

Received: 3 ม.ค. 2568

Revised: 26 มี.ค. 2568

Accepted: 28 มี.ค. 2568

การศึกษาประสิทธิภาพของ Tesseract OCR สำหรับการประมวลผลภาพ
ในการตรวจสอบธุรกรรมทางการเงิน

The Study of Tesseract OCR Performance for Image Processing
in Financial Transaction Verification

ณรงค์เกียรติ นามห้วยทอง¹, ไตรภาค สิทธิแก้ว¹, อรรถวิท ชังคมานนท์¹, สมนึก สิ้นรूपวน¹ และ

ก้องกาญจน์ ดุลยไชย^{1*}

¹สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้

Narongkiad Namhuaithong¹, Traiphakh Sitthieaw¹, Attawit Changkamanon¹,

Somnuek Sinthupuan¹ and Kongkarn Dullayachai^{1*}

¹Computer Science Department, Faculty of Science, Maejo University

*Corresponding author: kongkarn@maejo.mju.ac.th

Abstract

This research aims to enhance Tesseract OCR efficiency for financial transaction verification by examining five image preprocessor codenames: Image Preprocessor Alpha (IPPA), Image Preprocessor 2 or IPP2, IPP6, IPP7, and IPP12. (Disintegration, n.d.) The study employed a two-phase methodology: accuracy testing on 116 bank transfer receipts and performance evaluation using datasets of 250 to 1,500 transactions. The findings revealed IPP12 (incorporating Grayscale, Resize, Contrast, and Sauvola Threshold) as the optimal preprocessing approach. The Sauvola Threshold method effectively consolidated preprocessing techniques into a single, accurate method. Accuracy metrics using Confusion Matrix demonstrated impressive results: Accuracy (81.03%), Precision (82.14%), Recall (90.79%), and F1-Score (86.26%), indicating balanced performance in reducing false positives and maintaining comprehensive receipt detection. Performance testing using GO programming language benchmarking tools showed the system achieved an average processing time of 5.313 seconds, utilizing 0.277 GB memory and 5.47M allocated space. The research highlights the significance of strategic image processing technique selection in improving OCR performance for financial document verification.

Keywords: Tesseract OCR; Image Processing; Sauvola Threshold; Financial Transaction Verification; Character Recognition Accuracy

บทคัดย่อ

งานวิจัยนี้มุ่งพัฒนาประสิทธิภาพของ Tesseract OCR สำหรับการตรวจสอบธุรกรรมทางการเงิน โดยศึกษาและเปรียบเทียบรหัสชุดประมวลผลภาพ 5 รูปแบบ ได้แก่ Image Preprocessor Alpha (IPPA), IPP2, IPP6, IPP7 และ IPP12 (Disintegration, n.d.) การวิจัยดำเนินการผ่านสองขั้นตอนหลัก คือ การทดสอบความแม่นยำบนใบเสร็จการโอนเงินจำนวน 116 รายการ และการประเมินประสิทธิภาพการประมวลผลกับชุดข้อมูล 250, 500, 1,000 และ 1,500 รายการ ผลการศึกษาพบว่ารหัสชุดประมวลผลภาพ IPP12 ซึ่งประกอบด้วยเทคนิค Grayscale, Resize, Contrast และ Sauvola Threshold มีประสิทธิภาพสูงสุด โดยเฉพาะเทคนิค Sauvola Threshold ที่สามารถรวมเทคนิคการประมวลผลอื่นๆ ให้อยู่ในวิธีเดียว การวัดประสิทธิภาพด้วย Confusion Matrix แสดงผลลัพธ์ที่อยู่ในเกณฑ์ที่ยอมรับได้ โดยประกอบด้วย Accuracy 81.03% Precision 82.14% Recall 90.79% และค่าดัชนี F1-Score 86.26% การทดสอบประสิทธิภาพด้วยภาษา GO พบว่าระบบใช้เวลาประมวลผลเฉลี่ย 5.313 วินาที ใช้หน่วยความจำ 0.277 GB และจัดสรรพื้นที่ 5.47M งานวิจัยนี้ชี้ให้เห็นความสำคัญของการเลือกเทคนิคประมวลผลภาพที่เหมาะสมในการเพิ่มประสิทธิภาพการตรวจสอบเอกสารทางการเงิน

คำสำคัญ: Tesseract OCR; การประมวลผลภาพ; Sauvola Threshold; การตรวจสอบธุรกรรมทางการเงิน; ความแม่นยำการรู้จำตัวอักษร

1. บทนำ

เทคโนโลยี Optical Character Recognition (OCR) โดยเฉพาะ Tesseract OCR (Google, n.d.) เป็นเครื่องมือยอดนิยมที่ถูกนำมาประยุกต์ใช้ในการแปลงข้อมูลจากรูปภาพให้เป็นข้อความ ดังนั้น Tesseract OCR จึงถูกนำมาประยุกต์ใช้อย่างแพร่หลายในการแปลงข้อมูลจากภาพใบเสร็จการโอนเงินโดยเฉพาะใบเสร็จการโอนเงินผ่านแอปพลิเคชันธนาคาร (Mobile Banking) ให้เป็นข้อความดิจิทัล เพื่อช่วยในการตรวจสอบและจัดเก็บข้อมูลการทำธุรกรรมอย่างเป็นระบบ อย่างไรก็ตาม ยังคงประสบปัญหาด้านความแม่นยำในการรู้จำตัวอักษรและตัวเลข เนื่องจากหลายปัจจัย ได้แก่ คุณภาพของภาพใบเสร็จที่มีความหลากหลาย ความแตกต่างของพื้นหลังลายน้ำ รูปแบบและฟอนต์ตัวอักษรในใบเสร็จของแต่ละธนาคาร รวมถึงธนาคารมีการเปลี่ยนพื้นหลังตามเทศกาลต่าง ๆ ส่งผลให้เกิดข้อผิดพลาดในการอ่านข้อมูลสำคัญ ได้แก่ จำนวนเงิน วันและเวลาที่ทำธุรกรรม ตลอดจนเลขที่บัญชี

ธนาคารของทั้งผู้ส่งและผู้รับ ซึ่งก่อให้เกิดผลกระทบร้ายแรงต่อการดำเนินธุรกิจ เช่น ความล่าช้าในการยืนยันการชำระเงิน หรือความเสี่ยงต่อการทุจริตจากการปลอมแปลงใบเสร็จ

ดังนั้น จึงมีการศึกษาและพัฒนาประสิทธิภาพ Tesseract OCR ด้วยเทคนิคการประมวลผลภาพสำหรับการตรวจสอบธุรกรรมทางการเงิน เพื่อเพิ่มความแม่นยำในการอ่านข้อมูลจากใบเสร็จการโอนเงิน และลดความเสี่ยงในด้านการถูกปลอมแปลงใบเสร็จการชำระเงิน ทั้งยังมีการทดสอบประสิทธิภาพสำหรับการจัดการข้อมูลขนาดใหญ่

