

การรับรู้ป้ายสาธารณะและการแปลภาษาไทย-จีน บนระบบแอนดรอยด์

Public Signage Recognition and Thai-Chinese Translation Based on Android

Zhou Hongying¹

ณรงค์เดช กীরติพรานนท์²

บทคัดย่อ

การรู้จำตัวอักษรภาษาไทยในป้ายสาธารณะเป็นเรื่องยากและท้าทายอันมีสาเหตุมาจากปัญหาทางเทคนิคได้แก่ ความซับซ้อนความสว่างความเอียง ยังรวมถึงปัญหาความหลากหลายของตัวอักษรภาษาไทยที่ใช้อยู่ในป้ายสาธารณะ จากปัญหาดังที่กล่าวมาแล้วงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาแอปพลิเคชัน Android สำหรับช่วยนักท่องเที่ยวและคนต่างชาติใช้แปลตัวอักษรไทยในป้ายสาธารณะ โดยขั้นตอนของแอปพลิเคชันเริ่มจากใช้ Leptonica C++ไลบรารีสำหรับการประมวลผลภาพ หลังจากนั้นใช้ Tesseract OCR ไลบรารีเพื่อรู้จำตัวอักษรภาษาไทย สาเหตุที่ใช้ Tesseract OCR engine เพราะเป็นไลบรารี open source ที่ได้รับความนิยมและเคยใช้รู้จำตัวอักษรไทยในงานวิจัยอื่นมาแล้ว ต่อจากนั้นใช้บริการแปลภาษาของ Baidu ทำการแปลภาษาไทยเป็นภาษาอังกฤษและจีน นอกจากนี้ในระหว่างการพัฒนาแอปพลิเคชันนี้ผู้วิจัยยังใช้เทคโนโลยีการพัฒนา SQLite JNI และ HTTP เพื่อทำการทดสอบและปรับปรุงการรู้จำตัวอักษรชนิดใหม่ๆ หลายชนิด ผลการทดสอบแสดงให้เห็นผลลัพธ์เป็นที่น่าพอใจ จากผลการทดลองป้ายที่อยู่บนถนนจำนวน 14 ป้าย (207 ตัวอักษร) ซึ่งมีความถูกต้องเฉลี่ย 97% ซึ่งในอนาคตอาจเพิ่มชนิดตัวอักษรแบบอื่นในการเรียนรู้ต่อไป

คำสำคัญ การรู้จำรูปภาพ, ภูมิทัศน์ธรรมชาติ, ป้ายสาธารณะ, การแปลภาษาด้วยคอมพิวเตอร์, แอปพลิเคชันแอนดรอยด์

ABSTRACT

Recognition of Thai characters in natural scene is challenge. The reason why Thai characters are difficult to recognize, is not only a technological problem, such as Thai characters are complicated, illumination, inclination and other factors could reduce the effect of recognition, but also a social problem, such as there are many kinds of fonts in Thailand public signage. The

¹ นักศึกษาปริญญาโท หลักสูตรวิศวกรรมคอมพิวเตอร์ วิทยาลัยนวัตกรรมการด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

² อาจารย์ที่ปรึกษา

purpose of this paper is to provide a development of Image recognition Android application for tourist or other foreigners in Thailand, which could use to recognize and translate Thai characters in the public signage board in natural scene. In this paper, it is selected Android as a mobile operating system, and Android Studio as development environment. The Leptonica C++ library used for image processing. The Tesseract OCR Engine, which one of most popular open source library of OCR, used to recognize Thai characters. After recognition, the BAIDU translation service used to translate Thai language to English. What is more, the development technology of SQLite, JNI, HTTP is used during developing application. In order to test and improve recognize effect of the application, this paper tried to train a new kind of font of Thai-language public signage. As a result of experiment, the application could recognize Thai-language public signage satisfied. Consequently, this paper success to develop a new Image recognition application for Thai-language public signage based on Android platform. In addition, it is improve effect of recognition by training new kind of font. Future work might continue to increase types of training font.

