

การตรวจับการทำงานของเครื่องปรับอากาศแบบไม่รุกรานโดยใช้โครงข่ายประสาทเทียมแบบ คอนโวลูชัน

NON-INTRUSIVE AIR-CONDITIONER USAGE DETECTION USING CONVOLUTIONAL NEURAL NETWORKS

เอกชัย เนาวนิช¹ ธนพร ปฏิกรณ์*²

Ekachai Naowanich¹, Thanaporn Patikorn*²

^{1,2} คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

^{1,2} Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi

* E-mail ผู้รับผิดชอบบทความ (Corresponding Author): thanaporn.pa@rmutsb.ac.th

บทคัดย่อ

ในปัจจุบันมีการสร้างและใช้งานเครื่องดูแลและตรวจับการทำงานของอุปกรณ์อิเล็กทรอนิกส์มากขึ้น เพื่อให้ผู้ใช้หรือผู้ดูแลระบบสามารถดูแลและตรวจับการทำงานระยะไกลผ่านอินเทอร์เน็ตได้ แต่ในการทำงานจริงนั้น ระบบเหล่านี้ยังมีข้อจำกัด โดยเฉพาะอย่างยิ่งสำหรับเครื่องปรับอากาศ เนื่องจากเครื่องปรับอากาศมีราคาสูง มีอายุการใช้งานยาวนาน และเมื่อมองภาพรวมระดับอาคารหรือองค์กร จะเห็นว่าการใช้เครื่องปรับอากาศหลายรุ่นหรือจากหลายผู้ผลิตมักเป็นสิ่งที่เลี่ยงไม่ได้ ทำให้การเปลี่ยนเครื่องเดิมให้เข้ากับระบบดังกล่าวมีความซับซ้อนและมีราคาสูง จึงทำให้เกิดการศึกษาและพัฒนาเครื่องมือตรวจับการทำงานของเครื่องปรับอากาศแบบไม่รุกรานโดยใช้ปัญญาประดิษฐ์จากภาพวงจรปิดขึ้น โมเดลที่ใช้เป็นโครงข่ายประสาทเทียมแบบคอนโวลูชัน โดยใช้การทรานสเฟอร์เลิร์นนิ่งจากโมบิลเน็ต พัฒนาโดยใช้เทนเซอร์ฟลว โดยมีการทดลองโมเดล 3 แบบจากการใช้ข้อมูลความละเอียดเชิงเวลาที่ต่างกัน คือ 1, 2, และ 30 เฟรมต่อวินาที ซึ่งผลการทดลองโมเดลที่พัฒนาพบว่า โดยรวมโมเดลสามารถแยกแยะการทำงานของเครื่องปรับอากาศได้ดีกว่าฐานการเปรียบเทียบ และในเครื่องปรับอากาศหลายรูปแบบ โมเดลสามารถแยกแยะการทำงานได้ 100%

คำสำคัญ: คอมพิวเตอร์วิทัศน์, โครงข่ายประสาทเทียมแบบคอนโวลูชัน, ระบบควบคุมและจัดการจากระยะไกล

ABSTRACT

Nowadays, there is an increasing number of monitoring tools for electronic systems. Such tools allow users or administrators to monitor such systems remotely via the internet. However, for many real systems, such tools see limited usage, especially for air conditioners since air conditioners are expensive and long-lasting. In addition, on a building-level or organization-level, having to use air conditioners of different models or from different manufacturers is often unavoidable or impractical. Thus, changing all air conditioners to fit a particular monitoring tool is complicated and expensive. In this work, a non-intrusive air conditioner monitoring tool from security camera footage using artificial intelligence was designed and developed. Models were developed using convolutional neural networks, with MobileNet v3 as the base for transferred learning, in TensorFlow. Three models were tested with different volumes of training data: 1, 2, and 30 frames per second. The result showed that, overall, models performed better than the baseline. In addition, models could detect air conditioner usages with 100% accuracy for models with specific characteristics.

Keywords: computer vision, convolutional neural network, remote monitoring systems

1. บทนำ

เทคโนโลยีสมัยใหม่ทำให้อุปกรณ์อิเล็กทรอนิกส์มีความฉลาดมากขึ้น หรือที่เรียกว่าอุปกรณ์อัจฉริยะ (smart devices) สามารถเข้าถึงกันมากขึ้นรวมเป็นอินเทอร์เน็ตของสรรพสิ่ง (Internet of Things) ทำให้ความต้องการในการดูแลและตรวจจัดการทำงานของอุปกรณ์อิเล็กทรอนิกส์ระยะไกลผ่านอินเทอร์เน็ตเป็นไปได้ง่ายขึ้น แต่สำหรับอุปกรณ์บางประเภทนั้น การตรวจจัดการทำงานระยะไกลผ่านอินเทอร์เน็ตยังมีข้อจำกัดมาก โดยเฉพาะอย่างยิ่งเครื่องปรับอากาศ ซึ่งมีความต้องการในการตรวจจัดการทำงานสูง เนื่องจากเป็นอุปกรณ์ที่ใช้พลังงานไฟฟ้ามาก [1] เป็นอุปกรณ์มักจะถูกลืมปิด และสามารถสร้างความเสียหายสูงได้ในกรณีเครื่องปรับอากาศในห้องที่ต้องรักษาระดับอุณหภูมิเกิดความขัดข้องและหยุดการทำงานกะทันหัน

