

การตรวจจับและจำแนกยานพาหนะด้วยอัลกอริทึม Deep Learning สำหรับระบบวิดีโอ รักษาความปลอดภัย

Detection and Classification of Vehicles using Deep Learning Algorithm for Video Surveillance Systems

พระระติชัย ไวรโรจนะพุทธร (Phonratichi Wairotchanaphuttha) ณรงค์ บุญศิริสัมพันธ์ (Narong Boonsirisumpun)

และวิชัย พัวร์งโรจน์ (Wichai Puarungroj)

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยราชภัฏเลย

sb5840248128@lru.ac.th, narong.boo@lru.ac.th, wichai@lru.ac.th

บทคัดย่อ

การตรวจจับและจำแนกยานพาหนะ (Vehicle Detection and Classification) จากระบบวิดีโอรักษาความปลอดภัยเป็นการประยุกต์ใช้งานเทคนิคทางวิทยาการคอมพิวเตอร์ที่มีความสำคัญอย่างยิ่งในปัจจุบัน โดยในมหาวิทยาลัยราชภัฏเลยมีจำนวนยานพาหนะที่เป็นรถจักรยานยนต์และทำผิดกฎจราจรเป็นจำนวนมากจนเกินความสามารถที่พนักงานรักษาความปลอดภัยจะตรวจตราได้ครบถ้วน การใช้ระบบตรวจจับอัตโนมัติจึงมีความจำเป็นที่จะช่วยให้การตรวจจับการกระทำผิดกฎจราจรและการรักษาความปลอดภัยมีประสิทธิภาพที่ดีขึ้นในงานวิจัยนี้มีจุดประสงค์ที่จะนำเทคนิคที่น่าสนใจทางวิทยาการคอมพิวเตอร์ที่เรียกว่า การเรียนรู้เชิงลึก (Deep Learning) มาประยุกต์ใช้ในการพัฒนาระบบอัตโนมัติสำหรับตรวจจับและแยกแยะรถจักรยานยนต์ออกจากยานพาหนะประเภทอื่นๆ เพื่อแก้ไขปัญหาดังกล่าวและนำข้อมูลที่ได้ไปใช้ในการตรวจจับการทำผิดกฎจราจรต่อไป

คำสำคัญ: การตรวจจับและแยกแยะยานพาหนะ การตรวจจับรถจักรยานยนต์ การเรียนรู้เชิงลึก วิธีรักษาความปลอดภัย

Abstract

Video Detection and Classification from Video Surveillance system is very important application from computer science techniques in nowadays. In Loei Rajabhat University (LRU), there are too many vehicles, especially motorcycles, those abused the traffic rules and

the security staffs were unable to detect all of them. Using an autonomous detection system become necessary to improve the performance of LRU's security and traffic system. In this paper, we proposed to apply some interesting computer science techniques call Deep Learning to implement the vehicle detection and classification autonomous system that can separate motorcycle from other vehicles to solve this problem and utilize the data into the traffic rules offence detection system in the next step.

Keyword: Vehicles Detection and Classification, Motorcycle Detection, Deep Learning, Video Surveillance

1. บทนำ

การตรวจจับและจำแนกวัตถุ (Object Detection and Classification) บนภาพนิ่งหรือภาพเคลื่อนไหวมีความจำเป็นอย่างยิ่งในปัจจุบัน ช่วยให้ผู้ใช้งานสามารถวิเคราะห์แยกแยะวัตถุต่างๆ ที่สนใจและนำผลลัพธ์ที่ได้ไปใช้ประโยชน์ต่อไป ตัวอย่างการใช้ประโยชน์จากการประยุกต์ใช้การตรวจจับและแยกแยะวัตถุ เช่น การจำแนกหน้าบุคคลจากภาพเคลื่อนไหว (Human Face Detection) [1] การจำแนกรอยโรค (Lesion) จากภาพถ่ายด้วยเทคนิคแมมโมแกรม (Mammogram) [2] การตรวจนับยานพาหนะที่ขับผ่านประตู (Vehicle Counting) [3] การนับจำนวนสัตว์ในฟาร์ม (Animal Counting) [4] การตรวจจับพฤติกรรมบุคคลจากกล้องวิดีโอรักษาความปลอดภัย [5] เป็น