Keywords Image Recognition, Convolutional Neural Networks, Public Signage, Machine Translation, Android Application

1. บทนำ

ประเทศไทยมีศักยภาพและสามารถสร้างโอกาสทางการท่องเที่ยว รวมถึงทำการค้าระหว่างประเทศ ซึ่งคนจีนนิยมมาท่องเที่ยวและทำการค้าขายกับประเทศไทย แต่ภาษาไทยก็เป็นข้อจำกัดของชาวต่างชาติที่เข้ามาทำธุรกิจหรือมาท่องเที่ยวในประเทศไทยทั้งการพูด อ่าน และเขียน โดยเฉพาะชาวจีน และถึงแม้ว่าจะมีเทคโนโลยีที่สามารถช่วยในการแปลภาษาเช่น Google Translate ที่คนไทยนิยมใช้ แต่สำหรับที่ประเทศจีนไม่สามารถใช้ google ได้ และคนจีนส่วนใหญ่ก็ไม่นิยมใช้ google translate คนจีนชอบใช้ Baidu translate แต่ Baidu translate ยังไม่มีการรู้จำตัวอักษรภาษาไทย[8] อย่างไรก็ตามมีเทคโนโลยีที่จะสามารถใช้ในการพัฒนาแอปพลิเคชันในการแปลภาษาที่รวดเร็วและแม่นยำได้แก่ซอฟต์แวร์การรู้จำตัวอักษร (Optical Character Recognition: OCR)[3]ซึ่งเป็นเทคโนโลยีสำหรับการแปลงภาพเอกสาร ป้าย[6] หรือรูปภาพให้กลายเป็นข้อความที่สามารถนำไปแปลเป็นภาษาจีนต่อไปได้ Abdurrahman (2012) ได้เสนอวิธีการ OCR Tesseract คือการจับภาพบนสมาร์ตโฟน ซึ่งประยุกต์ใช้กับกล้องโทรศัพท์มือถือ ภาษาที่ใช้ทดสอบในครั้งนี้มีสามภาษาคือ ชุนดา อินโดนีเซีย และอังกฤษ ผลลัพธ์ที่ได้มีความถูกต้องร้อยละ 70[2]

ดังนั้น ผู้วิจัยจึงได้ดำเนินการพัฒนาโมบายแอปพลิเคชันสำหรับแปลภาษาไทยเป็นภาษาจีน บนระบบปฏิบัติการแอนดรอยด์ เพื่อให้นักท่องเที่ยวและนักธุรกิจชาวจีนได้ใช้แอปพลิเคชันดังกล่าวในการแปลภาษาไทยเป็นภาษาจีน ได้อย่างสะดวกสบาย เนื่องจากเป็นอุปกรณ์พกพาขนาดเล็ก ซึ่งคำที่นำมาพัฒนาในแอปพลิเคชันดังกล่าวเป็นคำที่ใช้ในชีวิตประจำวัน เช่น ชื่อสถานที่ ชื่อจังหวัด หอพัก โรงแรม ร้านอาหาร เป็นต้น ซึ่งสาเหตุที่เลือกพัฒนาแอปพลิเคชันดังกล่าวขึ้นเพราะในปัจจุบันเทคโนโลยีมีความเจริญก้าวหน้าคนส่วนใหญ่ใช้สมาร์ทโฟนในชีวิตประจำวันการพัฒนาแอปพลิเคชันเพื่อใช้ผ่านโมบายจะทำให้สะดวกและรวดเร็วต่อผู้ใช้ในปัจจุบัน

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 Tesseract OCR

Tesseract OCR เป็น Engine ที่ใช้สำหรับการรู้จำอักขระทางภาพ ซึ่งมีกระบวนการหลัก 2 ขั้นตอน [5] ได้แก่

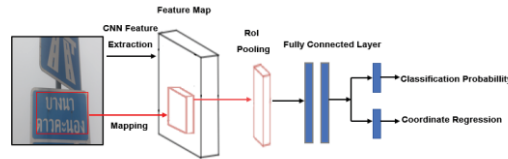
ขั้นตอนที่ 1 การวิเคราะห์โครงร่างรูปภาพ

การวิเคราะห์โครงร่างรูปภาพเป็นขั้นตอนแรกที่ระบบ OCR จัดจํารูปภาพ ความสำคัญก็คือการเตรียมความพร้อมสำหรับการรับรู้ภาพโดยวิธีการเฉพาะดำเนินการวิเคราะห์รูปแบบ แยกระหว่างพื้นที่ข้อความและพื้นที่ที่ไม่ใช่ข้อความเพื่อกรองส่วนที่ไม่ใช่ข้อความของภาพ วิธีการวิเคราะห์ที่ใช้ Tesseract เป็นวิธีการวิเคราะห์เค้าโครงหน้าของ Hybrid tab-based detection[2] วิธีนี้สามารถวิเคราะห์เลย์เอาต์ที่ซับซ้อนมากขึ้นโดยไม่เพิ่มจำนวนการคำนวณ