2. วัตถุประสงค์

1. เพื่อเพิ่มความแม่นยำในการรู้จำข้อมูลที่สำคัญจากรูปภาพใบเสร็จการโอนเงิน
2. เพื่อประเมินประสิทธิภาพในสภาพแวดล้อมที่แตกต่างกันของแต่ละทิศการประมวลผลภาพ และข้อมูลประกอบการตัดสินใจในการนำผลการประมวลผลดังกล่าวไปประยุกต์ใช้ในตามสภาพแวดล้อมของภาคอุตสาหกรรม อย่างมีประสิทธิภาพ

3. งานวิจัยที่เกี่ยวข้อง

Jaakko Sauvola, Matti Pietikäinen (2000) ได้ศึกษาเรื่อง การแปลงภาพเอกสารเป็นภาพสองระดับแบบปรับเปลี่ยนได้ (Adaptive Document Image Binarization) นำเสนอวิธีการใหม่สำหรับการแปลงภาพเอกสารให้เป็นภาพขาวดำแบบปรับตัวได้ โดยมองว่าหน้าเอกสารเป็นการรวมกันของส่วนประกอบย่อยต่างๆ เช่น ตัวอักษร พื้นหลัง และรูปภาพ ปัญหาที่เกิดจากสัญญาณรบกวน ความสว่าง และความเสื่อมสภาพที่เกี่ยวข้องกับแหล่งข้อมูลหลายประเภท ได้รับการแก้ไขด้วยวิธีนี้ มีการนำเสนออัลกอริทึมใหม่สองตัวเพื่อกำหนดค่าความเข้มของสีในระดับท้องถิ่นสำหรับแต่ละพิกเซล การประเมินประสิทธิภาพของอัลกอริทึมใช้ภาพทดสอบที่มีข้อมูลอ้างอิง (Ground-truth) เมตริกสำหรับการประเมินผลการแปลงเป็นขาวดำของภาพตัวอักษรและภาพสังเคราะห์ รวมถึงขั้นตอนการจัดอันดับตามน้ำหนักเพื่อแสดงผลลัพธ์สุดท้าย อัลกอริทึมที่นำเสนอได้รับการทดสอบกับภาพที่มีส่วนประกอบของเอกสารและการเสื่อมสภาพหลายรูปแบบ ผลลัพธ์ถูกเปรียบเทียบกับเทคนิคที่เป็นที่รู้จักหลายวิธีในเอกสารงานวิจัย ผลการเปรียบเทียบแสดงให้เห็นว่าวิธีการนี้สามารถปรับตัวและทำงานได้ดีทั้งในเชิงคุณภาพและเชิงปริมาณในแต่ละกรณี

Lily Rojabiyati Mursari, Antoni Wibowo (2021) ได้ศึกษาเรื่อง ประสิทธิภาพของการประมวลผลภาพเบื้องต้นในการรู้จำลายมือเขียนแบบดิจิทัลด้วยการใช้ OCR Tesseract (The Effectiveness of Image Preprocessing on Digital Handwritten Scripts Recognition with The Implementation of OCR Tesseract) งานวิจัยนี้ได้นำเสนอ การปรับปรุงประสิทธิภาพของ OCR ด้วยการประมวลผลภาพล่วงหน้า (Image Preprocessing) ซึ่งรวมถึงการแปลงภาพเป็น

ระดับสีเทา (Grayscale Conversion) การดำเนินงานทางสัณฐานวิทยา (Morphological Operations) และการกำจัดสัญญาณรบกวน (Noise Removal) ผลลัพธ์เฉลี่ยของการอ่านอักขระจากภาพมีความแม่นยำอยู่ที่ 79.26% ซึ่งถือว่าเป็นความสำเร็จที่ช่วยเพิ่มประสิทธิภาพของ OCR ได้อย่างชัดเจน

Meiling Lu, Yingxia Yang (2024) ได้ศึกษาเรื่อง การประยุกต์ใช้การประมวลผลภาพในสัญญาอัจฉริยะ: การตรวจสอบธุรกรรมทางการเงินแบบอัตโนมัติ (Image Processing Applications in Smart Contracts: Automated Financial Transaction Verification) ด้วยการนำสัญญาอัจฉริยะ (Smart Contract) มาใช้อย่างแพร่หลายในการทำธุรกรรมทางการเงินแบบอัตโนมัติ การประมวลผลข้อมูลภาพที่เกี่ยวข้องกับธุรกรรมทางการเงินอย่างแม่นยำและมีประสิทธิภาพได้กลายเป็นความท้าทายที่สำคัญ การดำเนินการสัญญาอัจฉริยะให้สำเร็งนั้นขึ้นอยู่กับ การตรวจสอบความถูกต้องของภาพใบเสร็จการทำธุรกรรม แต่เทคโนโลยีการประมวลผลภาพที่มีอยู่ในปัจจุบันยังคงมีข้อจำกัดในการจัดการกับความซับซ้อนของพื้นหลัง การรบกวนของสัญญาณรบกวน และความแม่นยำในการสกัดข้อความ การศึกษานี้จึงนำเสนอแนวทางการประมวลผลภาพแบบครอบคลุมที่มุ่งเน้นการเพิ่มประสิทธิภาพการตรวจสอบธุรกรรมทางการเงินแบบอัตโนมัติ โดยงานวิจัยมุ่งเน้นไปที่สี่ด้านหลัก ได้แก่ การแยกเส้นตารางและบริเวณข้อความในภาพ การใช้การแปลงภาพเป็นขาว-ดำแบบ Sauvola ที่ปรับตัวได้เฉพาะที่ การตรวจจับและสร้างตารางขึ้นใหม่ และเทคนิคการสกัดข้อความและการฟื้นฟูส่วนที่แตกหัก ผ่านความพยายามเหล่านี้ การศึกษามุ่งหวังที่จะให้การสนับสนุนทางเทคนิคที่มีประสิทธิภาพและน่าเชื่อถือมากขึ้นสำหรับการตรวจสอบธุรกรรมทางการเงินในสัญญาอัจฉริยะ อันจะนำไปสู่การพัฒนาเทคโนโลยีสัญญาอัจฉริยะต่อไป

Shijian Lu, Bolan Su, Chew Lim Tan (2010) ได้ศึกษาเรื่อง การแปลงภาพเอกสารเป็นภาพสองระดับโดยใช้การประเมินพื้นหลังและขอบเส้น (Document Image Binarization Using Background Estimation and Stroke Edges) ได้นำเสนอเทคนิคการแปลงภาพเอกสารเป็นไบนารีที่ช่วยแยกข้อความจากภาพเอกสารที่เสื่อมสภาพอย่างรุนแรงได้อย่างแม่นยำ โดยอาศัยการสังเกตว่าพื้นหลังเอกสารมักมีสีและลวดลายสม่ำเสมอ ขณะที่ข้อความมีระดับความเข้มที่แตกต่างออกไป เทคนิคนี้เริ่มด้วยการประมาณพื้นหลังเอกสารผ่านกระบวนการปรับเรียบด้วยสมการพหุนามแบบวนซ้ำ เพื่อลดผลกระทบจากการเสื่อมสภาพของเอกสาร จากนั้นตรวจจับขอบข้อความด้วยความชันของภาพแบบ L1-norm และใช้ค่าเรโซลต์แบบพื้นที่ (Local Threshold) เพื่อแยกข้อความออกมา เทคนิคนี้ได้รับการประเมินว่าให้ผลลัพธ์ที่ดีที่สุดในการแข่งขัน DIBCO 2009 เมื่อเปรียบเทียบกับ 43 อัลกอริทึมจากกลุ่มวิจัยนานาชาติ 35 กลุ่ม

