การพัฒนาโมบายแอพพลิเคชั่นแปลภาษาไทย-จีน จากป้ายสาธารณะ บนระบบปฏิบัติการแอนดรอยด์

Zhou Hongying

DPU

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม

วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์

มหาวิทยาลัยธุรกิจบัณฑิตย์

พ.ศ. 2563

Thai-Chinese Translation from Public Signage Based on Android Mobile Application

Zhou Hongying

A Thematic Paper Submitted in Partial Fulfillment of the Requirements

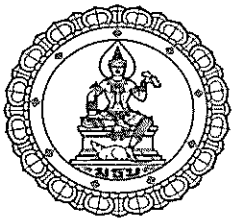
for the Degree of Master of Engineering

Department of Computer and Telecommunication Engineering

College of Innovative Technology And Engineering

Dhurakij Pundit University

2020



ใบรับรองสารนิพนธ์

วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์

มหาวิทยาลัยธุรกิจบัณฑิตย์

ปริญญา วิศวกรรมศาสตรมหาบัณฑิต

หัวข้อสารนิพนธ์ การพัฒนาโมบายแอพพลิเคชั่นแปลภาษาไทย-จีน จากป้ายสาธารณะบนระบบปฏิบัติการแอนดรอยด์

เสนอโดย นาง Zhou Hongying

สาขาวิชา วิศวกรรมคอมพิวเตอร์และโทรคมนาคม

อาจารย์ที่ปรึกษาสารนิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.ณรงค์เดช กีรติพรานนท์

ได้พิจารณาเห็นชอบโดยคณะกรรมการสอบสารนิพนธ์แล้ว

.....ประธานกรรมการ

(รองศาสตราจารย์ ดร.ลัญฉกร วุฒิสิทธิกุลกิจ)

.....กรรมการและอาจารย์ที่ปรึกษาสารนิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร.ณรงค์เดช กีรติพรานนท์)

.....กรรมการ

(อาจารย์ ดร.ชัยพร เขมะภาตะพันธ์)

วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์รับรองแล้ว

.....คณบดีวิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์

(ผู้ช่วยศาสตราจารย์ ดร.ณรงค์เดช กีรติพรานนท์)

วันที่ ...19... เดือน ...สิงหาคม... พ.ศ. ...2563...

หัวข้อสารนิพนธ์	การพัฒนาโมบายแอพพลิเคชั่นแปลภาษาไทย-จีน จากป้ายสาธารณะบนระบบปฏิบัติการแอนดรอยด์
ชื่อผู้เขียน	Zhou Hongying
อาจารย์ที่ปรึกษา	ผศ.ดร.ณรงค์เดช กีรติพรานนท์
สาขาวิชา	วิศวกรรมคอมพิวเตอร์และโทรคมนาคม
ปีการศึกษา	2562

บทคัดย่อ

การจดจำตัวอักษรภาษาไทยเป็นเรื่องยากและท้าทาย ซึ่งมีมาจากหลายสาเหตุ ถึงแม้จะมีเทคโนโลยีช่วยแต่ก็ยังเป็นปัญหาเพราะมีความซับซ้อน เช่นเรื่องแสงสว่าง ความเอียงของตัวอักษร และปัจจัยอื่นๆ อีก รวมทั้งตัวอักษรนั้นก็มีฟอนต์หลายชนิดที่ใช้ในประเทศไทย การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาแอพพลิเคชั่น Android สำหรับการรู้จำตัวอักษรภาพสำหรับนักท่องเที่ยวต่างชาติ ที่มาเที่ยวในประเทศไทยซึ่งสามารถใช้ในการจดจำและแปลตัวอักษรภาษาไทยในป้ายต่างๆ ในประเทศไทย ในงานวิจัยนี้เลือกพัฒนาบนระบบปฏิบัติการ Android

Tesseract OCR เป็น Engine หนึ่งในไลบรารี่โอเพ่นซอร์สที่ได้รับความนิยมมากที่สุดในการรู้จำตัวอักษร และเคยใช้เพื่อจดจำตัวอักษรไทย หลังจากรับรู้แล้วบริการแปล BAIDU จะใช้ในการแปลภาษาไทยเป็นภาษาอังกฤษ ยิ่งไปกว่านั้นเทคโนโลยีการพัฒนาของ JNI, HTTP ถูกใช้ในระหว่างการพัฒนาแอพพลิเคชั่น เพื่อทดสอบและปรับปรุงการรับรู้ผลกระทบของแอปพลิเคชั่น และงานวิจัยนี้พยายามทดสอบการใช้ฟอนต์รูปแบบใหม่ของป้ายสาธารณะภาษาไทย จากผลการทดสอบแอพพลิเคชั่นสามารถรับรู้ป้ายสาธารณะภาษาไทยที่พึงพอใจ ผลที่ตามมาคือความสำเร็จในการพัฒนาแอปพลิเคชั่นจดจำภาพใหม่สำหรับป้ายประชาสัมพันธ์ภาษาไทยบนแพลตฟอร์ม Android นอกจากนี้ยังปรับปรุงผลของการจดจำโดยการทดสอบกับตัวอักษรแบบใหม่ๆ ซึ่งในอนาคตอาจเพิ่มประเภทของแบบอักษรการทดสอบต่อไป

Thematic Paper Title	Thai-Chinese Translation from Public Signage Based on Android Mobile Application
Author	Zhou Hongying
Thematic Paper Advisor	Asst. Prof. Narongdech Keeratipranon, Ph.D
Department	Computer and Telecommunication Engineering
Academic Year	2019

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 purpose of this paper is to provide a development of Image recognition Android application for tourist or other foreigners in Thailand, which could used to recognize and translate Thai characters in the pubic 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.

กิตติกรรมประกาศ

ความสำเร็จของการจัดทำสารนิพนธ์เรื่อง “การพัฒนาโมบายแอพพลิเคชั่นแปลภาษาไทย-จีนจากป้ายสาธารณะบนระบบปฏิบัติการแอนดรอยด์ ” ได้รับความกรุณาอย่างยิ่งจาก ผศ.ดร.ณรงค์เดช กีรติพรานนท์ ที่ปรึกษาสารนิพนธ์ ที่ได้ให้ความรู้ คำปรึกษา ตลอดระยะเวลาของการวิจัยอันเป็นประโยชน์ต่องานสารนิพนธ์ รวมถึงได้รับความกรุณาข้อเสนอแนะจากคณะกรรมการสอบสารนิพนธ์ทุกๆ ท่าน ผู้วิจัยของขอบพระคุณเป็นอย่างสูง

นอกจากนี้ขอขอบคุณบิดา มารดา และบุคคลอีกหลายท่านที่ให้การสนับสนุนและเป็นกำลังใจมาโดยตลอด

Zhou Hongying

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ฆ
บทคัดย่อภาษาอังกฤษ ..... ง
กิตติกรรมประกาศ ..... จ
สารบัญตาราง ..... ซ
สารบัญภาพ ..... ฌ
บทที่
1. บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ 2
1.3 ขอบเขตของงานวิจัย 3
1.4 ประโยชน์ที่คาดว่าจะได้รับ 3
1.5 วิธีการทดสอบ 3
1.6 ภาพรวมของระบบ 4
1.7 อุปกรณ์และซอฟต์แวร์ 4
1.8 องค์ความรู้ใหม่ 5
2. แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง 6
2.1 การติดต่อสื่อสารของมนุษย์ 6
2.2 การแปลความหมายข้ามภาษา 7
2.3 เทคนิคการแปลภาษาจากภาพ 8
2.4 การแปลภาษาไทย-จีน 18
2.5 Mobile application สำหรับการแปลภาษา 19
2.6 MVC 24
2.7 พื้นฐานการพัฒนา JAVA NATIVE INTERFACE(JNI) 26
2.8 สรุป 27

สารบัญ (ต่อ)

3. การออกแบบและพัฒนา 28
3.1 การออกแบบระบบแอพมือถือ 28
3.2 การใช้งานการรับภาพ 30
3.3 การรู้จำภาพ 33
3.4 Baidu Translation 42
3.5 อินเตอร์เฟซ JNI 45
3.6 OCR ระบบปรรจุภัณฑ์ 47
3.7 สรุป 50
4 ผลการศึกษา 51
4.1 สภาพแวดล้อมในการทดสอบ 51
4.2 หน้าจอแสดงผลการทำงานของโปรแกรม 52
4.3 Thai OCR มีปุ่มฟังก์ชั่นคือVocabulary Translate OCR 52
4.4 ผลการทดสอบ 53
5 สรุปผลและข้อเสนอแนะ 72
5.1 สรุปผลการวิจัย 72
5.2 อภิปรายผลการวิจัย 72
5.3 ข้อจำกัดและแนวทางแก้ไขของงานวิจัย 73
5.4 ข้อเสนอแนะและแนวทางการพัฒนาในอนาคต 73
บรรณานุกรม 74
ประวัติผู้เขียน 77

สารบัญภาพ

ภาพที่ **หน้า**

1.1 ภาพรวมการออกแบบระบบ 4
2.1 ฮิสโทแกรมของGlobal Thresholding 9
2.2 กระบวนการหาภาพถ่าย 10
2.3 การแยกตัวอักษรออกจากป้าย 11
2.4 การแก้ไขตัวละอักษร 11
2.5 เริ่มการฝึก 12
2.6 เซลล์ประสาทเทียมที่ถูกจำลองขึ้น 13
2.7 โครงข่ายประสาทเทียมแบบชั้นเดียว 15
2.8 โครงข่ายประสาทเทียมแบบหลายชั้น 16
2.9 Supervised Learning 16
2.10 Unsupervised Learning 17
2.11 Baidu Translate 18
2.12 กระบวนการการแปลภาษา 19
2.13 แอพมีฟังก์ชั่นให้เลือกใช้งานตัวอย่างโดยใช้Google Translate and Baidu Translate .. 20
2.14 กล้องสแกนตัวอย่างการแปลป้ายภาษาไทยเป็นภาษาจีนโดยใช้ Google translate ในโหมดถ่ายภาพ 21
2.15 กล้องสแกนตัวอย่างการแปลป้ายภาษาไทยเป็นภาษาจีนโดยใช้ Google translate ในนิ้วลาก 21
2.16 แอพมีฟังก์ชั่นให้เลือกใช้งานตัวอย่างโดยใช้ Baidu Translate 22
2.17 กล้องการหาคำตัวอย่างการแปลป้ายภาษาอังกฤษเป็นภาษาจีนโดยใช้ Baidu Translate ในโหมดถ่ายภาพ 22
2.18 กล้องสแกนตัวอย่างการแปลป้ายภาษาอังกฤษเป็นภาษาจีนโดยใช้ Baidu Translate ในโหมดถ่ายภาพ 23
2.19 การพูดตัวอย่างโดยใช้ Baidu Translate 23
2.20 Baidu Translate อื่นๆ 23
2.21 สถาปัตยกรรมของ MVC 24
2.22 แสดงการสื่อสารระหว่างโปรแกรมJavaและNative C 26

สารบัญภาพ (ต่อ)

ภาพที่ **หน้า**

3.1 การออกแบบระบบแอพมือถือ 28
3.2 กระบวนการถ่ายภาพ 30
3.3 code ของกระบวนการถ่ายภาพ 31
3.4 กระบวนการรับภาพจากอัลบั้มรูป 31
3.5 code ของกระบวนการรับภาพจากอัลบั้มรูป 32
3.6 codeของตัดภาพ 33
3.7 ติดตั้ง Tesseract-OCR เสร็จ 34
3.8 การรู้จำอักขระ 34
3.9 หลังการรู้จำอักขระ 34
3.10 สร้างไฟล์tif 35
3.11 สร้างbox 35
3.12 ภาพการรู้จำอักขระต้นฉบับจากป้าย 36
3.13 ภาพการรู้จำอักขระต้นฉบับจากการพิมพ์ 36
3.14 ภาพการรู้จำอักขระจากป้ายหลังการแก้ไข 37
3.15 ภาพการรู้จำอักขระจากการพิมพ์หลังการแก้ไข 37
3.16 เริ่มการฝึก unicharset 38
3.17 ทำFont characteristics กับ Aggregating font features 39
3.18 สร้างtraineddataใหม่ 40
3.19 ผลทดสอบใหม่ 41
3.20 กระบวนการการปรับแต่งเมื่อการรู้จำผิดพลาด 41
3.21 Baidu access process 42
3.22 Translation process 42
3.23 code ต่อnetwork 43
3.24 code ทำการขอร้อง 43
3.25 code แพ็คเกจ json 45
3.26 code Get The Returned Data 45
3.27 code ผลการแปล 45

สารบัญภาพ (ต่อ)

ภาพที่ **หน้า**

3.28 กระบวนการระบบอินเตอร์เฟซ JNI 46
3.29 code Java declare local methods 46
3.30 code Dynamic link library 47
3.31 code run Tesseract 47
3.32 code delete 47
3.33 กระบวนการ OCR ระบบบรรจุภัณฑ์ 48
3.34 OCR ระบบบรรจุภัณฑ์ใช้งานจริง 49
4.1 โทรศัพท์มือถือแอนดรอยด์ รุ่น Huawei 9 ที่ใช้ในการทดสอบโปรแกรม 51
4.2 แสดงThaiOCR APPบนเครื่องโทรศัพท์มือถือ 52
4.3 หน้าหลักของโปรแกรม 52
4.4 แสดงการถ่ายรูปจาก Snapshot 53
4.5 แสดงการถ่ายรูปจาก SelectFromAlbum 53
4.6 แสดงการตัดภาพ 54
4.7 แสดงภาษาไทยแปลเป็นภาษาอังกฤษกับจีน 54
4.8 แสดงการ Vocabulary 55
4.9 ตัวอย่างของป้ายถนน 55
4.10 ป้ายดั้งเดิมที่ใช้ในการทดสอบ 56
4.11 ป้ายเอียงที่ใช้ในการทดสอบ 59
4.12 ป้ายแสงน้อยที่ใช้ในการทดสอบ 62
4.13 ป้ายตัวพิมพ์ที่ใช้ในการทดสอบ 65
4.14 ผลการเปรียบเทียบความแม่นยำ 68
4.15 ผลการทดลองป้ายเอียง 69
4.16 ผลการเปรียบเทียบความแม่นยำของ Google translation 70

สารบัญตาราง

ตารางที่ **หน้า**

2.1 Baidu Translate and Google Translate เปรียบเทียบการทำงาน 24
3.1 Mobile hardware standards..... 29
3.2 data type..... 44
4.1ป้ายดั้งเดิมที่ยังไม่ได้เพิ่มความแม่นยำ..... 57
4.2 ป้ายดั้งเดิมที่เพิ่มความแม่นยำ..... 58
4.3 ป้ายเอียงที่ไม่ได้เพิ่มความแม่นยำ..... 60
4.4 ป้ายเอียงที่เพิ่มความแม่นยำ..... 61
4.5 ป้ายแสงน้อยที่ยังไม่ได้เพิ่มความแม่นยำ..... 63
4.6 ป้ายแสงน้อยที่เพิ่มความแม่นยำ..... 64
4.7 ป้ายตัวพิมพ์ที่ยังไม่ได้เพิ่มความแม่นยำ..... 66
4.8 ป้ายตัวพิมพ์ที่เพิ่มความแม่นยำ..... 67
4.9 ผลการเปรียบเทียบความแม่นยำ..... 68
4.10 ผลการเปรียบเทียบความแม่นยำ..... 65
4.11ผลการทดลองป้ายเอียง 66