ต้น การประยุกต์ใช้เทคนิคต่างๆ ทางคอมพิวเตอร์ช่วยให้การตรวจจับและแยกแยะวัตถุสามารถทำได้อย่างอัตโนมัติ โดยไม่ต้องอาศัยแรงงานมนุษย์ในการดำเนินการ ซึ่งมีโอกาสที่จะเกิดความผิดพลาดได้สูง งานวิจัยนี้มีวัตถุประสงค์ที่จะประยุกต์ใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ในการตรวจจับและแยกแยะรถจักรยานยนต์จากยานพาหนะประเภทอื่นๆ เพื่อนำข้อมูลที่ได้นำไปประยุกต์ใช้ในการตรวจรักษาความปลอดภัยและการทำผิดกฎจราจรจากระบบกล้องวงจรปิดภายในมหาวิทยาลัยราชภัฏเลย อาทิ เช่น การจับที่รถจักรยานยนต์โดยไม่ใส่หมวกกันน็อก การขับที่รถย้อนศร หรือการจอดจบละเมิดระเบียบของนักศึกษา มหาวิทยาลัยราชภัฏเลย เป็นต้น โดยจากผลการทดลองเบื้องต้นในงานวิจัยนี้พบว่าเทคนิคการเรียนรู้เชิงลึกมีความสามารถในการตรวจจับและแยกแยะรถจักรยานยนต์ออกจากยานพาหนะอื่นๆ ทั้งจากภาพนิ่งและภาพวิดีโอโดยมีความถูกต้องค่อนข้างสูงถึง 90%

2. วรรณกรรมที่เกี่ยวข้อง

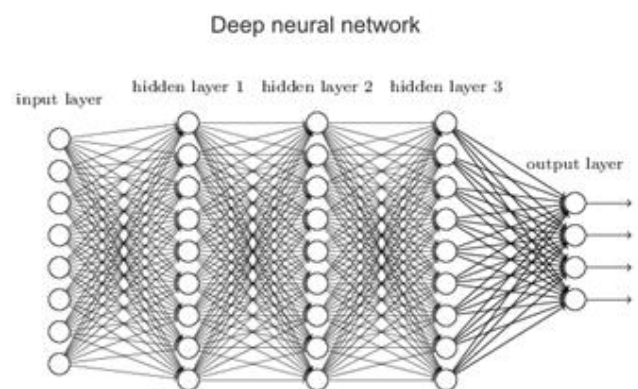
2.1 การตรวจจับและแยกแยะวัตถุในภาพ

การตรวจจับและแยกแยะวัตถุในภาพ (Object Detection) คือเทคนิคอย่างหนึ่งทางด้านคอมพิวเตอร์วิทัศน์ (Computer Vision) และ การประมวลผลภาพ (Image Processing) [6] ที่ถูกนำมาใช้ในการตรวจจับขอบเขตของวัตถุหรือสิ่งที่เราสนใจในรูปภาพดิจิทัล วัตถุในภาพแต่ละชิ้นจะผ่านกระบวนการหลายขั้นตอน ก่อนที่จะถูกจะวิเคราะห์และสรุปผลออกมาว่าเป็นวัตถุประเภทใดๆ กระบวนการต่างๆ เหล่านี้อาจประกอบไปด้วย การกรองภาพ (Image Filtering) การแบ่งส่วนรูปภาพ (Image Segmentation) การสกัดคุณลักษณะ (Feature Extraction) การรู้จำรูปภาพ (Image Recognition) หรือ การเรียนรู้ของเครื่อง (Machine Learning) เป็นต้น

โดยเทคนิคเกี่ยวกับการตรวจจับและแยกแยะวัตถุในภาพได้ถูกนำไปประยุกต์ใช้ในงานหลายๆ ด้าน อาทิเช่น การตรวจจับการเกิดอุบัติเหตุบนท้องถนน [7] การจำแนกภาพแมวโมแกรมโดยใช้การประมวลผลภาพ [8] ซึ่งการประยุกต์ใช้งานเหล่านี้ทำให้คอมพิวเตอร์เข้าใจข้อมูลของวัตถุต่างๆ รวมทั้งระบุตำแหน่งของวัตถุนั้นๆ ในรูปภาพได้อีกด้วย

