

การสกัดข้อมูลจากบัตรประชาชนด้วยวิธีการรู้จำอักขระด้วยแสง

EXTRACTING INFORMATION FROM ID CARDS USING OPTICAL CHARACTER RECOGNITION

ภาณุวัฒน์ จันทพร ทิพย์อักษร¹

ดร.ธนภัทร มั่งคะจิตร²

บทคัดย่อ

ในอดีต การจัดเก็บข้อมูลบัตรประชาชนมักใช้การสำเนาเอกสารแนบกับแบบฟอร์มเพื่อประกอบการขอใช้บริการจากหน่วยงานภาครัฐและเอกชน ซึ่งเป็นการเก็บข้อมูลในรูปแบบกระดาษที่ไม่เอื้อต่อการนำไปใช้งานต่อ ปัจจุบันแม้จะมีการใช้ Smart Card (Dip Chip) ซึ่งสามารถอ่านข้อมูลด้วยอุปกรณ์เฉพาะทาง (Reader) และชุดโปรแกรมเฉพาะแต่ยังคงมีข้อจำกัดในการใช้งาน ดังนั้นการประยุกต์ใช้เทคนิคการรู้จำอักขระด้วยแสง (OCR) จึงเป็นทางเลือกที่เหมาะสมเนื่องจากความสะดวกในใช้งาน และจากความก้าวหน้าของปัญญาประดิษฐ์ในงานด้านทางด้านคอมพิวเตอร์วิทัศน์ ส่งผลให้เทคนิค OCR มีประสิทธิภาพสูงในการสกัดข้อมูลข้อความจากรูปภาพ อย่างไรก็ตามพบว่าลายน้ำบนหน้าบัตรประชาชนส่งผลต่อความแม่นยำในการสกัดข้อมูลด้วย OCR ดังนั้นงานวิจัยนี้นำเสนอวิธีการแก้ไขปัญหานี้ด้วยการใช้เทคนิค CLAHE (Contrast Limited Adaptive Histogram Equalization) เพื่อปรับปรุงคุณภาพของภาพให้เหมาะสมก่อนนำไปใช้สกัดข้อมูลด้วย OCR จากผลการศึกษาพบว่าการใช้เทคนิค CLAHE สามารถเพิ่มความแม่นยำให้กับการใช้งาน OCR ได้จริง ดังนั้นจึงเหมาะสำหรับการนำไปใช้งานจริงซึ่งช่วยลดค่าใช้จ่ายให้กับองค์กรจากการใช้ API ของภาคเอกชนที่มีค่าใช้จ่ายค่อนข้างสูงในระยะยาว

คำสำคัญ : การรู้จำอักขระด้วยแสง, คอมพิวเตอร์วิทัศน์, ปัญญาประดิษฐ์, บัตรประจำตัวประชาชน, ลายน้ำ, วิธีการปรับฮิสโตแกรมในแบบที่ปรับตัวเองได้

¹ นักศึกษาหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมข้อมูลขนาดใหญ่

² ที่ปรึกษาสารนิพนธ์หลัก

ABSTRACT

In the past, the storage of citizen identification card primarily involved attaching photocopies to forms for service requests from public and private institutions. This paper-based method was inefficient for further data utilization. Although the current use of Smart Cards (Dip Chips) enables data extraction through specialized readers and software, significant limitations persist. As a more practical alternative, Optical Character Recognition (OCR) has gained traction due to its convenience and the advancements in computer vision and artificial intelligence, which have notably improved the accuracy of text extraction from images. However, watermarks on ID cards remains a key issue that reduces OCR accuracy. This study proposes a solution by leveraging the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique to enhance image quality prior to OCR processing. Experimental results indicate that applying CLAHE significantly improves OCR accuracy, making it a viable option for real-world applications while reducing long-term expenses typically incurred from using proprietary APIs.

Keywords: OCR, Computer vision, Artificial Intelligence, Citizen identification card, watermark, CLAHE

บทนำ

ในอดีตการจัดเก็บข้อมูลบัตรประชาชนจะใช้การสำเนาเอกสารแนบไปกับแบบฟอร์มเอกสารต่าง ๆ โดยใช้ยื่นประกอบการขอใช้บริการต่าง ๆ ของหน่วยงานภาครัฐและเอกชน ซึ่งทำได้เพียงการจัดเก็บสำเนากระดาษ ไม่เอื้อต่อการนำข้อมูลไปใช้งานด้านอื่น ๆ ต่อได้

ปัจจุบันมีการปรับมาใช้การอ่าน Smart Card ที่เรียกว่า Dip Chip แล้ว แต่ยังต้องอาศัยอุปกรณ์ตัวอ่าน (Reader) และชุดโปรแกรมร่วมด้วยเพื่อให้สามารถนำข้อมูลไปใช้งานได้ มักพบเจอในการลงทะเบียนเข้ารับบริการของภาครัฐ เอกชน ธนาคาร ห้าง ร้าน ฯลฯ ที่รองรับบริการดังกล่าว

ในการอ่านข้อมูลจากบัตรจะต้องมีเครื่องอ่านที่รองรับเสมอ จึงนำมาสู่การเปลี่ยนมาใช้งานสกัดข้อมูลจากหน้าบัตรที่มีความสามารถใกล้เคียงกับ Dip Chip ในปัจจุบัน โดยจะพบเห็นได้ในหลาย Application เช่น เป๋าตัง และ E-Wallet ต่าง ๆ ที่จะต้องทำการยืนยันตนก่อนสมัครใช้บริการทางธุรกรรมตามกฎหมายของธนาคารแห่งประเทศไทย ที่ว่าด้วยการป้องกันการฟอกเงิน