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของปัญหา

ประเทศไทยมีศักยภาพและโอกาสทางการท่องเที่ยวมากมาย ทั้งด้านทำเลที่ตั้งในการเป็นจุดศูนย์กลางของภูมิภาคเอเชียตะวันออกเฉียงใต้ ความหลากหลายของทรัพยากรธรรมชาติ ศิลปวัฒนธรรม ประเพณีที่เป็นเอกลักษณ์ อัธยาศัยไมตรีของคนไทย (สำนักงานปลัดกระทรวงการท่องเที่ยวและกีฬา, 2560) ส่งผลให้มีนักท่องเที่ยวเดินทางเข้ามาท่องเที่ยวที่ประเทศไทยเป็นจำนวนมาก โดยเฉพาะนักท่องเที่ยวชาวจีนซึ่งในปี 2560 มีจำนวนชาวจีนเดินทางเข้ามาท่องเที่ยวในประเทศไทยจำนวน 9,805,753 คน (การท่องเที่ยวและกีฬา, 2561) แต่นักท่องเที่ยวชาวจีนส่วนใหญ่เป็นนักท่องเที่ยวที่มีข้อจำกัดในการรู้ภาษาไทยทั้งในการพูด อ่าน และเขียน ซึ่งจะทำให้มีความเข้าใจข้อมูลต่างๆ คลาดเคลื่อน รวมทั้งทำให้มีความกังวลใจในการเดินทางท่องเที่ยวในประเทศไทย และอาจจะส่งผลต่อการกลับมาเที่ยวซ้ำที่เมืองไทยอีกครั้งของนักท่องเที่ยว ซึ่งเป็นผลสืบเนื่องมาจากการไม่รู้ภาษาไทย

ปัจจุบันถึงแม้ว่าจะมีเทคโนโลยีที่สามารถช่วยในการแปลภาษาเช่น Google Translate ที่สามารถแปลได้หลายภาษาแต่ก็มีข้อจำกัดในเรื่องของความไม่สะดวกในการป้อนตัวอักษรเข้าสู่เครื่องคอมพิวเตอร์หรือ smart phone เพราะมีความซับซ้อนและยุ่งยากสำหรับผู้ไม่รู้ภาษาไทย นอกจากนี้ที่ประเทศจีนก็ไม่สามารถใช้ google ได้ และคนจีนส่วนใหญ่ก็ไม่นิยมใช้ google translate และประกอบกับในปัจจุบันยังไม่มีเครื่องมือที่ใช้ในการแปลภาษาไทยให้เป็นภาษาจีนที่รวดเร็วและแม่นยำในการแปลภาษา อย่างไรก็ตามมีเทคโนโลยีที่จะสามารถใช้ในการพัฒนาแอพพลิเคชั่นในการแปลภาษาที่รวดเร็วและแม่นยำได้แก่ซอฟต์แวร์การรู้จำตัวอักษร (Optical Character Recognition: OCR) ซึ่งเป็นเทคโนโลยีสำหรับการแปลงภาพเอกสาร ป้าย หรือรูปภาพให้กลายเป็นข้อความที่สามารถนำไปแปลเป็นภาษาจีนต่อไปได้

Fragoso,Kleban(2011) เสนอวิธีการ Singleclick Augmented reality translator บนโทรศัพท์มือถือสมาร์ทโฟน โนเกีย N 900 โดยประยุกต์ใช้ความสามารถในการซ้อนทับข้อความที่แปลโดยอัตโนมัติ OCR และ online translation services (Google Translation API) ได้ใช้การวิเคราะห์ภาพและจับคู่เฟรมกับมุมสุดท้ายโดยเปลี่ยนขนาดของภาพให้มีขนาด 9x9 พิกเซลผลลัพธ์

ที่ได้จาก online translation services มีความถูกต้องมากขึ้น ในขณะที่ Martinez Carballido (2011) เสนอวิธีการรู้จำเฉพาะตัวเลขโดยการใช้แม่แบบ(Template) ของตัวเลขขนาด 5x7 พิกเซล แต่ละพิกเซลมีค่า 0 และ 1 การรู้จำทำได้โดยการลดขนาดของตัวเลขลงให้เหลือขนาด 5x7 พิกเซล แล้วนำไปเปรียบเทียบกับแม่แบบ วิธีนี้มีความถูกต้องในการรู้จำตัวเลขร้อยละ 97.3 หลังจากนั้นไม่นาน Abdurrahman (2012) ได้เสนอวิธีการ OCR Tesseract คือการจับภาพบนสมาร์ทโฟน ซึ่งประยุกต์ใช้กับกล้องโทรศัพท์มือถือ ภาษาที่ใช้ทดสอบในครั้งนี้มีสามภาษาคือ ซุนดา อินโดนีเซีย และอังกฤษ ผลลัพธ์ที่ได้มีความถูกต้องร้อยละ 70 และ Ravina(2013) ได้ประยุดต์ใช้ OCR บนโทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์ มีวัตถุประสงค์การทำงานของ OCR นี้คือ รับรู้แสงและสังเคราะห์เสียงพูดให้มีประสิทธิภาพการใช้งานง่ายระบบการทำงานนี้จะทำการป้อนข้อมูลที่ได้รับจากข้อความภาพแล้วแปลงเป็นการพูด ระบบนี้มีประโยชน์ในการใช้งานต่าง ๆ เช่น ธนาคาร สำนักงานฯ ข้อจำกัดของงานวิจัยนี้ คือ ความถูกต้องของ OCR จะขึ้นอยู่กับคุณภาพการป้อนข้อมูลเอกสาร

ดังนั้นผู้วิจัยจึงได้ดำเนินการพัฒนาโมบายแอพพลิเคชั่นสำหรับแปลภาษาไทยเป็นภาษาจีน บนระบบปฏิบัติการแอนดรอยด์ เพื่อให้นักท่องเที่ยวชาวจีนได้ใช้แอพพลิเคชั่นดังกล่าวในการแปลภาษาไทยเป็นภาษาจีน ได้อย่างสะดวกสบาย เนื่องจากเป็นอุปกรณ์พกพาขนาดเล็ก ซึ่งคำที่นำมาพัฒนาในแอพพลิเคชั่นดังกล่าวเป็นคำที่ใช้ในชีวิตประจำวัน เช่น ชื่อสถานที่ ชื่อจังหวัด ห้องน้ำ โรงแรม ร้านอาหาร เป็นต้น ซึ่งสาเหตุที่เลือกพัฒนาแอพพลิเคชั่นดังกล่าวขึ้นเพราะในปัจจุบันเทคโนโลยีมีความเจริญก้าวหน้าคนส่วนใหญ่ใช้สมาร์ทโฟนในชีวิตประจำวันการพัฒนาแอพพลิเคชั่นเพื่อใช้ผ่านโมบายจะทำให้สะดวกและรวดเร็วต่อผู้ใช้ในปัจจุบัน

1.2 วัตถุประสงค์

1. เพื่อพัฒนาเทคโนโลยีที่จะใช้ในการระบุขอบเขตของป้าย (Sign image segmentation)
2. เพื่อแยกตัวอักษรที่ละตัวจากภาพป้าย
3. เพื่อแปลงรูปภาพตัวอักษรให้เป็นตัวอักษรที่สามารถเปลี่ยนแปลงได้
4. เพื่อแปลข้อความภาษาไทยเป็นข้อความภาษาจีน
5. เพื่อพัฒนาโมบายแอพพลิเคชั่นบนระบบปฏิบัติการแอนดรอยด์สำหรับแปลภาษาไทยเป็นภาษาจีน

1.3 ขอบเขตของการวิจัย

1. การพัฒนาเทคนิคการรับรู้อักษรไทย (Thai character recognition)
2. พัฒนาแอพพลิเคชั่นในการแปลภาษาไทยที่ใช้ในชีวิตประจำวัน จำนวน 50 คำ เพื่อแปลเป็นภาษาจีน
3. งานวิจัยชิ้นนี้ศึกษาขึ้นเพื่อแปลภาษาไทยเป็นภาษาจีน (Thai character recognition) โดยการสแกนตัวอักษร แบบ off-line
4. เพิ่มความแม่นยำการรู้จำภาษาไทยและพัฒนาเป็นแอปพลิเคชั่นบนมือถือ
5. เลือกขอบเขตภาพถ่าย เพื่อข้อความที่จะรับรู้ถูกต้องอยู่ในภาพที่ถ่าย
6. ใช้เทคโนโลยีJNIเพื่อเชื่อมต่อแอปพลิเคชั่นมือถือ
7. ใช้baidu server เพื่อแปลภาษาไทยเป็นภาษาจีนหรืออังกฤษ
8. ใช้เทคโนโลยี APKเพื่อดาวน์โหลดแอปพลิเคชั่นในมือถือ
9. โปรแกรมใช้ได้ทั้งถาบในอาคารและกลางแจ้งที่มีแสงสว่างมาก
10. ป้ายที่ระบบวิเคราะห์ได้ต้องมีขนาดไม่ต่ำกว่า 1440*1080 Pixels

1.4 ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่เกิดขึ้นจากงานวิจัยนี้ประกอบด้วย

1. แอพพลิเคชั่นการแปลภาษาไทยเป็นภาษาจีนบนระบบแอนดรอยด์สามารถใช้งานได้สะดวกทำให้ผู้ใช้บริการสามารถนำไปใช้ได้ง่ายด้วยวิธีการสแกนรูปภาพตัวอักษรโดยไม่ต้องพิมพ์ทำให้รวดเร็วและไม่ต้องกังวลว่าจะพิมพ์ผิดเพราะไม่ถนัดในการพิมพ์

2. ช่วยให้ผู้ใช้บริการหรือนักท่องเที่ยวชาวจีนเข้าใจภาษาไทยส่งผลให้รับรู้สถานที่หรือผลิตภัณฑ์ต่างๆ ในประเทศไทยทำให้มีความมั่นใจในการท่องเที่ยวในประเทศไทย

3. แอพพลิเคชั่นการแปลภาษาไทยเป็นภาษาจีนบนระบบแอนดรอยด์ที่เป็นแบบออฟไลน์นี้สามารถนำไปต่อยอดเพื่อเป็นระบบออนไลน์ซึ่งจะทำให้ผู้ใช้สะดวกมากขึ้น

1.5 วิธีการทดสอบ

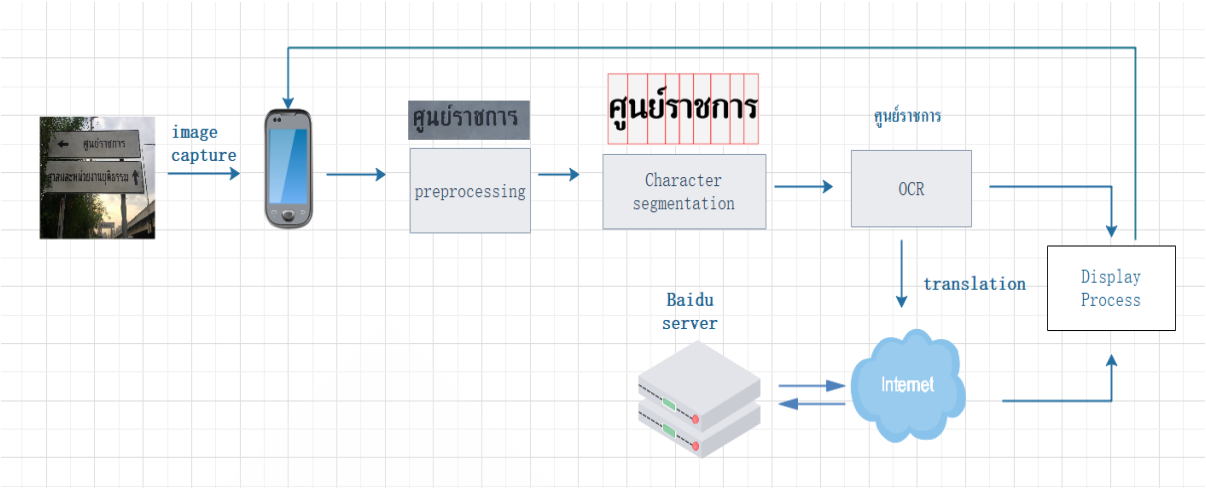
การทดสอบด้วยตนเองการทำงานของAPP จะทำการทดสอบเป็น 3 ระยะ โดยจะแบ่งตามลักษณะการทำงาน

ระยะแรกทดสอบVocabulary Translate OCR แต่ละฟังก์ชั่นในAndroid Studio

ระยะที่สองทดสอบ Vocabulary Translate OCR แต่ละฟังก์ชั่นในมือถือ

ระยะที่สามทดสอบป้ายที่สภาพแวดล้อมจริง

1.6 ภาพรวมของระบบ



ภาพที่ 1.1 ภาพรวมการออกแบบระบบ

งานวิจัยออกแบบเป็น 3 ส่วน ส่วนที่ 1 ถ่ายรูปจากป้าย ส่วนที่ 2 ประมวลผล image ส่วนที่ 2 translate

1.7 อุปกรณ์และซอฟต์แวร์

1.7.1 ส่วนของ(Hardware)

1) เครื่องคอมพิวเตอร์ Notebook

2) Android mobile simulator

3) โทรศัพท์มือถือที่Android

1.7.2 ส่วนของ(software)ซอฟต์แวร์

1) Java, Eclipse

2) Android studio

3) Android virtual device

4) Tesseract

5) baidu API

1.8 องค์ความรู้ใหม่

การพัฒนาแอพพลิเคชั่นการแปลภาษาไทยเป็นภาษาจีนบนระบบแอนดรอยด์ในครั้งนี้ทำให้มีแอพพลิเคชั่นในการแปลภาษาไทยเป็นภาษาจีนด้วยการสแกนรูปภาพให้เป็นข้อความทำให้สะดวกและรวดเร็วต่อผู้ใช้บริการ

DPU

บทที่ 2
เอกสาร แนวคิด และทฤษฎีที่เกี่ยวข้อง

การทบทวนเอกสาร แนวคิด และทฤษฎีที่เกี่ยวข้องผู้วิจัยได้ทบทวน

2.1 การติดต่อสื่อสารของมนุษย์

2.1.1 การฟัง

การฟังเป็นพฤติกรรมการรับสารผ่านโสตประสาทอย่างตั้งใจเชื่อมโยงกับกระบวนการคิดในสมอง โดยสมองแปลความหมายของเสียงจนเกิดความเข้าใจและมีปฏิกิริยาตอบสนอง การฟังจึงเป็นกระบวนการที่เกิดขึ้นภายในตัวบุคคล การฟังมีความสำคัญมากต่อการติดต่อสื่อสารของมนุษย์ในชีวิตประจำวัน ดังจะเห็นว่ามนุษย์ใช้เวลาไปกับการฟังมากที่สุดหากเปรียบเทียบกับการพูด การอ่านและการเขียน จอห์น ดับบลิว เคล์ทเนอร์ พบว่า ผู้ที่สื่อสารนั้น มีอัตราส่วนของการใช้ทักษะทางภาษา คือ ใช้เวลาในการฟัง 42% การพูด 32% การอ่าน 15% และการเขียน 11% ซึ่งทำให้เห็นว่า การฟังมีความสำคัญในการกำหนดความล้มเหลวหรือความสำเร็จของการสื่อสารอย่างมาก โดยการฟังทำให้ได้รับความรู้ การฟังทำให้รู้ข้อมูลข่าวสารต่างๆ ที่เป็นประโยชน์ต่อตนเองและสังคม เป็นกระบวนการเรียนรู้อย่างหนึ่งของมนุษย์ ทั้งที่เกิดจากการฟังจากบุคคลโดยตรงหรือฟังผ่านสื่ออิเล็กทรอนิกส์ ช่วยยกระดับจิตใจ ทำให้ได้รับความบันเทิง ช่วยผ่อนคลายความเครียด ช่วยพัฒนาทักษะการพูดให้มีประสิทธิภาพได้ และสร้างความเข้าใจอันดีระหว่างคนในสังคม