2.2 การเรียนรู้เชิงลึก

การเรียนรู้เชิงลึก (Deep Learning) เป็นเทคนิคหนึ่งทางด้านการเรียนรู้ของเครื่อง (Machine Learning) [9, 10, 11] ที่ถูกพัฒนาขึ้นมาจากอยู่บนพื้นฐานของโครงข่ายประสาทเทียมแบบลึก (Deep Artificial Neural Networks) หรือ DNN (ภาพที่ 1) ซึ่งอาจจะอยู่ในรูปของโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks) หรือ CNN หรือว่าจะเป็นโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) หรือ RNN ก็ได้



ภาพที่ 1: โครงข่ายประสาทเทียมแบบลึก

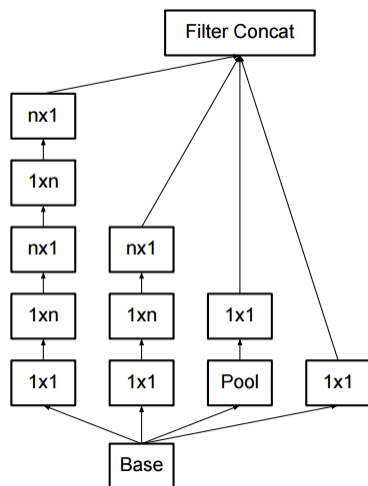
ที่มา: <http://neuralnetworksanddeeplearning.com/>

เทคนิคของการเรียนรู้เชิงลึกได้ถูกนำไปใช้แก้ไขปัญหาในหลายๆ ด้าน โดยเฉพาะอย่างยิ่งด้านที่เกี่ยวกับปัญญาประดิษฐ์ (Artificial Intelligence) และ การประมวลผลรูปภาพ (Image Processing) ซึ่งรวมทั้งการตรวจจับและแยกแยะวัตถุในรูปภาพด้วย

2.3 Inception_v3 และ MobileNets

นับตั้งแต่การเรียนรู้เชิงลึกและโครงข่ายประสาทเทียมแบบลึกได้ถูกพัฒนาขึ้น การประยุกต์ใช้การเรียนรู้เชิงลึกในด้านการตรวจจับและแยกแยะวัตถุในภาพหรือการรู้จำรูปภาพ (Image Recognition) ได้ถูกพัฒนาไปอย่างกว้างขวาง มีผู้พัฒนาหลายกลุ่มได้สร้างโมเดลของโครงข่ายประสาทเทียมแบบลึกในแบบของตนเองขึ้นมา อาทิเช่น “inception_v3” ของทีมผู้พัฒนา ชื่อ

ว่า “GoogLeNet” [12] เป็นโครงข่ายประสาทเทียมแบบลึกที่ถูกพัฒนาให้มีจำนวน 22 ชั้น (Layers) นอกจากนี้ในหน่วยย่อยของบางชั้นจะประกอบไปด้วย โมดูลตั้งต้น (Inception Module) ที่จะทำให้หน้าที่ในการก่อสร้างหน่วยสกัดคุณลักษณะที่ดีที่สุดในแต่ละชั้นของโครงข่าย (ภาพที่ 2) เพื่อให้สามารถทำการตรวจสอบและแยกแยะองค์ประกอบโดยรวมของวัตถุแต่ละชนิดได้ดียิ่งขึ้น



ภาพที่ 2: inception module ของ inception_v3 [12]

ยังมีโครงข่ายประสาทเทียมแบบลึกอีกรูปแบบหนึ่งที่ได้ถูกพัฒนาขึ้นมามีชื่อว่า “MobileNets” ซึ่งเป็นโครงข่ายประสาทเทียมแบบลึกที่ถูกพัฒนาขึ้นโดย Andrew G. Howard และคณะ ในปี 2560 โดยรูปแบบของ MobileNets ถูกพัฒนาให้เป็นโครงข่ายประสาทเทียมแบบลึกแต่มีขนาดเล็กกว่าโครงข่ายประสาทเทียมแบบลึกทั่วไป เพื่อให้สามารถนำโมเดลไปใช้งานบนอุปกรณ์คอมพิวเตอร์ขนาดเล็ก อาทิเช่น สมาร์ทโฟน หรือ แท็บเล็ต และยังสามารถรักษาประสิทธิภาพในการทำงานไว้ได้ในระดับที่ใกล้เคียงกับโครงข่ายประสาทเทียมแบบลึกขนาดใหญ่ โดยจากการทดสอบของผู้พัฒนา MobileNets มีความถูกต้องในการแยกแยะวัตถุในรูปภาพที่ใกล้เคียงกับ inception_v3 คือ 83.3% และ 84% ตามลำดับ [13]