จากแนวทางข้างต้น ผู้วิจัยได้ศึกษาหาข้อมูลจากผู้ให้บริการภาครัฐ พบว่าการขอใช้บริการผ่าน API จะต้องยื่นโดยหน่วยงานของรัฐ และลงนาม MOU ร่วมกับกรมการปกครอง ส่วนภาคเอกชนจะมีค่าบริการในการใช้งานที่ค่อนข้างสูงเมื่อเทียบกับ Dip Chip

แนวคิดของงานวิจัยนี้ คือการพัฒนาและปรับปรุงกระบวนการสกัดข้อมูลจากหน้าบัตรประชาชน ด้วยการใช้เทคโนโลยีการรู้จำอักขระด้วยแสง (OCR) ซึ่งสามารถประยุกต์ใช้งานได้โดยไม่ต้องอาศัยอุปกรณ์พิเศษเพิ่มเติม โดยมีเป้าหมายเพื่อสร้างความสะดวกและยืดหยุ่นในการใช้งาน ทั้งยังพยายามแก้ไขปัญหาลายน้ำบนหน้าบัตรประชาชนที่ส่งผลต่อการสกัดข้อมูลด้วยการนำเทคนิค CLAHE มาช่วยในการปรับปรุงคุณภาพของภาพอีกด้วย

วัตถุประสงค์ของงานวิจัย

เพื่อพัฒนาระบบสำหรับอ่านข้อความจากภาพบัตรประชาชน ที่มีประสิทธิภาพใกล้เคียงกับ Solutions เชิงพาณิชย์ ที่ให้ความแม่นยำสูง แม้มีย่านับบนตัวอักษรที่อ่านได้ยาก ซึ่งอยู่ในรูปแบบ API Services ที่ง่ายต่อการประยุกต์ใช้งาน

ขอบเขตงานวิจัย

ภาพถ่ายต้องไม่มีแสงสะท้อน ตรงบริเวณข้อความที่จะอ่าน

ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อนำเสนอการปรับปรุงกระบวนการสกัดข้อมูลจากหน้าบัตรประจำตัวประชาชน เพื่อนำไปประยุกต์ใช้งานกับระบบต่าง ๆ ได้อย่างมีประสิทธิภาพมากขึ้น ในอดีตการจัดเก็บข้อมูลบัตรประชาชนมักใช้การสำเนาเอกสารแนบกับแบบฟอร์มเพื่อประกอบการขอใช้บริการจากหน่วยงานภาครัฐและเอกชน ซึ่งเป็นการเก็บข้อมูลในรูปแบบกระดาษที่ไม่เอื้อต่อการนำไปใช้งานต่อ ปัจจุบันแม้จะมีการใช้ Smart Card (Dip Chip) ซึ่งสามารถอ่านข้อมูลด้วยอุปกรณ์เฉพาะทาง (Reader) และชุดโปรแกรมเฉพาะ แต่ยังคงมีข้อจำกัดในการใช้งาน

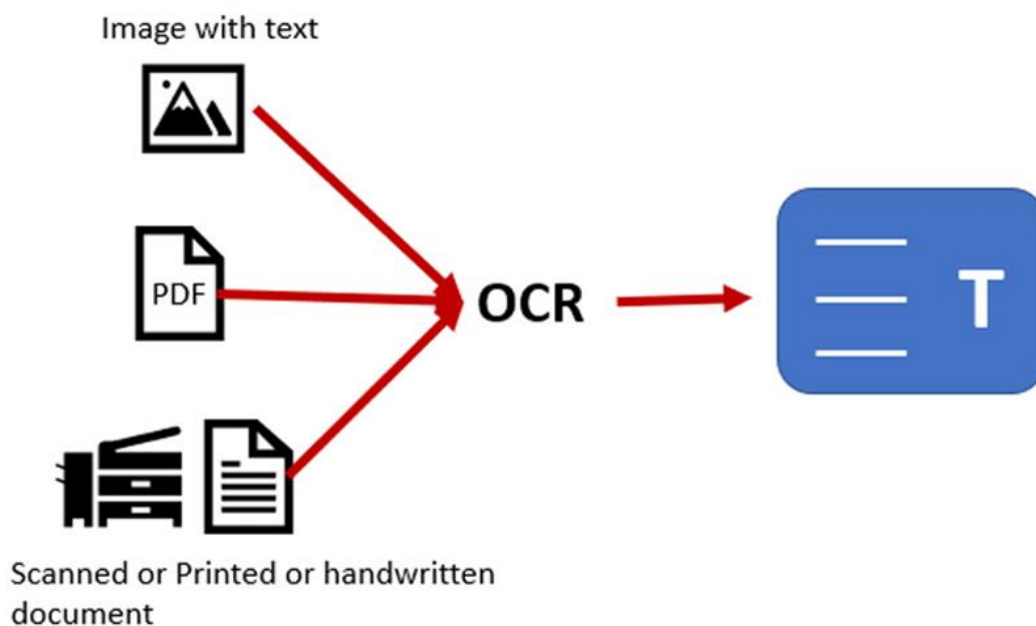
งานวิจัยนี้จึงมุ่งเน้นไปที่การใช้เทคโนโลยีการรู้จำอักขระด้วยแสง (OCR) เพื่อสกัดข้อมูลจากหน้าบัตรประจำตัวประชาชน ซึ่งสามารถทำได้โดยไม่ต้องใช้อุปกรณ์พิเศษเพิ่มเติม เพียงแค่นำรูปถ่ายหน้าบัตรส่งเข้าสู่ระบบ ระบบจะประมวลผลและสกัดข้อความออกมาอย่างสะดวกและรวดเร็ว นอกจากนี้ยังสามารถประยุกต์ใช้งานร่วมกับ Web Application ได้อย่างง่ายดาย

ทฤษฎีที่เกี่ยวข้อง

1. Contrast Limited Adaptive Histogram Equalization (CLAHE): CLAHE เป็นเทคนิคการปรับปรุงคุณภาพของภาพที่ใช้การแบ่งภาพออกเป็นช่วงย่อยหลายส่วนแล้วทำการกระจายค่าสีและแสงอย่างจำกัด เพื่อ

