่มที่ 9 Bin ซึ่งจะเท่ากับ 20 องศา ต่อ Bin, Pixels per Cell และ Cell per Block ซึ่งจะเป็นค่าที่นำมาสกัดให้ได้ Feature โดยหลักการวิเคราะห์ในเบื้องต้นจะเหมือนของโมเดล CNN คือใช้การประเมินด้วยวิธีตรวจสอบด้วยสายตา (Visual Check) ควบคู่กับค่าความแม่นยำ (Accuracy) และค่า Loss เพื่อทำการปรับโมเดลแบบหยาดก่อนจะไปพิจารณาปรับแบบละเอียดโดยการดูค่า Precision, Recall และค่า F1 Score อีกครั้งจะพบว่า Feature ที่สกัดได้ค่อนข้างจะหยาดเกินไปและสังเกตเห็นว่ามุมมองจะไม่ชัดเจนซึ่งจาก Dataset เมื่อทำการสำรวจข้อมูล (Data Explore) จะพบว่ารูปบ้านส่วนใหญ่จะมีรูปทรงสี่เหลี่ยมไม่ได้มีองศาที่หลากหลาย ดังนั้นเพื่อให้ได้ Feature ที่เหมาะสมกับชุดข้อมูลจึงพิจารณาทำการลดจำนวน Bin ลงและเพื่อให้ได้ Feature ที่ชัดเจนขึ้นพร้อมกับทำการลดจำนวน Pixel per Cell และ Cell per Block ลง



ภาพที่ 17 ตัวอย่าง Feature Extraction โดย Parameter Orientation = 5, Pixel per Cell = (5,5), Cell per Block = (2,2)

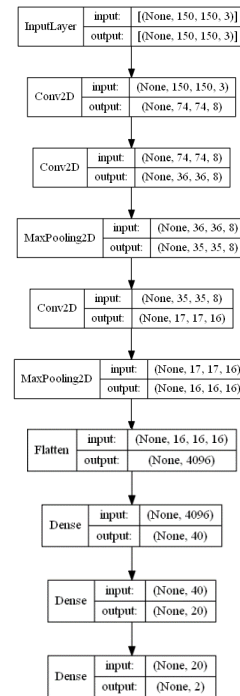
จากผลลัพธ์จะเห็นลักษณะของ Feature ที่ชัดเจนขึ้น ซึ่งคาดว่าจะส่งผลทำให้การ Detect ได้ดีขึ้นจะใช้ค่านี้เป็นค่าตั้งต้นเพื่อนำไปปรับ Parameter โดยดูการประเมินผลรวมกับค่าอื่นๆ ควบคู่กันต่อไป

4.2 ผลการสร้าง Model และการประเมินผลการทำนายของ Model

4.2.1 Model CNN

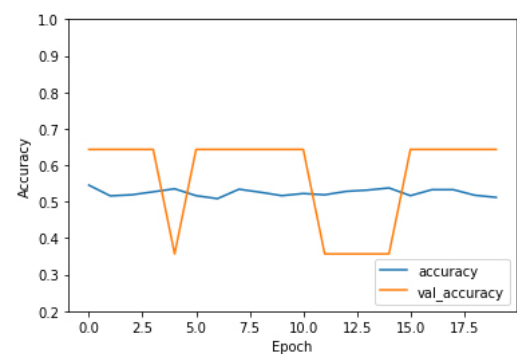
โมเดลตั้งต้นที่ได้จากการขั้นตอนการทำสกัดคุณลักษณะจะประกอบไปด้วยชั้น Convolution 2 ชั้น แต่ละชั้นจะสกัด 8 Feature Map, Kernel Size = 4,4 และ Stride = 2,2 จากนั้นตามด้วยชั้นด้วย Max Pooling โดยที่ Pool Size = 2,2 ต่อด้วยชั้น Convolution 1 ชั้น สกัด 16 Feature Map, Kernel Size = 4,4 และ Stride = 2,2 และตามด้วยชั้น Max pooling Pool Size = 2,2 ก่อนทำการ Flatten และเข้าช่วง Fully Connected