ขั้นตอนที่ 2 การแยกและการรู้จำตัวอักษร

วัตถุประสงค์ของการออกแบบ Tesseract [3] คือการจดจำตัวอักษรอย่างถูกต้องเนื่องจาก Tesseract รู้จักอักขระแต่ละตัวที่ละตัว สิ่งนี้ต้องใช้อักขระในภาพเพื่อแยกทีละตัว ขั้นตอนของการแบ่งส่วน Tesseract คือการแบ่งข้อความให้คร่าว ๆ เพราะการแบ่งคร่าว ๆ ส่วนใหญ่แยกอักขระได้ จากนั้นใช้การแบ่งส่วนแบบละเอียดและแก้ไขข้อผิดพลาด หลักการทำงานคือขั้นแรกให้บล็อกอักขระในแนวตั้งเพื่อให้มีช่องว่าง แบ่งอย่างหยาบ ๆ ตามขนาดของช่องว่าง แม้ว่าการแบ่งส่วนแบบคร่าว ๆ ส่วนใหญ่สามารถแยกอักขระได้ อย่างไรก็ตามหลังจากการแบ่งส่วนแบบคร่าว ๆ ก็จะมีอักขระที่ติดกันหรืออักขระที่แบ่งกลุ่มไม่ถูกต้อง ในขั้นตอนนี้การแบ่งส่วนแบบละเอียดถูกใช้การแบ่งส่วนแบบละเอียดจะเปรียบเทียบอักขระกับไลบรารีอักขระเพื่อตัดสินอักขระเดียวซึ่งทำให้การจดจำข้อความมีความแม่นยำมากขึ้น

2.2 Fast R-CNN



รูปที่ 1 Fast R-CNN Architecture

2.2.1 กระบวนการการรู้จำป้ายถนนของ Fast R-CNN

กระบวนการการรู้จำป้ายถนนของ Fast R-CNN มีกระบวนการหลัก 2 ขั้นตอน [4] ได้แก่

ขั้นตอนที่ 1: ขั้นตอนการฝึกอบรม ขั้นตอนการฝึกอบรมใช้ font: DB-Bork เพราะไลบรารีโอเพนซอร์ซของ Tesseract มีพารามิเตอร์เริ่มต้นของ Convolutional Neural Networks สามารถฝึกอบรมได้ ในที่สุดจะได้การรู้จำของ Model

ขั้นตอนที่ 2: ขั้นตอนการรู้จำ ได้การรู้จำของ Model และก็ป้อนตัวอักษรของป้ายถนนเพื่อให้ได้ผลการรู้จำ

2.2.2 กระบวนการทำงานของ Fast R-CNN มีประกอบด้วย 5 ขั้นตอนดังต่อไปนี้:

ขั้นตอนที่ 1 ให้นำเข้าทั้งภาพ

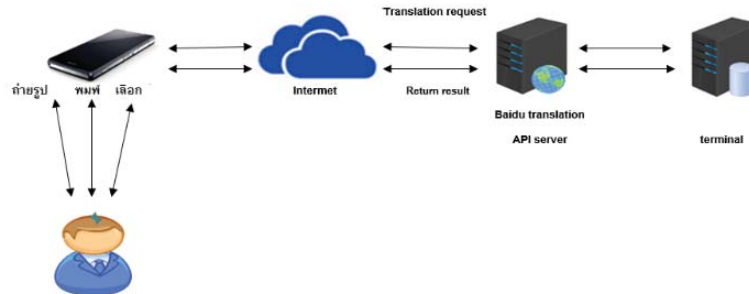
ขั้นตอนที่ 2 Extract feature map ให้นำเข้าทั้งภาพไปผ่าน CNN กับ ROI Pooling Layer จะได้ผล feature map

ขั้นตอนที่ 3 Suggest region proposal ใช้ algorithm Selective Search เพื่อ mapping จะได้ region proposal ประมาณ 2,000 ภาพ[7]

ขั้นตอนที่ 4 region proposal Region Normalization region proposal ไปผ่าน ROI Pooling Layer เพื่อให้ได้ feature vector ที่มีขนาดคงที่

ขั้นตอนที่ 5 algorithm softmax กับ Bounding-box feature vector ไปผ่าน Fully connected layer จะได้เอาต์พุตสองรายการก็คืออัลกอริทึม softmax กับอัลกอริทึม Bounding-box ผลลัพธ์ที่อัลกอริทึม softmax เพื่อตรวจสอบว่าเป็นคลาสที่ระบุหรือไม่ (เช่นเป็นคลาสตัวอักษร ก หรือ ข ถ้าเป็นตัวอักษร ก จะไปหาคลาสตัวอักษร ก) ผลลัพธ์ที่อัลกอริทึม Bounding-box เพื่อปรับตำแหน่งของ region proposal และลบพื้นที่ที่ทับซ้อนกัน (เช่นตำแหน่งของ region proposal ซ้อนกัน จะลบส่วนที่ซ้อนกันทำให้ region proposal ที่มีขนาดใหญ่และก็ปรับขนาดใหม่)