แม้เครื่องที่ผลิตมาใหม่จะมีระบบตรวจจัดการทำงานระยะไกลมาพร้อมกับตัวเครื่อง เครื่องรุ่นเก่าที่มีอยู่แล้วมักจะไม่สามารถเชื่อมต่อเข้ากับระบบใหม่ได้ โดยเฉพาะหากเป็นเครื่องที่ผลิตจากผู้ผลิตที่ต่างกัน อีกทั้งการรวมระบบตรวจจัดการทำงานเครื่องปรับอากาศในทุกเครื่องในอาคารหรือองค์กรเข้าเป็นระบบเดียวกันนั้นเป็นไปได้ยากหากไม่ได้ติดตั้งทั้งระบบพร้อมกัน หรือจำเป็นต้องใช้ผู้มีความชำนาญการสูงและมีค่าใช้จ่ายสูง

จากปัญหาข้างต้นจึงทำให้เกิดความต้องการเครื่องมือตรวจจัดการทำงานของเครื่องปรับอากาศ ที่สามารถทำให้ผู้ใช้ตรวจจากระยะไกลได้ว่าเครื่องปรับอากาศกำลังทำงานอยู่หรือไม่ โดยไม่จำเป็นต้องเปลี่ยนแปลงเครื่องปรับอากาศมาก และสามารถใช้ได้กับเครื่องปรับอากาศทุกรุ่นทุกผู้ผลิต ซึ่งโดยปกติแล้วหากผู้ใช้อยู่ในบริเวณเครื่องปรับอากาศ สามารถตรวจได้เองโดยการสังเกตที่ลักษณะภายนอก เช่น แผงควบคุมหรือรีโมทควบคุม การเคลื่อนไหวของใบพัดและบานสวิง หรือจากฟลูส์ ซึ่งเป็นแถบที่ทำจากกระดาษหรือพลาสติก ตัดเป็นแถบ

ยาว ใช้ติดเครื่องปรับอากาศหรือพัดลมเพื่อใช้ในการวัดว่ามีลมออกมาหรือไม่ ซึ่งการตรวจจับวิธีนี้มีราคาสูงและไม่ต้องการความชำนาญในการติดตั้ง

วัตถุประสงค์ของงานวิจัยนี้ คือ ปรับเปลี่ยนส่วนของการตรวจจับการทำงานของเครื่องปรับอากาศด้วยสายตา ทดแทนด้วยการใช้เทคนิคทางคอมพิวเตอร์วิทัศน์ (Computer Vision) ที่เรียกว่าโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks, CNNs) [2] และใช้ภาพจากกล้องวงจรปิด ทำให้การตรวจจับด้วยวิธีนี้สามารถตรวจจับได้แม้ผู้ใช้จะไม่ได้อยู่ในบริเวณของเครื่อง ซึ่งวิธีนี้เป็นวิธีที่ไม่รุกราน (non-intrusive) คือ ไม่รบกวนการทำงานของเครื่องปรับอากาศที่มีอยู่แล้ว และผู้ใช้ไม่จำเป็นต้องเปลี่ยนแปลงระบบเครื่องปรับอากาศที่มี ดังนั้น วิธีนี้จึงใช้ได้กับระบบควบคุมเครื่องปรับอากาศได้ทุกรุ่นทุกผู้ผลิต

ด้วยเหตุนี้จึงเกิดงานวิจัยเพื่อพัฒนาโมเดลทางคอมพิวเตอร์วิทัศน์เพื่อตรวจจับการทำงานของเครื่องปรับอากาศจากภาพของเครื่องปรับอากาศ ซึ่งสามารถใช้ได้ในทั้งแบบส่วนบุคคลในบ้านพักที่อยู่อาศัย ในอาคารสำนักงาน และสามารถใช้เป็นส่วนหนึ่งของระบบดูแลและตรวจสอบการทำงานของอุปกรณ์อิเล็กทรอนิกส์ในองค์กร

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

เป็นที่ทราบกันว่าโครงข่ายประสาทเทียมแบบคอนโวลูชัน [2] นั้นถูกใช้ในการตรวจจับวัตถุอย่างแพร่หลายเนื่องจากเป็นโมเดลที่มีประสิทธิภาพสูง เช่น การตรวจจับผู้บุกรุกสถานที่ [3] การตรวจจับชนิดพันธุ์นก คัมครอง [4] และการตรวจจับชนิดสัตว์ป่าในอุทยานแห่งชาติในรัสเซีย [5] รวมไปถึงการตรวจจับสถานะ เช่น รอยยิ้มบนใบหน้า [6] อารมณ์ที่แสดงบนใบหน้า [7] และการใส่หน้ากากของผู้คน [8]