ประกอบไปด้วย Node 3 ชั้น โดยชั้นแรกมี 40 Node ชั้นที่ 2 20 Node และชั้นที่ 3 มี 2 Node โดยมี Activation Function คือ Softmax ในชั้นที่ 3 เพื่อ Transform ผลลัพธ์ออกมาให้รวมกับเท่ากับ 1 เพื่อง่ายต่อการวิเคราะห์



ภาพที่ 18 Process ของ CNN ตั้งต้น

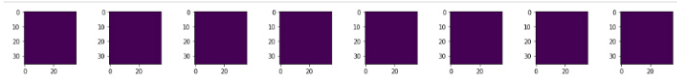
ผลลัพธ์ที่ได้จากการประมวลผลค่าความแม่นยำ มีการเปลี่ยนแปลงในรอบ 50-95% ในทุกรอบการทดลอง โดยมีแนวโน้มค่อยๆดีขึ้น เช่นเดียวกับค่า Loss ลดลงอย่างต่อเนื่องผลสุดท้ายค่าอยู่ที่ 5 หมื่นซึ่งเป็นค่า Loss ที่สูงมากอยู่ จึงพิจารณาเพิ่ม Activation Function ReLu เพื่อลดความซับซ้อนของ Feature Map ลดลงเมื่อให้ไปประมวลผลพบว่าค่าความแม่นยำลดลงและมีค่าเปลี่ยนแปลงอยู่ในรอบ 50-60% โดยมีแนวโน้มที่เริ่มค่าคงที่แต่ที่เป็นจุดน่าสังเกตคือค่า Validate Accuracy ที่มีการเปลี่ยนแปลงไปมาอยู่ 2 ค่า มีลักษณะกราฟที่ผิดปกติ แต่ค่า Loss ลดลงอย่างต่อเนื่องมาอยู่ที่ หลักร้อยเท่านั้นซึ่งเป็นไปตามที่คาดไว้



ภาพที่ 19 Accuracy Result Trend after 20 Round Training with ReLu Function

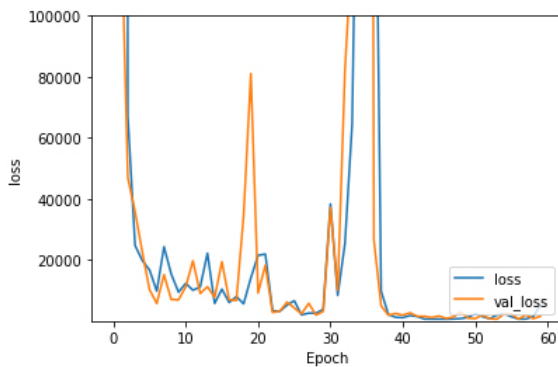


แต่จากการที่ค่าความแม่นยำที่ลดลงอย่างมากและค่า Validate Accuracy ที่มีลักษณะที่ผิดปกติจึงได้ทำการตรวจสอบ Feature Map ที่สกัดออกมาในชั้นที่มี Activation Function ReLu พบว่าผลลัพธ์ที่ออกมาเป็นภาพสีเหมือนกันทั้งหมดจึงต้องพิจารณาไม่ใช้การใช้ Function ReLu เนื่องจากไม่เหมาะกับ Dataset ชุดนี้



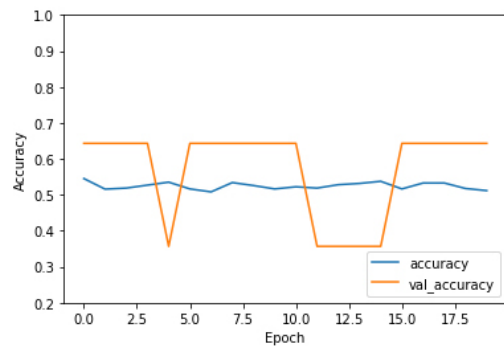
ภาพที่ 20 Feature Mapping ผ่าน Function ReLu