2.3 Baidu translation



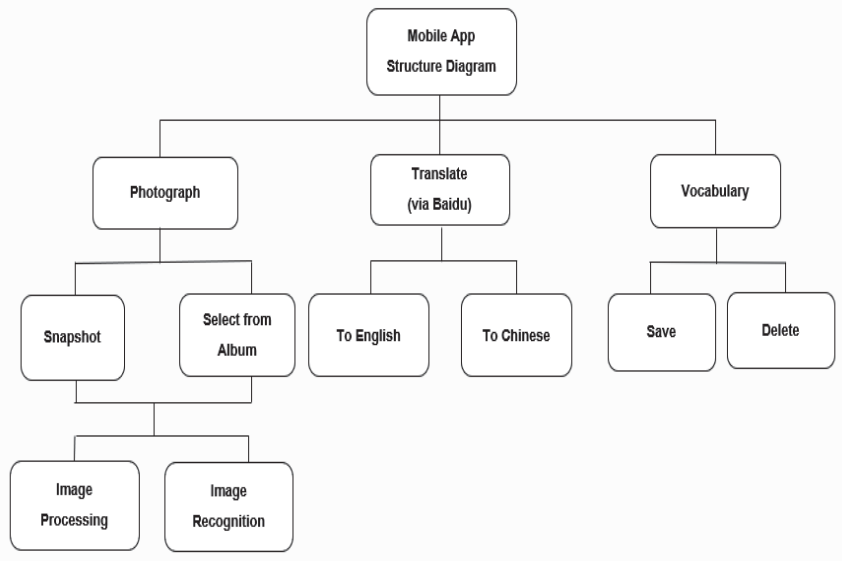
รูปที่ 2 กระบวนการการแปลภาษา

ผู้ใช้เปิดแอปพลิเคชัน Thai OCR ขึ้นมา จะปรากฏเมนูการแปลภาษาของ Thai OCR โดยจะให้เลือกรายการถ่ายภาพ หรือการพิมพ์ เมื่อถ่ายภาพหรือพิมพ์เสร็จแล้ว ต่ออินเทอร์เน็ตไปขอ Baidu translation API server โดย Baidu translation API server จะไป Translation terminal เพื่อทำการแปล พอแปลเสร็จจะส่งผลให้ Baidu translation API server ให้แสดงผลออกมา

3.วิธีดำเนินการวิจัย

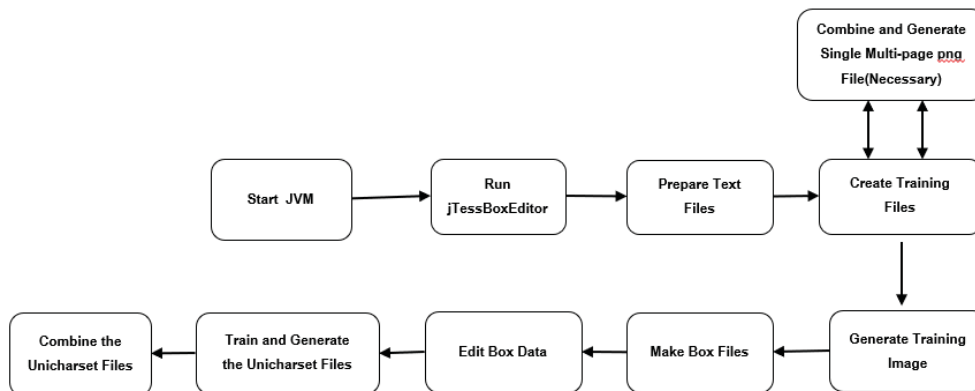
3.1 การออกแบบระบบแอปมือถือ

ระบบแอปมือถือประกอบด้วย 3 ส่วน ประกอบด้วย ส่วนที่ 1 photograph ส่วนที่ 2 Translate and internet และส่วนที่ 3 Vocabulary โดยในส่วนที่ 1 photograph หมายถึงภาพที่ถ่ายจากกล้องโทรศัพท์มือถือหรือเลือกภาพที่มีข้อความเป้าหมายจากอัลบั้มโทรศัพท์มือถือ การประมวลผลภาพหมายถึงการประมวลผลภาพเพื่อเพิ่มอัตราการรู้จำภาพตัวอย่างเช่นการกรองค่ามัธยฐาน การแปลงภาพระดับเทาให้เป็นภาพขาว-ดำ เป็นต้น การรู้จำตัวอักษร หมายถึง การใช้เทคนิค OCR เพื่อระบุเป้าหมายภาพและแยกข้อความออกมา ส่วนที่ 2 การแปลและระบบเครือข่าย (Translate and internet) หมายถึง การใช้เทคนิค OCR ได้ตัวอักษรและส่งข้อความไปถึง Baidu เพื่อแปลภาษาไทยเป็นภาษาจีนหรือภาษาอังกฤษ และส่วนที่ 3 Vocabulary หมายถึงการคงไว้และลบข้อความ จึงมีภาพรวมของ Mobile app structure ดังรูปที่ 2



