JASET KMUTNB
Research Article

Journal of Applied Science and Emerging Technology (JASET)
Published by Faculty of Applied Science, KMUTNB, Thailand
doi: 10.14416/JASET.KMUTNB.2024.02.007
Vol. 23, No. 2 [2024]: e256364

การตรวจจับตำแหน่งเครื่องหมายกากบาทที่เขียนด้วยลายมือบนกระดาษคำตอบแบบปรนัยด้วยเทคนิคการตรวจจับวัตถุแบบ YOLO

YOLO-Based Object Detection for Locating Handwritten X-Marks on Multiple-Choice Answer Sheets

พิศิษฐ์ นาคใจ[1] สุชานาถ พัดเทพ[1], วราภรณ์ ปานอูป[2], และพัชรี มณีรัตน์[1*]

Pisit Nakjai, Suchanard Padthep, Waraporn Panup and Patcharee Maneerat*

[1]สาขาวิทยาการข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏอุตรดิตถ์, อุตรดิตถ์, ประเทศไทย

[2]สาขาคณิตศาสตร์ประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏอุตรดิตถ์, อุตรดิตถ์, ประเทศไทย

[1]Applied Mathematics, Faculty of Science and Technology, Uttaradit Rajabhat University, Uttaradit, Thailand

[2]Data Science, Faculty of Science and Technology, Uttaradit Rajabhat University, Uttaradit, Thailand

บทคัดย่อ

การประเมินความรู้ของผู้เรียนถือเป็นการวัดผลสัมฤทธิ์ในการเรียน สำหรับเครื่องมือที่นิยมใช้วัดผลสัมฤทธิ์ทางการเรียน คือ การสอบแบบปรนัย ซึ่งการตอบคำถามทำได้ทั้งวิธีการฝนและกากบาทลงในกระดาษคำตอบ หากแต่การตอบคำถามโดยวิธีการกากบาทนั้นทำให้เกิดประเด็นปัญหาที่สำคัญคือ หากผู้เข้าสอบมีจำนวนมากอาจจำเป็นต้องใช้บุคลากรเพิ่มขึ้นและมีความล่าช้าในการตรวจคำตอบ นอกจากนี้อาจเกิดข้อผิดพลาดในตรวจคำตอบซึ่งจะส่งผลทำให้การสรุปผลคะแนนเกิดความคลาดเคลื่อนได้ ในการศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลทางคณิตศาสตร์ที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยด้วยกระบวนการตรวจจับวัตถุและการเรียนรู้ของเครื่องเชิงลึก สำหรับการตรวจจับวัตถุอาศัยเทคนิค You Only Look Once version 8 (YOLOv8) และการสร้างโมเดลที่สามารถตรวจจับกากบาทแบบอัตโนมัติอาศัยเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network : CNN) พร้อมทั้งการวัดประสิทธิภาพโมเดลที่พัฒนาขึ้นซึ่งพิจารณาจากค่าความแม่นยำ (Precision) ค่าค้นคืน (Recall) และ F1-score ผลการศึกษา พบว่า เมื่อกำหนดค่าของ Threshold เท่ากับ 0.496 โมเดลสำหรับการตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยให้ค่าความแม่นยำสูงสุด และให้ค่า F1-score สูงสุดเท่ากับ 0.989 ซึ่งส่งผลทำให้โมเดลมีประสิทธิภาพดีที่สุดสำหรับการตรวจจับและระบุตำแหน่งของกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัย

คำสำคัญ: การตรวจจับกากบาทแบบอัตโนมัติ, กระบวนการตรวจจับวัตถุ, การเรียนรู้ของเครื่องเชิงลึก, เทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน, Yolo

ABSTRACT

Student assessment is considered a measure of learning achievement. Among the various assessment tools available, multiple-choice tests are commonly used, allowing students to mark their answers by filling or marking crosses on provided answer sheets. However, this method presents significant challenges, particularly when dealing with a large number of questions, as it may require additional personnel for grading, leading to delays in the assessment process. Additionally, manual grading can introduce errors, resulting in inaccuracies in score calculation. This study aims to address these challenges by developing a mathematical model capable of automatically detecting crosses on multiple-choice answer sheets using object detection techniques and deep learning. The model is based on the You Only Look Once version 8 (YOLOv8) technique for object detection. It employs a Convolutional Neural Network (CNN) for cross detection. The performance of the model is evaluated using metrics such as Precision, Recall, and F1-score. The results indicate that at a Threshold value of 0.496, the model achieves the highest F1-score of 0.989. This suggests that the model is effective in automatically detecting and identifying cross markings on multiple-choice answer sheets.

KEYWORDS: Automatic X-mark detection, Object detection, Deep learning, Convolution neural network, Yolo

***Corresponding Author:** m.patcharee@uru.ac.th

Received: 09/04/2024; Revised: 28/05/2024; Accepted: 14/06/2024

1. บทนำ

การวัดผลสัมฤทธิ์ทางการเรียนนั้นเป็นส่วนหนึ่งสำหรับประเมินความรู้ความเข้าใจของผู้เรียน ซึ่งข้อสอบถือเป็นเครื่องมือวัดผลสัมฤทธิ์ทางการเรียนอย่างหนึ่ง สำหรับข้อสอบที่นิยมใช้สามารถแบ่งออกเป็น 2 ประเภท คือ ข้อสอบแบบเขียนตอบ (Supply Type) และข้อสอบแบบปรนัยหรือข้อสอบแบบเลือกตอบ (Multiple Choice Questions: MCQ) ซึ่งข้อสอบแบบ MCQ เป็นวิธีที่ง่ายในการวัดความเข้าใจมากกว่าการใช้ข้อสอบแบบเขียนตอบ ในปัจจุบันมีเครื่องที่ช่วยในการตรวจข้อสอบแบบ MCQ ซึ่งจะมีอุปกรณ์พิเศษในการตรวจคำตอบคือ เครื่องโอเอ็มอาร์ (Optical Mark Reader) สร้างโดยอาศัยกระบวนการรู้จำเครื่องหมายด้วยแสง (Optical Mark Recognition: OMR) ซึ่งเครื่องโอเอ็มอาร์จะทำการอ่านสัญลักษณ์โดยอาศัยหลักการอ่านค่าการสะท้อนของแสง ซึ่งผู้เข้าสอบจำเป็นต้องทำการฝนทึบตัวเลือกที่ต้องการ หากใช้ดินสอที่มีความเข้มต่ำกว่าระดับที่กำหนดอาจทำให้เครื่องไม่สามารถอ่านได้ชัดเจน และที่สำคัญจำเป็นต้องใช้กระดาษคำตอบที่เครื่องดังกล่าวรองรับเท่านั้น ซึ่งจะเห็นได้ว่าอาจเกิดค่าใช้จ่ายทั้งในส่วนอุปกรณ์และการเปลี่ยนพฤติกรรมการเลือกคำตอบจากเดิมที่ใช้การทำเครื่องหมายกากบาทเป็นการฝนทับแทน นอกจากนี้หากมีการเปลี่ยนคำตอบที่ผู้เข้าสอบจะต้องทำให้ตัวเลือกที่ยกเลิกมีความสะอาดเพียงพอเพื่อให้แสงสะท้อนผ่านได้