Supriya Sunil Lokhande, Dawande (2015) ได้ศึกษาเรื่อง การสำรวจเทคนิคการแปลงภาพเอกสารเป็นภาพสองระดับ (A Survey on Document Image Binarization Techniques) พบว่า เทคนิคนี้ถูกใช้ในการแยกข้อความส่วนหน้า (Foreground text) ออกจากพื้นหลัง (Background text) ในเอกสารที่มีคุณภาพต่ำหรือเสื่อมสภาพอย่างรุนแรง ในงานวิจัยนี้ ได้มีการสำรวจอย่างครอบคลุมเกี่ยวกับเทคนิคการแปลงภาพเอกสารเป็นภาพสองระดับหรือภาพขาว-ดำที่ล้ำสมัยหลายเทคนิค หลังจากอธิบายเทคนิคการแปลงภาพเอกสารเหล่านี้แล้ว ได้มีการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิค โดยใช้ตัวชี้วัดด้านประสิทธิภาพต่าง ๆ ที่ถูกใช้กันอย่างแพร่หลายสำหรับการวิเคราะห์และรู้จำภาพเอกสาร จากการเปรียบเทียบ พบว่าวิธีการปรับคอนทราสต์เชิงปรับตัว (Adaptive Contrast) เป็นวิธีที่มีประสิทธิภาพดีที่สุด พร้อมทั้งได้อธิบายรายละเอียดของแบบจำลองทางคณิตศาสตร์และแผนภาพบล็อกของวิธีนี้อย่างครบถ้วน

4. วิธีการดำเนินการ

จัดเตรียมข้อมูล ชุดข้อมูลที่ใช้ในการทดสอบครั้งนี้แบ่งออกเป็น 2 ประเภท ได้แก่

4.1 การทดสอบความแม่นยำ

4.1.1 ชุดข้อมูลสำหรับทดสอบความแม่นยำ คือ สลิปอิเล็กทรอนิกส์ จำนวน 116 รายการจากธนาคาร 9 แห่ง ได้แก่ ธนาคารกรุงไทย ธนาคารกรุงศรีอยุธยา ธนาคารกรุงเทพ ธนาคารทีเอ็มบีธนชาต MAKE ธนาคารออมสิน ธนาคารไทยพาณิชย์

4.1.2 แนวทางการทดสอบความแม่นยำ ผู้วิจัยได้ทำการตั้งสมมติฐานการทดสอบความแม่นยำ โดยนำข้อมูลที่ถูกการรู้จำตัวอักษร (OCR) จากสลิปอิเล็กทรอนิกส์ เปรียบเทียบกับข้อมูลจำลองจริง โดยใช้อัลกอริทึมระยะทางเลเวนสไตน์ (Levenshtein Distance) เมตริกสำหรับวัดความแตกต่างระหว่างสองข้อมูล โดยวัดว่าเกณฑ์ความแตกต่างอยู่ในนัยยะสำคัญที่ยอมรับได้หรือไม่ โดยสูตรที่ใช้ในการวัดเกณฑ์ความแตกต่าง (Threshold) ที่ยอมรับได้

$$\text{เกณฑ์ความแตกต่างที่ยอมรับ} = \left(\frac{\text{เกณฑ์ร้อยละขั้นต่ำ}}{100} \right) \times \text{จำนวนตัวอักษรที่ใช้เปรียบเทียบ}$$

4.1.3 รายละเอียดสำหรับเกณฑ์การประเมินความแม่นยำ

ในการประเมินประสิทธิภาพของแบบจำลองเพื่อให้ได้ผลลัพธ์ที่มีความแม่นยำและเป็นมาตรฐาน สามารถนำไปใช้ในการวัดผลได้อย่างมีประสิทธิภาพ ผู้วิจัยได้ใช้วิธีการประเมินผลผ่านเมทริกซ์ความสับสน (Confusion Matrix) ซึ่งเป็นเครื่องมือที่ใช้เปรียบเทียบระหว่างค่าที่ทำนาย (Predicted) กับค่าจริง (Actual) เพื่อวิเคราะห์ประสิทธิภาพของแบบจำลอง โดยเมทริกซ์ดังกล่าวประกอบด้วยค่าหลัก 4 ประเภท

1. ถูกต้องเป็นบวก True Positive (TP) โมเดลทำนายถูกว่าเป็นคลาสบวก (Positive) และค่าจริงก็เป็นคลาสบวก เช่น ทำนายว่าสลิปถูกต้อง (Positive) และสลิปถูกต้องจริง
2. ถูกต้องเป็นลบ True Negative (TN) โมเดลทำนายถูกว่าเป็นคลาสลบ (Negative) และค่าจริงก็เป็นคลาสลบ เช่น ทำนายว่าสลิปไม่ถูกต้อง (Negative) และสลิปไม่ถูกต้องจริง
3. ผิดพลาดเป็นบวก False Positive (FP) โมเดลทำนายผิดว่าเป็นคลาสบวก (Positive) แต่ค่าจริงเป็นคลาสลบ (Negative) เช่น ทำนายว่าสลิปถูกต้อง (Positive) แต่สลิปไม่ถูกต้องจริง
4. ผิดพลาดเป็นลบ False Negative (FN) โมเดลทำนายผิดว่าเป็นคลาสลบ (Negative) แต่ค่าจริงเป็นคลาสบวก (Positive) เช่น ทำนายว่าสลิปไม่ถูกต้อง (Negative) แต่สลิปถูกต้องจริง

หลังจากที่ได้ทดสอบความแม่นยำตามกระบวนการ เมทริกซ์ความสับสน ผู้วิจัยได้นำข้อมูลเหล่านี้มาประเมินจากการแสดงผลการทำนายของแบบจำลองการจำแนกประเภท (Classification Model) เพื่อวัดประสิทธิภาพของโมเดลดังนี้

1. ค่าความถูกต้อง (Accuracy) คือ อัตราส่วนของการทำนายถูกต้องทั้งหมด

$$\text{สูตร} \quad \frac{\text{จำนวนกรณีทดสอบที่ผ่าน}}{\text{จำนวนกรณีทดสอบทั้งหมด}} \times 100$$

2. ความแม่นยำ (Precision) คือ อัตราส่วนของการทำนาย Positive ที่ถูกต้อง

$$\text{สูตร} \quad \frac{\text{True Positives (TP)}}{\text{True Positives (TP) + False Positives (FP)}}$$

