# บทความวิจัย

Computer Science Department
Faculty of Informatics, Mahasarakham University

# การตรวจจับหลายคิวอาร์โค้ด สำหรับการขนส่งโดยการเรียนรู้เชิงลึก

# Multi QR Code detection for logistic by deep learning

อัยยรัชฌ์ ภูพันธ์เว่อ , นัฐธริยา เหล่าประชา

สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม

**บทคัดย่อ**

การพัฒนาระบบการตรวจจับหลายคิวอาร์โค้ด สำหรับการขนส่งโดยการเรียนรู้เชิงลึก มีจุดมุ่งหมายเพื่อที่จะลดต้นทุนการใช้เครื่องแสกนบาร์โค้ดและคิวอาร์โค้ดโดยหันมาใช้สมาทโฟนที่เป็นอุปกรณ์เคลื่อนที่มีทุกคนในปัจจุบัน โดยเพื่อให้สามารถใช้ได้ทั้งบนมือถือและบนคอมพิวเตอร์โดยสามารถใช้สแกนคิวอาร์โค้ดพร้อมกันได้มากกว่า 1 แต่ไม่เกิน 10 คิวอาร์โค้ดได้

**คำสำคัญ :** การตรวจจับวัตถุ **,** การรู้จำภาพ

**1.บทนำ**

นับตั้งแต่ที่การซื้อของออนไลน์ในไทยเริ่มเป็นที่รู้จักกัน ส่งผลให้ธุรกิจการบริการขนส่งพัสดุมีการเติมโตและขยายกิจการไปอย่างรวดเร็ว โดยบริษัทขนส่งในไทยได้มีการพัฒนารูปแบบของการบริการให้ดีขึ้น เช่น การเก็บเงินปลายทาง ส่งพัสดุไม่เว้นวันหยุดเสาร์อาทิตย์ หรือแม้แต่การรับประกันพัสดุถึงมือผู้รับ แต่การพัฒนาการขนส่งไม่ได้ทำให้การส่งพัสดุดีขึ้นอย่างเดียวยังมีผลเสียเช่น มีความล่าช้าในการจัดส่ง พัสดุมีความเสียหายหรือแม้แต่การส่งพัสดุให้ผิดคน เพราะว่าการส่งพัสดุที่มากขึ้น พัสดุที่รับเข้ามาในคลังที่มากขึ้น เวลาในการตรวจสอบพัสดุที่น้อยลง และการทำงานที่ล้าช้าลงไปมากและค่าใช้จ่ายที่ต้องเพิ่มตามขึ้นมา

ผู้พัฒนาได้เล็งเห็นปัญหาที่เกิดจากการทำงานที่ล้าช้าเพราะพัสดุที่มากขึ้นไปและค่าอุปกรณ์ที่ต้องจ่ายงบประมาณที่เพิ่มขึ้น จึงได้พัฒนาระบบที่สามารถช่วยลดระยะเวลาในการทำงานลงและยังลดค่าใช้จ่ายไปในตัวด้วยระบบที่สามารถสแกนพัสดุได้หลากหลายชิ้นในครั้งเดียว ไม่ว่าจะเป็นบาร์โค้ด หรือคิวอาร์โค้ด โดยใช้เพียงสมาร์ทโฟนที่ใช้กันอย่างแพร่หลาย เพื่อลดเวลาในการอ่านบาร์โค้ดและคิวอาร์โค้ดโดยถ่ายรูปที่ละหลายๆภาพแล้วบาร์โค้ดจะได้รหัสวัสดุ และคิวอาร์โค้ดจะได้ชื่อ-นามสกุล เบอร์โทรศัพท์เราจะทำการอ่านเพื่อรู้จำ QR Code และบุคลลภายนอกสามารถนำไปต่อยอดโดบประยุกต์ใช้กับงานพัสดุ คลังสินค้าได้ต่อไป