เนื่องด้วยวิธีการดั้งเดิมเป็นการเลือกคำตอบโดยวิธีการกากบาทลงในกระดาษคำตอบซึ่งเป็นวิธีการเลือกคำตอบที่ไม่ซับซ้อน แต่อย่างไรก็ตามทำให้เกิดประเด็นปัญหาที่สำคัญคือ การจำเป็นต้องใช้บุคลากรในการตรวจหาตำแหน่งตัวเลือกที่ทำเครื่องหมายกากบาทเทียบกับตัวเลือกที่ทำเครื่องหมายในข้อสอบชุดเฉลยคำตอบ หากผู้เข้าสอบมีจำนวนมากหรือข้อคำถามมีจำนวนมาก อาจส่งผลให้เกิดความล่าช้าในการตรวจคำตอบ และที่สำคัญอาจเกิดข้อผิดพลาดในตรวจคำตอบซึ่งจะส่งผลทำให้การสรุปผลคะแนนเกิดความคลาดเคลื่อนได้ จากข้างต้นจะเห็นได้ว่าการเลือกคำตอบด้วยวิธีกากบาทเป็นวิธีที่ไม่ซับซ้อนแต่อาจจะเกิดข้อผิดพลาดจากบุคคลากรที่ทำการระบุตำแหน่งของการทำเครื่องหมายกากบาท

ด้วยเหตุนี้ผู้วิจัยจึงมีวัตถุประสงค์เพื่อพัฒนาโมเดลทางคณิตศาสตร์ที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบแบบปรนัย ด้วยเทคนิคการเรียนรู้ของเครื่อง (Deep Machine Learning: DML) เพื่อทำการรู้จำเครื่องหมายกากบาท และอาศัยกระบวนการตรวจจับวัตถุ (Object Detection) สำหรับการตรวจจับและระบุตำแหน่งของเครื่องหมายกากบาท

2.เอกสารงานวิจัยที่เกี่ยวข้อ

ในระบบการตรวจข้อสอบแบบ MCQ ผู้เข้าสอบจำเป็นต้องทำการเลือกตอบหรือทำเครื่องหมายลงในช่องกระดาษคำตอบ ในปัจจุบันได้มีการพัฒนาเครื่องตรวจกระดาษคำตอบแบบหลายตัวเลือกซึ่งจำเป็นต้องมีอุปกรณ์และซอฟแวร์ที่ใช้ในการตรวจโดยเฉพาะ สำหรับกระบวนการ OMR อาจจำเป็นต้องใช้อุปกรณ์ต่อพ่วง (Peripheral devices) ประเภทต่าง ๆ เข้าด้วยกัน อาทิเช่น เครื่องแสกนเพื่อนำข้อมูลกระดาษคำตอบให้อยู่ในรูปแบบที่เครื่องโอเอ็มอาร์สามารถนำไปประมวลผลได้ โดยเครื่องโอเอ็มอาร์จะทำการสแกนตำแหน่งของช่องคำตอบโดยอาศัยหลักการสะท้อนแสงของช่องคำตอบเพื่อประมวลผลการสอบแบบอัตโนมัติ โดยปกติแล้วเครื่องสแกนโอเอ็มอาร์จะใช้งานร่วมกับซอฟแวร์ที่พัฒนาขึ้น (de Elias et al., 2021; Hussmann & Deng, 2005; Kumar et al., 2018) ซึ่งอุปกรณ์และซอฟแวร์เหล่านี้อาจมีค่าใช้จ่ายค่อนข้างสูง ในงานวิจัยที่ผ่านมาได้พัฒนากระบวนการและเปลี่ยนจากเครื่องแสกน เป็นกล้องเว็บแคมและกล้องมือถือแทน หลังจากนั้นจะใช้กระบวนการตรวจจับหารูปแบบของการทำ เครื่องหมายในกระดาษคำตอบโดยอาศัยการประมวลผลภาพ (Image Processing: IP)(Tinh & Minh, 2024)

ในยุคเริ่มต้นของการพัฒนาเครื่อง OMR จะเป็นการพัฒนาอุปกรณ์ที่ใช้ทำระบุตำแหน่งของการเลือกคำตอบ (Hussmann & Deng, 2005; the School of Engineering, Edith Cowan University, Perth, Australia et al., 2018) ซึ่งสามารถตรวจจับได้ทั้งการทำเครื่องหมาย วงกลม วงรี และสี่เหลี่ยม โดยการทำเครื่องหมายแบบฝนทึบอาศัยหลักการของการผ่านของลำแสงร่วมกับค่าเทรชโฮลด์ (Threshold) ของค่าแสงที่ส่องผ่านเพื่อระบุตำแหน่งของการเลือกคำตอบ การใช้หลักการของเทรชโฮลด์นี้เป็นการคำนวณที่ไม่ซับซ้อน ทำให้ในงานวิจัยที่ผ่านมาได้ประยุกต์ใช้เทคนิคการใช้ค่าเทรชโฮลด์นี้กับหลักการอื่น ๆ ในการระบุตำแหน่งของการเลือกคำตอบในกระดาษคำตอบ (Deng et al., 2008)

ต่อมาได้มีเปลี่ยนอุปกรณ์ที่ใช้หลักการของแสงสองผ่านเปลี่ยนเป็นการใช้ระบบการประมวลภาพดิจิตอลมากขึ้นร่วมการการใช้ค่าเทรชโฮลด์ในการระบุตำแหน่งของคำตอบโดยการใช้ความแตกต่างของค่าสีดำและสีขาว