ในการที่จะต้องลดค่า Loss ลงจึงได้ทำการเพิ่มจำนวนรอบการฝึกจาก 20 รอบเป็น 60 รอบ ผลลัพธ์ที่ได้คือ ค่า Loss ลดลงอย่างต่อเนื่องจนเหลือค่าเพียงหลักร้อยเท่านั้น แต่เมื่อพิจารณารูปภาพว่าค่ามีการเปลี่ยนแปลงไปมา โดยที่ค่าความแม่นยำในรอบสุดท้ายอยู่ที่ประมาณ 90%

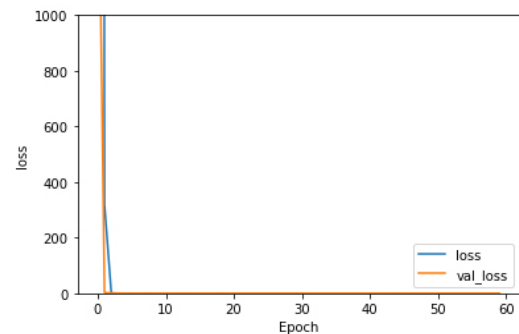


ภาพที่ 21 Loss Result Trend after 60 Round Training

จากกราฟ จะพบว่าค่า Loss มีช่วงที่ค่าเปลี่ยนแปลงกลับไปมาที่ค่าสูงและเมื่อขยายเข้าไปที่จุดช่วงท้ายเพื่อพิจารณาว่ารอบการ Training เพียงพอหรือไม่ พบว่าช่วงการฝึกตั้งแต่รอบที่ 40 มีค่าที่เปลี่ยนแปลงไปมาในรอบที่ต่ำและไม่ได้มีแนวโน้มที่จะดีขึ้น ดังนั้นจากผลลัพธ์จึงจะคงจำนวนรอบการฝึกไว้ที่ 60 รอบและทำการปรับค่า Learning Rate พบว่าค่า Learning Rate ที่ 0.1 ได้ผลลัพธ์ที่ดีที่สุดกล่าวคือเป็นการปรับค่าตัวแปรที่ละเล็กละน้อย จะพบว่ากราฟความแม่นยำกับ Val_accuracy มีความใกล้เคียงกันในเกือบทุกรอบการฝึกและค่าความแม่นยำมีการเปลี่ยนแปลงในรอบที่ต่ำ ตั้งแต่รอบการฝึกประมาณรอบที่ 10 และค่า Loss ลดลงอย่างรวดเร็ว ซึ่งค่า Loss กับค่า Val_loss มีค่าใกล้เคียงกันตั้งแต่รอบการฝึกที่ 10 เช่นเดียวกัน ดังนั้นจึงพิจารณาลดจำนวนรอบการฝึกจาก 60 รอบ ให้เหลือเพียง 20 รอบซึ่งเพียงพอเพื่อให้สามารถประมวลผลโปรแกรมได้รวดเร็วขึ้น



ภาพที่ 22 Accuracy result trend after 60 rounds training with learning rate 0.1



ภาพที่ 23 Loss result trend after 60 rounds training with learning rate 0.1

ทดลองปรับ Loss function จาก Category Cross Entropy เป็น Binary Cross Entropy ที่ถูกออกแบบมาให้ใช้กับชุดข้อมูลที่มี Label จำนวน 2 ประเภทโดยเฉพาะเพื่อให้การเปรียบเทียบชัดเจนยิ่งขึ้นจะทำการพิจารณาค่า Precision, Recall และค่า F1 Score ประกอบการพิจารณาด้วย

	precision	recall	f1-score	support
0	0.98	0.98	0.98	335
1	0.96	0.96	0.96	197
micro avg	0.97	0.97	0.97	532
macro avg	0.97	0.97	0.97	532
weighted avg	0.97	0.97	0.97	532
samples avg	0.97	0.97	0.97	532

ภาพที่ 24 Classification report after 20 rounds training with categorical loss function