2.1.2 การพูด

การพูดหมายถึงพฤติกรรมในการสื่อความหมายของมนุษย์ โดยการเปล่งเสียงเป็นถ้อยคำ (วัจนภาษา) ขณะเดียวกันก็มีการใช้ภาษาที่ไม่ใช้ถ้อยคำ (อวัจนภาษา) ประกอบ โดยการพูดแบ่งตามลักษณะการพูดได้ 2 ประเภท คือ 1) การพูดอย่างไม่เป็นทางการ และ 2) การพูดอย่างเป็นทางการ ซึ่งการพูดมีองค์ประกอบพื้นฐานที่สำคัญคือ 1. ผู้พูด คือ ผู้ส่งสาร (sender) 2. เรื่องที่พูด คือ สาร หรือเนื้อหาสาระ (message) 3. ภาษา คือสื่อ (media) และ4. ผู้ฟัง คือ ผู้รับสาร (receiver)

2.1.3 การเขียน

การเขียน คือทักษะการใช้ภาษาชนิดหนึ่ง เป็นการถ่ายทอดความรู้ ความคิด จินตนาการ ประสบการณ์ต่าง ๆ รวมทั้งอารมณ์และความรู้สึกกับข่าวสาร เป็นการสื่อสารหรือสื่อความหมายโดยมีตัวหนังสือตลอดจนเครื่องหมายต่าง ๆ เป็นสัญลักษณ์แทนถ้อยคาในภาษาพูด เพื่อให้ผู้อ่าน

เข้าใจได้ตามความมุ่งหมายของผู้เขียน การเขียนจึงเป็นทักษะที่มีหลักฐานถาวรปรากฏอยู่นาน และการเขียนจะเกิดผลดีหรือผลเสียนั้น ขึ้นอยู่กับคุณภาพของเนื้อหาและกลวิธีการเขียนของผู้เขียน" (กองเทพ เคลือบพณิชกุล, 2542)

2.1.4 การอ่าน

การอ่านคือการรับรู้ความหมายจากถ้อยคำที่ตีพิมพ์จากสิ่งพิมพ์ชนิดต่าง ๆ โดยที่ผู้อ่านต้องเริ่มทำความเข้าใจวลี ประโยค ซึ่งรวมอยู่ในย่อหน้า แต่ละย่อหน้า แล้วรวมเป็นเรื่องเดียวกัน การอ่านเปรียบเหมือนการถอดรหัส อันเป็นผลจากการเห็นสัญลักษณ์หรือข้อความ การอ่านเน้นกระบวนการทางสมองที่ซับซ้อน ซึ่งการอ่านนั้นเกี่ยวข้องกับพฤติกรรม 3 ลักษณะ คือ 1. การรับรู้ ได้แก่การรับรู้คำ คือแปลสัญลักษณ์ที่เน้นลายลักษณ์อักษร 2. การมีความเข้าใจ 3. การมีปฏิกิริยาต่อสิ่งที่อ่าน

2.2 การแปลความหมายข้ามภาษา

2.2.1 ล่ามแปลภาษา

ล่ามแปลภาษาเป็นผู้ทำหน้าที่เรียบเรียงถ่ายทอดเนื้อความจากภาษาหนึ่ง โดยรับข้อความจากผู้พูดและถ่ายทอดเป็นคำพูดไปเป็นอีกภาษาหนึ่ง Basnett (1991) ได้กำหนดหลักการแปลไว้ 5 ประการ คือ 1) ผู้แปลต้องทำความเข้าใจอย่างถ่อยแท้ ทั้งความรู้สึกและการสื่อความหมายของเจ้าของต้นฉบับ ถึงแม้ว่ามีอิสระสามารถแปลเพิ่มเติมให้ชัดเจนกว่านี้ได้ 2) ผู้แปลควรมีความรู้ดีเยี่ยมทั้งภาษาต้นฉบับและภาษาเป้าหมาย 3) ผู้แปลควรหลีกเลี่ยงการแปลคำต่อคำ 4) ผู้แปรควรใช้รูปแบบภาษาที่นิยมใช้กันทั่วไป และ 5) ผู้แปลควรเลือกและเรียงคำให้เหมาะสมทำให้เกิดน้ำเสียงที่ถูกต้อง

2.2.2 ระบบอัตโนมัติในการแปล

แนวทางในการพัฒนาโปรแกรมแปลภาษาแบ่งได้เป็น 2 แนวทาง คือ การแปลด้วยคอมพิวเตอร์โดยอาศัยกฎไวยากรณ์ ซึ่งเป็นแนวทางที่เริ่มพร้อมกับการแปลภาษาด้วยคอมพิวเตอร์ยุคต้นๆ อีกแนวทางคือการแปลภาษาด้วยคอมพิวเตอร์โดยอาศัยบทความ มาช่วยในการแปล (รัฐติกร วรากูลศิริพันธ์ สมศักดิ์ จันวัน และอรัญญา ปรีชาไว, 2532) โดยมีรายละเอียดดังนี้

1) การแปลด้วยคอมพิวเตอร์โดยอาศัยกฎไวยากรณ์ ประกอบด้วย

1.1) การแปลด้วยวิธีการถ่ายโอนกระบวนการแปลประกอบด้วยสามขั้นตอน คือ การวิเคราะห์ภาษาต้นทางไปเป็นรูปแสดงแทนภาษาต้น การย้ายข้างไปเป็นรูปแสดงแทนของภาษาเป้าหมาย และการสร้างหรือสังเคราะห์เป็นข้อความของภาษาเป้าหมาย

1.2) การแปลด้วยวิธีภาษากลาง วิธีนี้มองกระบวนการแปลว่าประกอบด้วย 2 ขั้นตอน คือ การวิเคราะห์ภาษาต้นทางไปเป็นรูปแทนกลาง ซึ่งไม่ขึ้นกับภาษาและการสังเคราะห์ภาษาเป้าหมายจากรูปแทนกลาง

1.3) Principle-based Machine Translation หลักการพื้นฐานของวิธีนี้คือ ถือว่ามีกฎสากลที่ใช้ได้กับทุกภาษา ส่วนความแตกต่างระหว่างภาษานั้นจัดการได้โดยการตั้งค่าพารามิเตอร์ทางวากยสัมพันธ์ (syntax) และอรรถลักษณ์ประจำคำ (lexical semantic) ที่แตกต่างกัน

2) การแปลภาษาด้วยคอมพิวเตอร์โดยอาศัยฐานข้อมูล (Compus-based Machine Translation) มีการวิจัยกลุ่มต่างๆ ดังนี้

2.1) การแปลโดยอาศัยข้อมูลเชิงสถิติ ลักษณะสำคัญคือใช้วิธีทางสถิติในการวิเคราะห์ภาษาต้นทาง และสังเคราะห์ภาษาเป้าหมาย จากการทดลองของบริษัทไอบีเอ็มในช่วยปลายทศวรรษที่ 90 ด้วยวิธีการนี้ผลการแปลที่ได้ดีกว่าที่ผู้วัยคาดไว้ จึงทำให้มีการทบทวนวิธีการในการแปลภาษาด้วยคอมพิวเตอร์กันใหม่

2.2) การแปลโดยอาศัยตัวอย่าง แนวคิดพื้นฐานของวิธีนี้มาจากการศึกษาการแปลของคน กล่าวคือ ผู้แปลมักจะหาลักษณะประโยคที่คล้ายกับที่เคยแปลมาแล้วมาเทียบในการแปลเป็นภาษาเป้าหมาย วิธีนี้จะใช้ฐานข้อความคู่ภาษา ของวลีหรือประโยคมาสนใจเป็นรูปแบบของตัวอย่างสำหรับการแปล

2.3 เทคนิคการแปลภาษาจากภาพ

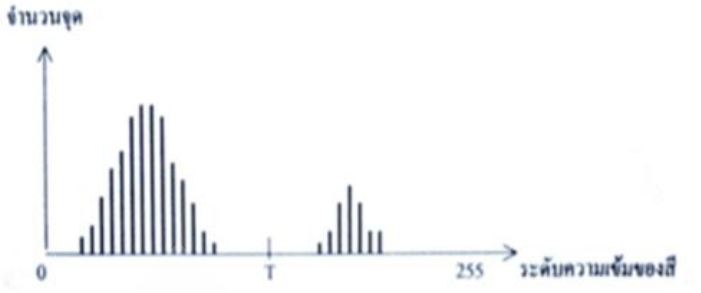
ในการแปลภาษาจากภาพนั้นมีหลายกระบวนการ โดยเริ่มจากส่วนการหาป้ายในภาพถ่าย แล้วจึงไปหาตัวอักษรบนป้ายนั้น เมื่อได้อักษรแล้วจึงส่งต่อไปยังระบบการตีความหมาย ซึ่งอาจจะใช้โครงข่ายปราสาทเทียมก็ได้ เมื่อได้ข้อความจึงส่งต่อไปส่วนการแปลภาษาต่อไป

2.3.1 ระบบรู้จำตัวอักษร (Optical character recognition: OCR)

ระบบรู้จำตัวอักษร [9,10,11] เป็นกระบวนการทางกลไกหรือทางอิเล็กทรอนิกส์เพื่อแปลภาพข้อความที่ได้จากการเขียนหรือจากการพิมพ์ไปเป็นข้อความที่สามารถแก้ไขได้โดยเครื่องคอมพิวเตอร์ โดยทั่วไปประกอบด้วยขั้นตอนต่อไปนี้

(1) การประมวลผลขั้นต้น (Pre-Processing) - เป็นขั้นตอนในการปรับแต่งและจัดเตรียมข้อมูลให้เหมาะสมกับกระบวนการรู้จำที่จะดำเนินต่อไปขั้นตอนนี้มีความสำคัญต่อประสิทธิภาพโดยรวมของระบบ หากมีความผิดพลาดเกิดขึ้นในส่วนนี้จะส่งผลกระทบไปยังส่วนถัดไปของระบบด้วย เทคนิคที่นิยมในการใช้งานมีหลายวิธีตามแต่ละวัตถุประสงค์ในการใช้งานเช่น Global Thresholding [12] ซึ่งเป็นเทคนิคในการจัดเตรียมภาพโดยจะใช้ค่า Thresholding ของฮิสโทแกรม

ในการจัดเตรียมภาพ ซึ่งจะใช้เพียงค่าเดียวในการหาจุดแบ่งกลุ่มของฮิสโทแกรม ดังแสดงในภาพที่ 2.1



ภาพที่ 2.1 ฮิสโทแกรมของGlobal Thresholding

(2) การรู้จำ (Recognition) [9,13] - ขั้นตอนนี้ถือเป็นหัวใจของกระบวนการทั้งหมด เพราะเป็นส่วนที่จะตัดสินว่ารูปตัวอักษรที่ส่งเข้าไปเป็นข้อมูลหรือตัวอักษรใด วิธีการหลากหลายได้ถูกนำเสนอขึ้นเพื่อปรับปรุงให้ได้ผลการทำงานที่ดีที่สุด สามารถรองรับลักษณะข้อมูลที่ซับซ้อนหรือมีเงื่อนไขในเอกสารได้ ซึ่งหากแบ่งตามรูปแบบหลักๆในทฤษฎีที่ใช้ในการรู้จำ สามารถแบ่งออกได้เป็นกลุ่มๆ ได้แก่ วิธีเข้าคู่รูปแบบ (Template Matching) วิธีทางสถิติ (Statistic Approach) วิธีการ วิเคราะห์ทางโครงสร้าง (Structural Analysis) และวิธีโครงข่ายประสาท เทียม (Neural Network)

(3) การประมวลผลขั้นปลาย (Post-Processing)–หลังจากที่ผ่านขั้นตอนการรู้จำแล้วจะได้ข้อมูลจากภาพ ซึ่งอาจเป็นข้อมูลในลักษณะของ Boolean (มีการท าเครื่องหมายหรือไม่) หรืออาจเป็นในรูปแบบตัวอักษร กระบวนการในส่วนนี้จะเป็นส่วนการตรวจสอบผลการแปลความหมาย (Data Validation) เช่น การตรวจสอบความถูกต้องของการสะกดคำและไวยากรณ์ [14] โดยอาจนำเครื่องมือต่าง ๆ มาช่วยในการตรวจสอบ เช่น พจนานุกรม เป็นต้น ซึ่งนอกเหนือจากการตรวจสอบความถูกต้องระดับคำแล้วยังอาจตรวจสอบไวยากรณ์ในระดับประโยคได้อีกด้วย

2.3.2 กระบวนการหาภาพถ่าย

1. การแปลงค่าสี RGB เป็น Gray Scale[17]

สำหรับขั้นตอนการเตรียมภาพเพื่อให้การประมวลผลมีความรวดเร็วและง่ายขึ้นจึงมีการเปลี่ยนแปลงภาพสีให้อยู่ในรูปของภาพที่มีระดับความเทามีค่าอยู่ระหว่าง0-255สามารถคำนวณได้ดังสมการที่ 1

$$\textbf{Gray value} = 0.3R + 0.59G + 0.11B \qquad (1)$$

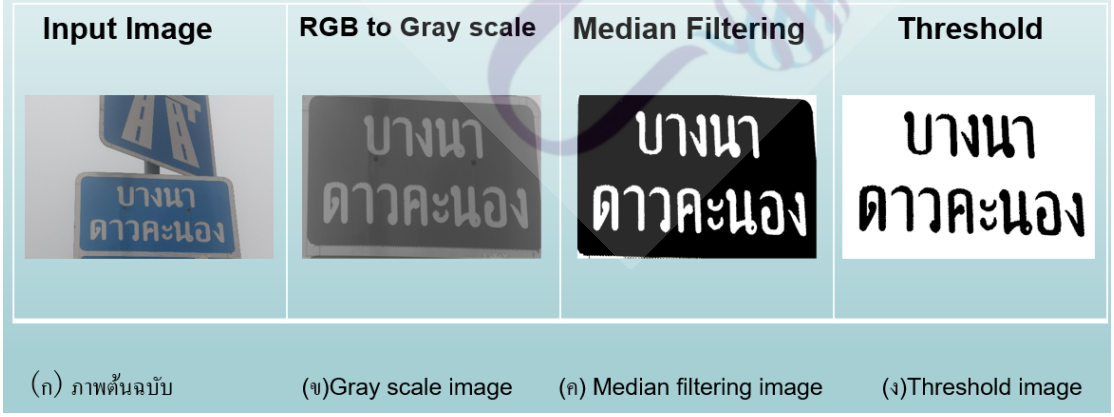
2. การกรองค่ามัธยฐาน

สัญญาณรบกวนเป็นสิ่งที่หลีกเลี่ยงไม่ได้สำหรับการใช้ภาพต้นฉบับ ระบบนี้ใช้การกรองค่ามัธยฐานเพื่อกำจัดสัญญาณรบกวนบนภาพต้นฉบับ การใช้การกรองค่ามัธยฐานไม่เพียงแต่สามารถลดสัญญาณรบกวนเท่านั้น แต่ยังทำให้ความถี่สูงขึ้นอีกด้วย โดยจะเป็นประโยชน์สำหรับใช้ตรวจจับขอบในภาพต้นฉบับ