3. ความครอบคลุม (Recall) คือ อัตราส่วนของค่าจริง Positive ที่ถูกทำนาย

$$\text{สูตร} \quad \frac{\text{True Positives (TP)}}{\text{True Positives (TP) + False Negatives (FN)}}$$

4. คะแนน F1 (F1-Score) คือ ค่าเฉลี่ยถ่วงน้ำหนักระหว่าง Precision และ Recall ซึ่งเป็นค่าที่บ่งบอกถึงความสมดุลของความแม่นยำโมเดล

$$\text{สูตร} \quad 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

การนำ เมทริกซ์ความสับสน มาใช้เป็นเกณฑ์การประเมินความแม่นยำ ถือเป็นมาตรฐานสากลที่ได้รับการยอมรับในการวัดประสิทธิภาพและวิเคราะห์ข้อผิดพลาดของชุดรหัส โดยช่วยให้สามารถระบุข้อเท็จจริงและความเอนเอียง (Bias) ที่อาจเกิดขึ้นจากชุดข้อมูลที่เป็นตัวเลขได้อย่างมีประสิทธิภาพ

4.2 การทดสอบประสิทธิภาพ

4.2.1 ชุดข้อมูลตัวดำเนินการสำหรับทดสอบประสิทธิภาพ โดยชุดข้อมูลที่ใช้ในการทดสอบมีจำนวนข้อมูลรูปภาพที่ใช้ตัวดำเนินการแตกต่างกันที่ 250 500 1,000 และ 1,500 ตามลำดับ เพื่อวัดความเร็ว หน่วยความจำ และการจัดสรรพื้นที่ โดยชุดข้อมูลตัวดำเนินการที่ผู้วิจัยได้ทำการเตรียมการเริ่มจากการนำจำนวนรูปภาพตั้งต้นที่แตกต่างกัน 10 รูปโดยผู้วิจัยได้ใช้เทคนิคการทำ Table-Driven Test เนื่องจากเทคนิคดังกล่าวช่วยให้สามารถรันการทดสอบแบบเดียวกันได้ด้วย อินพุตและเอาต์พุตที่คาดหวังหลายรายการ อีกทั้งยังสามารถใช้วิธีนี้ในการแบ่งชุดข้อมูลออกเป็นหลายขนาดชุดข้อมูล โดยในแต่ละชุดข้อมูลสามารถเพิ่มจำนวนรูปภาพตั้งต้นให้อยู่ในจำนวนชุดข้อมูลที่ผู้วิจัยต้องการ อีกทั้งการทดสอบครั้งนี้มีจุดมุ่งหมายเพื่ออย่างทดสอบประสิทธิภาพทางด้านความเร็ว หน่วยความจำ และการจัดสรร ดังนั้นวิธีนี้จึงเป็นวิธีที่เหมาะสมที่สุดทั้งเตรียมการชุดข้อมูลและแนวทางทดสอบสำหรับจุดมุ่งหมายเพื่อสังเกตการณ์ประสิทธิภาพการประมวลผลต่อความถี่ของชุดข้อมูลที่ใกล้เคียง และขอขวดยการรับโหลดสูงสุดของแต่ละรหัสการประมวลผลภาพ

รหัสการประมวลผลภาพ มีเทคนิคแตกต่างกัน 5 ชุดรหัส ได้แก่:

- IPPA: มีการใช้เทคนิค Grayscale ซึ่งเป็นเทคนิคพื้นฐานสำหรับการทำ การประมวลผลภาพ ไม่มีการใช้เทคนิคใด ๆ แบบพิเศษ เนื่องจากผู้วิจัยต้องการทดสอบและทราบถึงประสิทธิภาพจากตัว Tesseract OCR โดยตรง
- IPP2: Grayscale, Resize (2 เท่าของขนาดรูปภาพ) , Contrast (+12%), Blur (+3.5%)
- IPP6: Grayscale, Resize (2.5 เท่าของขนาดรูปภาพ), Brightness (+10%), Sharpen (+3%), Image Binarization (Threshold ที่ 171), Blur (+0.5%)
- IPP7: Grayscale, Resize (2.5 เท่าของขนาดรูปภาพ), Brightness (+10%), Sharpen (+15%), Image Binarization (Threshold ที่ 141), Blur (+0.5%), Dilate (Threshold ที่ 10), Gamma (+1%)
- IPP12: Grayscale, Resize (2 เท่าของขนาดรูปภาพ), Contrast (+20%), Sauvola Threshold (ใช้ thresholding_kfactor ที่ 0.2)

โดย IPP ย่อมาจาก Image Preprocessor

เครื่องมือและเทคโนโลยี การทดสอบในครั้งนี้ เนื่องจากผู้วิจัยได้ใช้ภาษา GO ในพัฒนาระบบการประมวลผลภาพในการตรวจสอบธุรกรรมทางการเงิน ดังนั้นภาษา GO จึงนำไปใช้ในการเขียนโค้ดทดสอบ โดยใช้ Standard library GO testing สำหรับการทดสอบความแม่นยำ และใช้ GO

Benchmarks สำหรับการประเมินประสิทธิภาพความเร็ว หน่วยความจำ และการจัดสรรพื้นที่ของกระบวนการประมวลผลภาพแต่ละครั้ง

5. ผลการศึกษา

ผลจากการทดลองได้ดำเนินการทดสอบความแม่นยำการประมวลผลภาพทั้งหมด 116 รายการ โดยผู้วิจัยได้นำผลลัพธ์การความแม่นยำโดยใช้เกณฑ์การประเมินแบบวิธี เมทริกซ์ความสับสน โดยการหาค่า TP, TN, FP และ FN จากนั้นนำค่าเหล่านี้มาคำนวณ ด้วยวิธีการทำ Classification Metrics ทั้ง 4 สูตร โดยมีรายละเอียดดังนี้

ตารางที่ 1 ผลการทดสอบความแม่นยำโดยชุดข้อมูล 116 รายการ

รหัสชุดประมวลผลภาพ	จำนวนใบเสร็จที่ถูกต้อง	TP	TN	FP	FN	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
IPPA	86	61	25	23	7	74.13	72.62	89.71	80.29
IPP2	83	57	26	27	6	71.55	67.86	78.08	72.62
IPP6	92	66	26	18	6	79.31	78.57	91.67	84.66
IPP7	92	65	27	19	4	79.31	77.38	94.20	84.99
IPP12	94	69	25	15	7	81.03	82.14	90.79	86.26