ลดปัญหาการเกิด noise หรือ artifact ที่อาจเกิดขึ้นจากการทำ Histogram Equalization แบบปกติ CLAHE ช่วยเพิ่มความชัดเจนของรายละเอียดในภาพโดยเฉพาะในส่วนที่มีความแตกต่างของแสงน้อย ซึ่งเหมาะสำหรับการปรับปรุงคุณภาพของภาพที่มีปัญหาแสงสะท้อนจากหน้าบัตรประชาชน ทำให้ OCR สามารถทำงานได้อย่างแม่นยำยิ่งขึ้น

2. Optical Character Recognition (OCR): หรือ การรู้จำอักขระด้วยแสง เป็นเทคโนโลยีที่ช่วยให้คอมพิวเตอร์สามารถแปลงรูปภาพของข้อความให้เป็นข้อความที่สามารถแก้ไขได้และจัดเก็บได้อย่างง่ายดาย ดังในภาพที่ 1 เทคโนโลยีนี้ได้รับความนิยมในหลายด้าน เช่น การสแกนเอกสาร การแปลงเอกสารทางกฎหมาย และการสกัดข้อมูลจากเอกสารที่ไม่เป็นดิจิทัล แม้ว่า OCR จะช่วยในการแปลงข้อความจากรูปภาพให้สามารถนำไปใช้ได้ แต่การอ่านข้อความอาจไม่สมบูรณ์ 100% เนื่องจากปัจจัยหลายประการ เช่น แบบจำลองที่ใช้คุณภาพของไฟล์รูปภาพ ความคมชัด และแสงสะท้อน ซึ่งอาจส่งผลต่อความถูกต้องในการอ่านข้อความโดยตรง



ภาพที่ 1 ขั้นตอนการทำงานของเทคโนโลยี การรู้จำอักขระด้วยแสง (OCR)

งานวิจัยที่เกี่ยวข้อง

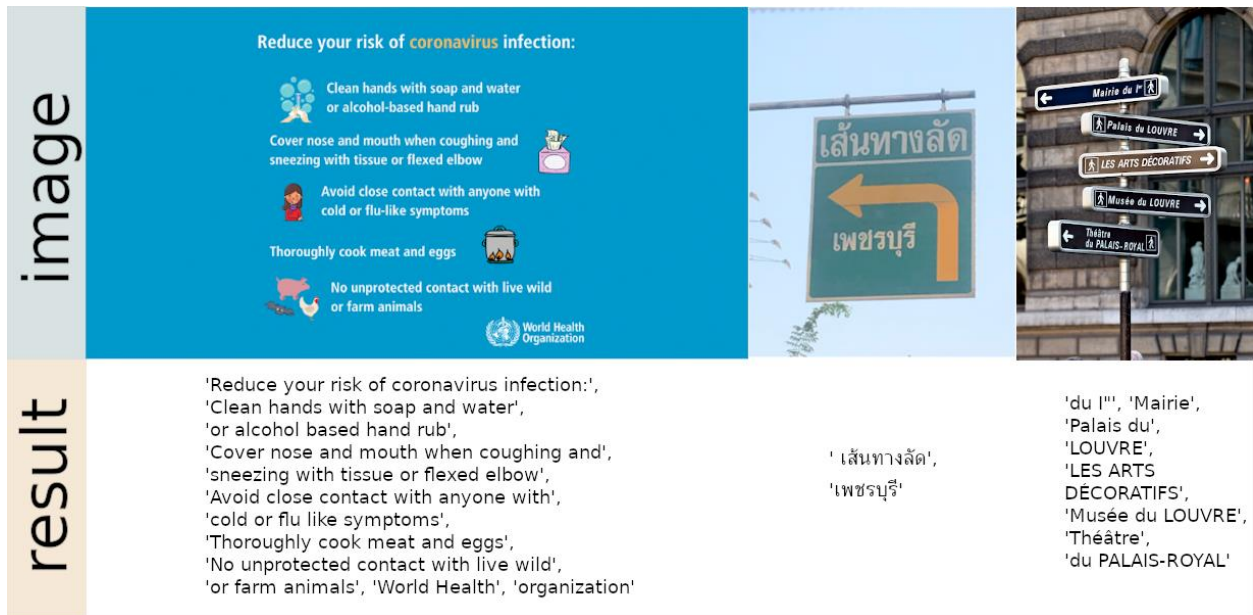
1. Smith et al. (2022) ได้กล่าวถึงงานวิจัยที่มุ่งเน้นการปรับปรุงความแม่นยำของ OCR โดยใช้เทคนิคการปรับปรุงภาพขั้นสูง โดยเฉพาะ CLAHE การศึกษานี้เน้นให้เห็นว่าการประมวลผลภาพก่อนใช้ OCR อย่างเหมาะสม สามารถเพิ่มความสามารถในการรู้จำอักขระได้อย่างมีนัยสำคัญภายใต้สภาพแสงที่ท้าทาย เช่น ความคมชัดต่ำหรือแสงที่แตกต่างกัน การวิจัยพบว่าการใช้ CLAHE ก่อนการประมวลผลด้วย OCR ช่วยลดข้อผิดพลาด ทำให้เป็นขั้นตอนสำคัญในการสกัดข้อมูลที่มีคุณภาพสูง

2. Kumar et al. (2020) ได้กล่าวไว้ในบทความสำรวจการใช้งาน CLAHE ในระบบ OCR เพื่อปรับปรุงคุณภาพของภาพก่อนการรู้จำอักขระ การศึกษานี้แสดงให้เห็นว่า CLAHE สามารถเพิ่มความสามารถในการอ่านข้อความในภาพที่มีแสงน้อยและคอนทราสต์ต่ำ ซึ่งส่งผลให้ประสิทธิภาพของ OCR ดีขึ้น ผลลัพธ์พบว่าการใช้ CLAHE ช่วยลดข้อผิดพลาดของ OCR อย่างมีนัยสำคัญ ทำให้เป็นเทคนิคการประมวลผลภาพที่มีประสิทธิภาพ