รูปที่ 3 Mobile App Structure

3.2 ฟังก์ชันการรู้จำภาพ



รูปที่ 4 Training Workflow for Tesseract with jTesseditorbox

Tesseract OCR เป็น Engine ที่ใช้สำหรับการรู้จำอักขระทางภาพซึ่งขั้นตอนการฝึกอบรมคือ

ขั้นตอนที่ 1 Start JAVA , run iTessBoxEditor.

ขั้นตอนที่ 2 สร้างภาพการฝึกอบรม (create training image) การแยกตัวอักษรจากภาพป้าย สร้าง font:DB-Bork.

ขั้นตอนที่ 3 สร้างไฟล์ Box Merge ภาพป้ายและสร้างไฟล์ Box หลังจากการแยกตัวอักษรจากภาพป้าย ขั้นตอนต่อไปคือ Merge ภาพป้ายและสร้างไฟล์ Box ซึ่งมันจะมีข้อผิดพลาดในการรู้จำ ในกระบวนการแก้ไขเราต้องแยกตัวอักษรให้ถูกต้องและป้อนตัวอักษรที่ถูกต้องในคอลัมน์ Char

ขั้นตอนที่ 4 สร้างไฟล์ font properties ไฟล์ font properties ทำหน้าที่ provide font style information รูปแบบของ font_properties คือ [DB-Bork]><italic><bold><fixed><serif><fraktur>โดยที่ <[DB-Bork]>เป็นเลือกที่ Font ต่อไปนี้ทำพารามิเตอร์เป็นรูปแบบของการฝึกอบรมตัวอักษร(Font style) ใส่ normal 0 และ 1 เพื่อระบุว่าไม่มีแตริวิต์เหล่านี้อยู่หรือไม่ 1 หมายถึงมีแตริวิต์ 0 หมายถึงไม่มีแตริวิต์

ขั้นตอนที่ 5 สร้างไฟล์ Clustering การฝึกอบรมไลบรารีต้องจัดกลุ่มหน้าการฝึกอบรมเพื่อสร้างต้นแบบ (Need to cluster training pages to create the prototypes)

1: การสร้างไฟล์ shapetable (ไฟล์ shapetable คือ shape table) ใช้คำสั่ง: shapeclustering -F font_properties -U unicharset [tha2].[DB-Bork].tr

2: การสร้างต้นแบบรูปร่างไฟล์ inttemp (ไฟล์ inttempคือ the shape prototypes) กับจำนวนของฟังก์ชันที่ต้องการต่อตัวอักษรไฟล์ pffmtable (ไฟล์ pffmtable คือ the number of expected feautres for each character) ใช้คำสั่ง: mftraining -F font_properties -U unicharset -O [tha2]. unicharset [tha2].[DB-Bork].tr

3: สร้างไฟล์ข้อมูล normproto (the character nomalization sensitivity prototypes) ใช้คำสั่ง: cntraining [DB-Bork].tr

ขั้นตอนที่ 6 สร้าง traineddata ใหม่ ขั้นตอนก่อนหน้าได้รับ 4 ไฟล์ก็คือ inttemp normproto pffmtable shapetable แก้ inttemp normproto pffmtable shapetable เป็น tha2.inttemp tha2.normproto tha2.pffmtable tha2.shapetable และใช้คำสั่ง combine_tessdata tha2. เพื่อสร้าง traineddata ใหม่