ผลจากการประมวลผลพบว่าค่าความแม่นยำของชุดข้อมูลที่ใช้ทดสอบ (Test Dataset) ที่ใช้ Loss Function Binary นั้นมีที่ต่ำกว่าเล็กน้อย แต่มีค่า Loss ที่ดีกว่า เมื่อพิจารณาค่า Precision, Recall กับค่า F1 Score พบว่าค่าค่อนข้างใกล้เคียงกันเมื่อดูกราฟระหว่างค่าความแม่นยำ (Accuracy) และ Val_accuracy ค่อนข้างที่จะมีค่าใกล้เคียงมากกว่าในรอบการฝึกและมีการเปลี่ยนแปลงที่น้อยกว่า

	precision	recall	f1-score	support
0	0.98	0.98	0.98	335
1	0.96	0.96	0.96	197
micro avg	0.97	0.97	0.97	532
macro avg	0.97	0.97	0.97	532
weighted avg	0.97	0.97	0.97	532
samples avg	0.97	0.97	0.97	532

ภาพที่ 25 Classification report after 20 rounds training with Binary loss function

เมื่อทดสอบประมวลผลโปรแกรมใหม่ทั้งหมดเพื่อพิจารณาค่าความสามารถในการทำซ้ำ (Repeatability) ของโมเดลผลลัพธ์ของทั้ง 2 Function มีค่าใกล้เคียงกันมากสามารถเลือกใช้ได้ทั้ง 2 Function แต่จากกราฟความแม่นยำและ Loss ของ Activation Function Binary มีค่าความเปลี่ยนแปลงที่น้อยกว่าจึงเลือก Binary ในการใช้วิเคราะห์ต่อไปซึ่งจะเป็นการปรับตัวแปรชั้น Pooling โดยการเปรียบเทียบระหว่าง Max Pooling และ Average Pooling ผลลัพธ์ที่ได้พบว่า Average Pooling ได้ค่าความแม่นยำที่น้อยกว่า และจากค่า Loss ที่มากกว่าจึงพิจารณาใช้ Max Pooling จากนั้นได้ทดลองปรับเพิ่มจำนวนชั้นของ Convolution เข้าไปอีกชั้นผลลัพธ์ที่ได้คือค่าความแม่นยำ ใกล้เคียงเดิมแต่ค่า loss เพิ่มขึ้นจึงไม่มีความจำเป็นต้องเพิ่มชั้น Convolution เข้าไปอีก

4.2.2 Model SVM

Model ตั้งต้นที่ได้จากการสกัดคุณลักษณะ (Feature Extraction) คือ $C = 1$ ซึ่งเป็นค่าตั้งต้น (Default) ของ Keras และ $\text{Gamma} = \text{scale}$ หลังจากประมวลผลพบว่าผลลัพธ์ที่ได้คือค่าความแม่นยำเท่ากับ 87% และได้ค่า Precision, recall และ F1 Score

	precision	recall	f1-score	support
0	0.87	0.94	0.90	335
1	0.88	0.75	0.81	197

ภาพที่ 26 ผลลัพธ์การประมวลผลโมเดล SVM ค่า $C=1$ $\text{Gamma} = \text{scale}$

ซึ่งจะพบว่าค่าออกมาดี จากนั้นดำเนินการปรับค่าตัวแปร C แบบหยาบซึ่งค่าตัวแปร C ยังมีค่าน้อยยิ่งจะส่งผลให้เพิ่มค่า Error ที่เกิดขึ้นได้ จะทำให้โมเดล ไม่เกิดการ Overfit