3. การแปลงภาพระดับเทาให้เป็นภาพขาว-ดำ

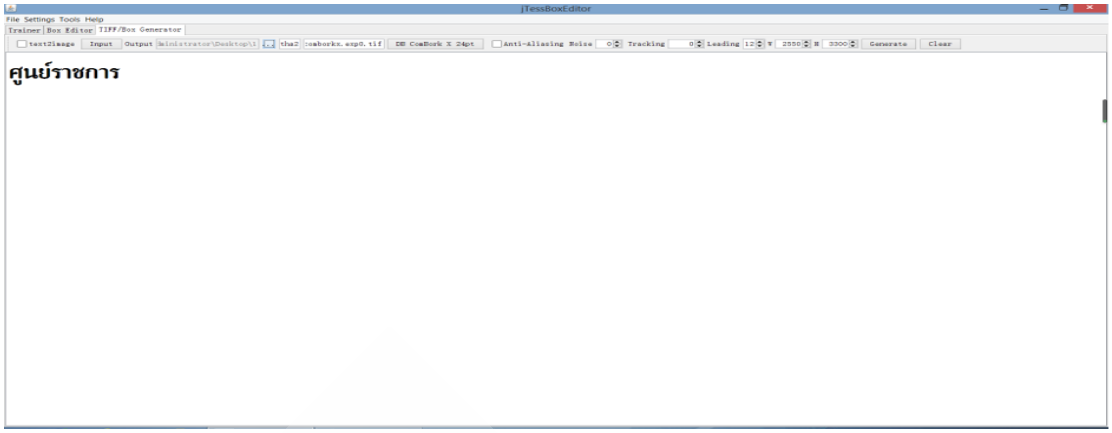
ความแม่นยำในการรู้จำป้ายอัตโนมัติของระบบที่ออกแบบขึ้นอยู่กับความถูกต้องของกระบวนการแบ่งส่วนตัวอักษรหากการแบ่งส่วนล้มเหลว ส่วนของตัวอักษรจะแบ่งออกเป็นสองชิ้นได้ไม่ถูกต้อง หรืออาจมีการรวมตัวสองตัวเข้าด้วยกัน อาจทำให้เกิดความไม่ถูกต้องในการประมวลผลภาพ หลังจากที่ภาพสีถูกแปลงเป็นค่าสีเทา(0-255) จะใช้กระบวนการกำหนดเกณฑ์การปรับตัว T ในการป้อนข้อมูลภาพระดับสีเทาเพื่อสร้างภาพไบนารีเทคนิคนี้สามารถแสดงดังสมการที่ 2 .2

$$f_{threshold}(x,y) = \begin{cases} 1, & f_f(x,y) < Threshold \\ 0, & f_f(x,y) \geq Threshold \end{cases} \quad (2)$$

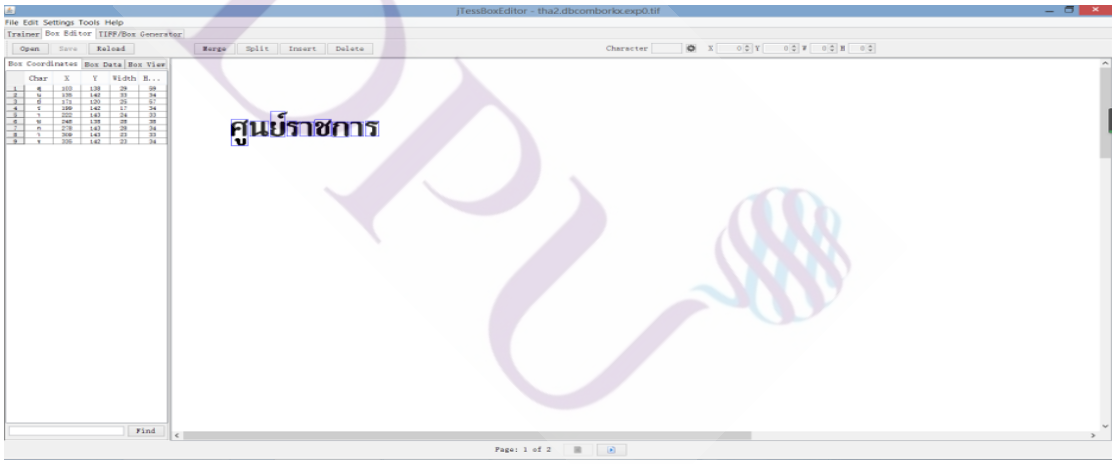


ภาพที่ 2.2 กระบวนการหาภาพถ่าย

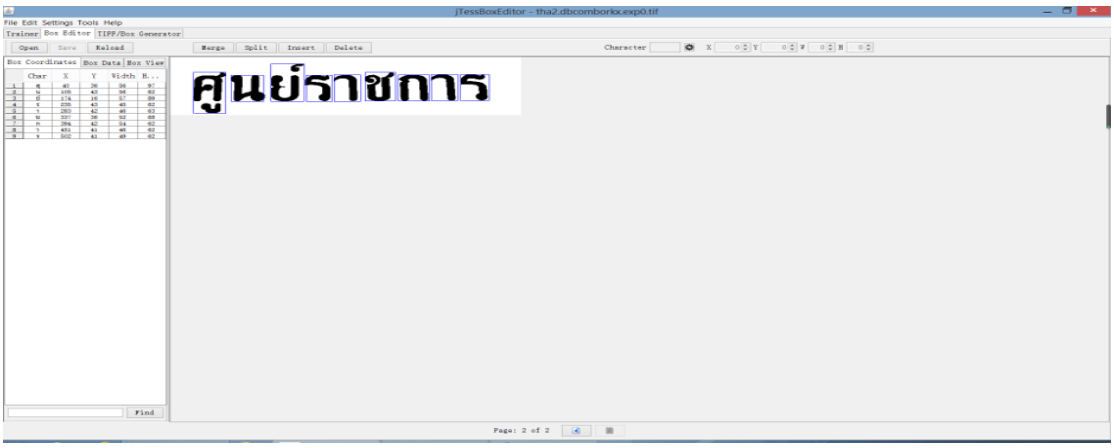
2.3.3 กระบวนการหารู้จำตัวอักษรบนป้าย



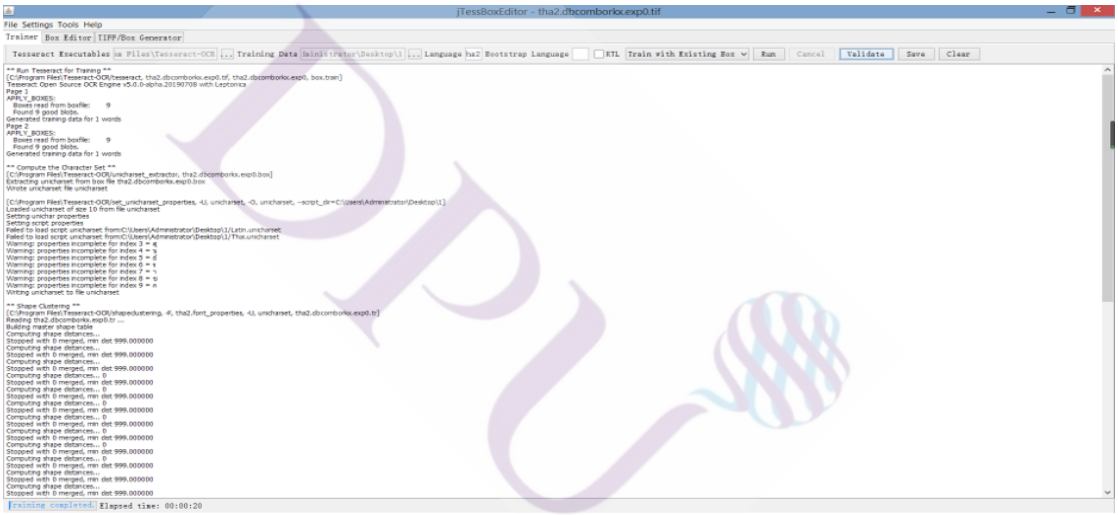
ภาพที่ 2.3 การแยกตัวอักษรออกจากป้าย

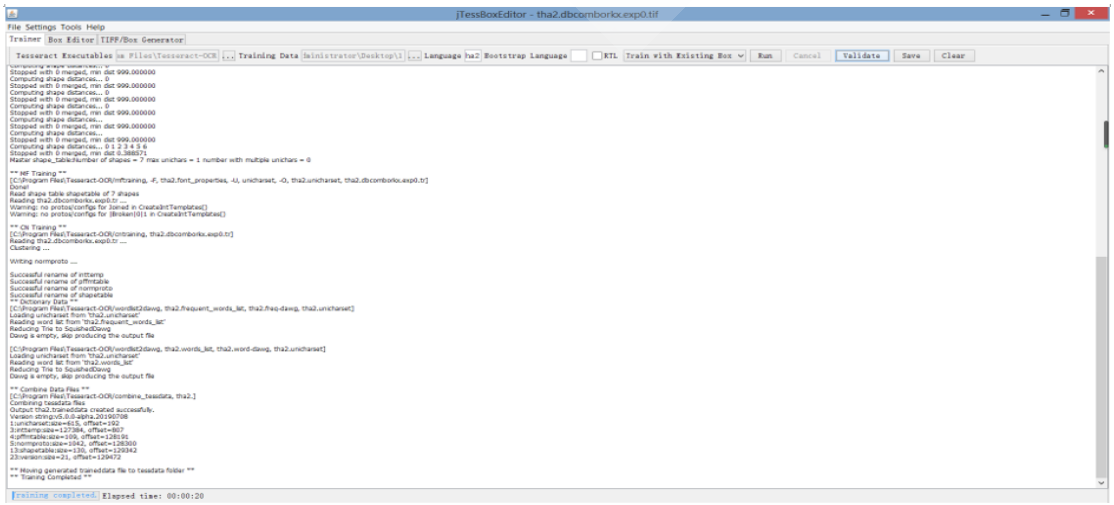


ภาพที่ 2.4 การแก้ไขตัวละอักษร



ภาพที่ 2.5 เริ่มการฝึก



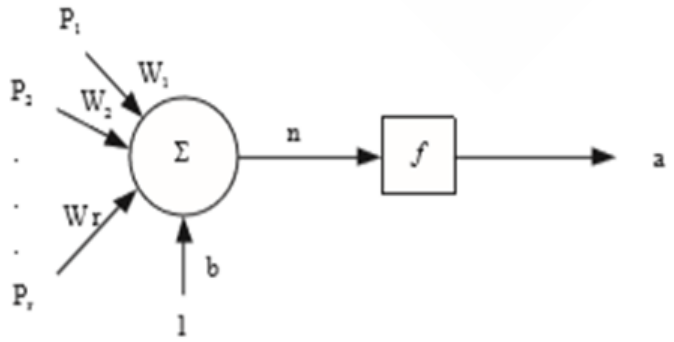


2.3.4 โครงข่ายประสาทเทียม

โครงข่ายประสาทประดิษฐ์ (Artificial Neural Network : ANN) [15]เป็นวิธีการการรู้จำหนึ่งที่ได้ ความนิยมอย่างสูง เนื่องจากสามารถฝึกการเรียนรู้ได้สามารถใส่ปัจจัยที่มีผลต่อผลลัพธ์ได้หลากหลาย โดยให้ผลลัพธ์เป็นที่น่าพอใจ ลักษณะการทำงานจะเป็นการจำลองการทำงานของสมองมนุษย์มาอยู่รูปแบบของคอมพิวเตอร์ เพื่อช่วยให้คอมพิวเตอร์รู้จักคิดจดจำได้ วิธีการของ Artificial Neural Network จะมีอยู่มากมายขึ้นอยู่กับการนำไปประยุกต์ใช้ ในแต่ละลักษณะงาน ซึ่งในงานวิจัยนี้เลือกการใช้โครงข่ายประสาทเทียมในการรู้จำภาพป้าย

1.โครงข่ายประสาทประดิษฐ์)Artificial Neural Network : ANN) โครงข่ายประสาทเทียนคือแบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นเพื่อจำลองการทำงานของโครงข่ายประสาทในสมองมนุษย์มีคุณลักษณะคล้ายกับการส่งผ่านสัญญาณประสาทในสมองของมนุษย์ กล่าวคือมีความสามารถในการรวบรวมความรู้ (Knowledge) โดยผ่านกระบานการเรียนรู้(Learning Process) และความรู้เหล่านี้จัดเก็บอยู่ในโครงข่ายในรูปค่านำหนัก(Weight) ซึ่งสามารถปรับเปลี่ยนค่าได้เมื่อมีการเรียนรู้ใหม่ๆ เข้าไป ค่าน้ำหนักเปรียบเสมือนความรู้ที่รวบรวมไว้เพื่อใช้ในการแก้ปัญหาเฉพาะอย่างของมนุษย์

การประมวลผลต่างๆ เกิดขึ้นในหน่วยประมวลผลย่อยเรียกว่าโหนด (Node) ซึ่งโหนดเป็นการจำลองลักษณะการทำงานมาจากเซลล์การส่งสัญญาณ(Signal)ระหว่างโหนดที่เชื่อมต่อกัน (Connection)จำลองมาจากการเชื่อมต่อของเดนไดรต์และแอกซอนในระบบประสาทของมนุษย์ ภายในโหนดจะมีฟังก์ชันกำหนดสัญญาณส่งออกที่เรียกว่าฟังก์ชันกระตุ้น(Activation Function) หรือฟังก์ชันการเปลี่ยนแปลง ซึ่งทำหน้าที่เสมือนกระบวนการทำงานในเซลล์ดังภาพที่ 2.6



ภาพที่ 2.6 เซลล์ประสาทเทียมที่ถูกจำลองขึ้น

โครงข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบดังนี้

1) ข้อมูลป้อนเข้า(Input)เป็นข้อมูลที่เป็นตัวเลขหากเป็นข้อมูลเชิงคุณภาพ ต้องแปลงให้อยู่ในรูปเชิงปริมาณที่โครงข่ายประสาทเทียมยอมรับได้

2) ข้อมูลส่งออก(Output)คือ ผลลัพธ์ที่เกิดขึ้นจริง(Actual Output)จากกระบวนการเรียนรู้ของโครงข่ายประสาทเทียน

3) ค่าถ่วงน้ำหนัก(Weights) คือสิ่งที่ได้จากการเรียนรู้ของโครงข่ายประสาทเทียนหรือเรียกอีกอย่างหนึ่งว่าค่าความรู้(Knowledge) ค่านี้จะถูกเก็บเป็นทักษะเพื่อใช้ในการจดจำข้อมูลอื่นๆ ที่อยู่ในรูปแบบเดียวกัน

4) ฟังก์ชันผลรวม(Transfer Functions)เป็นผลรวมของข้อมูลป้อนเข้า และค่านำหนัก

5) ฟังก์ชันการเปลี่ยนแปลง(Transfer Function)เป็นการคำนวณการจำลองการทำงานของโครงข่ายประสาทเทียม เช่น ซิกมอยด์ฟังก์ชัน(Sigmoid Function)ฟังก์ชันไฮเปอร์โบลิกแทนเจนต์ (Hyperbolic Tangent Function) เป็นต้น