3. Musa et al. (2019) ได้กล่าวไว้ในบทความที่เน้นการใช้งาน OCR แบบเรียลไทม์ในสภาพแวดล้อมที่เปลี่ยนแปลง โดยมุ่งเน้นการสกัดข้อความจากภาพโดยไม่ต้องใช้อุปกรณ์เฉพาะทาง การวิจัยแสดงให้เห็นว่าการผสาน OCR กับเทคนิคการปรับปรุงภาพ เช่น CLAHE ทำให้ระบบ OCR สามารถทำงานได้อย่างมีประสิทธิภาพในแอปพลิเคชันแบบเรียลไทม์ แม้ในสภาพแวดล้อมที่มีแสงน้อยหรือคุณภาพภาพที่แตกต่างกัน

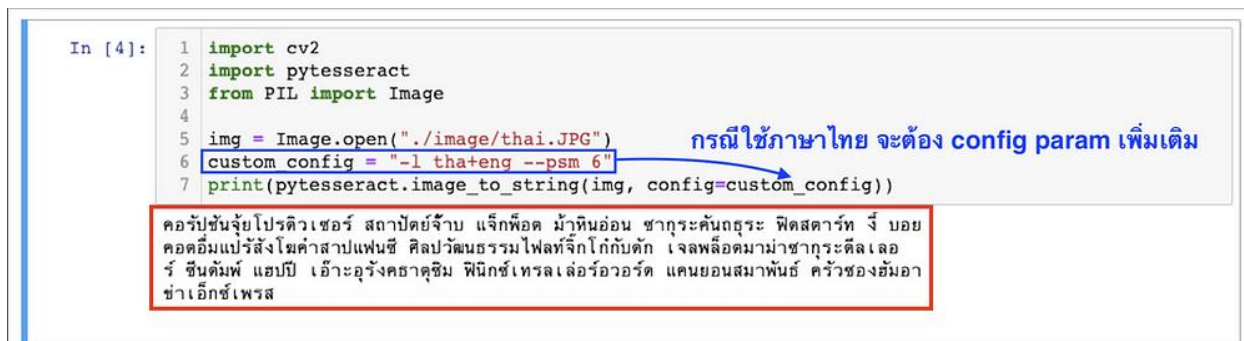
4. CLAHE Histogram Equalization - OpenCV เนื้อหาจากบทความใน GeeksforGeeks (2023) อธิบายถึง CLAHE ซึ่งเป็นเทคนิคที่ใช้ในการปรับปรุงความคมชัดของภาพ โดยเฉพาะในพื้นที่ที่มีความคอนทราสต์ต่ำเป็นเทคนิคการปรับปรุงความคมชัดของภาพที่ใช้ในการแก้ปัญหาคอนทราสต์ต่ำในบางพื้นที่ของภาพ ซึ่งช่วยลดความรบกวนในพื้นที่ที่มีความคอนทราสต์สูง โดยการจำกัดความยาวของฮิสโตแกรมเพื่อป้องกันการขยายที่เกินจริงในพื้นที่ที่มีความสว่างหรือมืด การใช้ CLAHE ช่วยให้ภาพมีรายละเอียดที่ชัดเจนทั้งในพื้นที่สว่างและมืดโดยไม่ทำให้ภาพบิดเบือน.

5. Easy OCR คือ ซอฟต์แวร์ที่แปลงภาพเป็นตัวอักษรที่พร้อมใช้งาน มาพร้อมภาษาที่รองรับมากกว่า 80 ภาษา และสคริปต์การเขียนขอลนิคม รองรับภาษาต่าง ๆ อาทิเช่น ละติน จีน อารบิก เทวนาครี ซิริลลิก ไทย ฯลฯ ตัวอย่างการทำงานของ Easy OCR แสดงในภาพที่ 2



ภาพที่ 2 ตัวอย่างการสกัดข้อความจากภาพ โดยใช้แบบจำลอง Easy OCR

6. Tesseract OCR คือ API ที่ถูกพัฒนาขึ้นมาโดย Google (Since 2006) ซึ่งใช้สำหรับทำ OCR (Optical Character Recognition) เป็น open-source OCR Engine ที่ใช้แปลง Image ให้เป็น Text เพื่อนำไปประยุกต์ใช้กับงานต่าง ๆ เช่น ตรวจสอบเช็คป้ายทะเบียนรถเข้าออกอาคาร, เช็คข้อมูลใน passport, digitizing invoices (menus, ID cards) เป็นต้น ดังตัวอย่างในภาพที่ 3



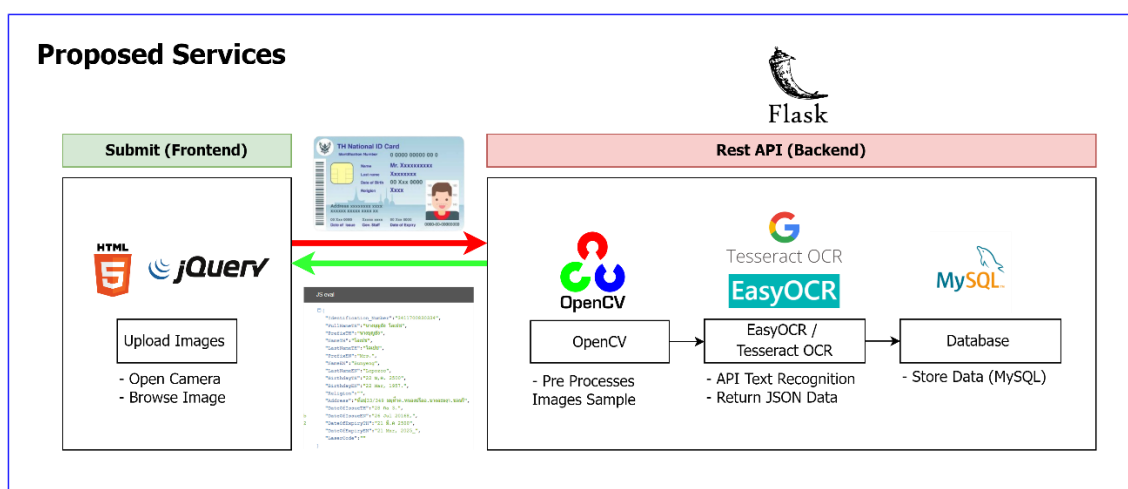
ภาพที่ 3 ตัวอย่างการสกัดข้อความจากภาพ โดยใช้แบบจำลอง Tesseract OCR