ขึ้นและรองรับ Dataset ที่สามารถเพิ่มขึ้นในอนาคตได้ดีกว่าวัตถุประสงค์ของการทดลองคือการลดค่า C ให้ได้น้อยที่สุดจะพบว่าเมื่อเพิ่มค่าตัวแปร C ขึ้นเป็น $C = 5$ ค่าความแม่นยำเพิ่มขึ้น แต่เมื่อเพิ่มไปถึงค่า $C=10$ ผลลัพธ์คือค่าทุกอย่างเท่าเดิมทั้งหมด ดังนั้นในขั้นตอนต่อมาจะทำการปรับละเอียดในช่วง 1-5 เพื่อหาค่าที่เหมาะสมที่สุดผลลัพธ์คือเลือกค่า $C = 2$ เนื่องจากมีค่าที่แตกต่างจากผลลัพธ์ที่ได้จากค่า $C=3$ เพียงเล็กน้อยเท่านั้น และชุดข้อมูลรูปภาพก่อนข้างมีความซับซ้อนจึงต้องมีการพิจารณาค่า Error เพื่อไว้ไม่ให้โมเดล Overfit จนเกินไปต่อไปจะเป็นการทดลองเปลี่ยน Parameter Gamma ยิ่งค่าสูงจะให้ขอบเขตการแบ่ง Class จะแคบลงซึ่งอาจจะทำให้ค่า Accuracy มากขึ้นแต่ก็ทำให้ Overfit ได้ง่ายเช่นกันเมื่อคำนวณตามสูตร Scale จะได้ค่า $\text{Gamma} = 0.00329$ จะพบว่าค่าที่คำนวณตาม Scale ที่เป็นตั้งต้น (Default) มีค่าความแม่นยำที่สูงที่สุดหลังจากลองทำการปรับทั้งหยาบและละเอียดรวมถึงให้ค่า Precision, Recall และค่า F1 Score ที่เหมาะสมที่สุด

4.3 อภิปรายผลการศึกษา

4.3.1 การทำ Understanding Data และ Data Preparation

จากกระบวนการศึกษา Data Mining โดยใช้ข้อมูลรูปภาพที่มาจากรายงานการจัดทำภาพถ่ายทางอากาศซึ่งรูปภาพต้นฉบับนั้นผ่านการกระบวนการต่างๆจนได้ภาพ Orthomosaic ซึ่งมีสัดส่วนที่ถูกต้องกับความเป็นจริงก่อนถูกตีกรอบมา (Crop) เฉพาะจุดที่สนใจมาใช้ในรายงานจัดทำภาพถ่ายทางอากาศซึ่งส่งผลต่อขนาดรูปภาพ, จำนวนรูปภาพที่ไม่มากและมีการเพิ่มตำแหน่งของท่อ, Right of Way ซึ่งเป็น Noise ของข้อมูลแต่ไม่สามารถจะกำจัด Noise นี้ไปได้ ซึ่งจะส่งผลต่อโมเดลที่ใช้สกัดคุณลักษณะออกมาเป็นรูปทรงอย่างเทคนิค HOG ได้รับ Noise นี้ไปด้วยโดยงานวิจัยนี้ได้ทำการ Data Selection มาเฉพาะภาพที่มีอาคารที่อยู่อาศัยเท่านั้นและทำการตีกรอบ (Crop) มาเฉพาะบริเวณที่อยู่อาศัยเพื่อใช้ในการฝึกและทดสอบโมเดลสำหรับ Label House และทำการตีกรอบ รูปภาพอื่นๆ ที่ไม่ใช่รูปบ้านซึ่งประกอบไปด้วย หญ้า, ต้นไม้, รถ, ถนน และอื่นๆ มาเพื่อเป็น Label สำหรับ Other ที่ไม่ใช่บ้าน จากนั้นได้ทำการ Exploration Data พบว่าข้อมูล



มีหลากหลาย Shape แต่มีนามสกุล File เดียวกันทั้งหมดทำให้ช่วง Data Preparation ได้มีขั้นตอนการปรับเปลี่ยน Shape ให้เหมือนกันทั้งหมดส่งผลให้สัดส่วนวัตถุในรูปจะเปลี่ยนไปซึ่งอาจจะส่งผลต่อการฝึกโมเดล ทั้ง 2 Model และข้อมูลรูปภาพเป็นภาพถ่ายมุมสูงที่จะเห็นเพียงหลังคาของอาคารที่อยู่อาศัย ซึ่งจากลักษณะทางสถาปัตยกรรมอาคารที่อยู่อาศัยในพื้นที่เขตตอนนั้นไม่ได้มีความซับซ้อนมากนักแต่จะมีความหลากหลายทางสีทำให้จำนวนรูปข้อมูลเมื่อเปรียบเทียบกับสิ่งที่ต้องตรวจจับนั้นน้อยจึงได้ทำการ Data Augmentation เพื่อเพิ่มจำนวนข้อมูลทำให้จำนวนข้อมูลไม่น้อยจนเกินไปจนทำให้เกิด Overfit