จากตารางที่ 1 ผลการประเมินความแม่นยำด้วยชุดข้อมูล 116 รายการ พบว่าชุดประมวลผลภาพ IPP12 ให้ผลลัพธ์สูงสุดในทุกเมตริกหลัก โดยมีค่าความแม่นยำ (Accuracy) อยู่ที่ 81.03% ค่าความแม่นยำเชิงบวก (Precision) 82.14% และค่าดัชนี F1-Score 86.26% ซึ่งสะท้อนถึงความสมดุลระหว่างการลดข้อผิดพลาดเชิงบวกเทียม (FP) และการตรวจจับใบเสร็จที่ถูกต้องได้ครบถ้วน (Recall 90.79%) สำหรับชุดประมวลผลอื่นๆ นั้น แม้ชุด IPP7 จะมีค่าการตรวจจับใบเสร็จถูกต้องสูงสุด (Recall 94.20%) แต่ค่าความแม่นยำเชิงบวกลดลงเล็กน้อย (77.38%) ขณะที่ชุด IPPA และ IPP2 แสดงค่า Precision ต่ำที่สุด (72.62% และ 67.86% ตามลำดับ) ส่งผลให้ดัชนี F1-Score ต่ำกว่า 80% และ 73% ตามลำดับ ส่วนชุด IPP6 มีความโดดเด่นด้านความสมดุลระหว่าง Precision (78.57%) และ Recall (91.67%) ด้วย F1-Score 84.66% โดยสรุป ชุด IPP12 ถือเป็นวิธีการ

ประมวลผลที่เหมาะสมที่สุด เนื่องจากไม่เพียงมี Accuracy สูงสุด แต่ยังรักษาสมดุลระหว่างการลดข้อผิดพลาดและการครอบคลุมผลลัพธ์ได้ดีกว่าชุดอื่นๆ อย่างมีนัยสำคัญ

ผลการทดสอบประสิทธิภาพ ทางด้านความเร็ว หน่วยความจำ และการจัดสรรพื้นที่ทั้งหมด 4 ชุดข้อมูล ได้แก่ 250, 500, 1,000 และ 1,500 รายการ ตามลำดับ โดยที่เวลาการประมวลผลที่น้อยลงแสดงถึงประสิทธิภาพของระบบ โดยผลจากการทดลองแสดงผลได้ตามตารางดังนี้

ตารางที่ 2 ผลการทดสอบประสิทธิภาพโดยรหัสการประมวลผลภาพ IPPA

ชุดข้อมูล	ค่าเฉลี่ยต่อการดำเนินการ (s)	หน่วยความจำ	การจัดสรร
250	3.531	0.088 GB	10.98M
500	3.454	0.088 GB	10.98M
1000	3.452	0.088 GB	10.98M
1500	3.598	0.088 GB	10.98M

จากตารางที่ 2 พบว่า การทดสอบประสิทธิภาพกับชุดข้อมูลทั้ง 4 ขนาด แสดงให้เห็นผลลัพธ์ที่มีความสม่ำเสมอ โดยชุดข้อมูลขนาด 1,000 รายการมีค่าเฉลี่ยเวลาการดำเนินการที่เหมาะสมที่สุดที่ 3.452 วินาที รองลงมาคือชุดข้อมูลขนาด 500 รายการที่ 3.454 วินาที ในขณะที่การใช้ทรัพยากรระบบที่คงที่ทั้งในด้านหน่วยความจำ (0.088 GB) และการจัดสรรพื้นที่ (10.98M) ซึ่งแสดงให้เห็นถึงประสิทธิภาพในการจัดการทรัพยากรที่ดีและความเสถียรของระบบในการประมวลผลข้อมูลที่มีขนาดแตกต่างกัน



ภาพที่ 1 ตัวอย่างผลลัพธ์การทำงานของ IPPA

ตารางที่ 3 ผลการทดสอบประสิทธิภาพโดยรหัสการประมวลผลภาพ IPP2

ชุดข้อมูล	ค่าเฉลี่ยต่อการดำเนินการ (s)	หน่วยความจำ	การจัดสรร
250	7.135	0.564 GB	43.71M
500	9.286	0.564 GB	43.71M
1,000	6.787	0.564 GB	43.71M
1,500	6.945	0.564 GB	43.71M

จากตารางที่ 3 พบว่ามีค่าเฉลี่ยเวลาการดำเนินการที่เหมาะสมที่สุดเมื่อใช้กับข้อมูลขนาด 1,000 และ 1,500 ชุด โดยมีค่าประมาณ 6.787 และ 6.945 วินาทีตามลำดับ อีกทั้งยังมีการใช้หน่วยความจำที่ต่ำ (578 MB) และจำนวนการจัดสรรที่คงที่ (43.71M) ซึ่งให้เห็นถึงความเหมาะสมสำหรับข้อมูลขนาดใหญ่ที่ไม่ต้องการการจัดสรรพื้นที่มากเกินไป

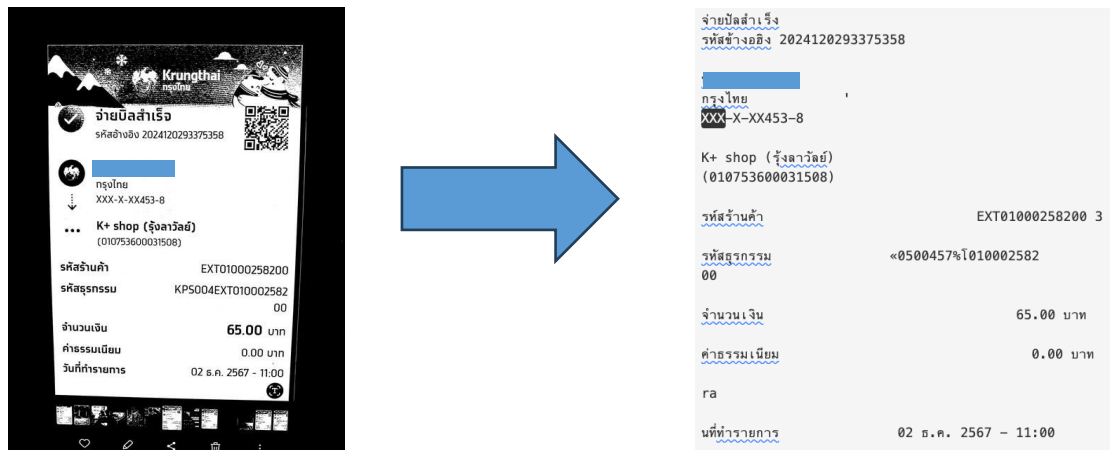


ภาพที่ 2 ตัวอย่างผลลัพธ์การทำงานของ IPP2

ตารางที่ 4 ผลการทดสอบประสิทธิภาพโดยรหัสการประมวลผลภาพ IPP6

ชุดข้อมูล	ค่าเฉลี่ยต่อการดำเนินการ (s)	หน่วยความจำ	การจัดสรร
250	15.968	1.801 GB	109.66M
500	21.450	1.801 GB	109.66M
1,000	15.338	1.801 GB	109.66M
1,500	21.096	1.801 GB	109.66M

จากตารางที่ 4 พบว่าการดำเนินการใช้เวลานานที่สุดเมื่อเปรียบเทียบกับรหัสอื่น โดยข้อมูลขนาด 1,500 ชุดมีค่าเฉลี่ยเวลาสูงสุดถึง 21.096 วินาที อีกทั้งยังใช้หน่วยความจำสูงถึง 1.801 GB ซึ่งสะท้อนถึงข้อจำกัดในการประมวลผลข้อมูลขนาดใหญ่ที่ต้องการทรัพยากรระบบมาก