**2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง**

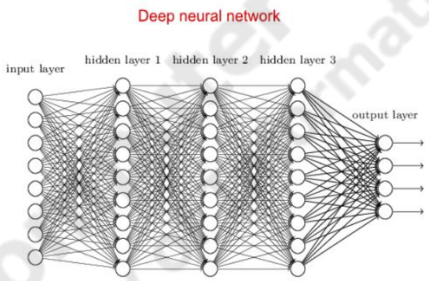
2.1 ปัญญาประดิษฐ์ คือ เครื่องจักร ที่มีฟังก์ชันที่มีความสามารถในการทำความเข้าใจ เรียนรู้องค์ความรู้ต่าง ๆ อาทิเช่น การรับรู้ การเรียนรู้ การให้เหตุผล และการแก้ปัญหาต่าง ๆ เครื่องจักรที่มีความสามารถเหล่านี้ก็ถือว่าเป็นปัญญาประดิษฐ์ นั่นเอง เพราะฉะนั้นจึงสามารถกล่าวได้ว่า AI ถือกำเนิดขึ้นเมื่อเครื่องจักรมีความสามารถที่จะเรียนรู้นั่นเอง ซึ่ง AI ก็ถูกแบ่งออกเป็นหลายระดับตาม

Computer Science Department Faculty of Informatics, Mahasarakham University

ความสามารถหรือความฉลาด โดยจะวัดจากความสามารถในการ ให้เหตุผล การพูด และทัศนคติของ AI ตัวนั้น ๆ เมื่อเปรียบเทียบกับมนุษย์อย่างเราๆ

2.2 การเรียนรู้เชิงลึก Deep Learning [2]

คือวิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาทเทียม (Neural Network) มาซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูลจะถูกนำไปใช้ในการตรวจจับรูปแบบหรือจัดหมวดหมู่ข้อมูล โดย Deep Learning การที่เครื่องจะสามารถเข้าใจสิ่งต่าง ๆ ได้จำเป็นที่จะต้องมี องค์ความรู้(Input layer) ก่อน จากนั้นจะประเมินชุดข้อมูล (Hidden layer) และนำเสนอหรือแทนองค์ความรู้นั้น (Output layer)




รูปภาพ 1 Deep Learning

2.3 การตรวจจับวัตถุ (Object Detection)

ตรวจจับวัตถุ หรือ Object Detection คือ การคาดเดาตำแหน่งของวัตถุพร้อมกับประเภทสัญลักษณ์เรียกว่าการตรวจจับวัตถุแทนที่การคาดเดาคลาสของวัตถุจากรูปภาพตอนนี้เราต้องคาดเดาคลาสรวมถึงสี่เหลี่ยม (เรียกว่ากล่องขอบเขต) ที่บรรจุวัตถุนั้น ใช้ตัวแปร 4 ตัวเพื่อระบุรูปสี่เหลี่ยมผืนผ้าโดยเฉพาะ Object Detection คือ (Classification) + (Localization)

2.4 R-CNN

R-CNN คือโครงข่ายเสนอพื้นที่ (Region Proposal Network) ที่ประกอบด้วยการทำงาน 4 ขั้นตอนดังนี้

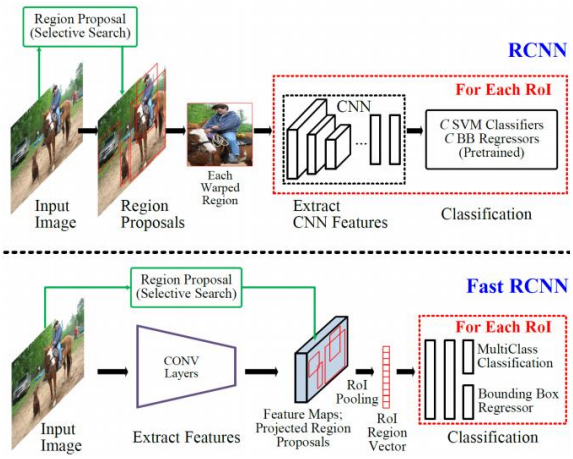
1. การเสนอพื้นที่ในภาพที่อาจจะมีวัตถุที่สนใจด้วย Selective Search
2. การฝึกและปรับแต่งอย่างละเอียดบนตัวแบบ CNN
3. การฝึกตัวจำแนกประเภทด้วย SVM แบบแยกทีละคลาส
4. การฝึกการตีกรอบวัตถุให้แม่นยำ