2.4 SSD (Single Shot multibox Detector)

SSD หรือ Single Shot multibox Detector [14] คือเทคนิคทางด้านการประมวลผลรูปภาพ ถูกใช้ในการวิเคราะห์ขอบเขตของวัตถุในรูปภาพโดยอาศัยการจับกลุ่มของคุณลักษณะย่อย

ต่างๆ ของแต่ละวัตถุในภาพแล้วทำการวิเคราะห์คะแนนเพื่อกำหนดขอบเขตของแต่ละวัตถุออกมา ซึ่งเทคนิค SSD สามารถนำมาประยุกต์ใช้ร่วมกับโครงข่ายประสาทเทียมแบบลึก อาทิเช่น “MobileNets” หรือ “Inception” เพื่อให้สามารถทำการตรวจจับวัตถุในรูปภาพ โดยสามารถที่จะบอกทั้งประเภทของวัตถุ ขนาดและตำแหน่งของวัตถุในภาพ รวมทั้งแสดงคะแนนความน่าจะเป็นที่แต่ละขอบเขตของแต่ละภาพจะถูกตัดสินว่าเป็นวัตถุชนิดใดๆ ด้วย (ภาพที่ 3)



ภาพที่ 3: ผลลัพธ์การวิเคราะห์วัตถุในภาพโดยใช้เทคนิค SSD และโครงข่าย MobileNets [13]

3. วิธีการดำเนินงานวิจัย

ทางผู้วิจัยมีจุดประสงค์ที่จะนำเทคนิคทางด้านการเรียนรู้เชิงลึกรวมทั้งเทคนิคต่างๆ ที่เกี่ยวข้องมาใช้ในการตรวจจับและแยกแยะวัตถุประเภทยานพาหนะภายในมหาวิทยาลัยราชภัฏเลย โดยเน้นไปที่การตรวจจับรถยนต์และรถจักรยานยนต์ที่มาจากรูปภาพทั้งภาพนิ่งและภาพวิดีโอ และต้องการให้ระบบสามารถแยกแยะภาพรถจักรยานยนต์ออกจากกรณยนต์ได้ จึงได้ออกแบบการดำเนินงานวิจัยไว้ดังนี้ 1. การรวบรวมข้อมูลภาพยานพาหนะภายในมหาวิทยาลัยราชภัฏเลย 2. การผสมข้อมูลภาพเข้ากับข้อมูลภาพมาตรฐานของต่างประเทศ 3. การทดสอบกับภาพนิ่ง 4. การทดสอบกับภาพวิดีโอ 5. การวิเคราะห์ผลการทดลอง

3.1 ข้อมูลภาพยานพาหนะภายในมหาวิทยาลัยราชภัฏเลย

ทางคณะผู้วิจัยได้ทำการรวบรวมรูปภาพยานพาหนะภายในมหาวิทยาลัยราชภัฏเลย (LRU_Vehicle) เพื่อนำมาใช้เป็นชุดข้อมูลในการเรียนรู้ (Training Data) ของโครงข่ายประสาทเทียมทั้งแบบ inception_v3 และ MobileNets โดยเน้นไปที่ภาพพาหนะประเภทรถยนต์และรถจักรยานยนต์ ซึ่งในส่วนของการ

LRU_Vehicle นั้น ทางคณะผู้วิจัยได้ทำการรวบรวมภาพรถยนต์ และรถจักรยานยนต์มาจำนวนทั้งสิ้นอย่างละ 100 ภาพ ตัวอย่าง ภาพบางส่วนที่ผู้วิจัยได้รวบรวมมาถูกแสดงอยู่ในภาพที่ 4



ภาพที่ 4: ตัวอย่างภาพที่ใช้ในการเรียนรู้ของโครงข่ายจาก LRU_Vehicle (ซ้าย รถจักรยานยนต์, ขวา รถยนต์)

3.2 ข้อมูลภาพจากต่างประเทศ (COCO)