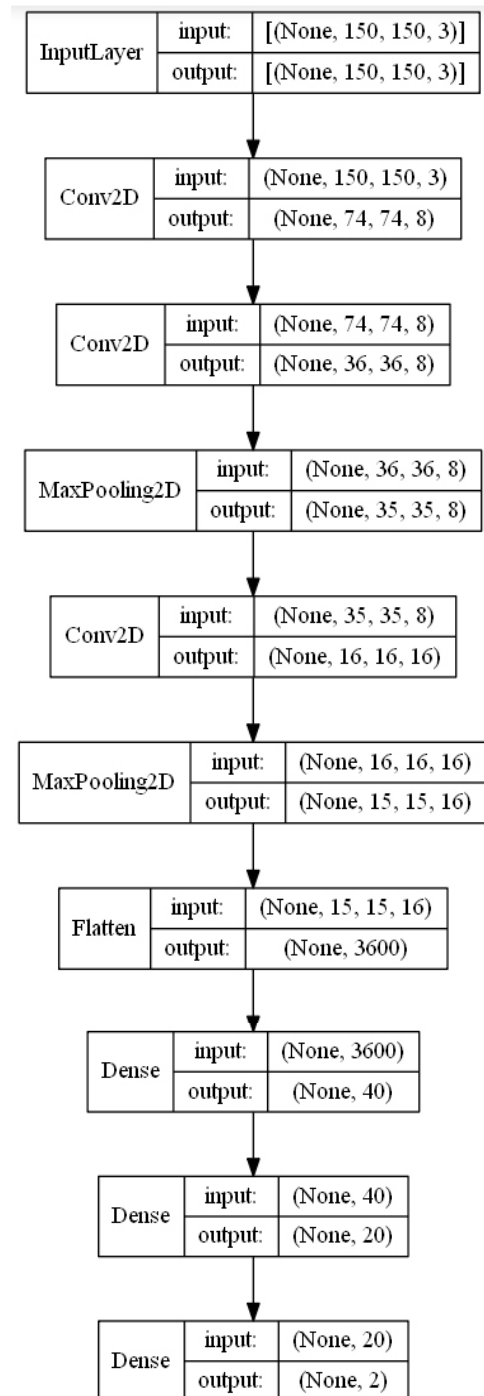
4.3.2 CNN Model

ผลจากการทดลอง CNN Model พบว่าไม่จำเป็นต้องใช้จำนวนชั้นของ Convolution มากและไม่ต้องการจำนวน Feature Map ที่มาก เพราะคุณลักษณะของสิ่งที่ตรวจจับไม่ได้

5. สรุปผลการศึกษา

จากการศึกษาการประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึกในการสำรวจแหล่งที่อยู่อาศัยด้วยภาพถ่ายจาก UAV ตลอดแนวท้องฟ้าธรรมชาติ และได้ทำการใช้เทคนิค SVM เพื่อเปรียบเทียบ Model ที่เหมาะสมที่สุด จากเทคนิค Deep Learning มีเทคนิคที่เรียกว่า Convolutional Neural Network (CNN) ที่ได้ถูกพัฒนาเพื่อใช้ในการทำ Image Processing กับเทคนิค Support Vector Machine (SVM) ที่มีการใช้เทคนิค Histogram of Oriented Gradients (HOG) ในการทำสกัดคุณลักษณะซึ่งเป็นขั้นตอนหนึ่งที่ใช้ในการทำ Image Processing เช่นเดียวกันผลลัพธ์โมเดลที่เหมาะสมกับชุดข้อมูลภาพถ่ายทางอากาศเพื่อตรวจจับแหล่งที่อยู่อาศัยคือโมเดล CNN โดยมีโครงสร้างที่ประกอบไปด้วยชั้น Convolution 8 Feature Map, Kernel Size = (4,4) , Stride = (2,2) จำนวน 2 ชั้น ตามด้วยชั้น Max Pooling Pool_Size = (2,2) จำนวน 1 ชั้นตามด้วยชั้น Convolution 16 Feature Map, Kernel Size = (4,4) , Stride = (2,2) จำนวน 1 ชั้น ตามด้วยชั้น Max Pooling Pool_Size = (2,2) 1 ชั้นต่อด้วยชั้น Flatten และเข้าช่วง Fully Connected ซึ่งประกอบไปด้วย Dense 40 Node 1 ชั้นตามด้วย Dense 20 Node และชั้น Dense