3.3 Baidu Translation วิธีใช้ baidu Translation API



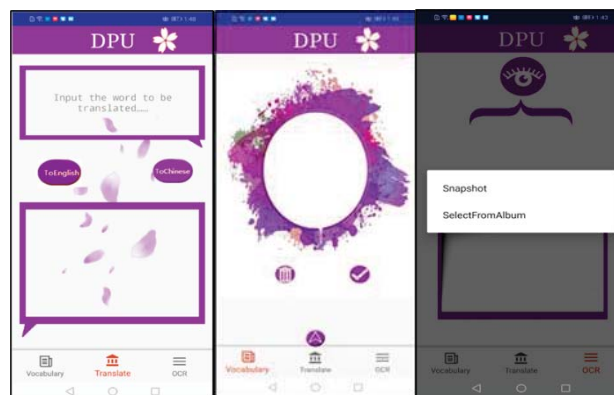
รูปที่ 5 Translation process

ฟังก์ชันการแปลจะดำเนินการผ่านทางอินเทอร์เน็ตที่ Baidu แปล ในขั้นแรกใช้วิธีการคีย์ API เพื่อรับส่วนต่อประสานข้อมูลแปลภาษาโดยต้องใช้ server จึงต้องเชื่อมต่อ network ซึ่งการแปล Baidu สามารถใช้วิธี get ของ HTTP เพื่อทำการร้องขอการแปล Baidu ข้อมูลที่ส่งกลับมาคือ แพ็คเกจ json และสุดท้ายก็จะส่งผลการแปลกลับมา

4 ผลและอภิปรายผล

ตอนที่ 1 ผลการทดสอบของหน้าจอเพื่อให้เป็นทำงานได้จริง

ThaiOCR App มี 3 ฟังก์ชันหลักคือการรู้จำรูปภาพภาษาไทย การแปล save; การออกแบบที่ด้านล่างของอินเทอร์เน็ตหลักของ app มี 3 ปุ่ม Tab ซึ่งก็คือ vocabulary Translate และ OCR ผู้ใช้สามารถคลิกและข้ามไปที่ UI Layer ของฟังก์ชันที่เกี่ยวข้อง ดังรูปที่ 6



(ก) vocabulary (ข) Translate (ค) OCR

ภาพที่ 6 หน้าหลักของโปรแกรม

ตอนที่ 2 ผลการทดสอบของการใช้งานโปรแกรม เพื่อให้เป็นทำงานได้จริง

(1) อินเทอร์เน็ต Vocabulary เป็นฟังก์ชันบันทึกและจัดเก็บประวัติการแปล สามารถบันทึกหรือลบเนื้อหาการแปลได้ ดังรูปที่ 7



ภาพที่ 7 แสดงการ Vocabulary

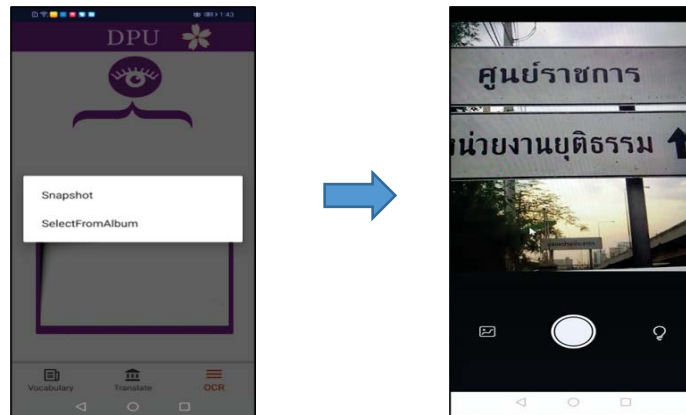
(2) อินเทอร์เน็ตเฟซ Translate เป็นฟังก์ชันการแปลโดยใช้คลาวด์ Baidu เพื่อแปลภาษา ปัจจุบันมีความสามารถแปลภาษาไทยเป็นภาษาอังกฤษกับภาษาจีน ดังรูปที่ 8



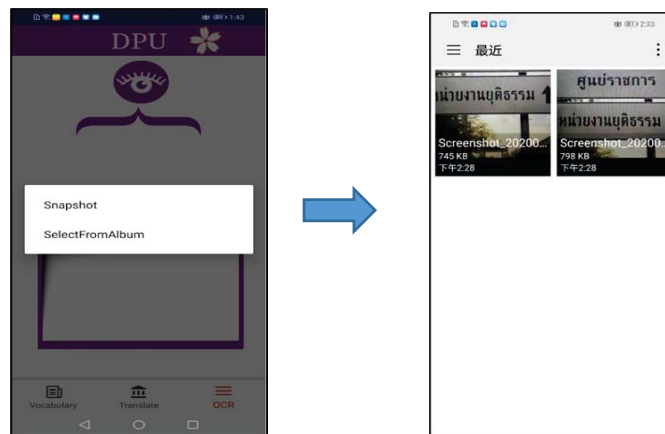
ภาพที่ 8 แสดงภาษาไทยแปลเป็นภาษาอังกฤษกับจีน