เซลล์ประสาทเทียมที่จำลองขึ้นมีหลักการโดยนำข้อมูลป้อนเข้า(p)คูณกับค่าถ่วงนำหนัก(w)แล้วนำมารวมกับผลคูณของไบแอส(Bias,b)กับตัวแปรต้นของไบแอสซึ่งมีค่าเท่ากับ 1 จะได้

$$n=wp+b$$

โดยที่ n คือ ตัวแปรต้นสุทธิ(Net Input) และค่าตัวแปรตามของเซลล์ประสาท(Neural Output, a) มีค่าเท่ากับ

$$a= f(wp+b)$$

โดยที่ f คือฟังก์ชันการเปลียนแปลง

โดยทั่วไปเซลล์ประสาทเทียมมีข้อมูลป้อนเข้าหรือตัวแปรต้นมากกว่าหนึ่งตัวแปร ดังแสดงในรูปที่ 2 ทำให้ได้สมการเซลล์ประสาทเทียมในกรณีนี้ดังนี้ฃ

$$n = \sum(wp+b)$$

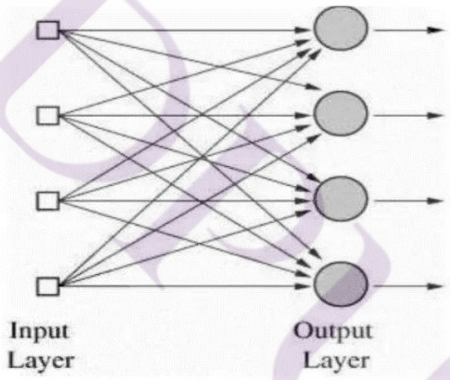
$$a= f(\sum wp+b)$$

2. ลักษณะของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมประกอบด้วยเซลล์ประสาทเทียมหรือโหนดจำนวนมากเชื่อมต่อกัน ซึ่งการเชื่อมต่อจะแบ่งออกเป็นกลุ่มย่อย ที่เรียกว่า‘ชั้น(Layer)’ชั้นแรกเป็นชั้นที่นำเอา

ข้อมูลเข้า เรียกว่า‘ชั้นรับข้อมูลป้อนเข้า(Input Later)’และชั้นสุดท้ายเรียกว่า‘ชั้นส่งข้อมูลออก (Output Layer)’ส่วนชั้นที่อยู่ระหว่างชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเรียกว่า‘ชั้นแอบแฝง (Hidden Layer)’ โดยทั่วไปชั้นแอบแฝงอาจมีมากกว่าหนึ่งชั้นก็ได้ด้วยเหตุนี้จึงสามารถแบ่งประเภทของโครงข่ายประสาทเทียมตามจำนวนชั้นของโครงข่ายได้2แบบคือ โครงข่ายแบบชั้นเดียว(Single Layer)และโครงข่ายแบบหลายชั้น(Multi Layer)

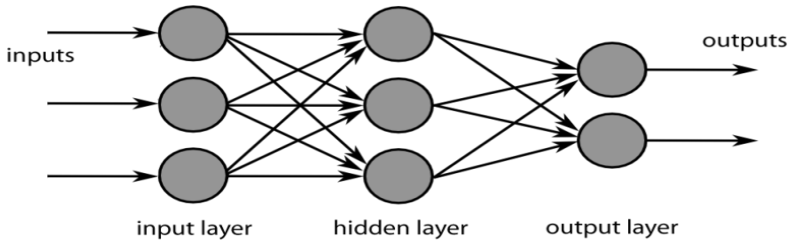
1)โครงข่ายแบบชั้นเดียวเป็นโครงข่ายประสาทเทียมอย่างง่ายที่มีแค่ชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเท่านั้น โหนดในชั้นรับข้อมูลผป้อนเข้าทำหน้าที่รับข้อมูลเข้าแล้วส่งข้อมูลผ่านเส้นเชื่อมโยงต่างๆ ไปให้โหนดชั้นส่งข้อมูลออก และ โหนดในชั้นนี้จะนำข้อมูลที่ได้รับมาคำนวณด้วยฟังก์ชันการเปลี่ยนเปลงที่เหมาะสมกับปัญหา แล้วส่งผลลัพธ์ที่ได้ออกมาเป็นข้อมูลส่งออกดังภาพที่ 2.7



ชั้นรับข้อมูลป้อนเข้า ชั้นส่งข้อมูลออก

ภาพที่ 2.7 โครงข่ายประสาทเทียมแบบชั้นเดียว

2) โครงข่ายแบบหลายชั้น เป็นโครงข่ายที่มีชั้นแอบแฝงตั้งแต่หนึ่งชั้นขึ้นไป โครงข่ายแบบนี้ใช้แก้ปัญหาที่มีความซับซ้อนที่โครงข่ายแบบชั้นเดียวแก้ไม่ได้ จึงต้องเพิ่มจำนวนโหนดที่มีการคำนวณหรือชั้นแอบแฝงให้กับโครงข่าย ดังภาพที่ 2.8

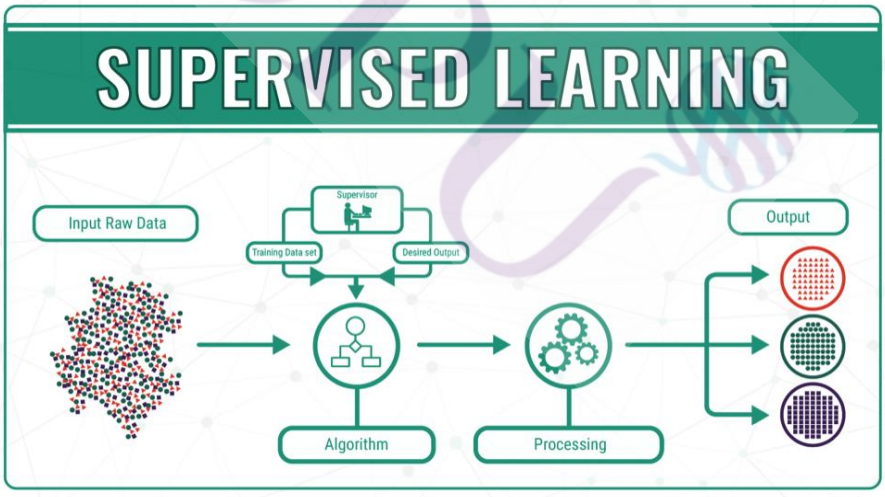


ชั้นรับข้อมูลป้อนเข้า (Hidden Layer) ชั้นส่งข้อมูลออก

ภาพที่ 2.8 โครงข่ายประสาทเทียมแบบหลายชั้น

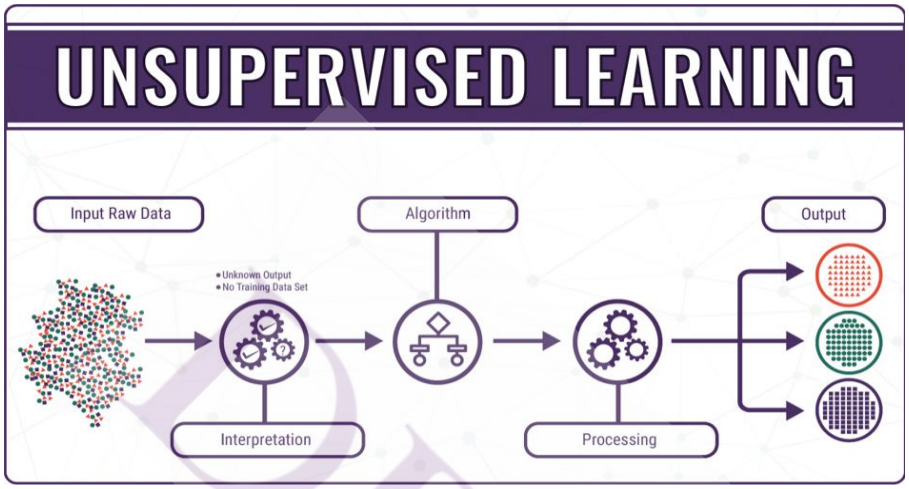
3. การเรียนรู้สำหรับโครงข่ายประสาทเทียม

1) การเรียนรู้แบบมีการสอน(Supervised Learning)เป็นการเรียนแบบมีการตรวจคำตอบเพื่อให้โครงข่ายปรับตัว ชุดวงจรที่ใช้สอนโครงข่ายจะมีคำตอบไว้คอยตรวจดูว่าโครงข่ายให้คำตอบที่ถูกต้องหรือไม่ ถ้าตอบไม่ถูก โครงข่ายก็จะปรับตัวเองเพื่อให้ได้คำตอบที่ดีขึ้นเปรียบเทียบกับคน เหมือนกับการสอนนักเรียนโดยมีครูผู้สอนคอยแนะนำ ดังภาพที่ 2.9



ภาพที่ 2.9 Supervised Learning

2) การเรียนแบบไม่มีการสอน (Unsupervised Learning) เป็นการเรียนแบบไม่มีผู้แนะนำ ไม่มีการตรวจคำตอบว่าถูกหรือผิด โครงข่ายจะจัดเรียงโครงสร้างด้วยตัวเองตามลักษณะของข้อมูล ผลลัพธ์ที่ได้โครงข่ายจะสามารถจัดหมวดหมู่ของข้อมูลได้ เปรียบเทียบกับคนเช่นเราสามารถแยกแยะพันธุ์พืชหรือพันธุ์สัตว์ ตามลักษณะรูปร่างของมันเองได้โดยที่ไม่ต้องมีใครสอน ดังภาพที่ 2.10



ภาพที่ 2.10 Unsupervised Learning

3) การเรียนรู้แบบถูกบังคับ (Reinforcement Learning) เป็นการเรียนรู้อีกวิธีหนึ่งเป็นลักษณะแม้ว่ามีครูคอยกำกับการทำงานอยู่ก็ตาม แต่คำตอบที่ถูกต้องหรือเป้าหมายไม่ได้ถูกนำมาแสดงในโครงข่าย ดังนั้นผลการคำนวณจากโครงข่ายจึงแสดงออกมาในรูปถูกหรือผิด โครงข่ายจึงต้องใช้ประโยชน์จากข้อมูลต่าง ๆ เพื่อนำไปปรับปรุงการทำงานให้ดีขึ้น ถ้าผลการคำนวณถูกต้องโครงข่ายจะได้รับรางวัล (Reward) ซึ่งเป็นการเพิ่มค่าถ่วงน้ำหนักของข้อมูลป้อนเข้าบางหน่วย ในทางตรงข้ามกันถ้าผลการคำนวณออกมาผิด โครงข่ายก็จะได้รับการลงโทษ (Penalty) โดยลดค่าถ่วงน้ำหนักลง

2.4 การแปลภาษาไทย-จีน

2.4.1 บริษัทและบริการที่มี

โดย Baidu นั้นก็ขยายกิจการเข้ามาในประเทศไทยแล้วโดยได้เน้นไปทางด้าซอฟต์แวร์และเว็บแอพพลิเคชั่นเป็นหลัก เช่น Baidu Antivirus, Baidu Spark, Hao123, Baidu IME, Simeji, Thailand Download, Baidu UEC แต่ก็มีบางแอพพลิเคชั่นและบางบริการ ที่ถูกพูดถึงอย่างหนักและเป็นกระแสในวงกว้าง เช่น Baidu PC Faster, Hao123, Baidu Spark, Baidu Translate.

ตัวอย่างโปรแกรมของ Baidu- Baidu Translate

ภาพที่ 2.11 Baidu Translate

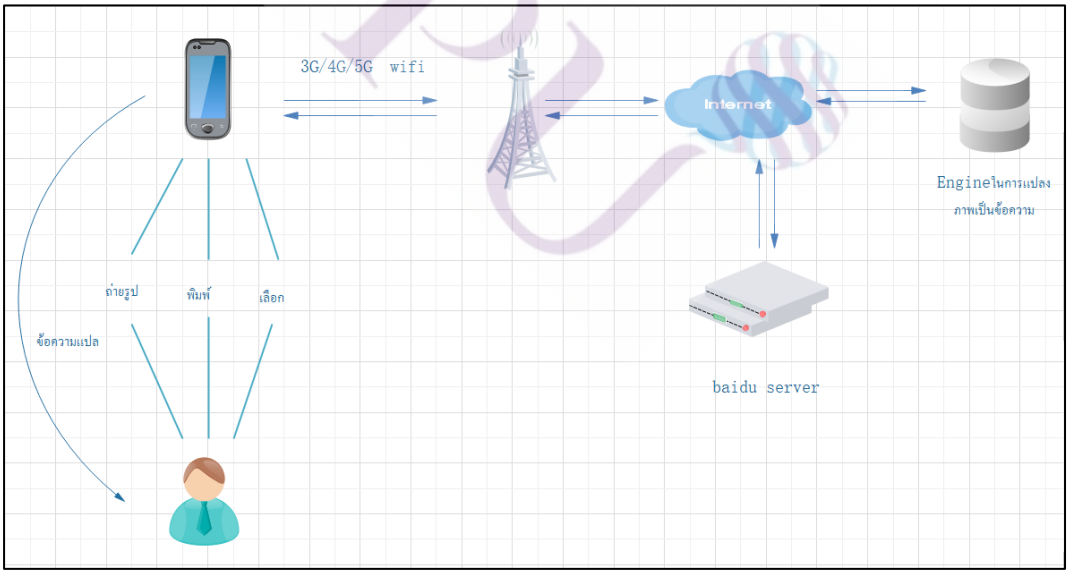
Baidu Translate เป็นหนึ่งการให้บริการของ Baidu ที่ใช้งานได้ดี และดูเหมือนว่าจะไม่มีโปรแกรมใดๆ แอบแฝงมาด้วย ข้อดีของแอพพลิเคชั่นตัวนี้คือ มันสามารถแปลภาษาได้ถูกต้องมากๆ

2.4.2 เทคนิคที่ใช้

Baidu มีการพัฒนาการแปลโดยนำ Neural Machine Translation (NMT) มาช่วยให้การแปลภาษาต่างๆ ทำได้ดีขึ้น จากเดิมที่ baidu Translateปลประโยคหรือข้อความภาษาไทยให้เข้าใจบ้าง ไม่เข้าใจบ้าง เรียงลำดับคำไม่รู้เรื่องบ้าง เนื่องจาก baidu คิดค้นการแปลภาษาด้วยเครื่องที่ใช้โมเดลทางสถิติจำนวนมหาศาลเพื่อระบุรูปแบบเนื้อหาของข้อมูลนับร้อยล้านแบบในเว็บ ซึ่งเมื่อ baidu Translate เปิดตัว Neural Machine Translation (NMT) ในหลายภาษากับจีน เช่น ภาษาไทย อังกฤษ เครื่องมือนี้จึงเข้ามาช่วยแปลและสื่อสารภาษาในรูปแบบที่เป็นธรรมชาติมากขึ้น โดยจะแปลประโยคทั้งประโยคในครั้งเดียว แทนที่จะแปลทีละส่วน ซึ่งจะใช้บริบทที่กว้างขึ้นเพื่อช่วยให้แปลได้ตรงกับความหมายที่สุด จากนั้นระบบก็จะจัดเรียงและปรับให้ตรงหรือใกล้เคียงกับภาษาพูดของมนุษย์มากที่สุด การแปลข้อความหลายๆ ย่อหน้าหรือบทความก็จะอ่านแล้วเข้าใจได้ง่ายขึ้น