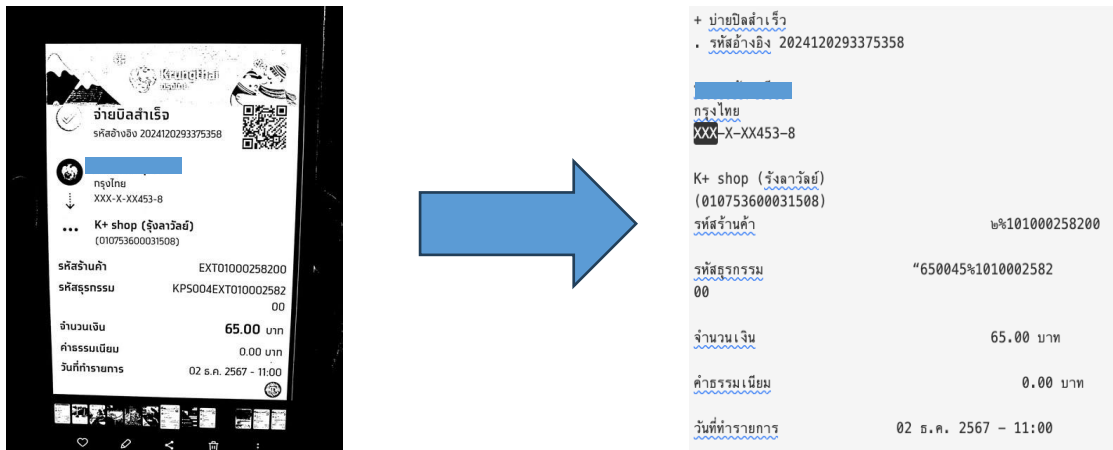


ภาพที่ 3 ตัวอย่างผลลัพธ์การทำงานของ IPP6

ตารางที่ 5 ผลการทดสอบประสิทธิภาพโดยรหัสการประมวลผลภาพ IPP7

ชุดข้อมูล	ค่าเฉลี่ยต่อการดำเนินการ (s)	หน่วยความจำ	การจัดสรร
250	13.099	2.227 GB	148.21M
500	22.682	2.227 GB	148.21M
1,000	20.408	2.227 GB	148.21M
1,500	30.306	2.227 GB	148.21M

จากตารางที่ 5 พบว่ามีการใช้งานหน่วยความจำสูงถึง 2.227 GB ซึ่งมากกว่ารหัสอื่นอย่างเห็นได้ชัด โดยมีค่าเฉลี่ยเวลาการดำเนินการที่เพิ่มขึ้นตามขนาดข้อมูล เช่น ข้อมูลขนาด 1,500 ชุด ใช้เวลาเฉลี่ย 30.306 วินาที การใช้เทคนิคการประมวลผลภาพที่มีความซับซ้อนมากเกินไปส่งผลให้ความเร็วในการประมวลผลภาพลดลง

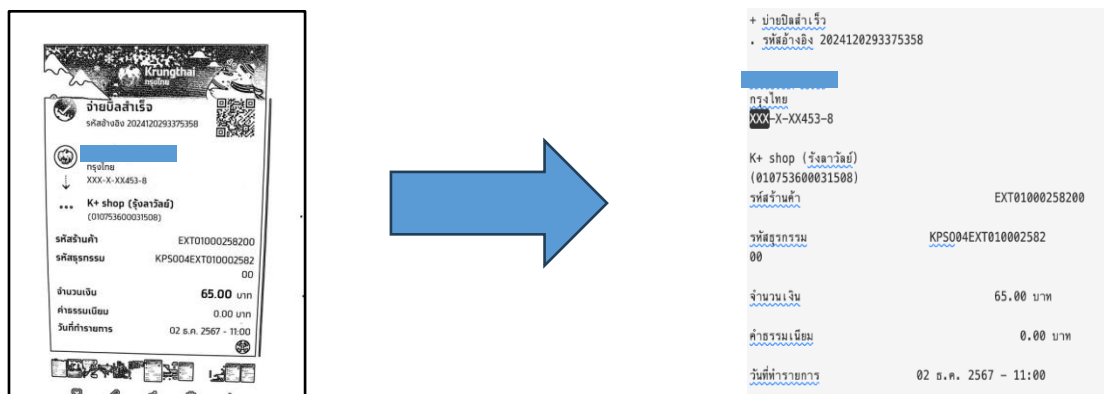


ภาพที่ 4 ตัวอย่างผลลัพธ์การทำงานของ IPP7

ตารางที่ 6 ผลการทดสอบประสิทธิภาพโดยรหัสการประมวลผลภาพ IPP12

ชุดข้อมูล	ค่าเฉลี่ยต่อการดำเนินการ (s)	หน่วยความจำ	การจัดสรร
250	4.303	0.277 GB	5.47M
500	6.499	0.277 GB	5.47M
1,000	5.021	0.277 GB	5.47M
1,500	5.782	0.277 GB	5.47M

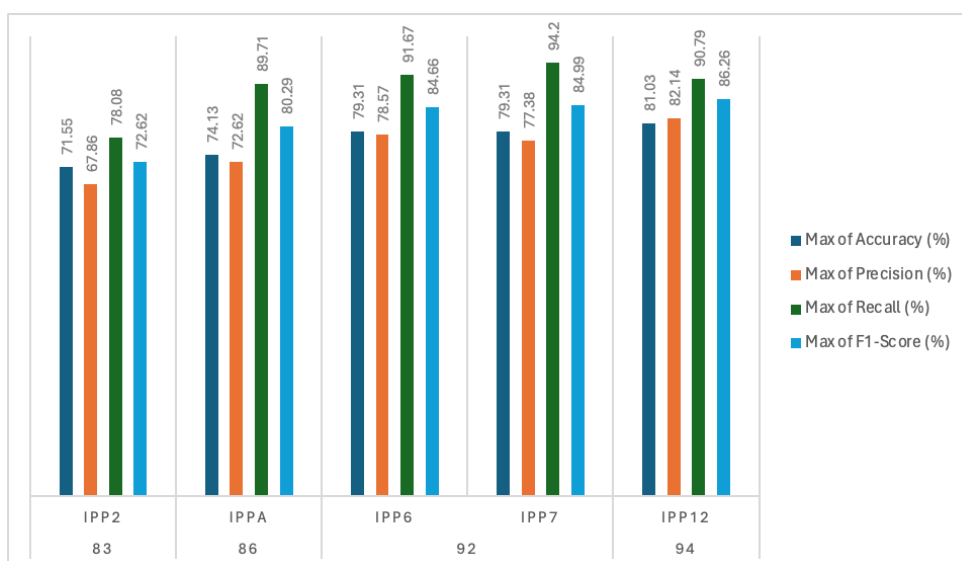
จากตารางที่ 6 พบว่าให้ผลลัพธ์ที่ดีที่สุดในด้านความเร็วและการใช้ทรัพยากร โดยข้อมูลขนาด 250 ชุด มีค่าเฉลี่ยเวลาการดำเนินการเพียง 4.303 วินาที และยังคงใช้หน่วยความจำต่ำสุดที่ 283.64 MB พร้อมกับจำนวนการจัดสรรที่ต่ำที่สุดที่ 5.47M ซึ่งเหมาะสมอย่างยิ่งสำหรับระบบที่มีข้อจำกัดด้านทรัพยากร