ระเบียบวิธีวิจัย

การศึกษาวิจัยครั้งนี้ เป็นการนำเสนอขั้นตอนการประมวลผลข้อความและดำเนินการสร้างระบบสำหรับตรวจสอบข้อความบนหน้าบัตรประชาชนให้ได้ตามวัตถุประสงค์ของงานจึงได้กำหนดกระบวนการการดำเนินการเพื่อศึกษาข้อมูลมีขั้นตอนดังนี้

1. แนวคิดของระบบ (System Flow)

แนวทางการสร้างระบบอัตโนมัติมีขั้นตอนดังภาพที่ 4

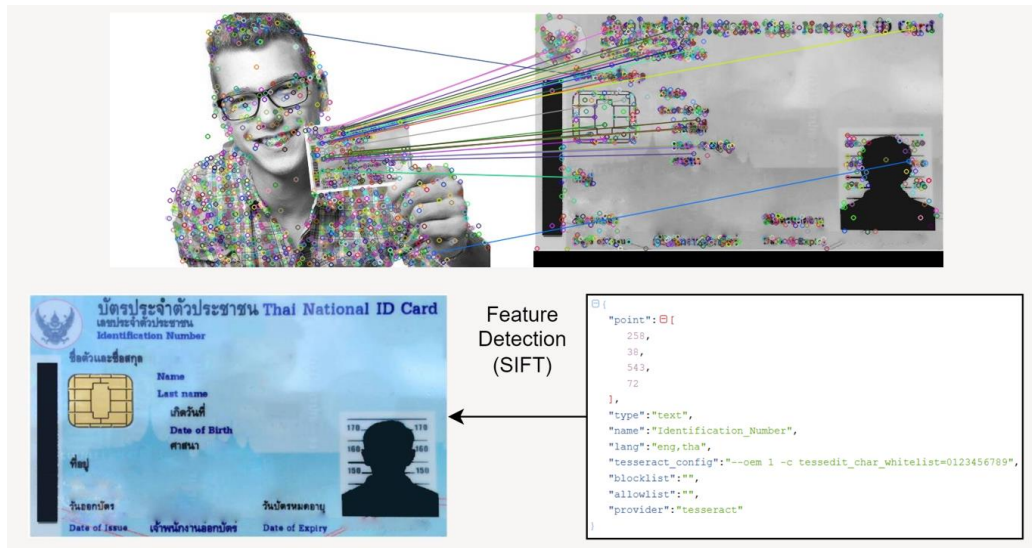


ภาพที่ 4 แผนผังการสร้างระบบสำหรับอ่านข้อมูลจากหน้าบัตรประชาชน

2. ขั้นตอนการสร้างระบบประมวลผลข้อความ

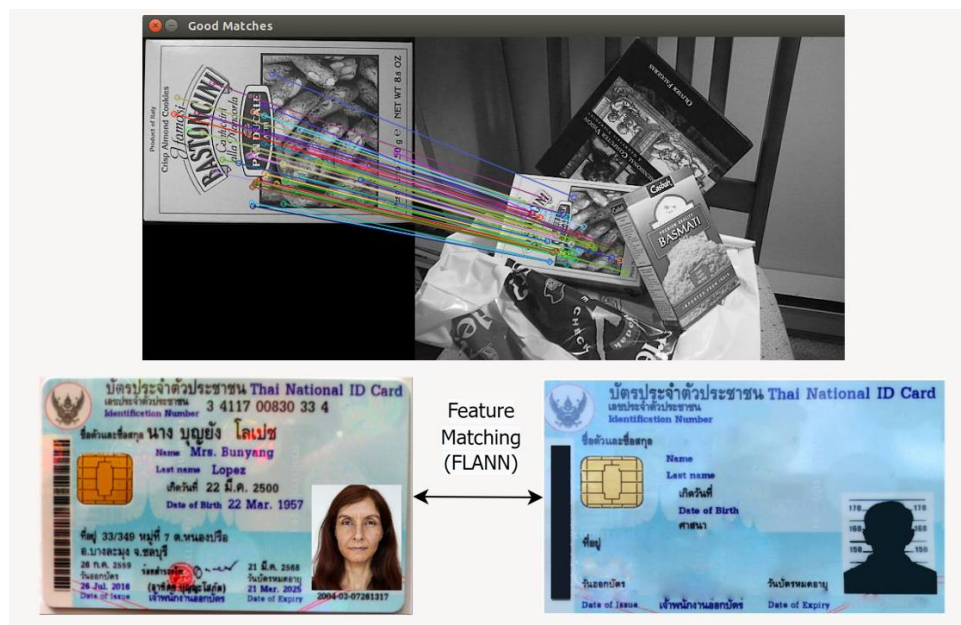
ขั้นตอนการประมวลผลข้อความและดำเนินการสร้างระบบสำหรับตรวจสอบข้อความบนหน้าบัตรประชาชนให้ได้ตามวัตถุประสงค์ของงานจึงได้กำหนดกระบวนการการดำเนินการเพื่อศึกษาข้อมูลมีขั้นตอนดังนี้