ของภาพถ่ายดิจิตอล ซึ่ง Nguyen et al. (Nguyen et al., 2011) ได้นำเสนอระบบที่นำเอากล้องดิจิตอลแทนการใช้เครื่องแสกนในการเก็บข้อมูลภาพถ่ายกระดาษคำตอบ หากแต่เกิดข้อจำกัดของการใช้กล้องดิจิตอลในการถ่ายภาพกระดาษคำตอบจะเกิดปัญหามุมเอียงของการตั้งกล้องดิจิตอล และได้เสนอวิธีการปรับมุมเพื่อแก้ไขภาพกระดาษคำตอบที่เอียงและใช้ค่าเทรชโฮลด์เพื่อระบุตำแหน่งคำตอบที่ถูกระบายทับได้ ต่อมา Spadaccini and Rizzo (Spadaccini & Rizzo, 2011) ได้ทำการนำภาพกระดาษคำตอบที่กล้องดิจิตอลมาประมวลผลเพื่อหาตำแหน่งของเครื่องหมายกากบาทในกระดาษคำตอบโดยใช้เทคนิค Gamera Framework (GF) ร่วมกับการกำหนดค่าเทรดโฮลด์ และที่สำคัญได้ระบบ JECT-OMR ที่นำเสนอสามารถตรวจจับการยกเลิกการเลือกคำตอบได้โดยการใช้วิธีการฝนทึบลงในช่องของคำตอบ เช่นเดียวกับงานวิจัยของ Fistesu et al. (Fisteus et al., 2013) ได้ใช้วิธีการยกเลิกตัวเลือกคำตอบด้วยวิธีการฝนทึบและได้นำเสนอวิธีการหาเครื่องหมายกากบาทด้วยระบบ Eyegrade ซึ่งประกอบไปด้วยการระบุค่าเทสโฮลด์และประมวลผลภาพกระดาษ คำตอบที่ได้จากกล้องเว็บแคมในการตรวจจับเครื่องหมายกากบาทซึ่งเป็นสิ่งที่ท้าทายสำหรับกระบวนการตรวจกระดาษคำตอบที่ใช้กระบวนการ IP โดย นอกจากนี้ Sanguansat (Sanguansat, 2015) ได้เสนอเทคนิคการตรวจจับการทำเครื่องหมายที่มีลักษณะหลากหลายรูปแบบโดยใช้วิธี Adaptive Threshold สำหรับประสิทธิภาพในการตรวจจับคิดเป็นร้อยละ 85.72

นอกจากนี้ได้มีนักวิจัยหลากหลายท่านที่ได้พัฒนากระบวนการในการตรวจข้อสอบด้วยเทคนิคต่าง ๆ ยกตัวอย่างเช่น กฤษณะและยุทธพงษ์ (กฤษณะ ชินสาร & ยุทธพงษ์ รังสรรค์เสรี, 1998) นำเสนอการใช้ฮีสโตรแกรม (Histogram) ซึ่งเป็นเทคนิควิธีการประมวลผลภาพ เบื้องต้นที่สามารถนำมาประยุกต์ใช้ในการหาตำแหน่งของการเลือกคำตอบโดยใช้วิธีการหาค่า ตำแหน่งสูงสุดของ Histogram ซึ่งจะได้ตำแหน่งของตัวเลือกคำตอบใน กระดาษคำตอบ ผลการศึกษา พบว่า เทคนิคนี้ทำงานได้ดีกรณีที่กระดาษคำตอบเป็นการตอบแบบฝน สำหรับการใช้วิธีเลือกคำตอบแบบกากบาทยังมีข้อจำกัดคือ หากกากบาททับวงกลมมีขนาดเล็กเกินไปทำให้เมื่อนับจำนวนจุดภาพดำที่เกิดขึ้นในวงกลมแต่ละวงมีความแตกต่างกันน้อยมาก ต่อมาพุทธินันท์ (พุทธินันท์ พัดกระจ่าง, 2010) ประยุกต์ ใช้กระบวนการหาขอบของส่วนที่เป็นกระดาษคำตอบและ ทำ Perspective Transformation ในกรณีที่ภาพกระดาษ คำตอบไม่เป็นสี่เหลี่ยมมุมฉาก และหาตำแหน่งของกากบาทโดยพิจารณาความเข้มของพิกเซลในภาพ ซึ่งให้ความแม่นยำในการระบุตำแหน่งคิดเป็นร้อยละ 90.47 อย่างไรก็ตามเทคนิคดังกล่าวยังมีข้อผิดพลาดเกิดจาก กรณีที่มีการตรวจพบกากบาทในข้อมากกว่า 1 คำตอบ

ต่อมา ภูมินทร์ (ภูมินทร์ ตันอุตม์, 2017) ได้ประยุกต์ใช้เทคนิคด้าน ปัญญาประดิษฐ์ (Artificial Intelligence: AI) และการ ประมวลผลภาพในการตรวจคำตอบ ซึ่งได้เสนอวิธีการตรวจหากากบาทบนภาพกระดาษคำตอบแบบปรนัย 5 ตัวเลือกด้วยกระบวนการ IP ร่วมกับทำการแบ่งกลุ่มของตำแหน่งกากบาทเพื่อตรวจคำตอบ ในการหากากบาทใช้กระบวนการประมวลผลแบบมอร์โฟโลยี(Morphology Technique) ร่วมกับสตัจเจอร์อิลิเมนต์ (Structure Element) แนว 45 และ -45 องศาทำให้ได้ตำแหน่งของกากบาทปรากฏขึ้นและใช้เทคนิคเคมีน (K-means) ในการตัดสินใจว่ากากบาทที่เลือกอยู่กลุ่มเดียวกับเฉลยของคำตอบหรือไม่ อย่างไรก็ตามยังมีข้อจำกัดคือ กระดาษ คำตอบที่มีการกากบาทจะต้องมีรูปแบบคือ กากบาทเพียง 1 ข้อ ไม่สามารถใช้ปากกาลบคำผิดทับเส้นและไม่ขีดทับคำตอบได้

จากข้างต้นจะเห็นได้ว่าเทคนิคต่าง ๆ ที่นำเสนอยังมีข้อจำกัดในการตรวจจับกากบาท ด้วยเหตุผลเหล่านี้ทำให้การศึกษาครั้งนี้มุ่งเน้นการโมเดลทางคณิตศาสตร์ที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยด้วยกระบวนการตรวจจับวัตถุและการเรียนรู้ของเครื่องเชิงลึก สำหรับการตรวจจับวัตถุอาศัยเทคนิค You Only Look Once version 8 (YOLOv8) และทำการสร้างโมเดลที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยอาศัยเทคนิคโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Network: CNN) พร้อมทั้งการวัดประสิทธิภาพโมเดลที่พัฒนาขึ้น ซึ่งคณะผู้วิจัยคาดหวังว่าในการศึกษาครั้งนี้จะช่วยแก้ปัญหาและลดความผิดพลาดในการตรวจคำตอบต่อไป

3. YOLO-BASED ARCHITECTURE

3.1 YOLO

ในการศึกษาครั้งนี้ระเบียบวิธีวิจัยที่ใช้คือ กระบวนการตรวจจับวัตถุ (Object Detection) (Liu et al., 2016) และการเรียนรู้ของเชิงลึก (Deep Machine Learning : DML) ซึ่ง DML เป็นส่วนหนึ่งของของการเรียนรู้ของเครื่อง (Machine Learning) ใน AI

ที่เลียนแบบการทำงานของ สมองมนุษย์ในการประมวลผลและการสร้างรูปแบบของข้อมูลสำหรับใช้ในการตัดสินใจ เนื่องด้วย DML สามารถรองรับการทำงานกับข้อมูลภาพที่เพิ่มขึ้นอย่างมหาศาล และใช้การประมวลผลแบบกลุ่มเมฆ (Cloud Computing Service) ที่สามารถเข้าถึงหน่วยการคำนวณ (Computation Unit) ได้ง่ายขึ้นในราคาที่ถูก (Deepak Kumar P, 2023)