(3) อินเทอร์เน็ตเฟซ OCR เป็นฟังก์ชันการรู้จำภาษาไทย เมื่อคลิกที่ปุ่ม OCR ตัวเลือกแหล่งรูปภาพสองตัวจะปรากฏขึ้น : Snapshot และ SelectFromAlbum หากเลือกฟังก์ชัน Snapshot จะใช้ฟังก์ชันกล้องของโทรศัพท์เพื่อถ่ายป้ายถนนที่รู้จำ หากเลือกฟังก์ชัน SelectFromAlbum จะสามารถเลือกภาพที่มีอยู่ในโทรศัพท์ของอัลบั้ม หลังจากกล้องสแกนแล้วปรับขนาดของกรอบการรับรู้ด้วยตนเอง, เพื่อปรับปรุงความเร็วและความแม่นยำของการรู้จำภาพป้าย , หลังจากเลือกช่วงการรับรู้ภาพ

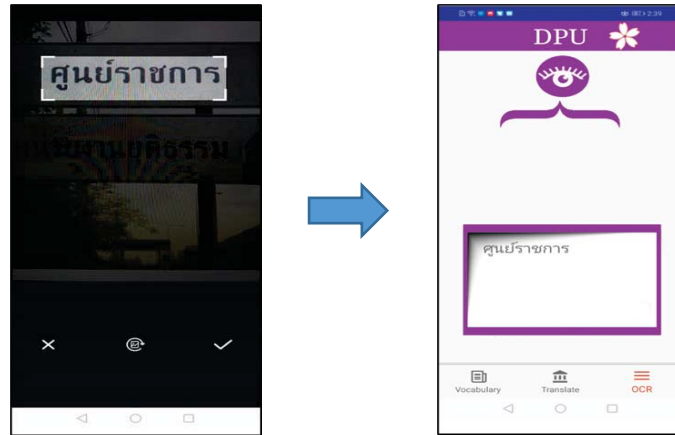
แล้วแอปจะทำการรู้จำภาษาไทย หลังจากการรู้จำสำเร็จแล้วผลการรู้จำจะปรากฏในรูปแบบของ Text มิฉะนั้นจะแสดงเป็น null ดังรูปที่ 9 กับรูปที่ 11



รูปที่ 9 แสดงการถ่ายรูปจาก Snapshot



ภาพที่ 10 แสดงการถ่ายรูปจาก SelectFromAlbum



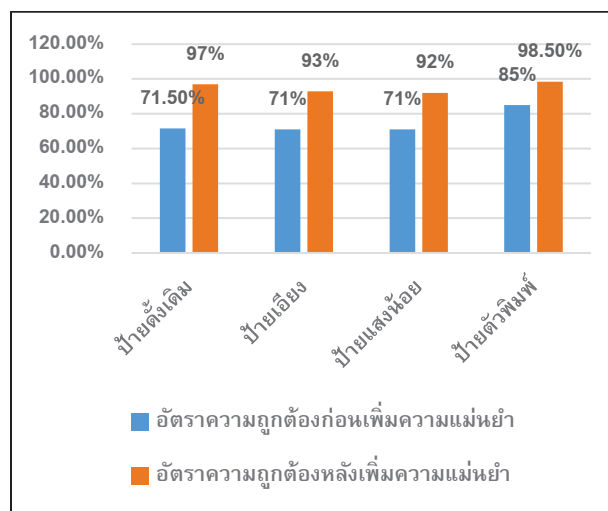
ภาพที่ 11 แสดงการตัดภาพ

ตอนที่ 3 ป้ายที่ใช้ในการทดสอบ

ป้ายที่ใช้ในการทดสอบเป็นป้ายที่ติดตั้งไว้ตามถนน จำนวน 14 ป้าย (207 ตัวอักษร) เป็น font:DB-Bork ต่อไปเป็นตัวอย่างของการทดสอบของป้าย



รูปที่ 12 Public Signage



ตารางที่ 1 Performance Improvements After Training



(ก)ป้ายดั้งเดิม (ข)ป้ายเอียง (ค)ป้ายแสงน้อย (ง)ตัวพิมพ์

ในการทดสอบนี้ นอกจากจะให้ทดลองกับป้ายแบบที่ถ่ายมาตั้งแต่ต้นแล้ว ยังได้ทำการปรับให้ภาพถ่ายเปลี่ยนไป คือมีการทำให้ป้ายเอียงเพิ่มขึ้น 10 องศา รวมทั้งป้ายแสงน้อยที่ลดปริมาณแสงลง นอกจากนี้ยังมีการพิมพ์คำตามป้ายเรียกว่าป้ายแบบตัวพิมพ์เพิ่มเติม

การทดสอบก่อนการเพิ่มความแม่นยำของข้อความที่พิมพ์อยู่ที่ 85% หลังจาก que เพิ่มความแม่นยำแล้วเพิ่มขึ้นเป็น 98.5% ส่วนป้ายดั้งเดิมก่อนการเพิ่ม