1. โหลดรูปหน้าบัตรประชาชนตัวอย่างที่ตัดข้อมูลออกแล้ว เพื่อใช้เป็นต้นแบบของบัตร โดยใช้ Feature Detection (SIFT) พร้อมกำหนดค่าของตำแหน่งแต่ละจุดที่จะใช้เปรียบเทียบ ดังในภาพที่ 5



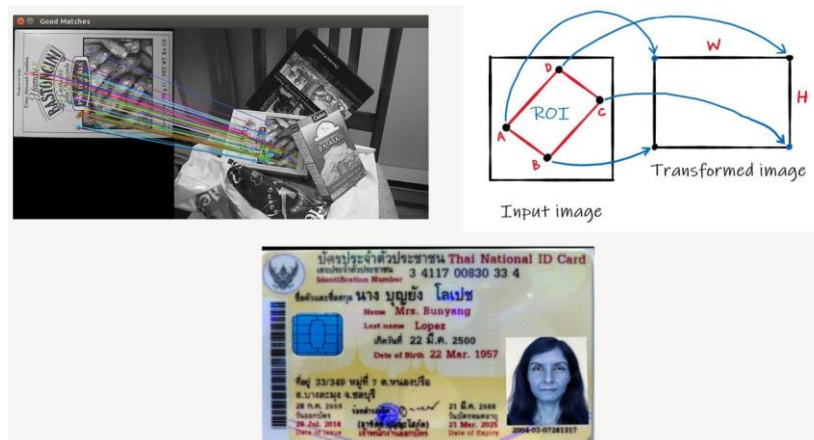
ภาพที่ 5 โหลดภาพต้นฉบับเข้าสู่แบบจำลองพร้อมกำหนดตำแหน่งแต่ละจุดที่จะใช้เปรียบเทียบ

2. โหลดภาพที่จะนำมาสกัดข้อความเข้าสู่แบบจำลองเพื่อเปรียบเทียบความเหมือนกับภาพต้นแบบในแต่ละจุดที่เกี่ยวข้องกัน โดยใช้แบบจำลอง Feature Matching (FLANN) ดังในภาพที่ 6



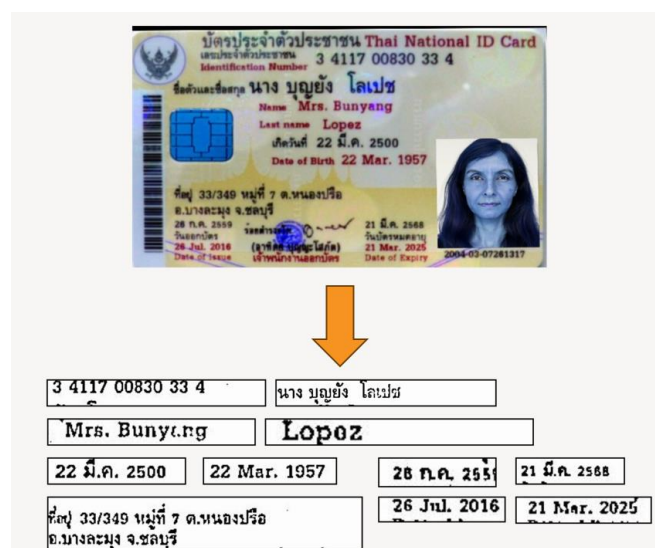
ภาพที่ 6 โหลดภาพเข้าสู่แบบจำลองเพื่อเปรียบเทียบความเหมือนกับภาพต้นแบบในแต่ละจุดที่เกี่ยวข้องกัน

3. ให้แบบจำลองทำการค้นหาจุดที่มีความคล้ายคลึงกันกับภาพต้นแบบ จากนั้นกำหนดจุดตัดแบ่งแต่ละส่วน เพื่อนำไปประมวลผลในขั้นตอนถัดไปโดยใช้ `cv2.findHomography` และ `cv2.warpPerspective` ดังในภาพที่ 7



ภาพที่ 7 ทำการค้นหาจุดที่มีความคล้ายคลึงกันกับภาพต้นแบบ และกำหนดจุดตัดแบ่งแต่ละส่วน เพื่อนำไปประมวลผลต่อไป

4. ทำการตัดแบ่งแต่ละส่วนที่ได้จากขั้นตอนก่อนหน้า โดยใช้ `cv2.adaptiveThreshold` ดังในภาพที่ 8