เนื่องจากโครงข่ายประสาทเทียมแบบคอนโวลูชันมีประสิทธิภาพสูงนั้นจำเป็นต้องมีขนาดใหญ่มาก ซึ่งต้องใช้ข้อมูลปริมาณมหาศาลและใช้เวลาในการเรียนรู้เป็นเวลานาน ในปัจจุบันจึงนิยมนำโมเดลขนาดใหญ่ที่ถูกรู้จักกันมาแล้ว (pre-trained models) มาใช้เป็นฐานในสร้างโมเดลที่ทำงานตรวจจับเฉพาะด้าน ซึ่งเทคนิคนี้เรียกว่าการส่งต่อการเรียนรู้ (transfer learning) [9] โดยที่โมเดลขนาดใหญ่ที่ถูกรู้จักกันมาแล้วสำหรับโครงข่ายประสาทเทียมแบบคอนโวลูชันที่นิยมใช้ในปัจจุบันมีมากมาย เช่น EfficientNet [10] MobileNet [11] และ ImageNet [12]

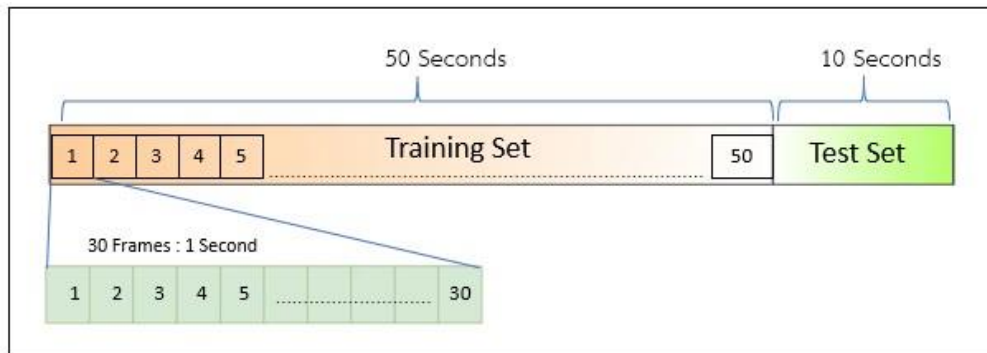
3. วิธีการวิจัย

3.1 การเก็บข้อมูล

ข้อมูลที่ใช้ในการสร้างโมเดล คือ ภาพของเครื่องปรับอากาศ ที่เก็บจากเครื่องปรับอากาศที่ใช้งานอยู่ในห้องต่าง ๆ จากมหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ ศูนย์นนทบุรี โดยที่กระบวนการเก็บใช้วิธีการบันทึกวิดีโอด้วยความละเอียด 1920×1080 pixels (1080p) ด้วยอัตรา 30 ภาพต่อวินาที (frames per second, fps) จากนั้นจึงนำวิดีโอที่ได้มาแยกเป็นภาพ เพื่อลดความลำเอียง (bias) จากข้อมูล แต่ละวิดีโอจะถูกดึงภาพจำนวนเท่า ๆ กัน คือ 1,800 ภาพ ($30 \text{ fps} \times 60 \text{ วินาที}$) ดังภาพที่ 1 การถอดภาพจากวิดีโอทำโดยใช้ imageio ซึ่งเป็น Python library สำหรับประมวลผลภาพและวิดีโอ

โดยภาพที่ได้มีทั้งหมด 7 ชุด ภาพชุดเดียวกันเป็นภาพจากเครื่องปรับอากาศเครื่องเดียวกันและมุมเดียวกันทั้งหมด โดยแต่ละชุดซึ่งแบ่งเป็น 4 ชุดย่อย ชุดย่อยละ 1,800 ภาพ (รวม 7,200 ภาพต่อชุด รวมทั้งหมด

50,400 ภาพ) คือ 1. เครื่องปรับอากาศปิด ขณะไฟสลัว 2. เครื่องปรับอากาศปิด ขณะไฟสว่างจ้า 3. เครื่องปรับอากาศเปิด ขณะไฟสลัว และ 4. เครื่องปรับอากาศเปิด ขณะไฟสว่างจ้า โดยที่ไฟสลัว (Dim, D) คือมีเพียงแสงจากแสงล้อมรอบเท่านั้น ในขณะที่ภาพที่มีไฟสว่างจ้า (Bright, B) มีการเปิดไฟส่องสว่างในบริเวณเครื่องปรับอากาศ