ภาพที่ 5 ตัวอย่างผลลัพธ์การทำงานของ IPP12

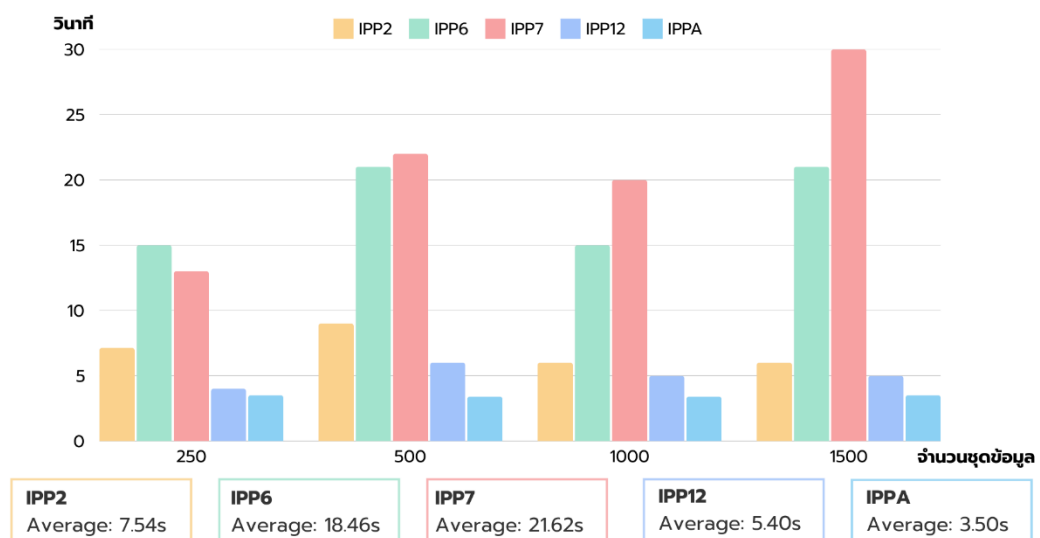
เมื่อเปรียบเทียบกับงานวิจัยของ Meiling Lu, Yingxia Yang (2024) ได้ศึกษาเรื่อง การประยุกต์ใช้การประมวลผลภาพในสัญญาอัจฉริยะ: การตรวจสอบธุรกรรมทางการเงินแบบอัตโนมัติ (Image Processing Applications in Smart Contracts: Automated Financial Transaction Verification) พบว่าบทความต้นฉบับศึกษาผลกระทบของจำนวนหน่วยประมวลผล (4-12 หน่วย) ต่อประสิทธิภาพของโมเดล โดยพบว่าการเพิ่มจำนวนหน่วยประมวลผลส่งผลให้ค่า F-measure เพิ่มขึ้นจาก 84.2% เป็น 90.6% แต่ค่า FPS ลดลงจาก 22.4 เป็น 10.6 ในขณะที่ยานวิจัยนี้นำเสนอแนวทางการพัฒนาที่แตกต่างโดยมุ่งเน้นการเลือกใช้เทคนิคการประมวลผลภาพที่เหมาะสม ผลการวิจัยแสดงให้เห็นว่า IPP12 ซึ่งใช้เทคนิคพื้นฐานเพียง Grayscale, Resize และ Contrast สามารถให้ความแม่นยำสูงถึง 91% พร้อมทั้งใช้ทรัพยากรระบบน้อยที่สุด (หน่วยความจำ 0.277 GB และเวลาประมวลผลเฉลี่ย 4.303-6.499 วินาที) เมื่อเทียบกับ IPP2, IPP6 และ IPP7 ที่มีความซับซ้อนมากกว่า ซึ่งสะท้อนให้เห็นว่าการออกแบบกระบวนการประมวลผลที่มีประสิทธิภาพสามารถให้ผลลัพธ์ที่ดีกว่าการเพิ่มความซับซ้อนของการประมวลผล โดยเฉพาะอย่างยิ่งในบริบทของการตรวจสอบธุรกรรมทางการเงินอัตโนมัติที่ต้องการทั้งความแม่นยำและความเร็วในการประมวลผลงานวิจัยนี้จึงนำเสนอแนวทางใหม่ที่น่าสนใจในการพัฒนาระบบประมวลผลภาพสำหรับการตรวจสอบธุรกรรมทางการเงิน โดยเน้นการเลือกใช้เทคนิคที่เหมาะสมแทนการเพิ่มความซับซ้อนของการประมวลผล ซึ่งสามารถให้ผลลัพธ์ที่ดีกว่าทั้งในแง่ความแม่นยำและประสิทธิภาพการใช้ทรัพยากร

6. อภิปรายผล



ภาพที่ 6 การเปรียบเทียบค่าความแม่นยำ

จากภาพที่ 6 พบว่าค่าความแม่นยำมีแนวโน้มลดลงในตอนแรกแต่เพิ่มขึ้นในตอนหลัง โดยรหัส IPPA มีค่าความแม่นยำร้อยละ 74 แต่รหัส IPP2 มีค่าความแม่นยำต่ำลงอยู่ที่ร้อยละ 71 แต่กลับมาสูงขึ้นอย่างมีนัยสำคัญใน IPP6 และ IPP7 มีค่าความแม่นยำร้อยละ 79 และ IPP12 มีค่าความแม่นยำสูงที่สุดที่ร้อยละ 81 แสดงให้เห็นว่าการปรับปรุงชุดการประมวลผลภาพ มีพัฒนาการที่ดีขึ้นหลังจาก IPP2 ตามลำดับ