ภาพที่ 8 ตัดแบ่งแต่ละส่วนที่ได้จากขั้นตอนก่อนหน้า

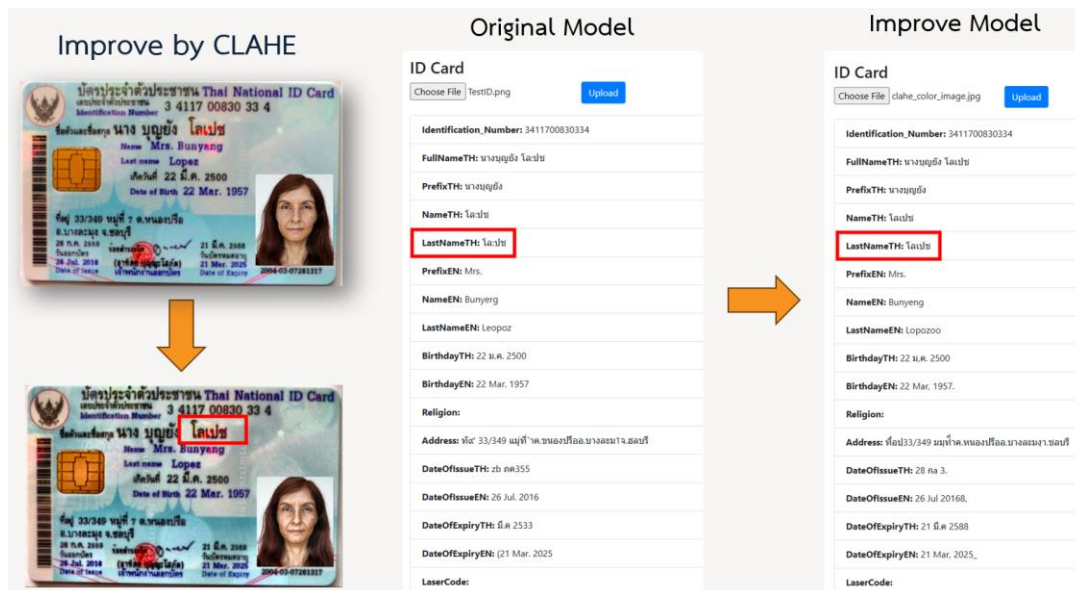
5. กำหนดให้แต่ละส่วนจะถูกสกัดข้อความโดยใช้แบบจำลองใด เพื่อให้มีความแม่นยำ จากนั้นจะให้คืนค่าข้อความที่สกัดได้กลับมาในรูปแบบอาร์เรย์ เพื่อให้สามารถนำไปกำหนดรูปแบบสำหรับส่งออกข้อมูลได้ดังในภาพที่ 9

Identification_Number : Tesseract
FullNameTH : EasyOCR
NameEN : Tesseract
LastNameEN : Tesseract
BirthdayTH : EasyOCR
Religion : EasyOCR
Address : EasyOCR
DateOfIssueTH : EasyOCR
DateOfIssueEN : Tesseract
DateOfExpiryTH : EasyOCR
DateOfExpiryEN : Tesseract
LaserCode : EasyOCR

```
Card(Identification_Number='3411700830334', FullNameTH='นางบุญ ียง โล:ปช', PrefixTH='นางบุญ ียง', NameTH='โล:ปช', LastNameTH='โล:ปช', PrefixEN='Mrs.', NameEN='Bunyang', LastNameEN='Leopoz', BirthdayTH='22 ม.ค. 2500', BirthdayEN='22 Mar. 1957', Religion='', Address='ทึ ๙ 33/349 หมู่ ที่ ๑ ค.หนองปรือ อ.บางละมิง.ชลบุรี ', DateOfIssueTH='zb กค355', DateOfIssueEN='26 Jul. 2016', DateOfExpiryTH='มี .ค 2533', DateOfExpiryEN='(21 Mar. 2025', LaserCode='')
```

ภาพที่ 9 กำหนดให้แต่ละส่วนจะถูกสกัดข้อความโดยใช้แบบจำลองใด เพื่อให้มีความแม่นยำ

ในขั้นตอนที่ 5 นี้ มักจะพบปัญหาแสงสะท้อนจากลายน้ำ จึงได้นำแนวคิดของ CLAHE (Contrast Limited Adaptive Histogram Equalization) ใน OpenCV ที่เป็นเทคนิคการปรับปรุงคอนทราสต์ของภาพที่ถูกพัฒนาให้สามารถแก้ปัญหาการปรับปรุงคอนทราสต์ที่มากเกินไปในพื้นที่ขนาดเล็ก เทคนิคนี้เป็นการขยายความสามารถของ Adaptive Histogram Equalization (AHE) โดยการจำกัดความคอนทราสต์ในพื้นที่ของภาพดังในภาพที่ 10



ภาพที่ 10 ภาพบัตรประชาชนที่ผ่านการปรับปรุงโดยเทคนิค CLAHE จะทำให้แบบจำลองสามารถสกัดข้อความได้แม่นยำมากขึ้น

ผลการศึกษา

จากผลการทดลองการนำภาพบัตรประชาชนที่ผ่านการปรับปรุงโดยเทคนิค CLAHE และเข้าสู่กระบวนการสกัดข้อความโดย OCR สามารถอธิบายผลลัพธ์ได้ดังนี้:

1. ความคมชัดของภาพบัตรประชาชน ภาพบัตรประชาชนที่มีความคมชัดและไม่มีแสงสะท้อนจะช่วยให้การสกัดข้อความทำได้แม่นยำยิ่งขึ้น เทคโนโลยี CLAHE ช่วยให้อาชีพมีความคมชัดและคอนทราสต์ที่ดีขึ้น ซึ่งส่งผลให้แบบจำลอง OCR สามารถสกัดข้อความได้ถูกต้องมากขึ้น
2. ผลการใช้ CLAHE การปรับปรุงภาพบัตรประชาชนด้วย CLAHE ทำให้แบบจำลอง OCR สามารถสกัดข้อความได้แม่นยำยิ่งขึ้น โดย CLAHE ช่วยเพิ่มความคมชัดและลดปัญหาแสงสะท้อนในภาพ ซึ่งช่วยให้ OCR ทำงานได้ดีขึ้น ดังภาพที่ 10
3. การเลือกแบบจำลอง OCR การเลือกแบบจำลอง OCR เช่น Easy OCR หรือ Tesseract OCR จำเป็นต้องทดสอบกับรูปภาพบัตรประชาชนหลายรูปแบบเพื่อประเมินความแม่นยำ อย่างไรก็ตาม ข้อมูลที่มีสำหรับการทดสอบมีข้อจำกัด เนื่องจากบัตรประชาชนเป็นข้อมูลส่วนบุคคลสำคัญ จึงทำให้ไม่สามารถประเมินผลการสกัดข้อความได้อย่างชัดเจนและถูกต้อง

จากผลการทดลองข้างต้นนี้สรุปได้ว่า การปรับปรุงภาพบัตรประชาชนด้วยเทคนิค CLAHE ช่วยเพิ่มความคมชัดและคอนทราสต์ ทำให้การสกัดข้อความโดย OCR มีความแม่นยำมากขึ้น ขณะเดียวกัน การเลือกแบบจำลอง OCR ต้องการการทดสอบที่หลากหลายซึ่งถูกจำกัดด้วยข้อกังวลด้านข้อมูลส่วนบุคคล