2 Node อีก 1 ชั้นซึ่งใช้ Function Softmax เป็น Model ผ่านการฝึกจำนวน 20 รอบ ด้วย Loss Function Binary Cross Entropy, Optimizer Adam, Learning Rate 0.1



ภาพที่ 27 โครงสร้างโมเดล CNN สุดท้าย

โดยประเมินจากค่าต่างๆ ซึ่งการประมวลผลแต่ละครั้งมีการเปลี่ยนแปลงค่าในรอบเล็กน้อยดังนี้

- Accuracy Training Dataset = 97%
- Accuracy Validate Dataset = 96.5%
- Loss Training Dataset = 0.18

- Accuracy Test Dataset = 96%
- Loss Test Dataset = 0.22
- Precision = 0.96
- Recall = 0.97
- F1 Score = 0.97

ส่วนเทคนิค SVM โมเดลที่เหมาะสมที่สุดประกอบด้วย Kernel = 'rbf', C=2, Gamma = 'scale' (ในการทดลองนี้ Gamma = 0.00329 ซึ่งผ่านการคำนวณสมการ $1/(n_feature * Xvar)$) โดยประเมินจากค่าดังนี้

- Accuracy Test Dataset = 88.34%
- Precision = 0.88
- Recall = 0.94
- F1 Score = 0.91

โดยการประมวลผลทุกครั้งได้ผลลัพธ์ที่เท่ากันหมด ทุกครั้งซึ่งแปลว่าเส้นแบ่ง Class ก่อนข้างชัดเจนและเหมาะสม ซึ่งผลการเปรียบเทียบทั้ง 2 โมเดลพบว่าโมเดล CNN ให้ค่าที่เหมาะสมกว่าโมเดล SVM ซึ่งผลการวิจัยได้ตอบโจทย์วัตถุประสงค์การวิจัยอย่างครบถ้วน

6. กิตติกรรมประกาศ

การที่ค้นคว้าอิสระฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี เนื่องจากได้รับความอนุเคราะห์และการชี้แนะจากอาจารย์ที่ปรึกษา ผศ.ดร.อัจฉรา ภู่อ่าง และคณะกรรมการสอบทุกท่าน ซึ่งประกอบด้วย รศ.ดร.จิรพล สังข์โพธิ์ รศ.ดร.ศราวุธ แรมจันทร์ และ ผศ.ดร.อัจฉรา ภู่อ่าง ที่คอยชี้แนะสิ่งที่เป็ประโยชน์และสิ่งที่ต้องปรับปรุงเพิ่มเติมตลอดระยะเวลาการศึกษาในหลักสูตรฯ

ขอขอบคุณทีมงานพนักงานบริษัท ปตท. จำกัด (มหาชน) ที่ให้ข้อมูลที่ใช้ในการวิจัยครั้งนี้

ผู้วิจัยหวังเป็นอย่างยิ่งว่า การศึกษานี้จะได้นำไปประยุกต์ใช้ให้เกิดประโยชน์ต่อองค์กร และประเทศชาติต่อไป ทั้งยังเป็นแนวทางในการนำ Machine Learning เข้ามาประยุกต์ใช้ร่วมกับเทคโนโลยี UAV เพื่อเป็นประโยชน์ในการนำไปพัฒนาปรับปรุงและต่อยอดให้ดียิ่งขึ้นต่อไป หากการศึกษานี้มีข้อบกพร่องประการใด ผู้วิจัยขออภัยมา ณ ที่นี้และขอน้อมรับไว้ไปปรับปรุงแก้ไขในภายภาคหน้าต่อไป ขอขอบคุณ