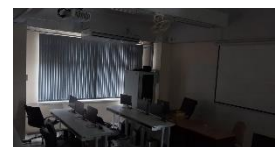


ภาพที่ 1 การสร้างชุดภาพ training set และ test set จากวิดีโอต้นฉบับความยาว 1 นาที

นอกจากลักษณะข้างต้นแล้ว ภาพแต่ละชุดมีความแตกต่างกันดังตารางที่ 1 และตัวอย่างในภาพที่ 2 โดยที่ชุดภาพย่อยจะใช้ชื่อชุดภาพเหมือนชุดภาพทุกประการ ยกเว้น XX จะถูกแทนที่ด้วยสถานะของเครื่องปรับอากาศ (ON OFF) และ Y จะถูกแทนที่ด้วยสถานะแสงไฟ (D = ไฟสลัว และ B = ไฟสว่างจ้า)

ตารางที่ 1 ลักษณะสำคัญของภาพแต่ละชุดภาพ

ชุดที่	ชื่อชุดภาพ	มุมมองภาพ	ลักษณะภายนอกเครื่องปรับอากาศที่สำคัญ	ตำแหน่งที่ตั้งของเครื่องปรับอากาศ	แสงล้อมรอบ
1	C_XX_Y_FL_COR	ใกล้	มีแผ่นบานสวิง	ทางเดิน	สว่างมาก
2	C_XX_Y_FLST_86	ใกล้	มีแผ่นบานสวิงและมีฟูลี	ห้องทำงานรวม	สว่าง
3	C_XX_Y_ST_86	ใกล้	มีฟูลี	ห้องทำงานรวม	สว่าง
4	F_XX_Y_FL_72	ไกล	มีแผ่นบานสวิง	ห้องเรียน1	ค่อนข้างมืด
5	F_XX_Y_FL_COR	ไกล	มีแผ่นบานสวิง	ทางเดิน	สว่างมาก
6	F_XX_Y_ST_86	ไกล	มีฟูลี	ห้องทำงานรวม	สว่าง
7	F_XX_Y_ST_75	ไกล	มีฟูลี	ห้องเรียน2	มืด





ภาพที่ 2 ตัวอย่างภาพจากชุดที่ 1, 2, 3, 4 (บน จากซ้ายไปขวา) และชุดที่ 5, 6, 7, 7 (ล่าง จากซ้ายไปขวา)

ทุกภาพเป็นภาพแสงสลัวยกเว้นภาพขวาล่างสุดซึ่งเป็นภาพแสงไฟสว่างจ้า

โดยที่ลักษณะสำคัญของชุดภาพมีรายละเอียดดังนี้

- มุมมองภาพ หมายถึง มุมมองจากกล้องเมื่อเทียบกับเครื่องปรับอากาศในภาพ
 - “ไกล” (far, F) หมายถึง เป็นมุมมองจากกล้องที่อยู่ไกลจากเครื่องปรับอากาศ ซึ่งนอกจากตัวเครื่องปรับอากาศแล้ว ในภาพอาจมีสิ่งของอื่นได้ เช่น โต๊ะ เก้าอี้ ตู้เก็บของ และทางเดิน
 - “ใกล้” (close-up, C) หมายถึง เป็นมุมมองจากกล้องที่อยู่ใกล้เครื่องปรับอากาศจากมุมตรง สามารถมองเห็นเครื่องปรับอากาศได้ชัดเจนและมองเห็นองค์ประกอบอื่นในภาพน้อย
- ลักษณะภายนอกเครื่องปรับอากาศที่สำคัญในงานวิจัยนี้แบ่งเป็นสองส่วน คือ
 - แผ่นบานสวิง (swing flap, FL) หมายถึง ส่วนประกอบภายนอกอย่างหนึ่งของเครื่องปรับอากาศบางรุ่น ซึ่งเป็นแผ่นพลาสติกหรือโลหะที่อยู่บริเวณช่องลมออกของเครื่องปรับอากาศ มีหน้าที่บังคับทิศทางลม ไกล-ใกล้โดยการหมุนแผ่นบานสวิงขึ้น-ลง (pitch movement)
 - พู่สี (paper/plastic strips, ST) หมายถึง ซึ่งเป็นแถบกระดาษหรือพลาสติกสีต่าง ๆ ที่ถูกตัดเป็นแถบยาว ใช้ติดเครื่องปรับอากาศหรือพัดลมเพื่อใช้ในการวัดว่ามีลมออกมาหรือไม่
- ตำแหน่งที่ตั้งของเครื่องปรับอากาศ ซึ่งส่วนนี้ไม่มีความสำคัญมากในตัวเอง แต่เป็นสิ่งที่ส่งผลกระทบต่อระดับแสงล้อมรอบ ซึ่งชุดภาพที่มีตำแหน่งตามตารางที่ 1 เหมือนกันเป็นเครื่องปรับอากาศเครื่องเดียวกัน
- แสงล้อมรอบ (ambient light) หมายถึง แสงแวดล้อมบริเวณภาพตามปกตินอกเหนือจากแสงไฟที่ตั้งใจส่องสว่างเพื่อการถ่ายภาพ ซึ่งปริมาณแสงจะส่งผลกระทบต่อคุณภาพของภาพที่ได้ โดยภาพที่เก็บทุกภาพเป็นภาพจากช่วงเวลาระหว่าง 16.00 น. ถึง 18.00 น. เพื่อควบคุมปริมาณแสง กรณีของห้องเรียนที่มีแสงล้อมรอบในระดับค่อนข้างมืดและมืดนั้น เกิดจากม่านกันแสง ทำให้แสงอาทิตย์เข้ามาในบริเวณถ่ายภาพได้น้อย