2.5 Mobile application สำหรับการแปลภาษา

1) Mobile application เป็นการพัฒนาโปรแกรมที่ใช้บนโทรศัพท์มือถือและแท็บเล็ต จะช่วยตอบสนองความต้องการของผู้บริโภค อีกทั้งยังสนับสนุน ให้ผู้ใช้โทรศัพท์ได้ใช้ง่ายยิ่งขึ้น ซึ่งในปัจจุบันโทรศัพท์มือถือ หรือสมาร์ทโฟน มีหลายระบบปฏิบัติการที่พัฒนาออกมาให้ผู้บริโภคใช้ ส่วนที่มีคนใช้และเป็นที่นิยมมากก็คือ iOS และ Android จึงทำให้เกิดการเขียนหรือพัฒนา Application ลงบนสมาร์ทโฟนเป็นจำนวนมาก เช่น แผนที่, เกมส์, โปรแกรมคุยต่างๆ รวมถึงโปรแกรมสำหรับแปลภาษา และหลายธุรกิจก็เข้าไปเน้นในการพัฒนา Mobile Application เพื่อเพิ่มช่องทางในการสื่อสารกับลูกค้ามากขึ้น ดังนั้นอาจสรุปได้ว่า Mobile Application เป็นการพัฒนาโปรแกรมบนโทรศัพท์มือถือและ แท็บเล็ตหรือสมาร์ทโฟน เพื่อตอบสนองความต้องการของผู้ใช้เพื่อให้เข้าถึงการใช้งาน Application ที่ต้องการใช้งานได้อย่างสะดวก ในการแปลภาษาก็เช่นเดียวกัน โดยงานวิจัยครั้งนี้ผู้วิจัยได้พัฒนา Application สำหรับแปลภาษาไทย เป็นภาษาจีน เพื่ออำนวยความสะดวกให้นักท่องเที่ยวชาวจีนที่เดินทางเข้ามาท่องเที่ยวในประเทศไทยได้มี Application สำหรับการแปลภาษาไทยเป็นภาษาจีนจากป้าย และฉลากผลิตภัณฑ์ต่างๆ ซึ่งจะทำให้นักท่องเที่ยวชาวจีนรู้สึกมั่นใจในการเดินทางมากขึ้นดังภาพที่ 2.12



ภาพที่ 2.12 กระบวนการการแปลภาษา

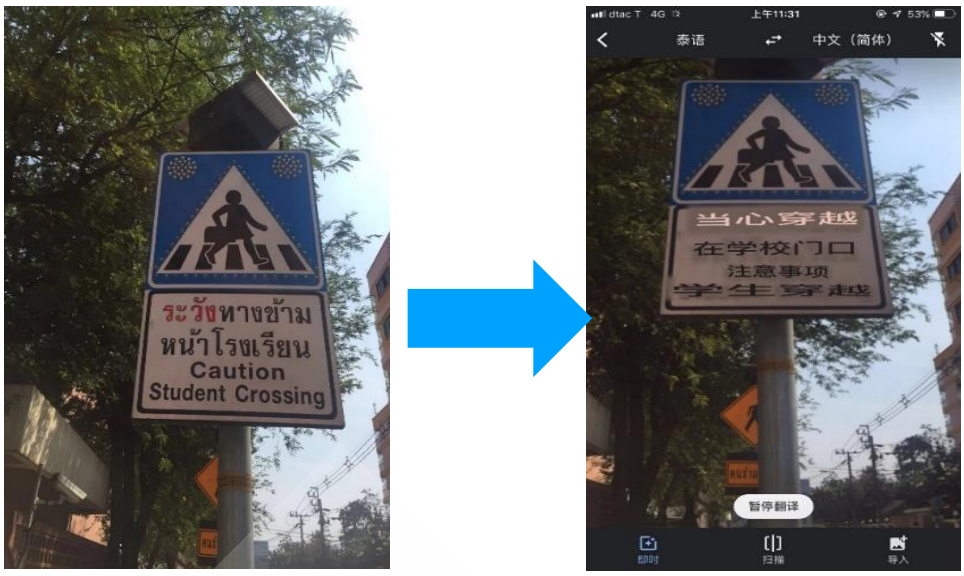
2) กระบวนการที่ผู้ใช้ป้อนอินพุทเข้าระบบ เช่นพูด พิมพ์ ถ่ายภาพ

1. การใช้งานเปิดเข้าแอป Google Translate ในแอพจะมีฟังก์ชั่นให้เลือกใช้งาน ทั้งการใช้กล้องสแกน การพิมพ์ การพูด การแปลโดยการใส่ Text เข้าไป รวมทั้งการใช้เสียง ดังภาพที่ 2.13

(ก) ภาพจาก Google Translate (ข)ภาพจาก Baidu Translate

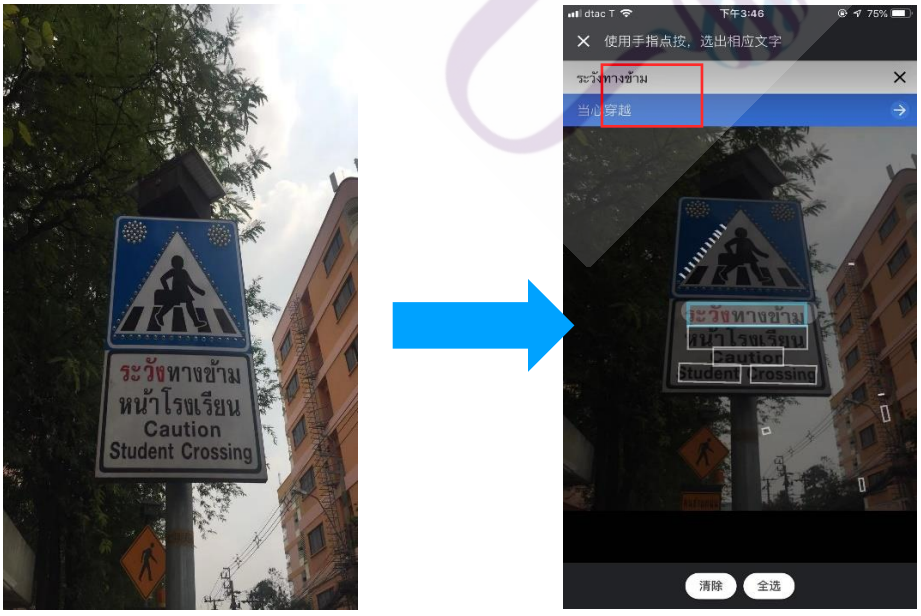
ภาพที่ 2.13 แอพมีฟังก์ชั่นให้เลือกใช้งานตัวอย่างโดยใช้ Google Translate and Baidu Translate

เมื่อกดเข้าใช้งานในรูปแบบกล้องและส่องไปยัง Text ระบบก็จะแปลให้แบบเรียลไทม์ ดังภาพที่ 2.14



ภาพที่ 2.14 กล้องสแกนตัวอย่างการแปลป้ายภาษาไทยเป็นภาษาจีน โดยใช้ Google translate ในโหมดถ่ายภาพ

ถ้าต้องการจะแปลข้อความแบบยาวก็กดไปที่เครื่องหมาย Scan แล้วใช้นิ้วลากเพื่อไฮไลท์ว่าอยากจะแปลตรงไหนบ้างระบบก็จะแปลให้ทันที ดังภาพที่2.15



ภาพที่ 2.15 กล้องสแกนตัวอย่างการแปลป้ายภาษาไทยเป็นภาษาจีน โดยใช้ Google translate ในนิ้วลาก

2. การใช้งานเปิดเข้าแอป baidu Translate ในแอพจะมีฟังก์ชั่นให้เลือกใช้งานทั้งการใช้กล้องสแกน การหาคำ การพูดและอื่นๆ ดังภาพที่ 2.16

ภาพที่ 2.16 แอพมีฟังก์ชั่นให้เลือกใช้งานตัวอย่างโดยใช้ Baidu Translate

การหาคำ ดังภาพที่ 2.17

Caution แปลเป็นภาษาจีนมีสองประเภทคือคำนามและคำกริยาแต่ความหมายเหมือนกัน ด้านล่างขวามี see more เข้าไปก็มีรายละเอียดในการการแปล

ภาพที่ 2.17 กล้องการหาคำตัวอย่างการแปลป้ายภาษาอังกฤษเป็นภาษาจีนโดยใช้ Baidu Translate ในโหมดถ่ายภาพ

การใช้กล้องสแกน แต่ไม่มีภาษาไทย ดังภาพที่ 2.18

ภาพที่ 2.18 กล้องสแกนตัวอย่างการแปลป้ายภาษาอังกฤษเป็นภาษาจีนโดยใช้ Baidu Translate ในโหมดถ่ายภาพ

การพูดแต่ไม่มีภาษาไทย ดังภาพที่ 2.19 และอื่นๆแต่ไม่มีภาษาไทย ดังภาพที่ 2.20

ภาพที่ 2.19 การพูดตัวอย่างโดยใช้ Baidu Translate **ภาพที่ 2.20** Baidu Translate อื่นๆ

3. Baidu Translate and Google Translate เปรียบเทียบการทำงาน

ตารางที่ 2.1 Baidu Translate and Google Translate เปรียบเทียบการทำงาน

ประเภท	Baidu แปล	Google แปล
ความจุคำ	รองรับ 28 ภาษา	รองรับ 103 ภาษา
การรูปภาพอักษร	ไม่มีแต่มีขอความ	มี
ใช้	การแปล Baidu ยังบ่งชี้ว่าคำนี้อยู่ในหมวดหมู่คำ เช่น CET4 ดังนั้นเหมาะสำหรับนักเรียนจีนและตลาดจีน	เหมาะสำหรับการการท่องเที่ยว

2.6 MVC

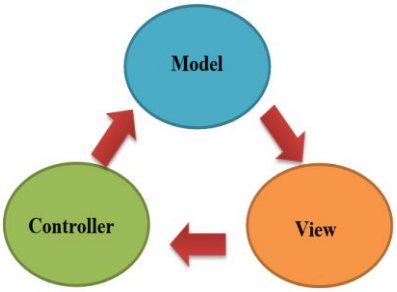
2.6.1 สถาปัตยกรรมของ MVC

Model-View-Controller (MVC) มี 3 ส่วนหลักๆ ได้แก่ model -view-controller

(1)Model (M) คือส่วนของ Business Logic และ ส่วนของ Entity ซึ่งเป็นส่วนทำหน้าที่ติดต่อกับฐานข้อมูล (Database)

(2) View (V) คือส่วนที่ใช้แสดงผล ซึ่งเป็นส่วนของ HTML, CSS และ JavaScript โดยทำเป็น Web Application ที่ใช้แสดงผล

(3) Controller (C) คือส่วนสมองที่ทำ หน้าที่ควบคุมการทำ งานของหน้าจอ และการดึงข้อมูลจากส่วนของ Model (M) มาแสดงผลในส่วนของ View (V)



ภาพที่ 2.21 สถาปัตยกรรมของ MVC

Android clientใช้สถาปัตยกรรมของระบบ MVC ส่วนของ View layer Controller layer ทำหน้าที่การรวบรวมและประมวลผลข้อมูล ส่วนของ Model layer ทำหน้าที่การเก็บรวบรวมข้อมูล ภาพที่ได้รับจาก View layer และส่งไปที่ Controller layer เพื่อประมวลผลเบื้องต้นแล้วส่งผ่านnetworkให้ Baidu เพื่อแปล และผลตอบรับผ่านnetworkถึง Controller layer Controller layerจะส่งข้อมูลตามข้อมูลModel layer ในที่สุดก็แสดงให้ผู้ใช้งานผ่าน View layer

2.6.2 ส่วนประกอบของแอพพลิเคชัน (Application Component)

คุณลักษณะอย่างหนึ่งของ Android คือเป็นแอพพลิเคชั่นที่สามารถใช้เป็นส่วนประกอบของแอพพลิเคชั่นอื่นๆได้ตัวอย่างเช่น ถ้าต้องการสร้างแอพพลิเคชั่นให้แสดงการเลื่อนของรายการรูปภาพ โดยที่อาจมีแอพพลิเคชั่นส่วนอื่นที่ได้พัฒนาไว้แล้ว สามารถเรียกใช้แอพพลิเคชั่นในส่วนที่มีอยู่มาพัฒนาต่อได้โดยที่ไม่จำเป็นต้องพัฒนาขึ้นเอง เพื่อความสะดวกและรวดเร็วในการพัฒนาแอพพลิเคชั่นซึ่งเรียกแอพพลิเคชั่นเหล่านี้ว่า Application Component ซึ่งสามารถแบ่งออกเป็น 4 ส่วนคือ Activity ,Service, Content Provider และ Broadcast and Intent Receiver

Activity คือหน้าจอที่ติดต่อกับผู้ใช้ (Graphic User - Interface) ทั้งนี้ในแต่ละแอพพลิเคชั่นอาจมีได้มากกว่า 1 หน้าจอ หรือ Activity เดียวซึ่งแต่ละ Activity จะทำหน้าที่เก็บสถานะการใช้งานส่วนต่าง ๆ ตัวอย่าง ในการแสดงรายการเมนู สามารถเลือกให้รายการเมนูที่แสดงออก กมามีภาพและคำบรรยายใต้ภาพได้สำหรับแอพพลิเคชั่นส่งข้อความอาจมี Activity หนึ่งจะเป็นส่วนของการเลือกการติดต่อ และ Activity อื่นๆจะทำหน้าที่ดูข้อความเก่าที่ถูกส่งมาแล้ว เป็นต้น

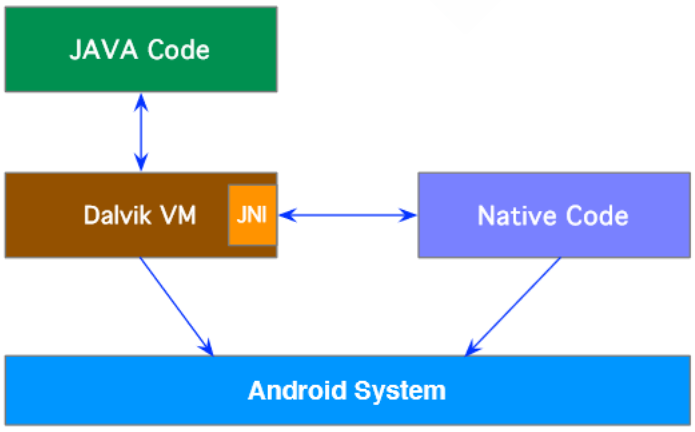
Service Provider เป็นส่วนที่ไม่มีการแสดงผล แต่ถูกเรียกใช้ให้รันอยู่ในลักษณะของback ground process โดย Service นั้นอาจจะมีการกระทำอะไรบางอย่าง เช่นติดต่อรับส่งข้อมูลผ่านเครือข่าย หรือคำนวณค่าต่าง ๆแล้วทำการส่งข้อมูลไปแสดงยัง Activity ก็ได้หรือการเปิดเพลงในขณะที่เรากำลังทำงานบนแอพพลิเคชั่นอื่น

Broadcast receiver และ Intent-Receiver หรือ Data Provider คือ ตัวที่ใช้สำหรับคอยรับและตอบสนองต่อเหตุการณ์ต่างๆที่เกิดขึ้น เช่น เมื่อแบตเตอรี่ต่ำ การเปลี่ยนภาษา มีการโทรออกมีข้อความเข้าและอื่นๆ ถึงแม้ Broadcast receiver จะไม่มีส่วนของการแสดงผลแต่สามารถที่จะเรียก Activity ขึ้นมาแสดงผลให้ผู้ใช้ทราบได้ เรียกว่า โนติฟิเคชันเมเนเจอร์ (Notification Manager) ซึ่งจะเป็นตัวที่แจงเตือนในรูปแบบต่าง ๆ เช่น การสั่น การแสดงไฟกระพริบที่หน้าจอ หรือการส่งเสียงออกมาโดยจะมี Icon แสดงอยู่บน Status Bar เพื่อแจ้งเหตุการณ์ที่เกิดขึ้น นอกจากนี้ Intent-Receiver คือเป็นส่วนทำให้แอพพลิเคชั่นอื่น ๆ เข้าถึงการทำงานของ Activity และ Service ซึ่งในการปฏิบัติงานแต่ละอย่างเป็นการตอบสนองการร้องขอจากข้อมูลหรือบริการของ Activity อื่น ๆ