เนื่องจากภาพที่ได้มาจากภายในมหาวิทยาลัยมีจำนวนน้อย และไม่หลากหลายพอ ทางผู้พัฒนาจึงได้ทำการหาคัดเลือกข้อมูลรูปภาพจากฐานข้อมูลรูปภาพที่เป็นมาตรฐานที่ถูกใช้ในการทดสอบกับโครงข่ายประสาทเทียมแบบลึกจากต่างประเทศมาผสมกับข้อมูลรูปภาพยานพาหนะจาก LRU_Vehicle เพื่อเพิ่มชุดข้อมูลในการเรียน โดยฐานข้อมูลรูปภาพที่ทางผู้วิจัยเลือกคือ ฐานข้อมูลที่มีชื่อว่า COCO (Common Objects in Context) [15] โดย COCO เป็นฐานข้อมูลรูปภาพขนาดใหญ่ที่มีจำนวนรูปภาพทั้งหมดมากกว่า 300,000 รูปภาพ จากประเภทของวัตถุทั้งสิ้น 90 ชนิด

ในการทดลองนี้ทางผู้วิจัยได้สุ่มเลือกรูปภาพรถยนต์และรถจักรยานยนต์ออกมาจากฐานข้อมูล COCO เป็นจำนวนทั้งสิ้นอย่างละ 100 ภาพ เพื่อให้มีขนาดจำนวนรูปภาพเท่ากับรูปที่มาจาก LRU_Vehicle ทำให้มีจำนวนรูปภาพที่จะนำมาเป็นข้อมูลในการเรียนของโครงข่ายรวมเป็น ภาพรถยนต์จำนวน 200 รูป และ ภาพรถจักรยานยนต์จำนวน 200 รูป โดยตัวอย่างภาพที่สุ่มเลือกมาจาก COCO ถูกแสดงไว้ในภาพที่ 5



ภาพที่ 5: ตัวอย่างภาพที่สุ่มเลือกมาจาก COCO (ซ้าย รถจักรยานยนต์, ขวา รถยนต์)

3.3 การทดสอบกับภาพนิ่ง

หลังจากที่ได้ทำการรวบรวมข้อมูลรูปภาพมาจำนวนทั้งสิ้น 400 รูปแล้ว ทางผู้วิจัยได้ทำการนำรูปที่ได้มาแบ่งเป็นชุดข้อมูลเรียน (Training Data) และ ข้อมูลทดสอบ (Testing Data) ของโครงข่ายประสาทเทียมแบบลึกทั้งสองแบบ คือ inception_v3 และ MobileNets โดยผู้วิจัยได้พัฒนาโปรแกรมด้วยภาษาไพธอน และไลบรารีมาตรฐานที่มีชื่อว่า เทนเซอร์โฟลว์ (TensorFlow) ซึ่งไลบรารีตัวนี้ได้ทำการแบ่งชุดข้อมูลรูปภาพออกเป็นสองส่วนคือชุดข้อมูลเรียนและชุดข้อมูลทดสอบให้โดยอัตโนมัติ โดยในแต่ละครั้งที่ทดสอบการเรียนรู้ของโมเดล โปรแกรมจะทำการแบ่งข้อมูลออกมาเป็นข้อมูลเรียนเป็นจำนวน 90% และข้อมูลทดสอบจำนวน 10% ของข้อมูลทั้งหมดโดยประมาณ ซึ่งผู้วิจัยได้ทำการทดสอบการเรียนรู้ของแต่ละโครงข่ายเป็นจำนวนทั้งสิ้น 3 ครั้ง เพื่อนำค่าความถูกต้องที่ได้มาเฉลี่ยและเปรียบเทียบกับประสิทธิภาพ ทั้งในแง่ของความถูกต้องและขนาดของโมเดลที่ได้จากการเรียนรู้

3.4 การทดสอบกับภาพวิดีโอ

หลังจากทดสอบการตรวจจับและแยกแยะยานพาหนะจากภาพนิ่งแล้ว ในขั้นตอนต่อไป ผู้วิจัยได้ทำการทดสอบการตรวจจับและแยกแยะวัตถุจากภาพวิดีโอ โดยผู้วิจัยจะเลือกโครงข่ายที่มีความเหมาะสมจากการทดสอบจากภาพนิ่งมาผสมรวมกับเทคนิคของ SSD เพื่อนำไปทดสอบกับภาพวิดีโอ ซึ่งเป็นภาพวิดีโอทดสอบที่ได้มาจากการตั้งกล้องถ่ายการจราจรบนถนนภายในมหาวิทยาลัยราชภัฏเลยในมุมต่างๆ จำนวนทั้งสิ้น 10 วิดีโอๆ ละ 10 วินาที รวมเวลาทั้งสิ้น 100 วินาที