3.2 การจัดทำชุดข้อมูลสำหรับสร้างโมเดล (dataset)

ภาพทั้งหมดนั้นถูกแบ่งเป็น 2 กลุ่มใหญ่ คือ training set และ test set โดยที่ training set คือ set ที่ใช้ในการสร้างและทดสอบโมเดล ซึ่งประกอบไปด้วยภาพจากทั้ง 7 ชุดภาพ เฉพาะภาพจาก 50 วินาทีแรกเท่านั้น (1,500 ภาพต่อชุดย่อย รวมทั้งสิ้น 42,000 ภาพ) และ test set ซึ่งเป็น set ที่ถูกกั้นไว้แต่แรกไม่ให้โมเดลได้เห็นหรือใช้เลย เก็บไว้เพื่อใช้ประเมินประสิทธิภาพของโมเดล ประกอบไปด้วยภาพจาก 10 วินาทีที่เหลือหลังจากภาพใน training set (300 ภาพต่อชุดย่อย รวมทั้งสิ้น 8,400 ภาพ) ดังภาพที่ 1

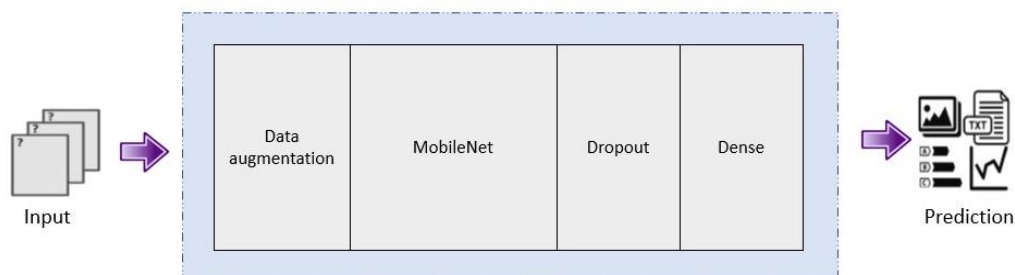
การแบ่งโดยใช้เวลาแบบนี้เหมาะสมกว่าการทำ cross-validation หรือ random sampling สำหรับงานวิจัยนี้ เนื่องจากแม้โมเดลจะไม่เห็นภาพใน validation/test set ในการทำ cross-validation แต่ในกรณีนี้โมเดลอาจได้เห็นภาพที่เกิดขึ้นหลังจากภาพใน test set ได้ เรียกว่า look-ahead bias [13] เช่น วินาทีที่ 5 ถูกสุ่มให้อยู่ใน test set แต่ใน training set มีภาพจากวินาทีที่ 6 ในทางตรงข้าม การกั้นโดยใช้เวลาสามารถรับรองว่าโมเดลไม่สามารถแอบดู (look-ahead) อนาคตได้เลยจาก training set ที่มี

เพื่อทดสอบผลของปริมาณต่อประสิทธิภาพของโมเดล ชุดข้อมูลจึงถูกจัดเป็น 3 ระดับ คือ 1. 30FPS คือ ชุดข้อมูลข้างต้น 2. 2FPS คือ ชุดข้อมูลที่มีภาพจากชุดข้อมูลข้างต้น แต่ใช้เพียง 2 ภาพต่อวินาที คือ ภาพขณะ frame หมายเลข 15 และ 30 ของแต่ละวินาที ทำให้มีภาพที่ใช้ใน training set ทั้งหมด 2,800 ภาพ จาก 42,000 ภาพ และ 3. 1FPS คือ ชุดข้อมูลที่มีภาพจากชุดข้อมูลข้างต้น แต่ใช้ภาพเพียง 1 ภาพต่อวินาที คือ ภาพขณะ frame หมายเลข 30 ของแต่ละวินาทีเท่านั้น ทำให้มีภาพที่ใช้ใน training set ทั้งหมด 1,400 ภาพ จาก 42,000 ภาพ

3.3 โครงข่ายประสาทเทียมแบบคอนโวลูชันและการส่งต่อการเรียนรู้

โมเดลที่เลือกใช้ในงานวิจัยนี้ คือ โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks, CNNs) ซึ่ง CNN นั้นเป็นรูปแบบเฉพาะแบบหนึ่งของ โครงข่ายประสาทเทียม (Artificial Neural Networks, ANNs) ที่มีประสิทธิภาพสูงเป็นพิเศษเมื่อใช้กับข้อมูลประเภทภาพและสัญญาณ [2] [10] [11] [12] โดยในงานวิจัยนี้ได้นำส่วนของ CNN model ขนาดใหญ่ที่ถูกสร้างขึ้นมาแล้วบนข้อมูลอื่น (pre-trained model) มาทำการส่งต่อการเรียนรู้ (transfer learning) เพื่อให้ได้โมเดลที่มีคุณภาพสูงและทำงานได้ดีกับชุดข้อมูลในงานวิจัยนี้