ความแม่นยำอยู่ที่ 71.5% หลังจาก que เพิ่มความแม่นยำแล้ว 97% สำหรับป้ายแสงน้อยพบว่าก่อนการเพิ่มความแม่นยำอยู่ที่ 71% หลังการเพิ่มความแม่นยำแล้ว 92% นอกจากนี้เมื่อนำโปรแกรมไปทดสอบกับป้ายที่มีลักษณะเอียงพบว่าก่อนเพิ่มความแม่นยำอยู่ที่ 71% และหลังเพิ่มความแม่นยำอยู่ที่ 93% ซึ่งถือว่ามีประสิทธิภาพสูง เพราะงว่า เมื่อใช้เทคนิคการแปลงค่าสี RGB เป็น Gray Scale , การกรองค่านัยฐาน,การแปลงภาพระดับเทาให้เป็นภาพขาว-ดำที่นำเสนอมาใช้ร่วมกันกับOCRสามารถเพิ่มความถูกต้องของการรู้จำตัวอักษรด้วยแสงสูงขึ้น เมื่อตัวอักษรยังมีความชัดเจน ตรง แสงปกติ ส่งผลให้ผลลัพธ์ที่ได้จากระบบรู้จำตัวอักษรมีประสิทธิภาพมากยิ่งขึ้น

5 อภิปรายผล

งานวิจัยนี้ทำการรู้จำตัวอักษร กระบวนการทำงานประกอบด้วยสามกระบวนการหลัก คือ 1. กระบวนการการออกแบบระบบแอปมือถือ 2. กระบวนการรู้จำภาพ 3. Baidu Translation โดยทำการทดสอบประสิทธิภาพของภาพป้าย จากการทดสอบกับภาพป้ายจริงจำนวน 14 ภาพ พบว่าป้ายดั้งเดิมมีอัตราความถูกต้องเท่ากับ 97% ป้ายเอียงมีอัตราความถูกต้องเท่ากับ 93% ป้ายแสงน้อยมีอัตราความถูกต้องเท่ากับ 92% ข้อความพิมพ์มีอัตราความถูกต้องเท่ากับ 98.5%

สรุปได้ว่าการรู้จำตัวอักษรภาษาไทยที่นำเสนอขึ้นให้ผลของการทดสอบมีความถูกต้องแม่นยำ สามารถนำไปประยุกต์ใช้งานได้จริง อย่างไรก็ตามวิธีที่นำเสนอยังมีข้อจำกัดในเรื่องของภาพป้าย เช่น(1) ฝึกการรู้จำภาพและ font จำนวนน้อยสิ่งที่จะต้องพัฒนาต่อคือ การเพิ่มขอบเขตของจำนวนภาพป้ายและ การเพิ่มจำนวน font ที่นำมาสร้าง Library (2) ถ้าภาพป้ายมี 2 ภาษาขึ้นไปและป้ายเอียง ป้ายแสงน้อยๆ ระบบสามารถรู้จำภาษาไทยได้เท่านั้นจึงทำให้การอ่านตัวอักษรมีความ

ผิดพลาดเกิดขึ้น สิ่งที่ต้องพัฒนาต่อคือใช้วิธี Multithreading กับใช้ฟังก์ชันการประมวลผลได้รับการปรับปรุงเพื่อเพิ่มความแม่นยำ

บรรณานุกรม

- [1] Kaothanthong N, Theeramunkong T , Chun J . Improving Thai Optical Character Recognition Using Circular-Scan Histogram[C] 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2018.
- [2] Sahil Badla. Improving the efficiency of Tesseract OCR Engine. Master's Projects. San Jose State University, 2014.
- [3] Lei Yin. Character Recognition System Based on Android System [D]. Beijing University of Posts and Telecommunications,2016.
- [4] Girshick, Ross. Fast R-CNN[J]. Computer Science, 2015.
- [5] Taolong Hong. Image Recognition and Voice Playing System Based on Android Platform [D], Nanjing University of Posts and Telecommunications.
- [6]Mengmeng Yang. Method for Traffic Sign Detection Based on Faster R-CNN[D]. Hubei University of technology 2019.
- [7]Selective Search for Object Recognition[J] . J. R. R. Uijlings,K. E. A. Sande,T. Gevers,A. W. M. Smeulders. International Journal of Computer Vision . 2013 (2)
- [8]He Z . Baidu Translate: Research and Products[C] . Proceedings of the Fourth Workshop on Hybrid Approaches to Translation (HyTra). 2015.