รูปภาพ 2 R-CNN

2.5 Fast R-CNN

คือการคำนวณ CNN ภาพใหญ่แบบรอบเดียว เมื่อมีการเสนอพื้นที่มา จะหยิบผลจาก CNN ภาพใหญ่นี้ไปใช้ โดยไม่คำนวณซ้ำ เมื่อหยิบฟีเจอร์มาแล้ว จึงค่อยทำการ Warp และสกัดเอาชุดฟีเจอร์แบบมีจำนวนเท่ากัน (fixed-length feature set) เพื่อประมวลผลต่อไป

Computer Science Department Faculty of Informatics, Mahasarakham University



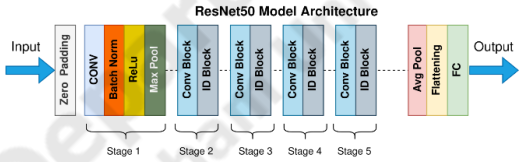

รูปภาพ 3 Fast R-CNN

2.6 Transfer Learning

เนื่องจาก Deep learning [4] มีปัญหาในเรื่องใช้ระยะเวลาในการเรียนรู้ที่ค่อนข้างนานและมีความซับซ้อนมี ตัวแปร (weight) จำนวนมากซึ่งในการเริ่มต้นฝึกโมเดล Deep learning ที่ซับซ้อนขนาดนี้ ในตั้งแต่ขั้นตอนแรกนั้นต้องใช้ทั้งข้อมูล dataset ขนาดใหญ่และใช้พลังการประมวลผลมหาศาลอีกทั้งต้องใช้เวลานานหลายวัน จนอาตถึงนานหลายสัปดาห์กว่าจะได้โมเดลที่มีประสิทธิภาพ ดั้งนั้นการใช้ Transfer learning จึงเทคนิคที่ช่วยลดเวลาการฝึกโมเดล Deep learning ด้วยการนำบางส่วนของโมเดลที่ฝึกเรียบร้อยแล้วกับงานที่ใกล้เคียงกันมาใช้เป็นส่วนหนึ่งของโมเดลใหม่ โดยโมเดลที่นำมาใช้ในงานวิจัยนี้คือ ResNet50 และ VGG16

2.6.1 Restnet50 [5]

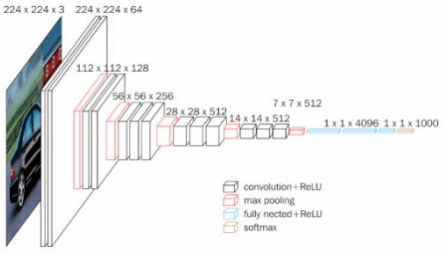
ResNet มาจาก Deep residual Network แก้ปัญหาเรื่อง vanishing gradient ซึ่งเกิดขึ้นกับโครงข่ายที่มี ความลึกค่อนข้างมากซึ่งมีจำนวนชั้นของ network ถึง 152 เลเยอร์ ( 8 เท่าของโมเดล VGG16) โดย ใช้เทคนิคการออกแบบ module ที่มีลักษณะทางลัดลงใน network ตัวโครงข่ายนี้ประกอบด้วยกัน 4 block โดยจำนวนที่มีพารามิเตอร์สำหรับฝึกทั้งหมดคือชั้นที่ใช้เรียกชื่อ เช่น ResNet50 จะหมายถึงจำนวน 50 เลเยอร์ซึ่งจะอธิบายขนาดว่า [3, 4, 6, 3] ซึ่งคือ (3 + 4 + 6 + 3) x 3 = 48 ชั้น + 2 ชั้น = 50 ซึ่ง ResNet ที่นิยมใช้จะเป็น ResNet18, ResNet34, ResNet50, ResNet101 และ ResNet152