โมเดลถูกสร้างด้วย Keras ซึ่งเป็น API สำหรับสร้างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึกในภาษา Python ซึ่งสร้างโดยใช้ TensorFlow เป็นฐานในการสร้างโครงข่ายประสาทเทียมแบบคอนโวลูชัน และใช้ pre-trained model คือ mobilenet_v3_large_100_224 (MobileNet [11] version 3 โมเดลขนาดใหญ่ 1.00 depth multiplier และขนาดภาพ 224x224px) ซึ่งเป็นหนึ่งในโมเดลที่ได้รับความนิยมมากที่สุดในปัจจุบันเนื่องจากมีประสิทธิภาพสูงและสามารถส่งต่อการเรียนรู้ได้ดี ซึ่งในกระบวนการสร้างโมเดลมีการแบ่ง validation set ออกจาก training set แบบสุ่ม โดย validation set นั้นจะถูกใช้เพื่อประเมินโมเดลในระหว่างการสร้างภาพรวมการออกแบบโมเดลเป็นดังภาพที่ 3



ภาพที่ 3 โครงสร้าง CNN สำหรับตรวจจับการทำงานของเครื่องปรับอากาศ

โมเดลที่ถูกสร้างจากชุดข้อมูลทั้ง 3 ชุดนั้น จะถูกทดสอบ (test) บน test set เต็มทุกโมเดล คือ ใช้ทั้ง 8,400 ภาพ แม้จะถูกสร้างจาก 1FPS 2FPS และ 30FPS ซึ่งการใช้ test set เดียวกันเป็นการลอคคิ (bias) ที่อาจเกิดขึ้นได้จากการใช้ test set ที่ต่างกัน โดยประสิทธิภาพของโมเดลจะประเมินจากค่าความแม่นยำ (accuracy) ค่าลอส (binary cross-entropy หรือ log loss) และเวลาที่ใช้ในการสร้างโมเดล

3.4 การจัดการการโอเวอร์ฟิต

หนึ่งในสิ่งที่ต้องระวังที่สุดในการสร้างโมเดล คือ อย่าให้โมเดลจดจำตัวข้อมูลมากเกินไป หรือที่เรียกว่าการโอเวอร์ฟิต (overfitting) ซึ่งส่งผลเสีย คือ ทำให้โมเดลมีประสิทธิภาพต่ำลงมากเมื่อถูกนำไปใช้กับข้อมูลที่ไม่เคยเห็นมาก่อน ซึ่งโครงข่ายประสาทเทียมนั้นสามารถโอเวอร์ฟิตได้ง่ายมาก เนื่องจากเป็นโมเดลที่มีจำนวนพารามิเตอร์เยอะ

มาก [14] โดยเฉพาะอย่างยิ่งในกรณีของการส่งต่อการเรียนรู้ ซึ่งมักจะมีข้อมูลปริมาณน้อยหรือมีความหลากหลายต่ำ ในงานวิจัยนี้จึงใช้ 2 เทคนิคในการลดการโอเวอร์ฟิต คือ

1. การออกเมนต์ข้อมูล (data augmentation) คือ การแปลงข้อมูลที่ใช้สร้างโมเดลในแบบที่ไม่เปลี่ยนความถูกต้องของข้อมูล เพื่อเพิ่มปริมาณข้อมูลและทำให้ข้อมูลมีความหลากหลายเพิ่มขึ้น ซึ่งในงานวิจัยนี้ได้ทำการออกเมนต์ด้วยการหมุนภาพ (rotation) การขยับภาพทั้งแกนตั้งและแกนนอน (x-translation, y-translation) การซูม (zooming) และการสลับซ้ายขวา (horizontal flipping)
2. การทำดรอปเอาต์ (dropout) คือ เทคนิคการสร้างโครงข่ายประสาทเทียมแบบหนึ่งซึ่งจะทำการสุ่มดับโหนดในโครงข่ายระหว่างการสร้างโมเดลโดยการทำให้โหนดนั้น ๆ เป็นศูนย์ จากงานวิจัยพบว่าการสุ่มดับแบบนี้ ทำให้โมเดลมีแนวโน้มโอเวอร์ฟิตน้อยลง [15] โดยในโมเดลนี้ใช้การแทรก dropout layer