สำหรับการสร้างโมเดลที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยอาศัยเทคนิค CNN ซึ่งเป็นกระบวนการหนึ่งใน DML ซึ่งให้ประสิทธิภาพที่ดีสำหรับข้อมูลภาพ และเป็นเทคนิคที่ใช้การเรียนรู้ตั้งแต่ต้นจนจบกระบวนการ (End-to-End Learning) เริ่มต้นตั้งแต่การสกัดและหาคุณลักษณะพิเศษจากภาพ (Feature Extraction) ตลอดจนการระบุตำแหน่งของวัตถุ ชนิดของวัตถุหรือการทำนายข้อมูลในรูปแบบต่าง ๆ ซึ่งเทคนิคนี้สามารถลดขั้นตอนที่มีความยุ่งยากในการตรวจจับภาพจากการใช้การเรียนรู้ของเครื่องแบบดั้งเดิมที่ต้องสกัดคุณลักษณะพิเศษจากภาพเพื่อหาเส้น สี รูปทรง และอื่น ๆ ด้วยตนเอง (Hand-crafted Features) (Lin et al., 2020) เพื่อให้ได้คุณลักษณะเฉพาะที่เหมาะสมกับปัญหาและชุดข้อมูลนั้น ๆ ก่อนที่นำไปสู่การสร้างโมเดล และที่สำคัญเทคนิค CNN มีประสิทธิภาพที่ให้ความแม่นยำสูงกว่าการเรียนรู้ของเครื่องแบบดั้งเดิม

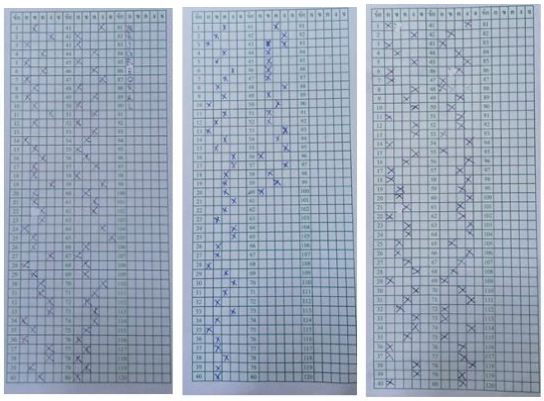
ในการตรวจจับกากบาทในกระดาษคำตอบปรนัยจะอาศัยเทคนิค YOLO (Redmon et al., 2016) ซึ่งถือว่าเป็นกระบวนการที่ถูกพัฒนามาอย่างต่อเนื่องสำหรับการตรวจจับแบบเรียลไทม์ (Real Time) ในการศึกษาครั้งนี้ได้ใช้ YOLOv8 (Jocher et al., 2023) ซึ่งแบ่งโครงสร้างออกเป็น 2 ส่วน คือ แกนหลัก (Backbone) และส่วนหัวทำนาย (Head Detection)

ส่วนแกนหลักมีหน้าที่ในการสกัดคุณลักษณะของภาพออกมา โดยกระบวนการการคอนโวลูชันซึ่งทำให้ในขณะที่ภาพถูกคอนโวลูชันจะมีมิติที่เล็กลงแต่จะได้คุณลักษณะที่แตกต่างกันมากขึ้น ในส่วนแกนหลักได้ใช้เทคนิคซีทู (C2 : CSP Bottleneck with 2 convolutions) โดยมีโครงสร้างแบบซีแอสพี (CSP: Cross Stage Partial) (Wang et al., 2020) มาพัฒนาต่อเป็นซีทูเอฟ (C2F) ทำให้ช่วยลดเวลาในการคำนวณลงแต่ยังคงประสิทธิภาพเดิมของโมเดลไว้ได้

ส่วนหัวทำนายของ YOLO จะทำการย่อยและขยายขนาดมิติของคุณลักษณะออกมาเพื่อช่วยในการตรวจจับวัตถุที่มีขนาดเล็กและขนาดใหญ่ ซึ่งอาจปรากฏในภาพเดียวกัน โดยส่วนหัวทำนายประกอบไปด้วยสองส่วนคือ ส่วนของการย่อยและขยายขนาดมิติของคุณลักษณะและส่วนของการตรวจจับ ในส่วนของย่อยและขยายขนาดมิติคุณลักษณะจะใช้เทคนิคเอสพีพีเอฟ (Spatial Pyramid Pooling Fast: SPPF) (He et al., 2015) เพื่อช่วยให้สามารถสกัดคุณลักษณะและขยายมิติให้มีความแตกต่างกัน เพื่อให้มิติของคุณลักษณะมีขนาดที่ใหญ่ขึ้นจำนวน 3 ขนาด ในส่วนสุดท้ายเป็นส่วนของการตรวจจับภาพ เมื่อทำการขยายมิติ คุณลักษณะ 3 ขนาดมาแล้ว แต่ละขนาดจะทำการตีกรอบที่มีลักษณะเป็นสี่เหลี่ยมล้อมรอบวัตถุเพื่อทำนายและระบุตำแหน่งของวัตถุ โดยจะทำการทำนายพารามิเตอร์ต่างๆ คือ ความกว้างและส่วนของกรอบ (w, h) จุดกึ่งกลางของกรอบ (b_x, b_y) ความน่าจะเป็นที่มีวัตถุอยู่ในกรอบ (P_c) และความน่าจะเป็นของการเป็นชนิดวัตถุที่ c_i ที่อยู่ภายในกรอบ $p(c_1, c_2, \ldots, c_n)$ โดยที่ $i = 1, .., n$ และ n แทนจำนวนชนิดของวัตถุที่ YOLO สามารถ ทำนายได้ สำหรับการฝึกสอนให้ YOLOv8 เพื่อตรวจจับกากบาทในกระดาษคำตอบนั้น จำเป็นจะต้องมีการเก็บรวบรวมข้อมูลเพื่อฝึกสอนให้กับ YOLOv8 ใหม่ เนื่องจาก YOLOv8 ไม่มีข้อมูลพื้นฐานของวัตถุที่เป็นลักษณะกากบาทโดยคณะผู้วิจัยได้เลือก YOLOv8 ขนาดกลาง (YOLOv8m) โดยทำการฝึกสอนแบบ Pre-trained โดยจะทำการปรับจูน พารามิเตอร์ต่อจากค่าพารามิเตอร์ที่ YOLOv8 ทำงานได้อยู่แล้ว

3.2 ข้อมูล

ในการศึกษาครั้งนี้ใช้ข้อมูลภาพกระดาษคำตอบแบบปรนัย 120 ข้อ แบบ 5 ตัวเลือก จำนวนทั้งหมด 163 ชุด แบ่งออกเป็น 100 ชุด สำหรับชุดฝึกสอน และ 63 ชุด สำหรับทำการทดสอบ ซึ่งภายในกระดาษคำตอบมีทั้งเครื่องหมายกากบาทและตัวอักษรอื่น ๆ ปรากฏร่วมด้วยในช่องตัวเลือกคำตอบ ดังรูปที่ 1a นอกจากนี้ภายในช่องตัวเลือกอาจมีการทำเครื่องหมายอื่นที่ไม่ใช่ลักษณะที่เป็นกากบาท ได้แก่ การเปลี่ยนตัวเลือกด้วยการฝนทับ ดังรูปที่ 1b และการใช้น้ำยาลบคำผิดแต่ยังปรากฏเครื่องหมายกากบาทเดิมอยู่ ดังรูปที่ 1c