รูปภาพ 4 Restnet50

2.6.2 VGG16

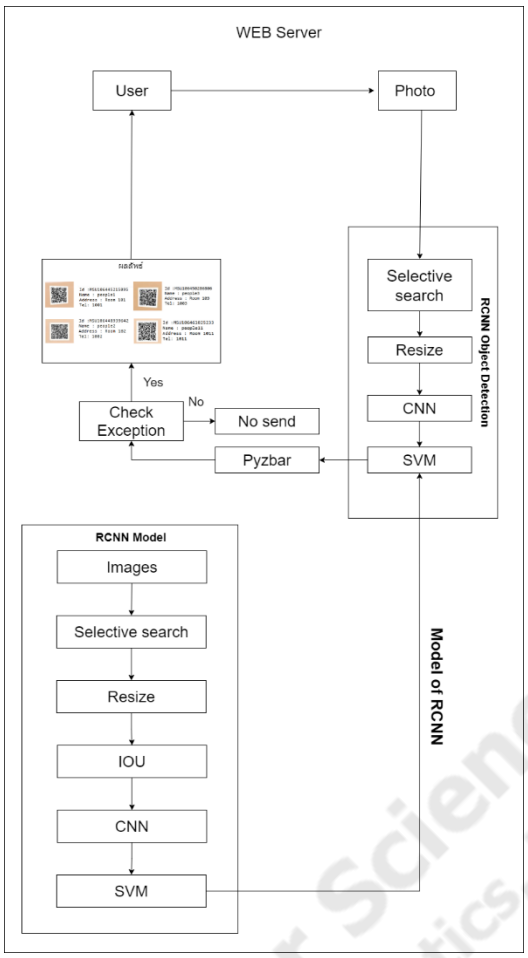
โมเดล VGG ย่อมาจาก Visual Geometry Group โดยสิ่งที่เป็นจุดเด่นของ VGG16 คือการแทนที่ hyperparameter จำนวนมาก เน้นไปที่การออกแบบชั้น (Layer) conv2D 3x3 pixels, 1 stride และการใช้ same padding และ maxpooling 2x2 pixels, 2 stride แบบเดียวกันตลอดทั้งโครงสร้าง โดยชื่อของ VGG16 หมายถึงมี 16 ชั้นที่มีน้ำหนักเครือข่ายนี้เป็นเครือข่ายที่ใหญ่และมีพารามิเตอร์ประมาณ 138 ล้าน




รูปภาพ 5 VGG16

Computer Science Department Faculty of Informatics, Mahasarakham University

3. แผนการดำเนินงาน




รูปภาพ 6 แผนการดำเนินงาน

**4.ผลการทดลอง**

จากการทดลองในการหาโมเดลที่มีประสิทธิภาพมากที่สุดพบว่าโมเดลชุดที่ 1 และ 2 มีค่ามากที่สุดและใกล้เคียงกันจึงเลือกใช้โมเดลแบบที่ 1 ในการใช้งาน และมีข้อจำกัดของโปรแกรมในการค้นหาคิวอาร์โค้ดไม่สามารถค้นคิวอาร์โค้ดที่มีความคลายภาพพื้นหลังจนเกินไปและไม่สามารถค้นหากับคิวอาร์โค้ดที่โดนแสงแฟลชของกล้องจนทำให้ภาพสว่างและมองไม่ชัดจนเกินไป และในการทดลองของการถอดรหัสโดนใช้ไลบรารี่ของ Pyzbar พบว่าไม่สามารถทำการถอดรหัสและแปลงเป็นข้อความกับคิวอาร์โค้ดที่มีความไม่ชัดเจนชัดพอได้

**5.สรุปผลและอภิปรายผลการทดลอง**

1. สรุปผลและอภิปรายผล

โครงการปริญญานิพนธ์ฉบับนี้นำเสนอการเว็บแอพพลิเคชันสำหรับการค้นหาคิวอาร์โค้ดที่มากกว่า 1 แต่ไม่เกิน 10 เพื่อ ซึ่งใช้การประมวลผลภาพ และการเรียนรู้เชิงลึก มีเป้าหมายในลดต้นทุนการใช้เครื่องอ่านคิวอาร์โค้ดโดยการใช้สมาทโฟนที่อยู่แล้วแทนและเป็นการพัฒนาทดสอบการออกแบบและเลือกใช้สถาปัตยกรรมการโครงข่ายประสาทเทียมที่เหมาะสมกับงาน ให้มีประสิทธิภาพ และสามารถใช้งานผ่านสมาทโฟนและคอมพิวเตอร์ได้