Content provider (System Event Listener) คือส่วนของการบริการข้อมูลสำหรับแต่ละแอพพลิเคชั่นทั้งนี้ข้อมูลสามารถเก็บอยู่ในรูปแบบของระบบไฟล์ หรือฐานข้อมูล เช่น Google สามารถเข้าใช้งานข้อมูลร่วมกันกับผู้ใช้งานได้ในแอพพลิเคชั่นที่ต้องการข้อมูลของผู้ใช้งานโดยปกติแล้วในแต่ละแอพพลิเคชั่นจะทำงานแยกกันในแต่ละ Process ซึ่งมีการทำงานที่แยกจากกันโดยชัดเจนในกรณีที่นักพัฒนาต้องการเข้าถึงการทำงาน หรือขอสิทธิ์เข้าใช้นั้นนักพัฒนาจะต้องกำหนดการขอเข้าใช้ที่ไฟล์ AndroidManifest.xml

2.7 พื้นฐานการพัฒนา JAVA NATIVE INTERFACE (JNI)

Java Native Interface หรือเรียกย่อๆว่า JNI เป็นส่วนหนึ่งที่อยู่ในชุดเครื่องมือพัฒนา JDK (Java Software Development Kit) ซึ่ง JNI นี้จะทำหน้าที่เป็นส่วนเชื่อมต่อระหว่างโปรแกรม native (C/ C++) และโปรแกรม JAVA ทำให้โปรแกรม JAVA สามารถเรียกโค้ดโปรแกรม หรือฟังก์ชันของโปรแกรมที่ ถูกพัฒนาด้วยโปรแกรมภาษา C/C++ หรือแม้แต่ภาษาอื่นๆได้ นอกจากนั้นโปรแกรมภาษาอื่นก็ยัง สามารถเรียกฟังก์ชันภายในโปรแกรมภาษาจาวาได้เช่นกัน โดยเรียกผ่าน Invocation API ที่อยู่ภายใน JVM (Java virtual machine) สำหรับระบบปฏิบัติการแอนดรอยด์โค้ดโปรแกรม JAVA (JAVA Code) ก็จะถูกรันบน Dalvik ซึ่ง เป็นเครื่องเสมือน (Virtual Machine) คล้ายกับ (แต่ไม่เหมือน) JVM ของบริษัท Oracle แต่ตัว Dalvik นั้นได้ถูกพัฒนาใหม่ทั้งหมดโดยนักพัฒนาของกูเกิลเอง ในขณะที่โค้ดโปรแกรม native (Native Code) จะถูกคอมไพล์เป็นโค้ดไบนารี ซึ่งพร้อมถูกเรียกได้ทันทีในระดับระบบปฏิบัติการ ดังนั้น JNI จึงทำตัวเอง เหมือนสะพานเชื่อมกันระหว่างโลกสองโลกนี้ ดังรูปแสดงความสัมพันธ์ระหว่าง JAVA code, Dalvik VM, native code, และ ระบบปฏิบัติการแอนดรอยด์ ข้างล่างนี้



ภาพที่ 2.22 แสดงการสื่อสารระหว่างโปรแกรม Java และ Native C

จากรูปข้างบนทั้ง Dalvik VM และ Native Code ต่างก็ทำงานอยู่บนระบบปฏิบัติการแอนดรอยด์ ดังนั้น ระบบปฏิบัติการจะต้องเตรียมสภาพแวดล้อมสำหรับการเรียกคำสั่งระหว่างกัน (execution environment) และจะเห็นว่า JNI เป็นส่วนหนึ่งที่อยู่ใน Dalvik VM ซึ่งจะอำนวยความสะดวกให้ Native Code ให้สามารถเข้าถึงตัวแปรภายในคลาส (Data fields) และสมาชิกฟังก์ชันภายในคลาส (Member functions) ได้จากการกำหนดสิทธิ์การเข้าถึง (Access Permissions) คลาสภายในโค้ด JAVA

ดังนั้นการแลกเปลี่ยนข้อมูลระหว่างโปรแกรม JAVA และโปรแกรม native (C/C++) นั้นจึงเป็น ประเด็นสำคัญที่นักพัฒนาควรทำความเข้าใจ ซึ่งใน JNI นั้นก็ได้มีการกำหนดตัวแปรที่สามารถรองรับ ตัวแปรมาตราฐานของแต่ละภาษา

2.8 สรุป

บทนี้แนะนำเกี่ยวกับเทคโนโลยีสำคัญที่ใช้ในวิทยานิพนธ์นี้ ซึ่งรวมถึงAndroid ส่วนประกอบของแอพพลิเคชัน MVC JNI แปลภาษา และเทคนิคการแปลภาษาจากภาพ โครงข่ายประสาทเทียม เป็นต้น

มนุษย์มีการติดต่อสื่อสารกันผ่านการฟัง พูด อ่าน เขียน แต่ภาษาที่ใช้ ย่อมแตกต่างกันไป ตามแต่ละเชื้อชาติ ทำให้ต้องมีการแปลข้ามภาษา การแปลดั้งเดิมจะเป็นการแปลด้วยล่ามแต่ในปัจจุบันมีการใช้คอมพิวเตอร์ช่วยแปลแบบอัตโนมัติโดยใช้เทคนิคการประมวลผลภาพและโครงข่ายประสาทเทียมขณะนี้

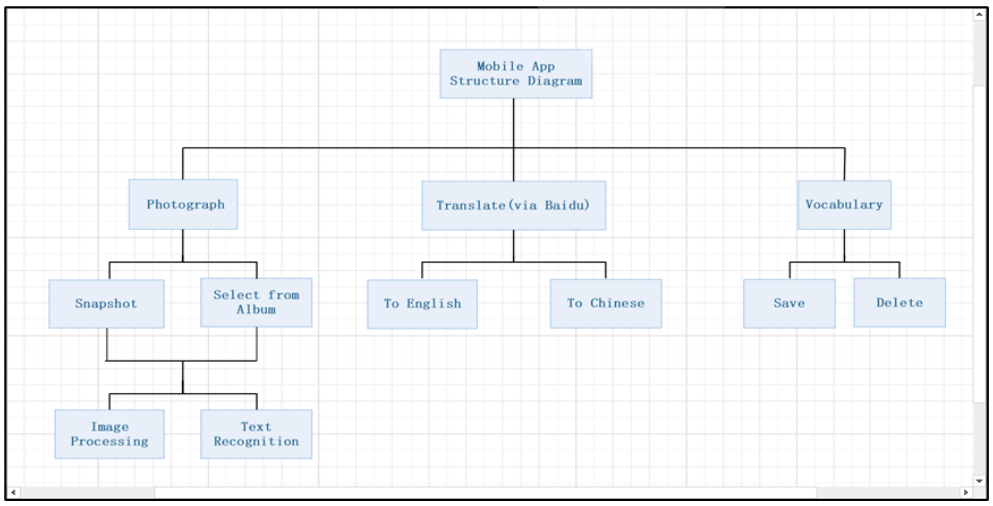
ปัจจุบันนี้ Google translation มีภาษาในการแปลประมาณ 103ชนิด เหมาะสำหรับการท่องเที่ยว Baidu translation มีภาษาในการแปลประมาณ 28ชนิด ยังไม่มีการรู้จำตัวอักษรไทยแต่เหมาะสำหรับคนจีนเรียนภาษาต่างประเทศ

บทที่ 3
การออกแบบและพัฒนา

เนื้อหาในบทนี้ จะกล่าวถึง Mobile app structure ซึ่งจะประกอบด้วย 3 ส่วนดังต่อไปนี้ ส่วนแรก คือ Photo Input ส่วนที่สอง คือ Translate (via baidu)และส่วนสุดท้ายเป็น Vocabulary

3.1 การออกแบบระบบแอพมือถือ

ระบบแอพมือถือประกอบด้วย 3ส่วน มี photograph, Translate and internet และ Vocabulary ในส่วน photograph หมายถึงภาพที่ถ่ายจากกล้องโทรศัพท์มือถือหรือเลือกภาพที่มีข้อความเป้าหมายจากอัลบั้มโทรศัพท์มือถือ การประมวลผลภาพหมายถึงการประมวลผลภาพเพื่อเพิ่มอัตราการรู้จำภาพตัวอย่างเช่นการกรองค่ามัธยฐาน การแปลงภาพระดับเทาให้เป็นภาพขาว-ดำ เป็นต้น การรู้จำตัวอักษรหมายถึงการใช้เทคนิคOCRเพื่อระบุเป้าหมายภาพและแยกข้อความออกมา ส่วนการแปลและระบบเครือข่ายหมายถึงใช้เทคนิค OCRได้ตัวอักษรและส่งข้อความไปถึง baidu เพื่อแปลภาษาไทยเป็นภาษาจีนหรือภาษาอังกฤษ ส่วนVocabulary หมายถึงคงไว้และลบข้อความ จึงมีภาพรวมของ Mobile app structure ดังภาพที่ 3.1



ภาพที่ 3.1 การออกแบบระบบแอพมือถือ

3.1.1 กระบวนการ Mobile app

การถ่ายภาพเป็นขั้นตอนแรกในกระบวนการทำงานของระบบถ่ายภาพเสร็จเรามีสองทางเลือก สิ่งแรกคือการใช้มือถือฟังก์ชั่นถ่ายภาพโดยตรงเพื่อให้ได้ภาพที่มีข้อความที่จะแปล ถ้าถ่ายภาพไม่ชัดเจนเราสามารถเลือกฟังก์ชั่นที่สองคือการอ่านภาพที่เก็บไว้ในมือถือ การประมวลผลภาพเป็นขั้นตอนที่สองในการทำงานของระบบ ขั้นตอนนี้ส่วนใหญ่เกี่ยวข้องกับการรู้จำตัวอักษร เช่นการแปลงค่าสี RGB เป็น Gray Scale การกรองค่ามัธยฐาน การแปลงภาพระดับเทาให้เป็นภาพขาว -ดำ ๆ การรู้จำตัวอักษรเป็นขั้นตอนที่สามในการทำงานของระบบ ขั้นตอนนี้ใช้ Tesseract ก่อนเพื่อฝึกอบรมการจดจำภาษาไทย หลังจากพื้นที่ข้อความที่ได้รับจากการประมวลผลภาพ ใช้ Tesseractขั้นตอนวิธีเพื่อการรับรู้ข้อความ การแปลเป็นขั้นตอนที่สี่ในการทำงานของระบบ ได้ข้อความแลวส่งไป Baidu แปลจะได้ภาษาจีนและภาษาอังกฤษ แต่ต้องต่อ network ดังนั้นระบบการรู้จำภาษาไทยสามารถสื่อสารกับ Baidu เพื่อแปล Vocabularyเป็นขั้นตอนที่ห้าในการทำงานของระบบ ขั้นตอนนี้ได้ข้อความที่จดจำสามารถคงไว้และลบได้

เนื่องจากโครงสร้างภาษาไทยมีความซับซ้อนมากกว่าและมีความคล้ายคลึงของตัวอักษรมากกว่า จึงต้องใช้อุปกรณ์ฮาร์ดแวร์พื้นฐานระดับสูงของโทรศัพท์มือถือ เนื่องจากพิกเซลของกล้องโทรศัพท์มือถือต่ำเกินไปจะทำให้ภาพที่ถ่ายไม่ชัดเจนซึ่งจะทำให้การจดจำข้อความในอนาคตมีลำบากมาก นอกจากนี้ระบบการรู้จำตัวอักษรต้องใช้หน่วยความจำมาก ยังการประมวลผลภาพต้องการการคำนวณและการประมวลผลข้อมูลจำนวนมาก จึงเสนอมาตรฐานฮาร์ดแวร์สำหรับการใช้โทรศัพท์มือถือ ดังตารางที่ 3.1

ตารางที่ 3.1 Mobile hardware standards

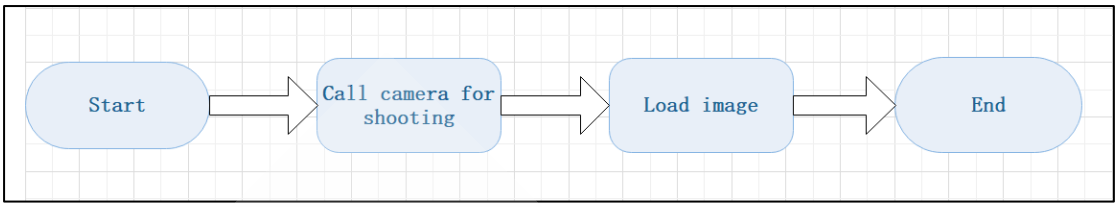
Operating system	Android
Minimum system version	9.0 or more
Camera pixel	8 million pixels
CPU	1GHz
RAM	2GB
Other	Higher versions of the Android system need dynamic authorization to keep the data network unblocked

3.2 การใช้งานการรับภาพ

การถ่ายภาพเป็นขั้นตอนแรกในกระบวนการทำงานของระบบ ภาพที่ชัดต่อประมวลผลภาพและการรับรู้ภาพสำคัญมาก จึงต่อไปนี้จะพูดถึงวิธีการรูปภาพ

3.2.1 ถ่ายภาพ

(1) กระบวนการถ่ายภาพดังภาพที่ 3.2



ภาพที่ 3.2 กระบวนการถ่ายภาพ

ที่มา: เมื่อผู้ใช้เข้าถึงMobile Clientและคลิกที่ปุ่มถ่ายภาพ View layer จะส่งคำสั่งให้ Controller layer เพื่อถ่ายภาพ Controller layerจะส่งคำสั่งให้Model layer เพื่อการประมวลผลภาพถัดไป

(2) กระบวนการถ่ายภาพของ code ดังภาพที่ 3.3

```
Uri PhotoUri;          //Photo Location 图片保存地址

//Click snapshot button 单击拍照按钮
public void getPicFromCamera() {

    //Check self permission 检查程序自身权限
    if (checkTokenStatus()) {
       infoPopText("Permission Denied!");
       return;
    }

    //Connect to Camera 准备启动摄像头
    Intent intent = new Intent(MainActivity.this, CameraActivity.class);
    intent.putExtra(CameraActivity.KEY_OUTPUT_FILE_PATH,
          FileUtil.getSaveFile(getApplication()).getAbsolutePath());
    intent.putExtra(CameraActivity.KEY_CONTENT_TYPE,
          CameraActivity.CONTENT_TYPE_GENERAL);
    Log.i(TAG, "Start Camera acitivity" );

    //Start启动摄像头
    startActivityForResult(intent, REQUEST_CODE_ACCURATE_BASIC);

  }


//After pick up photo 选择图片后
protected void onActivityResult(int requestCode, int resultCode, Intent data) {
    super.onActivityResult(requestCode, resultCode, data);

          //Succeed to pick up photo 成功选择图片
          if (resultCode == RESULT_OK){
             // Pick Up From Album 从相册中选择
             if ( requestCode == ALBUM_CODE_FOR_RESULT )
      {
        PhotoUri = null;
        if (data != null) {
           PhotoUri = data.getData();
        }
      // Pick up from snapshot通过拍照选择图片
      if else(requestCode == CAMERA_CODE_FOR_RESULT) {
                  RecognizeService.recAccurateBasic(this, FileUtil.getSaveFile(getApplicationContext()).getAbsolutePath(),new RecognizeService.ServiceListener() {
             @Override
             public void onResult(String result) {
                  OcrResultString = result;
               }
             });
          OCRPage.setText(OcrResultString);
          PhotoUri = getFileName();;
             }
```