(a) (b) (c)

รูปที่ 1 ภาพตัวอย่างกระดาษคำตอบแบบปรนัย: (a) กระดาษคำตอบมีทั้งกากบาทและตัวอักษรอื่น ๆ (b) การเปลี่ยนตัวเลือกด้วยการฝนทับ และ (c) การใช้น้ำยาลบคำผิดแต่ยังปรากฏเครื่อง หมายกากบาทเดิมอยู่

3.3 การฝึกสอนโมเดล

สำหรับงานวิจัยนี้ได้ใช้กระบวนการตรวจจับวัตถุโดยใช้เทคนิค YOLOv8m (Jocher et al., 2023) ซึ่งจะมีค่าของพารามิเตอร์ที่ต้องการปรับจูนทั้งหมด สามล้านสองแสนพารามิเตอร์ โดยที่ YOLOv8m จะรับรูปภาพขนาด 640 พิกเซล เพื่อใช้สำหรับการตรวจจับกากบาทในภาพ โดยจำนวนเครื่องหมายกากบาทจากภาพสำหรับฝึกสอน 100 ภาพ มีจำนวนเครื่องหมายกากบาทจำนวน 7,396 เครื่องหมายและทำการเพิ่มจำนวนภาพด้วยการทำ การเสริมข้อมูล(Augment Data) เพื่อให้โมเดลได้ข้อมูลของเครื่องหมายกากบาทที่หลากหลายสำหรับการฝึกสอนเพิ่มขึ้น ในการฝึกสอน YOLOv8m นั้นเพื่อให้ YOLOv8m สามารถทำการตรวจจับเครื่องหมายกากบาทในภาพได้นั้นจะทำการฝึกสอนโมเดลด้วยการกำหนด Loss function ดังนี้

$$Loss = Loss_{\{CIoU\}} + Loss_{\{Conf\}} + Loss_{\{Class\}}$$

โดยที่

$$Loss_{\{CIoU\}} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha\nu$$

$$Loss_{\{Conf\}} = \sum_{i=0}^{S^2}\sum_{j=0}^{B} I_{ij}^{obj}\left[\hat{C}_i \log(C_i) + (1-\hat{C}_i)\log(1-C_i)\right] - \lambda_{noobj}\sum_{i=0}^{S^2}\sum_{j=0}^{B} I_{ij}^{noobj}\left[\hat{C}_i \log(C_i) + (1-\hat{C}_i)\log(1-C_i)\right]$$

$$Loss_{\{Class\}} = \sum_{i=0}^{S^2}\sum_{c\in classes}\{\hat{p}_i(c)\log p_i(c) + [1-\hat{p}_i(c)]\log[1-p_i(c)]\}$$

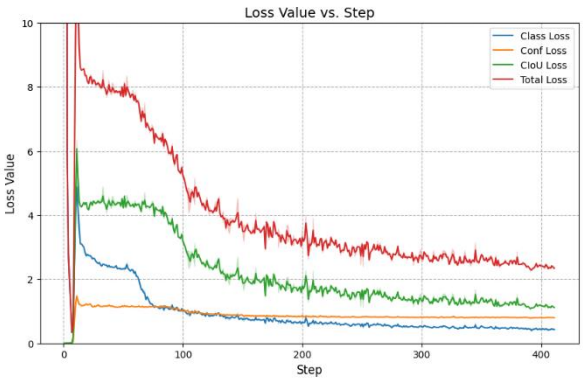
$$\alpha = \frac{\nu}{1 - \mathrm{IoU} + \nu}$$

$$\nu = \frac{4}{pi^2}\left[\tan^{-1}\left(\frac{w^{gt}}{h^{gt}}\right) - \tan^{-1}\left(\frac{w}{h}\right)^2\right]$$

จากข้างต้น Loss function ประกอบด้วย 3 ส่วนคือ $Loss_{CIoU}$ แทนส่วนของการทำนายขนาดและตำแหน่งของกรอบล้อมรอบวัตถุ ขณะที่ $\rho^2(.)$ เป็นการหาระยะห่างของจุดกึ่งกลางของวัตถุกับกรอบการทำนายโดยใช้ระยะทางแบบยุคลิด (Euclidean Distance) และ c เป็นความยาวเส้นทแยงมุมของกล่องที่เล็กที่สุดระหว่างกรอบวัตถุจริงกับกรอบการทำนาย นอกจากนี้ $Loss_{Conf}$ แทนส่วนทำนายการปรากฏของวัตถุและ $Loss_{Class}$ แทนส่วนของการทำนายชนิดวัตถุ

สำหรับภาพที่นำเข้าของ YOLO จะถูกปรับขนาดจากเดิม 4000 x 6000 พิกเซล เป็นขนาด 640 x 640 พิกเซลก่อนและจะแบ่งส่วนของรูปภาพออกเป็นเป็นตารางทั้งหมด S^2 ส่วน $(S \times S)$ แต่ละส่วนสามารถทำนายว่า มีวัตถุอยู่ในส่วนนั้น ๆ ได้จำนวน B กล่อง หาก $I^{obj} = 1$ เกิดขึ้นในกรณีที่มีวัตถุปรากฏอยู่ภายในกล่องนอกเหนือจากนั้นมีค่าเป็น 0 ในทางตรงกันข้าม หาก $I^{noobj} = 1$ เกิดขึ้นในกรณีที่ไม่มีมีวัตถุปรากฏอยู่ภายในกล่อง สำหรับ λ ทำหน้าที่ควบคุมสมดุลของวัตถุที่สนใจ (Positive sample) และวัตถุที่ไม่สนใจ (Negative sample) ในที่นี้กำหนดให้มีค่าเท่ากับ 0.5 สำหรับค่า C เป็นค่าความมั่นใจที่จะมีวัตถุปรากฏอยู่ภายในกล่องซึ่งได้จากการทำนายของ YOLO สำหรับส่วนสุดท้าย $p(c)$ เป็นค่าความน่าจะเป็นของแต่ละชนิดของวัตถุที่จะปรากฏภายในกล่องได้มาจากการทำนายของ YOLO เช่นกัน

ในกระบวนการฝึกสอนจะทำการปรับพารามิเตอร์ของโมเดลด้วย Adam Optimizer (Kingma & Ba, 2017) ทำการฝึกสอนทั้งหมด 500 รอบ ใช้ค่าดอปเอ้าต์ (Dropout) เท่ากับ 0.5 ใช้ละครั้งของการฝึกสอนทำการนำรูปเข้าจำนวน 4 ภาพ (batch size เท่ากับ 4) โดยมีเงื่อนไข หากมีการเปลี่ยนแปลงของค่าผิดพลาดของ Loss ทั้งหมดไม่เกิน 0.001 เมื่อทำการเปลี่ยบเทียบกับรอบติดต่อกันมากกว่า 50 ครั้งจึงจะการหยุดการฝึกสอน จากการฝึกสอนทำให้ได้ค่าของ Loss ต่ำสุดอยู่ที่ 2.355 ที่รอบ 411 และทำการหยุดการฝึกสอนก่อนครบ 500 รอบ รายละเอียดของค่า Loss ของแต่ละรอบแสดงดังรูปที่ 2