จากการทดลองวัดประสิทธิภาพของโมเดลทั้ง 4 ชุดได้เลือกใช้โครงข่ายประสาทเทียมที่ผ่านการฝึกอบรมล่วงหน้าของโครงข่ายประสาทเทียม VGG16 มาใช้กับข้อมูลในการเรียนรู้ทั้งหมด 1000 ภาพและภาพทดสอบ 100 ภาพ ได้ผลการทดลองค้นหาคิวอาร์โค้ดได้ได้ 527 คิวอาร์โค้ดจากทั้งหมด 550 และสามารถถอดรหัสโดยใช้ ไลบรารี่ของ Pyzbar ได้ 271 จากทั้งหมด 527 คิดเป็นร้อยละ 51.42 โดยคิวอาร์โค้ดที่ไม่สามารถอ่านได้คือคิวอาร์โค้ดที่มีรายละเอียดและความชัดเจนไม่พอ

2. ปัญหาและอุสรรคในการนำเดินงาน

2.1 เวลาในการประมวลผลค่อนข้างนานเนื่องจากจำนวนภาพที่นำมาใช้ในการเรียนรู้แลโครงข่ายประสาทเทียม VGG16 ที่มี

Computer Science Department
Faculty of Informatics, Mahasarakham University

พื้นที่ค่อนข้างใหญ่ถึง 512 และหลังจากทำการเรียนรู้ขั้นสุดท้ายเสร็จมีขนาดถึง 1.44 GB

2.2 อุปกรณ์ในการทดสอบ ภาพที่ถ่ายได้จากสมาทโฟนมีความละเอียดไม่พอและบ้างภาพอาจจะมีขนาดใหญ่เกินไปส่งผลให้ใช้เวลาในการประมวลผลค่อนข้างนานถึงนานมาก

2.3 ข้อจำกัดของฮาร์ดแวร์คอมพิวเตอร์ที่ใช้ในการพัฒนาเนื่องจากมีรูปภาพจำนวนมากส่งผลให้หน่วยความจำ(RAM)ไม่พอ เป็นต้น

3. ข้อเสนอแนะ

3.1 ปรับปรุงอัลกอลิทึมในการเรียนรู้เพื่อลดขนาดของโมเดลและลดระยะเวลาการรู้จำโมเดลในการใช้งานได้

3.2 ควรพัฒนาอัลกอลิทึมในการค้นหาคิวอาร์โค้ดเนื่องจากใช้เวลาในการค้นหานานเกินไป

3.3 ใช้การกรองภาพในการทำให้คิวอาร์โค้ดมีความชัดเจนขึ้นเพื่อให้คิวอาร์โค้ดถอดรหัสได้

**อ้างอิง**

1. Nessessence , “ปัญญาประดิษฐ์ (AI : Artificial Intelligence) คืออะไร”, https://www.thaiprogrammer.org/2018/12/whatisai/ , Accessed Mar 20, 2020.

2. Athiwat, “Deep learning คืออะไร” , https://medium.com/@athivvat/deep-leaning-คืออะไร -785e16d01773 , Accessed Mar 21 ,2020

3. Koteshwar Surga , “Introduction to Object Detection in Deep Learning” ,https://medium.com/@skoteshwar930/introduction-to-object-detection-in-deep-learning-5f0c3b9862f0, Accessed Mar 23, 2020.

4. Natthawat Phongchit , “ย้อนรอย Object Detection และเจาะลึก RetinaNet” ,https://medium.com/@natthawatphongchit/ย้อนลอย-object-detection-เจาะลึก-retinanet-ea1407ad7325 Accessed Mar 23, 2020.

Computer Science Department
Faculty of Informatics, Mahasarakham University