ในขั้นตอนนี้ผู้วิจัยจะได้ทำการตรวจสอบความถูกต้องในแต่ละวิดีโอด้วยว่า ระบบตรวจจับสามารถตรวจจับยานพาหนะทั้งรถยนต์และรถจักรยานยนต์ได้ถูกต้องหรือไม่ และสามารถแยกแยะประเภทรูปภาพรถจักรยานยนต์ออกมาจากรถยนต์ได้ถูกต้องหรือไม่

3.5 การวิเคราะห์ผลการทดลอง

หลังจากทำการทดลองตาม 4 ขั้นตอนที่กล่าวมาข้างต้นแล้ว ทางคณะผู้วิจัยจะได้ทำการสรุปผล และวิเคราะห์ผลที่ได้จากการทดลอง โดยผลที่ได้จากการทดลองทั้งหมดถูกนำเสนออยู่ในหัวข้อถัดไป

4. ผลการดำเนินงาน

4.1 การทดสอบกับภาพนิ่ง

ตารางที่ 1 แสดงผลการทดลองที่ได้จากการนำรูปภาพทั้งหมด 400 ภาพ มาทำการเรียนรู้กับโครงข่ายประสาทเทียมแบบ inception_v3 และ MobileNets เพื่อเปรียบเทียบประสิทธิภาพในการตรวจจับและแยกแยะยานพาหนะจากรูปภาพ

ตารางที่ 1: ผลการตรวจจับและแยกแยะยานพาหนะจากรูปภาพนิ่ง

ชื่อโครงข่าย	ความถูกต้อง (%)	ขนาดของโมเดล (KB)
inception_v3	91.50	85,383
MobileNets	91.33	16,723

จากตารางที่ 1 ผู้วิจัยพบว่าความถูกต้องในการตรวจจับและแยกแยะรูปภาพรถยนต์และรถจักรยานยนต์ในภาพนิ่งมีความถูกต้องที่ใกล้เคียงกันในทั้งสองโครงข่าย โดย inception_v3 มีค่าความถูกต้อง 91.50% และ MobileNets มีค่าความถูกต้อง 91.33% ซึ่งทางผู้วิจัยถือว่าทั้งสองโครงข่ายมีประสิทธิภาพใกล้เคียงกัน แต่ถ้ามองจากขนาดของโมเดลที่ได้จากการเรียนรู้ของโครงข่ายจะเห็นว่า MobileNets มีขนาดของโมเดลเล็กกว่า inception_v3 อย่างเห็นได้ชัด โดยมีขนาดเพียงแค่ 16,723 กิโลไบต์ (KB) หรือว่าเพียงหนึ่งในห้าส่วนของ inception_v3 เท่านั้น

ดังนั้นทางผู้วิจัยจึงได้เลือกที่จะนำโครงข่าย MobileNets ไปผสมกับเทคนิค SSD เพื่อไปทดสอบกับข้อมูลที่เป็นภาพวิดีโอต่อไป