รูปที่ 2 ค่า Loss ที่ทำการฝึกสอนโมเดลในแต่ละรอบโดยแสดงแบ่งออกเป็น Class Loss, Conf Loss, และ CIou Loss

3.4 การวัดประสิทธิภาพของโมเดล

การวัดประสิทธิภาพของโมเดลสำหรับตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยนั้น พิจารณาจากค่า F_{1score} ซึ่งสร้างขึ้นมาเพื่อเป็นการวัดประสิทธิภาพของโมเดล โดยพิจารณาจากค่าความแม่นยำ (Precision) และความสามารถในการค้นคืน (Recall) ค่าของ F_{1score} จะอยู่ในช่วงระหว่าง 0 ถึง 1 หากให้ค่าเข้าใกล้ 1 นั่นแสดงว่าโมเดลมีประสิทธิภาพในการทำงานที่ดี สำหรับค่าของ F_{1score} แสดงดังนี้

$$F_1\text{score} = 2\left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\right)$$

โดยที่

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

ในขณะที่

True Positive (TP) คือ จำนวนตำแหน่งที่โมเดลระบุว่ามีกากบาท และ ณ ตำแหน่งนั้นมีกากบาทจริง

False Positive (FP) คือ จำนวนตำแหน่งที่โมเดลระบุว่ากากบาท แต่ ณ ตำแหน่งนั้นไม่มีกากบาท

False Negative (FN) คือ จำนวนตำแหน่งที่โมเดลระบุว่าไม่มีกากบาท แต่ ณ ตำแหน่งนั้นมีกากบาท

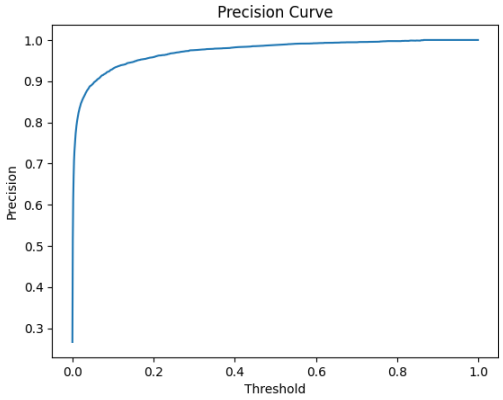
รูปที่ 3 ภาพแสดงการพิจารณาการเกิด TP, FP และ FN ของการตรวจจับกากบาทบนกระดาษคำตอบปรนัย

โดยตัวอย่างการพิจารณาว่าเป็น TP FP และ FN แสดงดังรูปที่ 3 ซึ่งจะเห็นได้ว่ากรอบสีที่ปรากฏขึ้นบนภาพเป็นการระบุตำแหน่งโดยที่โมเดลคาดว่ามีกากบาทปรากฏอยู่ ณ ตำแหน่งนั้น ๆ ภายในกรอบ โดยกรอบสีแดงแสดงถึงกรณีที่โมเดลระบุว่าตำแหน่งนั้น ๆ มีกากบาทปรากฏอยู่และมีกากบาทปรากฏขึ้นจริง (TP) สำหรับกรอบน้ำเงินแสดงถึงกรณีที่โมเดลระบุว่าตำแหน่งนั้น ๆ มีกากบาทปรากฏอยู่ แต่ไม่มีกากบาทปรากฏอยู่หรืออาจเป็นสัญลักษณ์อื่น ๆ ที่ไม่ใช่กากบาท จึงนับว่าเป็นการทำนายผิด (FP) สำหรับกากบาทที่ไม่ถูกตีกรอบจะเป็นกรณีที่โมเดลทำนายว่าไม่มีกากบาทบริเวณนั้นจึงนับว่าเป็นการทำนายที่ผิด (FN)

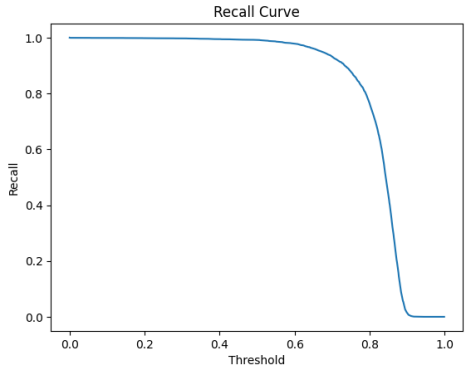
4. ผลการวิจัย

ผลการทดสอบประสิทธิภาพของโมเดลสำหรับการตรวจจับเครื่องหมายกากบาทภายในกระดาษคำตอบปรนัย โดยจะทำการกำหนดค่า IOU อยู่ที่ 0.7 สำหรับการทดสอบและทำการปรับเทรสโฮลด์ (Threshold) โดยกำหนดค่าตั้งแต่ 0.0 ถึงค่า 0.99 หากตำแหน่งใดมีความมั่นใจว่ามีตำแหน่งของกากบาท ปรากฏ (Confidence) อยู่มากกว่าค่า Threshold จะทำการแสดงกรอบหรือตำแหน่งนั้นขึ้นมา และในทางตรงกันข้ามหากค่าความมั่นใจของตำแหน่งนั้น ๆ น้อยกว่าค่า Threshold จะไม่นำมาแสดงบนภาพหลังจากทำการฝึกสอนโมเดลกับชุดฝึกสอน 100 ภาพ และทดสอบกับภาพชุดทดสอบ 63 ภาพแล้ว ผลการทดลอง พบว่าเมื่อกำหนดค่า Threshold ที่แตกต่างกัน จากนั้นทำการประเมินค่าของ Precision และ Recall แสดงดัง

รูปที่ 4



(a) Precision

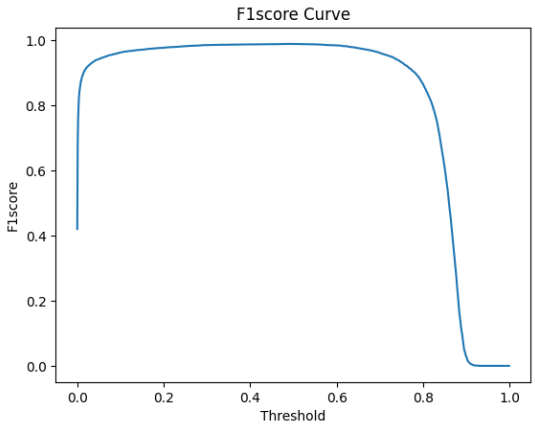


(b) Recall

รูปที่ 4 ภาพความสัมพันธ์ของ (a) Precision และ (b) Recall เมื่อกำหนดค่า Threshold ที่แตกต่างกัน