เอกสารอ้างอิง

- [1] Shubhra Aich และ Ian Stavness.(2018). Improving Object Counting with Heatmap Regulation. ResearchGate 2018. arXiv:1803.05494v1
- [2] The American Society of Mechanical Engineers (2020). Gas Transmission and Distribution Piping Systems. ASME B31.8, 2020.
- [3] Sumit Saha, (2018). A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. สืบค้นเมื่อวันที่ 1 มีนาคม 2565, จาก <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.
- [4] Cosmos Team, (2022). Convolutional Neural Network. สืบค้นเมื่อวันที่ 1 มีนาคม 2565, จาก <https://www.cosmos.esa.int/web/machine-learning-group/convolutional-neural-networks-introduction>.
- [5] Wikipedia Team, (2022). Support-vector machine. สืบค้นเมื่อวันที่ 10 พฤษภาคม 2565, จาก https://en.wikipedia.org/wiki/Support-vector_machine.
- [6] Rishabh Jain, (2020). Support Vector Machine (SVM). สืบค้นเมื่อวันที่ 10 พฤษภาคม 2565, จาก <https://medium.com/@rishabhjain9440/support-vector-machine-svm-f23cece26419>.
- [7] Mrinal Tyagi, (2021). HOG (Histogram of Oriented Gradients): An Overview. สืบค้นเมื่อวันที่ 10 พฤษภาคม 2565, จาก <https://towardsdatascience.com/hog-histogram-of-oriented-gradients-67ecd887675f>.
- [8] Scikit-image Team. Histogram of Oriented Gradients. สืบค้นเมื่อวันที่ 10 พฤษภาคม 2565, จาก https://scikit-image.org/docs/stable/auto_examples/features_detection/plot_hog.html
- [9] Andreas C.Muller, (2020). Support Vector Machine. สืบค้นเมื่อวันที่ 10 พฤษภาคม 2565, จาก <https://towardsdatascience.com/support-vector-machine-introduction-to-svm>

amueller.github.io/aml/02-supervised-learning/07-support-vector-machines.html.

[10] Savaram Ravindra. How Convolutional Neural Networks Accomplish Image Recognition. สืบค้นเมื่อวันที่ 10 มีนาคม 2565, จาก <https://www.kdnuggets.com/2017/08/convolutional-neural-networks-image-recognition.html>.

[11] Cube Analytics Consulting Co., Ltd., (2018). กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM และตัวอย่างการประยุกต์ใช้ทางการศึกษา. สืบค้นเมื่อวันที่ 10 มีนาคม 2565, จาก <https://datacubeth.ai/crisp-dm/>.

[12] LOOKNorth Company. About Pipeline Class Change Location (CCL) Service. สืบค้นเมื่อวันที่ 10 มีนาคม 2565, จาก <https://www.looknorthservices.com/pipeline-ccl.html>.

[13] TIBCO Software Inc. What is a Neural Network. สืบค้นเมื่อวันที่ 10 มีนาคม 2565, จาก <https://www.tibco.com/reference-center/what-is-a-neural-network>

[14] Mehmet Tekman. Support Vector Machine Object Detection. สืบค้นเมื่อวันที่ 10 พฤษภาคม 2565, จาก <https://www.kaggle.com/code/mehmetlaudatekman/support-vector-machine-object-detection>

[15] บริษัท ปตท. จำกัด (มหาชน) (2564). รายงานผลการปฏิบัติงานการจัดทำภาพถ่ายทางอากาศเพื่อประเมินการเปลี่ยนแปลงของพื้นที่แนวท่อด้วยอากาศยานไร้คนขับ (UAV) ครั้งที่ 2 ประจำปี 2564 ณ พื้นที่ศูนย์ปฏิบัติการเขต