4. ผลการวิจัย

จากการทดลองสร้างโมเดลโดยใช้ข้อมูลที่ต่างกัน 3 ระดับ พบว่าการสร้างโมเดลเป็นไปด้วยดี คือ ผลการวัดค่าความแม่นยำ (accuracy) ระหว่างการสร้างมากกว่าการสุ่ม (chance) มีค่าลอสอยู่ในเกณฑ์ที่รับได้เมื่อทดลองกับ validation set ตามที่แสดงในตารางที่ 2 เมื่อเปรียบเทียบแล้วพบว่า การเพิ่มปริมาณข้อมูลทำให้โมเดลดีขึ้นเพียงเล็กน้อย แต่ใช้เวลาเพิ่มขึ้นอย่างมีนัย

ตารางที่ 2 ผลการสร้าง (train) model วัดบน training set และ validation set

Model	Training Time (seconds)	Training Set		Validation Set	
		Loss	Accuracy	Loss	Accuracy
1FPS	475	0.1671	0.6964	0.1150	0.7096
2FPS	857	0.1767	0.6938	0.1805	0.7714
30FPS	6026	0.1649	0.7039	0.1420	0.7043

ในการวัดประสิทธิภาพโดยการทดลองตรวจจับการทำงานของเครื่องปรับอากาศในภาพของ test set ได้มีการคำนวณค่าความแม่นยำสำหรับแต่ละชุดภาพ ดังตารางที่ 3 เพื่อให้สามารถตีความผลการวิจัยได้อย่างตรงจุด ซึ่งจากผลการทดลองพบว่า ทั้ง 3 โมเดลมีความเหมือนกัน คือ สามารถตรวจจับการทำงานในชุดภาพแบบใกล้เคียงได้มากกว่า 90% และมากถึง 100% ได้ในหลายชุดภาพแบบใกล้เคียง ยกเว้นเพียงสองกรณี คือ C_XX_D_ST_86 จากโมเดล 1FPS และ C_XX_D_ST_86 จากโมเดล 30FPS ที่ได้ค่าความแม่นยำ 72% และ 85% ตามลำดับ

ในทางกลับกัน ทุกโมเดลทำงานได้ไม่ดีในชุดภาพแบบไกล คือ ได้ค่าความแม่นยำเท่ากับหรือใกล้เคียงกับการสุ่ม (chance = 50%) ยกเว้น F_XX_B_ST_86 จากโมเดล 2FPS ที่ตรวจจับได้ถูกต้อง 100%

ตารางที่ 3 ประสิทธิภาพการทำงานของโมเดลบน test set แยกตามชุดภาพและระดับแสง

Model	C_XX_D_FL_COR	C_XX_B_FL_COR	C_XX_D_FLST_8	C_XX_B_FLST_8	C_XX_D_ST_86	C_XX_B_ST_86	F_XX_D_FL_72	F_XX_B_FL_72	F_XX_D_FL_COR	F_XX_B_FL_COR	F_XX_D_ST_86	F_XX_B_ST_86	F_XX_D_ST_75	F_XX_B_ST_75
1FPS	1.00	1.00	1.00	1.00	0.72	1.00	0.50	0.56	0.50	0.50	0.50	0.52	0.50	0.50
2FPS	1.00	1.00	1.00	1.00	1.00	0.99	0.50	0.50	0.50	0.50	0.64	1.00	0.68	0.50
30FPS	1.00	1.00	1.00	0.95	0.85	1.00	0.50	0.50	0.40	0.50	0.50	0.50	0.50	0.50

5. อภิปรายผลและข้อเสนอแนะการวิจัย

จากผลการทดลองสรุปได้ว่า การสร้างใช้โมเดลโครงข่ายประสาทเทียมในการตรวจจับสถานะเปิด-ปิดของเครื่องปรับอากาศนั้นสามารถทำได้ดีมากหากภาพที่ได้เป็นภาพมุมใกล้ คือ มีสัดส่วนของเครื่องปรับอากาศในภาพสูง ซึ่งสามารถใช้ได้กับเครื่องปรับอากาศแบบที่มีบานสวิงอยู่แล้ว หรือสามารถติดฟูลี่เพิ่มในกรณีของเครื่องที่ไม่มีบานสวิง ในทางตรงข้ามโมเดลจะทำงานได้ไม่ดีสำหรับภาพมุมไกล ซึ่งสันนิษฐานว่าเกิดจากการที่ MobileNet นั้นรับ input ขนาด 224 x 224px ทำให้ภาพต้นฉบับที่ขนาดใหญ่กว่าต้องถูกย่อให้มีขนาดเท่ากับ 224 x 224px ซึ่งในกรณีมูไกลนั้น สัดส่วนของบานสวิงและฟูลี่นั้นน้อยอยู่แล้ว เมื่อถูกย่อขนาดอาจทำให้ถูกสืรอบข้างหรือความผันผวน (noise) กลืนหายไป ทำให้ไม่สามารถตรวจจับได้