สรุปและอภิปรายผลการวิจัย

1. **ความแม่นยำของระบบสกัดข้อความ** ระบบที่ใช้สำหรับสกัดข้อความจากหน้าบัตรประชาชนได้แสดงถึงความแม่นยำในระดับที่ยอมรับได้ เทคโนโลยี CLAHE ช่วยในการปรับปรุงความคมชัดของภาพ ซึ่งทำให้การสกัดข้อความโดย OCR มีประสิทธิภาพดีขึ้น อย่างไรก็ตาม การปรับปรุงแบบจำลองด้วยเทคนิคอื่น ๆ อาจช่วยเพิ่มความแม่นยำของระบบในอนาคต

2. **ประสิทธิภาพของ OCR** การใช้เทคนิค OCR ช่วยให้สามารถสกัดข้อมูลจากรูปภาพหน้าบัตรประชาชนได้อย่างมีประสิทธิภาพ แต่การทำงานของ OCR ยังต้องการข้อกำหนดเบื้องต้น เช่น คุณภาพของภาพและความคมชัด ซึ่งอาจจำเป็นต้องปรับปรุงเพื่อให้ระบบทำงานได้อย่างเต็มประสิทธิภาพ

3. **การตรวจสอบความถูกต้องของข้อมูล** ก่อนการนำข้อมูลที่ได้จากการสกัดข้อความไปใช้งาน จำเป็นต้องมีกระบวนการตรวจสอบความถูกต้องของข้อมูลอีกครั้งเพื่อให้ข้อมูลที่ได้นั้นมีความถูกต้องและเชื่อถือได้มากที่สุด

4. **ความสะดวกในการใช้งาน** การใช้เทคโนโลยี OCR และ CLAHE ช่วยลดระยะเวลาและเพิ่มความสะดวกในการป้อนข้อมูลจากหน้าบัตรประชาชน โดยไม่จำเป็นต้องพึ่งพาเพียงระบบอ่านบัตร Dip Chip เท่านั้น เทคโนโลยี CLAHE ยังช่วยในการปรับปรุงความคมชัดของภาพ ทำให้การสกัดข้อความมีความแม่นยำและสะดวกยิ่งขึ้น.

อภิปรายผลจากการใช้เทคนิค CLAHE ร่วมกับ OCR ทำให้การสกัดข้อความจากหน้าบัตรประชาชนมีประสิทธิภาพและแม่นยำมากขึ้น CLAHE ช่วยปรับปรุงคุณภาพของภาพในขณะที่ OCR ใช้ข้อมูลที่ได้รับการปรับปรุงเพื่อให้การอ่านข้อความมีความแม่นยำสูงสุด การตรวจสอบความถูกต้องของข้อมูลหลังการสกัดเป็นขั้นตอนสำคัญในการรับรองความถูกต้องของข้อมูลที่ได้จากระบบดังกล่าว

ข้อเสนอแนะ

1. ควรพัฒนาแบบจำลองให้สามารถทำงานที่หลากหลายรูปแบบมากยิ่งขึ้น ทำให้สามารถนำแบบจำลองประยุกต์ใช้กับข้อมูลจากหน้าบัตรรูปแบบอื่นได้ในอนาคต

2. ควรมีการกำหนดมาตรฐานของรูปภาพบัตรประชาชนที่เหมาะสม เพื่อให้ระบบสามารถสกัดข้อความได้อย่างแม่นยำ และถูกต้องมากขึ้น

3. กระบวนการ OCR จะใช้ระยะเวลาในการประมวลผลค่อนข้างนาน ในขั้นตอนติดตั้ง Library ของ Easy OCR และ Tesseract OCR ควรเตรียมเครื่องคอมพิวเตอร์มีประสิทธิภาพเพียงพอ และมี GPU เป็นอย่างน้อย พร้อมทั้งติดตั้ง CUDA ให้มีประมวลผลผ่าน GPU ซึ่งตัว Library มักจะแนะนำว่าหากมี GPU จะช่วยให้สามารถประมวลผลได้รวดเร็วขึ้นกว่าเดิม

บรรณานุกรม

Adrian Rosebrock. (2024). *Open CV Histogram Equalization and Adaptive Histogram Equalization*

(CLAHE)”. [Online]. Retrieved from <https://pyimagesearch.com/2021/02/01/opencv-histogram-equalization-and-adaptive-histogram-equalization-clahe>

GeeksforGeek. (2024). *CLAHE Histogram Equalization – OpenCV*”. [Online]. Retrieved from

<https://www.geeksforgeeks.org/clahe-histogram-equalization-opencv>

Jaied AI. GitHub. (2024). *Easy OCR*” [Online]. Retrieved from <https://github.com/JaiedAI/EasyOCR>

Kumar et al. (2024). *OCR Image Enhancement & Implementation by using CLAHE algorithm*” [Online].

Retrieved from https://www.researchgate.net/publication/342872425_OCR_Image_Enhancement_Implementation_by_using_CLAHE_algorithm

Musa et al. (2024). *A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition*” [Online]. Retrieved from

<https://ieeexplore.ieee.org/document/8780492>

Nattapol Krobklang. (2024). *Library for extract information from Thai personal identity card*” [Online].

Retrieved from <https://pypi.org/project/thai-personal-card-extract>

Samuel Hoffstaetter. (2024). *Python-tesseract is a python wrapper for Google's Tesseract-OCR*” [Online].

Retrieved from <https://pypi.org/project/pytesseract>

Smith et al. (2024). *Image Enhancement and Implementation of CLAHE Algorithm and Bilinear Interpolation*” [Online]. Retrieved from <https://doi.org/10.1080/01969722.2022.2147128>