ภาพที่ 7 การเปรียบเทียบค่าเฉลี่ยความเร็วในการประมวลผล

จากภาพที่ 7 เมื่อทำการเปรียบเทียบพบว่ารหัส IPP12 ใช้เวลาต่อการดำเนินการน้อยรองลงมาจาก IPPA ที่เป็นโมเดลพื้นฐานใช้ในการทำการประมวลผลรูปภาพที่ใช้เวลาต่อการดำเนินการน้อยที่สุดที่ โดยอ้างอิงจากตารางที่ 6 พบว่า โดย IPP12 มีค่าเฉลี่ยต่อการดำเนินการอยู่ที่ 5.40 วินาทีต่อรายการ และมีความเสถียรภาพสูง โดยสังเกตจากเส้นกราฟ ขณะที่รหัส IPP7 ใช้เวลาต่อการดำเนินการมากที่สุดที่ 21.62 วินาที และมีความเสถียรภาพต่ำ ซึ่งสะท้อนให้เห็นถึงพัฒนาการที่มีประสิทธิภาพของรหัส IPP12 ที่สามารถลดเวลาการประมวลผลได้อย่างมีประสิทธิภาพ และจากการเปรียบเทียบประสิทธิภาพหน่วยความจำ และการจัดสรร พบว่ารหัส IPP7 ใช้หน่วยความจำและการจัดสรรพื้นที่มากที่สุด ขณะที่ IPP2 และ IPP12 ใช้หน่วยความจำและการจัดสรรพื้นที่น้อยกว่าอย่างมีนัยสำคัญ สะท้อนให้เห็นถึงการใช้ทรัพยากรที่มีการบริหารจัดการอย่างมีประสิทธิภาพและเหมาะสมตามความต้องการในแต่ละช่วงเวลา

7. สรุปผลการวิจัย

จากการศึกษาประสิทธิภาพ Tesseract OCR สำหรับการตรวจสอบธุรกรรมทางการเงิน พบว่ารหัส IPP12 ซึ่งประกอบไปด้วยเทคนิค Grayscale, Resize, Contrast และ Sauvola Threshold เหมาะสมที่สุดสำหรับการประมวลผลภาพ เนื่องจากมีค่า Accuracy อยู่ที่ 81.03% ค่า Precision 82.14% ค่า Recall อยู่ที่ 90.79% และค่าดัชนี F1-Score 86.26% และการทดสอบประสิทธิภาพด้วยภาษา GO พบว่าระบบใช้เวลาในการดำเนินการทั้ง 4 ชุดโดยเฉลี่ยข้อมูลเพียง 5.313 วินาทีต่อรายการ ใช้หน่วยความจำ 0.277 GB และจัดสรรพื้นที่ 5.47M ซึ่ง ซึ่งต่ำที่สุดเมื่อเทียบกับรหัสอื่น ๆ แสดงให้เห็นว่าการใช้เทคนิคการประมวลผลภาพที่เหมาะสมสามารถเพิ่มประสิทธิภาพของ Tesseract OCR ในการตรวจสอบธุรกรรมทางการเงิน ทั้งในด้านความแม่นยำ ความเร็ว และการใช้ทรัพยากร

8. ข้อเสนอแนะ

ในการพัฒนาความแม่นยำและประสิทธิภาพของการประมวลผลของรหัสชุดการประมวลผลภาพ ผู้พัฒนาควรให้ความสำคัญกับการออกแบบและพัฒนารหัสชุดการประมวลผลภาพการประมวลผลภาพ (Image Processing) ที่สามารถทำงานร่วมกับเทคโนโลยี Leptonica (Bloomberg, D. n.d.) ได้อย่างเหมาะสม เนื่องจาก Leptonica เป็นเทคโนโลยีสำคัญที่ช่วยเตรียมสภาพแวดล้อมของภาพให้เหมาะสมต่อการประมวลผลด้วย OCR ของ Tesseract การปรับปรุงกระบวนการประมวลผลภาพให้เหมาะสมจึงเป็นสิ่งสำคัญเพื่อเพิ่มความแม่นยำในการตรวจจับและแปลงข้อความจากภาพ

สำหรับการเพิ่มประสิทธิภาพด้านความเร็วในการประมวลผล ผู้พัฒนาควรพิจารณาเลือกใช้ อัลกอริทึมที่เหมาะสมกับเป้าหมายของงาน เนื่องจากอัลกอริทึมแต่ละประเภทมีผลกระทบต่อความเร็วในการประมวลผลและการใช้ทรัพยากรระบบ เช่น หน่วยความจำและพื้นที่จัดเก็บข้อมูล ที่แตกต่างกันอย่างมีนัยสำคัญ ดังนั้น ควรมีการประเมินผลกระทบทั้งในด้านความแม่นยำและประสิทธิภาพควบคู่กัน เพื่อให้สามารถบรรลุวัตถุประสงค์ของการพัฒนาระบบได้อย่างสมดุลและมีประสิทธิภาพสูงสุด

9. กิตติกรรมประกาศ

บทความนี้สำเร็จลุล่วงได้ด้วยดี เนื่องจากได้รับความกรุณาอย่างสูงจาก ผู้ช่วยศาสตราจารย์ ก่องกาญจน์ ดุลยไชย ประธานที่ปรึกษาโครงการ อาจารย์ อรรถวิทย์ ชังคมานนท์ และผู้ช่วยศาสตราจารย์ ดร.สมนึก สินธุปวน กรรมการที่ปรึกษาโครงการ ที่ให้คำแนะนำและข้อคิดเห็นต่าง ๆ อันเป็นประโยชน์อย่างยิ่งต่องานวิจัย ขอขอบคุณบริษัท เพย์ โซลูชั่น จำกัด ที่ให้โอกาส และการ

สนับสนุนรวมถึงการให้คำปรึกษา ขอขอบคุณบุคลากรสาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้ ที่ให้ความรู้และคำปรึกษาด้านการจัดทำเอกสารโครงการ รวมถึงการสนับสนุนตลอดจนการปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ จนทำให้งานวิจัยครั้งนี้สำเร็จลุล่วงด้วยดี ผู้วิจัยขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้

10. เอกสารอ้างอิงอิเล็กทรอนิกส์

- Bloomberg, D. (n.d.). **Leptonica: An open source C library for image processing and analysis.** <http://leptonica.org/>
- Disintegration. (n.d.). **Imaging: Image processing package for Go.**
<https://pkg.go.dev/github.com/disintegration/imaging>
- Google. (n.d.). **Tesseract OCR.**
<https://github.com/tesseract-ocr/tesseract>
- Lokhande, S. S., & Dawande, N. A. (2015). **A survey on document image binarization techniques.** In *2015 International Conference on Computing Communication Control and Automation* (pp. 742-746). IEEE.
<https://doi.org/10.1109/ICCUBEA.2015.148>
- Lu, M., & Yang, Y. (2024). **Image processing applications in smart contracts: Automated financial transaction verification.** *Technical Sciences*, 41(4), 22.
<http://dx.doi.org/10.18280/ts.410422>
- Lu, S., Su, B., & Tan, C. L. (2010). **Document image binarization using background estimation and stroke edges.** *International Journal on Document Analysis and Recognition*, 13(4), 303-314. <http://dx.doi.org/10.1007/s10032-010-0130-8>
- Mursari, L. R., & Wibowo, A. (2021). **The effectiveness of image preprocessing on digital handwritten scripts recognition with the implementation of OCR Tesseract.** *Communications in Computer and Information Science*, 10(3).
<https://doi.org/10.18495/comengapp.v10i3.386>
- Sauvola, J., & Pietikäinen, M. (2000). **Adaptive document image binarization.** *Pattern Recognition*, 33(2), 225-236. [https://doi.org/10.1016/S0031-3203\(99\)00055-2](https://doi.org/10.1016/S0031-3203(99)00055-2)