แม้ว่าโมเดลในงานวิจัยนี้จะไม่สามารถตรวจจับการทำงานจากภาพมุมไกลได้ดี แต่ในมุมมองคอมพิวเตอร์วิทัศน์แล้ว สามารถใช้กระบวนการตรวจจับวัตถุ (object detection) ได้ ซึ่งเป็นกระบวนการในการตรวจจับว่า ในภาพนี้มีวัตถุที่สนใจอยู่ส่วนไหนของภาพ ดังนั้น หากนำกระบวนการนี้มาใช้เป็นส่วนหนึ่งของการเตรียมข้อมูลก่อนเข้าโมเดล (preprocessing) ตรวจจับการทำงาน คือ เริ่มจากการตรวจจับว่าเครื่องปรับอากาศอยู่ส่วนไหนของภาพ แล้วตัดเฉพาะส่วนนั้นของภาพมาใช้ในการตรวจจับว่าเครื่องปรับอากาศกำลังทำงานอยู่หรือไม่ วิธีนี้จะทำให้สัดส่วนของบานสวิงและฟูลี่มากขึ้นจากการตัดส่วนอื่นของภาพออก ทำให้ถูกกลืนหายได้ยากขึ้นเมื่อภาพถูกย่อขนาด

เนื่องจากภาพเครื่องปรับอากาศที่ใช้มีข้อมูลเชิงเวลา (time series) ดังนั้น การใช้ข้อมูลเชิงเวลาอาจช่วยให้โมเดลมีความแม่นยำมากขึ้นและสามารถนำไปใช้กับกรณีอื่น ๆ ได้ดีขึ้น เช่น โดยใช้โครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Networks, RNNs) ร่วมกับโครงข่ายประสาทเทียมแบบคอนโวลูชัน

คำขอบคุณ

ขอขอบคุณ ศูนย์นวัตกรรมเทคโนโลยีดิจิทัลอัจฉริยะ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีสุวรรณภูมิที่สนับสนุนการดำเนินการ ทำให้งานวิจัยในครั้งนี้สำเร็จลุล่วงไปได้ด้วยดี

6. เอกสารอ้างอิง

- [1] McNeil, M. A., & Letschert, V. E. (2008). **Future air conditioning energy consumption in developing countries and what can be done about it: the potential of efficiency in the residential sector.**
- [2] Luo, H., Xiong, C., Fang, W., Love, P. E., Zhang, B., & Ouyang, X. (2018). **Convolutional neural networks: Computer vision-based workforce activity assessment in construction.** *Automation in Construction*, 94, 282-289.
- [3] Srijakkot, K., Kanjanasurat, I., Wiriakrieng, N., Lartwatechakul, M., & Benjangkaprasert, C. (2020). **Intruder Detection by Using Faster R-CNN in Power Substation.** In *Recent Advances in Information and Communication Technology 2020: Proceedings of the 16th International Conference on Computing and Information Technology (IC2IT 2020)* (pp. 159-167). Springer International Publishing.
- [4] Harjoseputro, Y., Yuda, I., & Danukusumo, K. P. (2020). **MobileNets: Efficient convolutional neural network for identification of protected birds.** *IJASEIT (International Journal on Advanced Science, Engineering and Information Technology)*, 10(6), 2290-2296.
- [5] Favorskaya, M., & Pakhirka, A. (2019). **Animal species recognition in the wildlife based on muzzle and shape features using joint CNN.** *Procedia Computer Science*, 159, 933-942.
- [6] Qu, D., Huang, Z., Gao, Z., Zhao, Y., & Song, G. (2018, December). **An automatic system for smile recognition based on CNN and face detection.** In *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (pp. 243-247). IEEE.
- [7] Akhand, M. A. H., Roy, S., Siddique, N., Kamal, M. A. S., & Shimamura, T. (2021). **Facial emotion recognition using transfer learning in the deep CNN.** *Electronics*, 10(9), 1036.
- [8] Kaur, G., Sinha, R., Tiwari, P. K., Yadav, S. K., Pandey, P., Raj, R., ... & Rakhra, M. (2022). **Face mask recognition system using CNN model.** *Neuroscience Informatics*, 2(3), 100035.
- [9] Torrey, L., & Shavlik, J. (2010). **Transfer learning.** In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI global.
- [10] Tan, M., & Le, Q. (2019, May). **Efficientnet: Rethinking model scaling for convolutional neural networks.** In *International conference on machine learning* (pp. 6105-6114). PMLR.
- [11] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). **Mobilenets: Efficient convolutional neural networks for mobile vision applications.** *arXiv preprint arXiv:1704.04861*.
- [12] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). **Imagenet: A large-scale hierarchical image database.** In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). IEEE.

- [13] Yae, J. (2022). **Unintended look-ahead bias in out-of-sample forecasting**. *Applied Economics Letters*, 1-5.
- [14] Ying, X. (2019, February). **An overview of overfitting and its solutions**. In *Journal of physics: Conference series* (Vol. 1168, p. 022022). IOP Publishing.
- [15] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). **Dropout: a simple way to prevent neural networks from overfitting**. *The journal of machine learning research*, 15(1), 1929-1958.