รูปที่ 4b แสดงค่าประสิทธิภาพของโมเดลในการค้นหาตำแหน่งของวัตถุ (Recall) พบว่า ค่า Recall มีค่าลดลงเมื่อเพิ่มค่า Threshold ขึ้น สำหรับการตรวจสอบค่าความแม่นยำในการระบุวัตถุนั้นว่าเป็นกากบาทหรือไม่จะพิจารณาจาก ค่าของ Precision ซึ่งแสดงใน

รูปที่ 4a หากค่าของ Precision เข้าใกล้ 1 แสดงถึง โมเดลมีความแม่นยำสูง ผลการทดลองพบว่าเมื่อค่า Threshold เพิ่มมากขึ้น ทำให้ค่า Precision ของโมเดล เพิ่มมากขึ้นเช่นกัน ซึ่งค่า Precision สูงสุดที่ 1 เมื่อ Threshold ที่ 0.891 นอกจากนี้หากพิจารณาความสัมพันธ์ระหว่างค่า Precision และ Recall แล้ว พบว่า Recall จะแสดงถึงประสิทธิภาพของโมเดลในการค้นหาตำแหน่งของวัตถุที่เป็นไปได้ทั้งหมด ตามค่าของ Threshold ที่กำหนดไว้ หากค่า Threshold มีค่าต่ำจะทำให้โมเดลแสดงทำนายตำแหน่งของวัตถุมากขึ้น และมีโอกาสมากที่จะสามารถระบุตำแหน่งของกากบาทที่มีในกระดาษคำตอบปรนัยได้ครบ แต่จะส่งผลทำให้ค่า Precision สำหรับระบุว่าตำแหน่งนั้นเป็นเครื่องหมายกากบาทหรือไม่ มีประสิทธิภาพลดลง ซึ่งเป็นเพราะมีการระบุตำแหน่งของวัตถุ เกินกว่าจำนวนของกากบาทที่มีอยู่จริง ในทางตรงกันข้ามหากค่า Threshold มีค่าสูงจะทำให้โมเดลทำ การทำนายตำแหน่งของวัตถุได้น้อยลง แต่จะมีความแม่นยำในสำหรับการระบุว่าตำแหน่งนั้นมีกากบาทหรือไม่ สูงขึ้นดังนั้นจึงมีความจำเป็นจะทำการหาจุดที่เหมาะสมที่สุดในการกำหนดค่า Threshold ซึ่งพิจารณาจากค่า F_{1score} แสดงดังรูปที่ 5

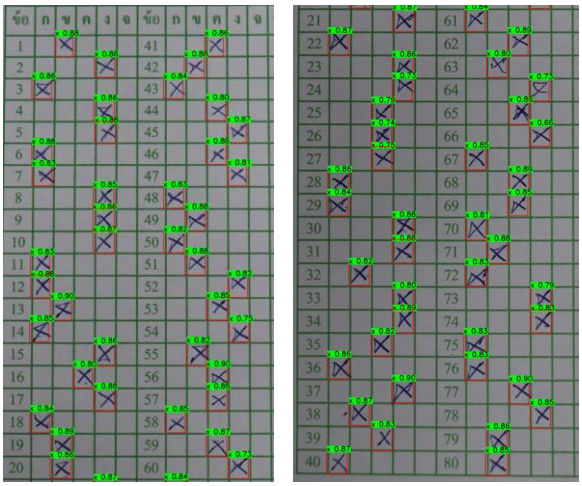


รูปที่ 5 ภาพแสดงค่าประสิทธิภาพ F_{1score} ของโมเดลเมื่อกำหนดค่า Threshold ที่แตกต่างกัน

จากรูปที่ 5 จะเห็นได้ว่าค่า Threshold อยู่ที่ตั้งแต่ 0.2 ถึง 0.6 จะให้ค่า F_{1score} ที่เข้าใกล้ 1 หากและให้ค่า F_{1score} สูงสุดอยู่ที่ 0.989 เมื่อกำหนดค่า Threshold ที่ 0.496 ดังนั้นหากทำการปรับค่า Threshold ที่ 0.496 จะทำให้โมเดลมีประสิทธิภาพดีที่สุด สำหรับการตรวจจับและระบุตำแหน่งของกากบาทในกระดาษคำตอบปรนัย โดยตัวอย่างการตรวจจับกากบาทในภาพของกระดาษคำตอบปรนัย เมื่อกำหนดค่า Threshold ที่ 0.496 แสดงดังรูปที่ 6

5. การวิเคราะห์ความผิดพลาด

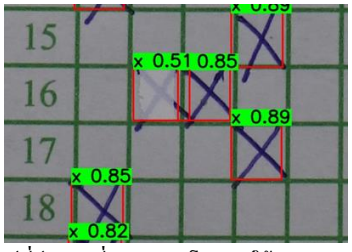
จากผลการทดลองจะเห็นได้ว่าโมเดลสามารถทำงานได้ดีในการตรวจจับและระบุตำแหน่งของกากบาท อย่างไรก็ตามพบว่าในการตรวจจับกากบาทนั้นโมเดลยังมีข้อผิดพลาดอยู่ซึ่ง ทางคณะผู้วิจัยได้ทำการตรวจสอบ



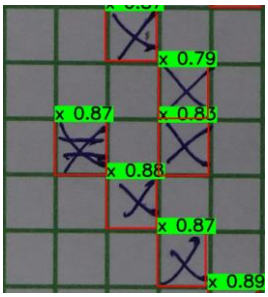
(a) ส่วนต้น (b) ส่วนท้าย

รูปที่ 6 แสดงการตรวจจับกากบาทของกระดาษคำตอบปรนัยแบ่งเป็น (a) ส่วนต้นและ (b) ส่วนท้ายของ กระดาษคำตอบ

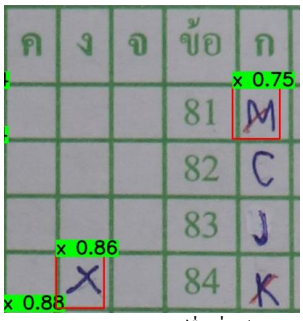
ข้อผิดพลาดของระบุตำแหน่งของกากบาทที่เกิดขึ้นแล้วพบว่า การกากบาทในตัวเลือกของกระดาษคำตอบบางชุดไม่ชัดเจน ซึ่งอาจส่งผลให้โมเดลทำการตรวจจับได้ว่าเป็นสัญลักษณ์กากบาท อาทิเช่น กรณีที่มีการเปลี่ยนคำตอบโดยการใช้ปากกาลบคำผิด แต่ยังคงปรากฏลวดลายที่ยังเป็นสัญลักษณ์ของกากบาทปรากฏอยู่ แสดงดังรูปที่ 7a และการเปลี่ยนแปลงคำตอบโดยการฝนทับตัวเลือกเดิม แต่ยังคงมีลักษณะเฉพาะของกากบาทปรากฏทำให้โมเดลทำนายว่าเป็นกากบาทแสดงดังรูปที่ 7b สำหรับในกรณีที่มีสัญลักษณ์อื่น ๆ ที่คล้ายคลึงกับกากบาท เช่นตัวอักษร R แสดงดังรูปที่ 7c ซึ่งโมเดลพยายามทำนายว่าตำแหน่งนั้นมีสัญลักษณ์กากบาทปรากฏอยู่