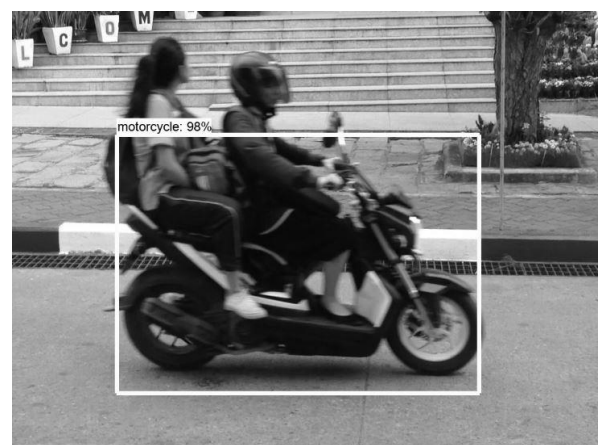
4.2 การทดสอบกับภาพวิดีโอ

ในตารางที่ 2 ได้แสดงผลการทดลองที่ได้จากทดสอบการตรวจจับและแยกแยะยานพาหนะในภาพวิดีโอด้วยโครงข่ายประสาทเทียม MobileNets ผสมกับเทคนิค SSD และได้แสดงภาพตัวอย่างการตรวจจับรถจักรยานยนต์ไว้ในภาพที่ 6 จากผลการทดลองจะเห็นว่ามีจำนวนรถยนต์ทั้งหมดที่อยู่ในวิดีโอทดสอบจำนวน 5 คัน และ จำนวนรถจักรยานยนต์ทั้งหมด 13 คัน โดยเมื่อทำการทดสอบภาพวิดีโอ ปรากฏว่ามีจำนวนรถยนต์ในภาพที่ถูกพบแบบถูกต้องตามประเภทจำนวนทั้งสิ้น 5 คัน คิดเป็นจำนวน 100% ในขณะที่ยานพาหนะประเภทรถจักรยานยนต์มีจำนวนที่ถูกพบแบบถูกต้องตามประเภทจำนวนทั้งสิ้น 8 คัน คิดเป็น 61.53% ในขณะที่ มีจำนวน

รถจักรยานยนต์ที่ถูกตรวจพบและระบุประเภทผิดคือระบุว่า เป็นรถยนต์อยู่ถึง 5 คัน คิดเป็น 38.47% ซึ่งถือว่าเป็นจำนวนที่สูง

ตารางที่ 2: ผลการตรวจจับและแยกแยะยานพาหนะจากภาพวิดีโอ

ยานพาหนะ	จำนวนที่มี (คัน)	จำนวนที่พบ ถูกประเภท (คัน)	จำนวนที่พบ ผิดประเภท (คัน)
รถยนต์	5	5	-
รถจักรยานยนต์	13	8	5

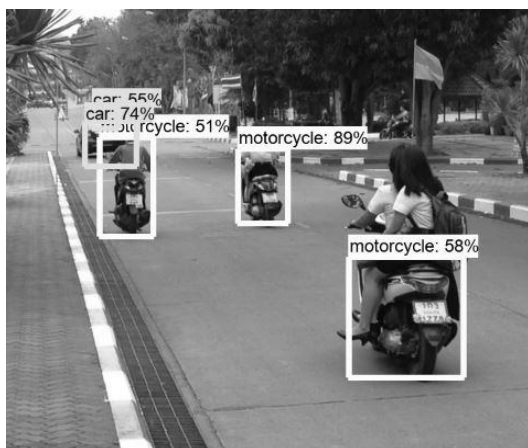


ภาพที่ 6: ตัวอย่างภาพการตรวจจับรถจักรยานยนต์

แต่เมื่อทำการตรวจสอบลงไปในรายละเอียดจะพบว่า จำนวนของรถจักรยานยนต์ที่ถูกตรวจจับผิดประเภททั้ง 5 คัน เป็นภาพของรถจักรยานยนต์ที่อยู่ในระยะไกลทั้งสิ้น (ภาพที่ 7) แต่รถจักรยานยนต์ที่อยู่ในระยะใกล้กลับมามีความสามารถตรวจจับได้อย่างถูกต้องทั้ง 8 ภาพ (ภาพที่ 8) ซึ่งอาจจะวิเคราะห์ได้ว่า ระบบตรวจจับที่พัฒนาขึ้นโดยโครงข่าย MobileNets รวมกับ SSD สามารถตรวจจับรถจักรยานยนต์ได้จากระยะใกล้ด้วยความถูกต้องที่สูง แต่มีข้อผิดพลาดเมื่อภาพรถจักรยานยนต์อยู่ไกล และจะถูกจำแนกว่าเป็นรถยนต์



ภาพที่ 7: ตัวอย่างภาพการตรวจจับรถจักรยานยนต์ระยะไกล



ภาพที่ 8: ตัวอย่างภาพการตรวจจับรถจักรยานยนต์ระยะใกล้

5. สรุป

จากการทดลองจะพบว่าเทคนิคของการเรียนรู้เชิงลึก โดยใช้โครงข่ายประสาทเทียมแบบ inception_v3 และ MobileNets มีประสิทธิภาพในการตรวจจับและแยกแยะวัตถุประเภทรถยนต์และรถจักรยานยนต์ที่ใกล้เคียงกัน แต่ MobileNets มีขนาดโมเดลที่เล็กกว่ามาก ในส่วนของการทดสอบกับวิดีโอ การจำแนกรถยนต์มีความถูกต้องค่อนข้างสูง แต่การแยกแยะรถจักรยานยนต์ยังมีความผิดพลาดอยู่สำหรับรถจักรยานยนต์ที่อยู่ในระยะไกลของภาพ