(a) กรณีที่มีการเปลี่ยนคำตอบโดยการใช้ปากกาลบคำผิด



(b) กรณีที่มีการเปลี่ยนคำตอบโดยทำสัญลักษณ์อื่น



(c) กรณีการตรวจจับสัญลักษณ์อื่นที่คล้ายกับกากบาท

รูปที่ 7 ตัวอย่างข้อผิดพลาดของโมเดลในการการตรวจจับกากบาทในกรณีดังนี้ (a) กรณีที่มีการเปลี่ยนคำตอบโดยการใช้ปากกาลบคำผิด (b) กรณีที่มีการเปลี่ยนคำตอบโดยทำสัญลักษณ์อื่น และ (c) กรณีการตรวจจับสัญลักษณ์อื่นที่คล้ายกับกากบาท

6. สรุปผลวิจัย

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลทางคณิตศาสตร์ที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยด้วยกระบวนการตรวจจับวัตถุและการเรียนรู้ของเครื่องเชิงลึก สำหรับการตรวจจับกากบาทในกระดาษคำตอบปรนัยจะอาศัยเทคนิค YOLOv8 และทำการสร้างโมเดลที่สามารถตรวจจับกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัยอาศัยเทคนิค CNN จากนั้นทำการวัดประสิทธิภาพโมเดลที่พัฒนาขึ้นจากค่าของ Precison, Recall และ F_{1score} ตามลำดับ

ผลการศึกษาพบว่า เมื่อกำหนดให้ Threshold เพิ่มมากขึ้น ทำให้ค่า Precision ของโมเดลเพิ่มมากขึ้น แต่สำหรับค่าของ Recall ของโมเดลมีแนวโน้มลดลงเมื่อเพิ่มค่าของ Threshold เพิ่มขึ้น และที่สำคัญการพิจารณาจุดที่เหมาะสมที่สุดในการกำหนดค่า Threshold ซึ่งจะพิจารณาจากค่าของ F_{1score} พบว่า เมื่อทำการปรับค่า Threshold ที่ 0.496 ทำให้ได้ ค่าของ F_{1score} สูงสุดเท่ากับ 0.989 ซึ่งส่งผลทำให้โมเดลมีประสิทธิภาพดีที่สุดสำหรับการตรวจจับและระบุตำแหน่งของกากบาทแบบอัตโนมัติในกระดาษคำตอบปรนัย

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาราชภัฏอุตรดิตถ์ที่สนับสนุนเครื่องแม่ข่ายสำหรับประมวลผลและชุดข้อสอบในการศึกษาครั้งนี้

เอกสารอ้างอิง

Alomran, M., & Chai, D. (2018). Automated scoring system for multiple choice test with quick feedback. *International Journal of Information and Education Technology*, *8*(8), 538-545. https://doi.org/10.18178/ijiet.2018.8.8.1096

Chinnasarn, K., & Rangsanseri, Y. (1998). A multiple choice test checking system that uses the principles of image processing. *The 36th Kasetsart University Academic Conference*, 218. (in Thai)

de Elias, E. M., Tasinaffo, P. M., & Hirata Jr, R. (2021). Optical mark recognition: Advances, difficulties, and limitations. *SN Computer Science*, *2*(5), 367.

Deepak Kumar P. (2023). Cloud computing. International Scientific Journal of Engineering and Management, 02. https://doi.org/10.55041/ISJEM00289

Deng, H., Wang, F., & Liang, B. (2008, December). A low-cost OMR solution for educational applications. *In 2008 IEEE international symposium on parallel and distributed processing with applications*. pp. 967-970. https://doi.org/10.1109/ISPA.2008.130

Fisteus, J. A., Pardo, A., & García, N. F. (2013). Grading multiple choice exams with low-cost and portable computer-vision techniques. Journal of Science Education and Technology, 22, 560-571. https://doi.org/10.1007/s10956-012-9414-8

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(9), 1904-1916.

Hussmann, S., & Deng, P. W. (2005). A high-speed optical mark reader hardware implementation at low cost using programmable logic. *Real-Time Imaging*, *11*(1), 19-30. https://doi.org/10.1016/j.rti.2005.03.001

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (8.0.0) [Computer software]. https://github.com/ultralytics/ultralytics

Kingma, D. P., & Ba, J. (2017). Adam: A Method for Stochastic Optimization, *Published as a conference paper at ICLR 2015*. https://doi.org/10.48550/arXiv.1412.6980

Kumar, A., Singal, H., & Bhavsar, A. (2018). Cost effective real-time image processing based optical mark reader. World Acad. Sci. Eng. *International Journal of Information Technology and Computer Engineering*, *12*(9), 787-791.

Lin, W., Hasenstab, K., Moura Cunha, G., & Schwartzman, A. (2020). Comparison of handcrafted features and convolutional neural networks for liver MR image adequacy assessment. *Scientific Reports, 10*(1), 20336. https://doi.org/10.1038/s41598-020-77264-y

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. *In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I* 14. pp. 21-37. Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

Nguyen, T. D., Manh, Q. H., Minh, P. B., Thanh, L. N., & Hoang, T. M. (2011). Efficient and reliable camera based multiple-choice test grading system. *In The 2011 International Conference on Advanced Technologies for Communications (ATC 2011)*. pp. 268-271. https://doi.org/10.1109/ATC.2011.6027482

Phatkrajang, P. (2010). Thesis database system, Department of Computer Engineering, Faculty of Engineering, Kamphaeng Saen. http://cpe.eng.kps.ku.ac.th/db_cpeproj/paperShowFile.php?id_pro=66 (in Thai)

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *In Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 779-788.

Sanguansat, P. (2015, June). Robust and low-cost Optical Mark Recognition for automated data entry. *In 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. pp. 1-5. https://doi.org/10.1109/ECTICon.2015.7206937

Spadaccini, A., & Rizzo, V. (2011). A multiple-choice test recognition system based on the gamera framework. arXiv preprint arXiv:1105.3834. https://doi.org/10.48550/arXiv.1105.3834

Tan-ut, P. (2017). Development of a program for checking multiple choice answer sheets using image processing. *Science and Technology Nakhon Sawan Rajabhat University Journal, 9*(10). (in Thai)

Tinh, P. D., & Minh, T. Q. (2024). Automated Paper-based Multiple Choice Scoring Framework using Fast Object Detection Algorithm. *International Journal of Advanced Computer Science & Applications, 15*(1). https://doi.org/10.14569/IJACSA.2024.01501115

Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. pp. 390-391.