ในงานวิจัยต่อไป ทางคณะผู้วิจัยจะได้ทำการเพิ่มรูปภาพรถจักรยานยนต์ที่อยู่ในระยะไกลของภาพเข้ามาอยู่ในชุดข้อมูลเรียนเพิ่มเติม เพื่อเพิ่มประสิทธิภาพในการตรวจจับ และจะได้ทำการเพิ่มชุดข้อมูลของผู้ขับขี่ที่กำลังโดยสารรถจักรยานยนต์อยู่ด้วย เพื่อจะได้ทำการตรวจจับผู้ขับขี่ที่สวมหมวกและไม่สวมหมวกกันน็อก เพื่อแยกแยะหาผู้ขับขี่รถจักรยานยนต์ที่กระทำผิดกฎจราจรในบริเวณมหาวิทยาลัยในขั้นตอนต่อไป

เอกสารอ้างอิง

- [1] Sung, K-K., and Tomaso Poggio. "Example-based learning for view-based human face detection." *IEEE Transactions on pattern analysis and machine intelligence* 20, no. 1, pp.39-51, 1998.
- [2] Al-Masni, M. A., M. A. Al-Antari, J. M. Park, G. Gi, T. Y. Kim, P. Rivera, E. Valarezo, S-M. Han, and T-S. Kim. "Detection and classification of the breast abnormalities in digital mammograms via regional Convolutional Neural Network." *In 39th Annual International Conference of Engineering in Medicine and Biology Society 2017 (EMBC)*, pp. 1230-1233. IEEE, 2017.
- [3] Liu, Xiaoming, Peter Henry Tu, Jens Rittscher, Amitha Perera, and Nils Krahnstoeber. "Detecting and counting people in surveillance applications." *In Conference on Advanced Video and Signal Based Surveillance, 2005. AVSS 2005*. pp. 306-311. IEEE, 2005.
- [4] Wichmann, Felix A., Jan Drewes, Pedro Rosas, and Karl R. Gegenfurtner. "Animal detection in natural scenes: critical features revisited." *Journal of Vision* 10, no. 4 (2010): 6-6.
- [5] Ovsenik, Luboš, A. Kažimírová Kolesárová, and Ján Turan. "Video surveillance systems." *Acta Electrotechnica et Informatica* 10, no. 4 (2010): 46-53.
- [6] Gonzalez, Rafael C., and Richard E. 1954- Woods. *Digital Image Processing. 3rd ed.* Upper Saddle River, N.J.: Prentice Hall, 2008.
- [7] ไกรศักดิ์ เกษร, "เทคนิคการวิเคราะห์วีดิทัศน์เพื่อตรวจจับการเกิดอุบัติเหตุบนถนน" *วารสารเทคโนโลยีสารสนเทศ* ปีที่ 11 ฉบับที่ 2 ก.ค.-ธ.ค. 2558 หน้า 24-33.
- [8] เกตุกาญจน์ ไชยจันทร์, "การจำแนกภาพแมวโมแกรมโดยใช้การประมวลผลภาพร่วมกับซอฟต์แวร์เวกเตอร์แมชชีน", *มหาวิทยาลัยเทคโนโลยีสุรนารี*, 2558.
- [9] Schmidhuber, J. "Deep Learning in Neural Networks: An Overview". *Neural Networks*. 61: pp.85–117. 2015.
- [10] Krizhevsky, A., Sutskever, I. and Hinton, G.E., Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105). 2012.
- [11] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z., Rethinking the inception architecture for computer vision. *In Conference on Computer Vision and Pattern Recognition* (pp. 2818-2826). 2016
- [12] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).
- [13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "Ssd: Single shot multibox detector". *arXiv:1512.02325*, 2015. 7
- [14] Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. "Microsoft coco: Common objects in context." *In European conference on computer vision*, pp. 740-755. Springer, Cham, 2014