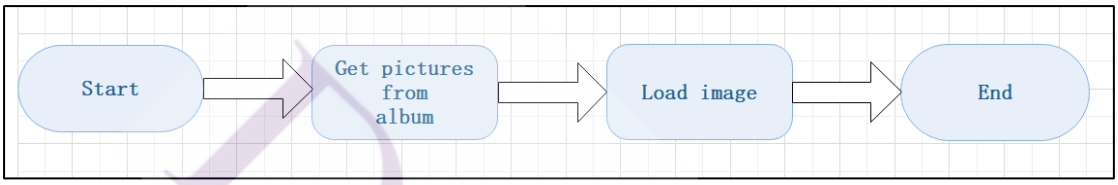
```
    } else { // Failed to pick up photo，go to the Translate Page选择照片失败，直接返回翻译界面
        PhotoUri = null;
        menuItem.setChecked(false);
        menuItem = bottomNavigationView.getMenu().getItem(TRANSLATE_PAGE);
        menuItem.setChecked(true);
        viewPager.setCurrentItem(TRANSLATE_PAGE);
    }

            //Go to OCR Page 跳转到OCR界面
    OCRPage.setImage(PhotoUri.toString());
}
```

ภาพที่ 3.3 code ของกระบวนการถ่ายภาพ

3.2.2 รับภาพจากอัลบั้มรูป

(1) กระบวนการรับภาพจากอัลบั้มรูป ดังภาพที่ 3.4



ภาพที่ 3.4 กระบวนการรับภาพจากอัลบั้มรูป

ที่มา: เมื่อผู้ใช้เข้าถึง Mobile Client และรับภาพจากอัลบั้มรูป View layer จะส่งคำสั่งให้ Controller layer เพื่อรับภาพ ผู้ใช้เลือกรูปภาพเสร็จแล้ว Controller layerจะส่งคำสั่งให้ Model layer เพื่อการประมวลผลภาพถัดไป

(2) กระบวนการรับภาพจากอัลบั้มรูปของ code ดังภาพที่ 3.5

```
Uri PhotoUri;           //Photo Location 图片保存地址

//Click SelectFromAlbum button 单击从相册中选择图片按钮
public void getPicFromAlbum() {
      Intent intent = new Intent();
      intent.setType("image/*");
      intent.setAction(Intent.ACTION_GET_CONTENT);
      startActivityForResult(intent, ALBUM_CODE_FOR_RESULT);
}

//After pick up photo 选择图片后
protected void onActivityResult(int requestCode, int resultCode, Intent data) {
      super.onActivityResult(requestCode, resultCode, data);

            //Succeed to pick up photo 成功选择图片
            if (resultCode == RESULT_OK){
               // Pick Up From Album 从相册中选择
               if ( requestCode == ALBUM_CODE_FOR_RESULT )
        {
          PhotoUri = null;
          if (data != null) {
             PhotoUri = data.getData();
          }
        // Pick up from snapshot通过拍照选择图片
        if else(requestCode == CAMERA_CODE_FOR_RESULT) {
                        RecognizeService.recAccurateBasic(this, FileUtil.getSaveFile(getApplicationContext()).getAbsolutePath(),new RecognizeService.ServiceListener() {
                  @Override
                  public void onResult(String result) {
                        OcrResultString = result;
                  }
                });
            OCRPage.setText(OcrResultString);
            PhotoUri = getFileName();;
               }
     } else {  // Failed to pick up photo，go to the Tranlsate Page选择照片失败，直接返回翻译界面
        PhotoUri = null;
        menuItem.setChecked(false);
        menuItem = bottomNavigationView.getMenu().getItem(TRANSLATE_PAGE);
        menuItem.setChecked(true);
        viewPager.setCurrentItem(TRANSLATE_PAGE);
     }

            //Go to OCR Page 跳转到OCR界面
     OCRPage.setImage(PhotoUri.toString());
  }
```

ภาพที่ 3.5 กระบวนการรับภาพจากอัลบั้มรูปของ code

3.2.3 ตัดภาพ

ในการรับภาพ เราต้องถ่ายรูปก่อนหรือรับภาพจากอัลบั้ม นอกจากนี้ยังต้องการที่จะเลือกขอบเขตของอักษรไทย เพราะในรูปมีข้อมูลมากมายที่ไม่ต้องแปล จึงต้องย้ายขนาดของกล่องตัดด้วยตนเอง ตัดเป็นภาษาไทยที่ต้องการแปลตัดภาพของ code ดังภาพที่ 3.6

```
//After snapshot,require to corp photo 拍照后，选定照片要识别的区域
private View.OnClickListener cropConfirmButtonListener = new View.OnClickListener() {
    @Override
    public void onClick(View v) {
      Rect rect;
            rect = overlayView.getFrameRect();
      Bitmap cropped = cropView.crop(rect);
      displayImageView.setImageBitmap(cropped);
      cropAndConfirm();
            doConfirmResult();
    }
};

//Save and pass the croped photo  保存并传递选定的照片识别区域
private void doConfirmResult() {

    CameraThreadPool.execute(new Runnable() {
      @Override
      public void run() {
        try {
          FileOutputStream fileOutputStream = new FileOutputStream(outputFile);
          Bitmap bitmap = ((BitmapDrawable) displayImageView.getDrawable()).getBitmap();
          bitmap.compress(Bitmap.CompressFormat.JPEG, 100, fileOutputStream);
          fileOutputStream.close();
        } catch (IOException e) {
          e.printStackTrace();
        }

                    // Pass the croped photo to OCR page 将选定的区域传递到OCR界面
        Intent intent = new Intent();
        intent.putExtra(CameraActivity.KEY_CONTENT_TYPE, contentType);
        setResult(Activity.RESULT_OK, intent);
        finish();
      }
    });
  }
```

ภาพที่ 3.6 code ของตัดภาพ

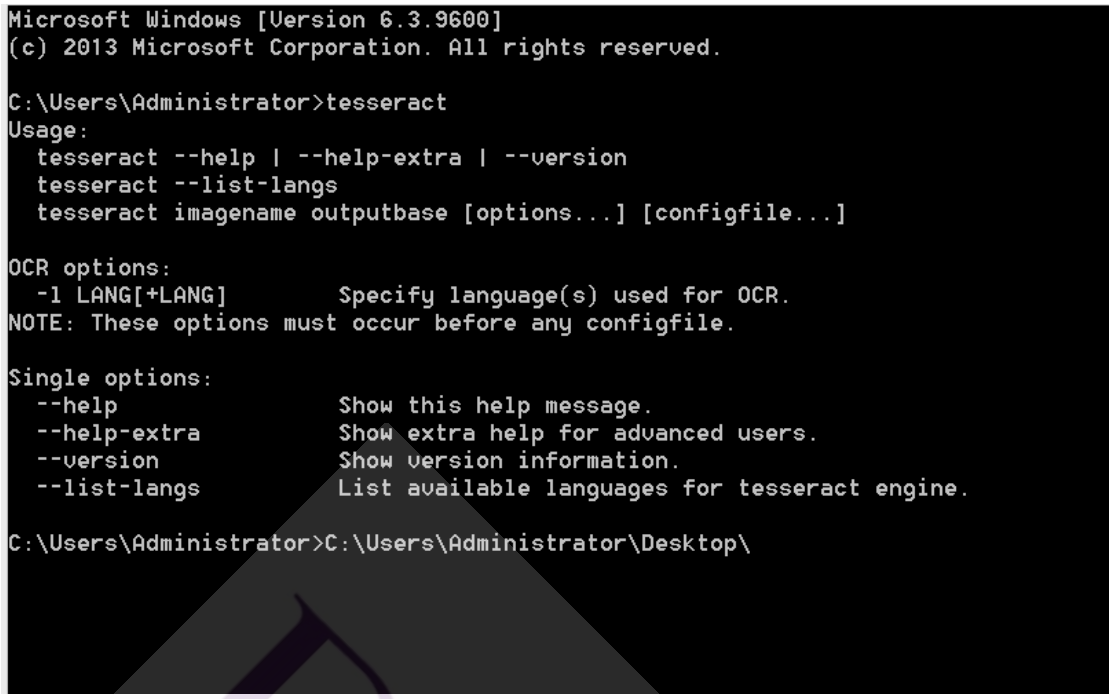
3.3 การรู้จำภาพ

3.3.1 การฝึกรู้จำอักษร

Tesseract OCR เป็น Engine ที่ใช้สำหรับการรู้จำอักขระทางภาพ พัฒนาขึ้นโดยบริษัท HP ระหว่างปี 1984-1985 โดยเริ่มต้นมาจากโปรเจควิจัยระดับปริญญาเอกในห้องปฏิบัติการHP โดยมีความตั้งใจเพื่อนำไปใช้กับงานเครื่องสแกนเนอร์เป็นหลัก ซึ่งในปี 2005 HP ก็ได้เผยแพร่ให้เป็น Open Source โดยมี google เป็นผู้สนับสนุน ซึ่งนักพัฒนาสามารถนำชุดคำสั่งนี้มาใช้งานได้ Tesseract นั้นถือว่าเป็นหนึ่งในOCR Open Sourceที่มีความแม่นยำสูงแต่ภาษาไทย font:DB-Bork ความแม่นยำไม่สูง ไม่มี Library OCR ซึ่งงานวิจัยนี้เสนอวิธีการสร้างฐานข้อมูลสำหรับการฝึกอบรมภาษาไทย

กระบวนการสร้างfont:DB-Bork Library OCR

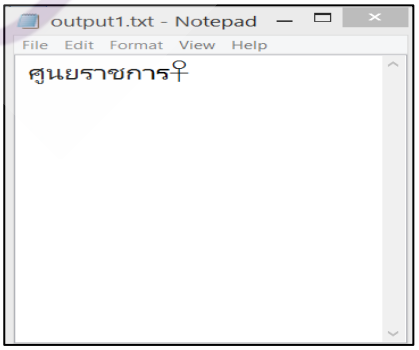
1) ติดตั้ง Tesseract-OCR เสร็จแล้วเปิด CMD inputคำสั่ง Tesseract ดังภาพที่ 3.7



ภาพที่ 3.7 ติดตั้ง Tesseract-OCR เสร็จ

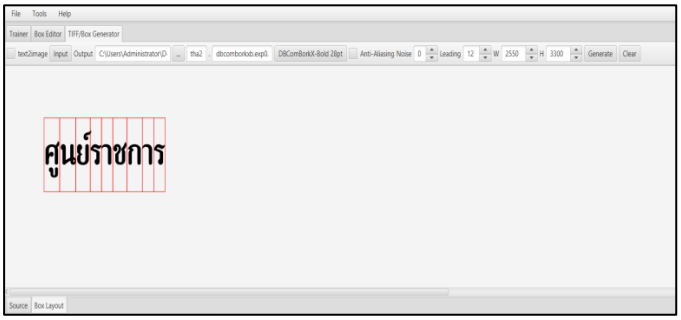
2) เมื่อเสร็จเรียบร้อยแล้วมาผลการทดสอบด้วยโมเดลจาก Tessdata_best โดยไม่ได้ปรับแต่งอะไรเพิ่ม แสดงผลก็คือความแม่นยำไม่สูงจึงงานวิจัยนี้จะเพิ่มความแม่นยำ

ภาพที่ 3.8 การรู้จำอักขระ

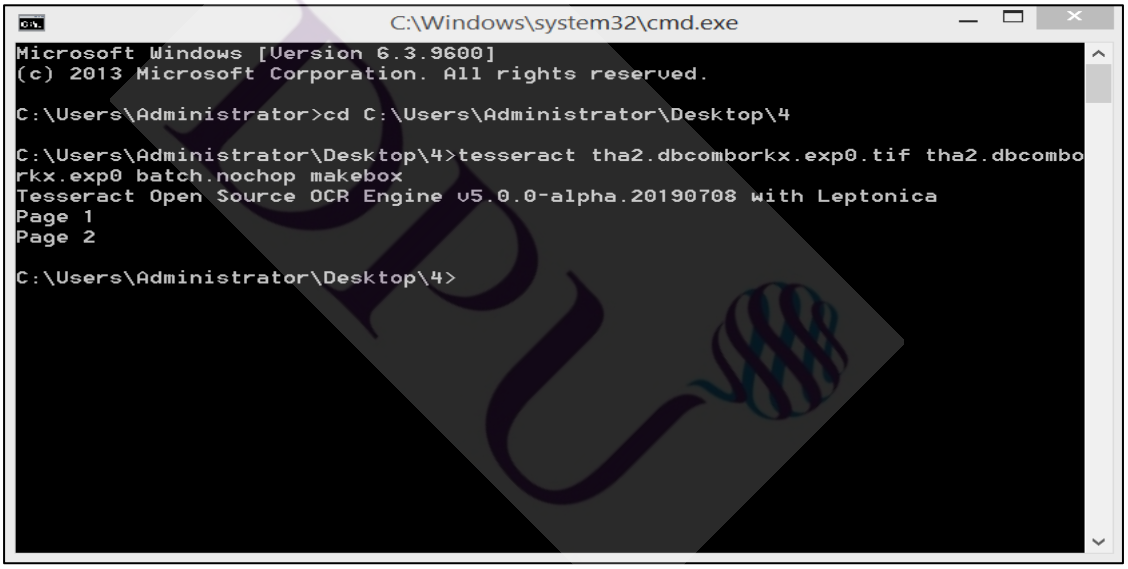


ภาพที่ 3.9 หลังการรู้จำอักขระ

3) สร้างไฟล์ tif และ box ดังภาพที่ 3.10-3.11

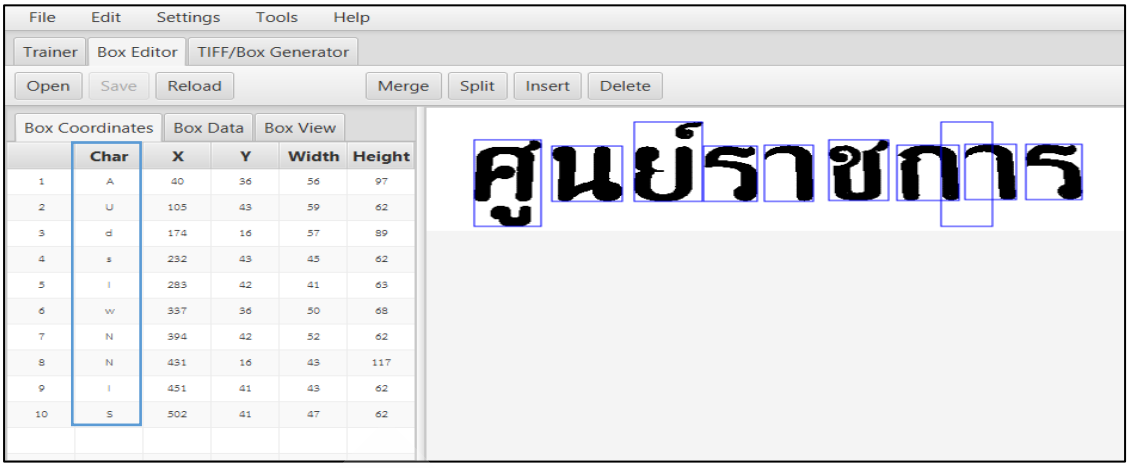


ภาพที่ 3.10 สร้างไฟล์ tif

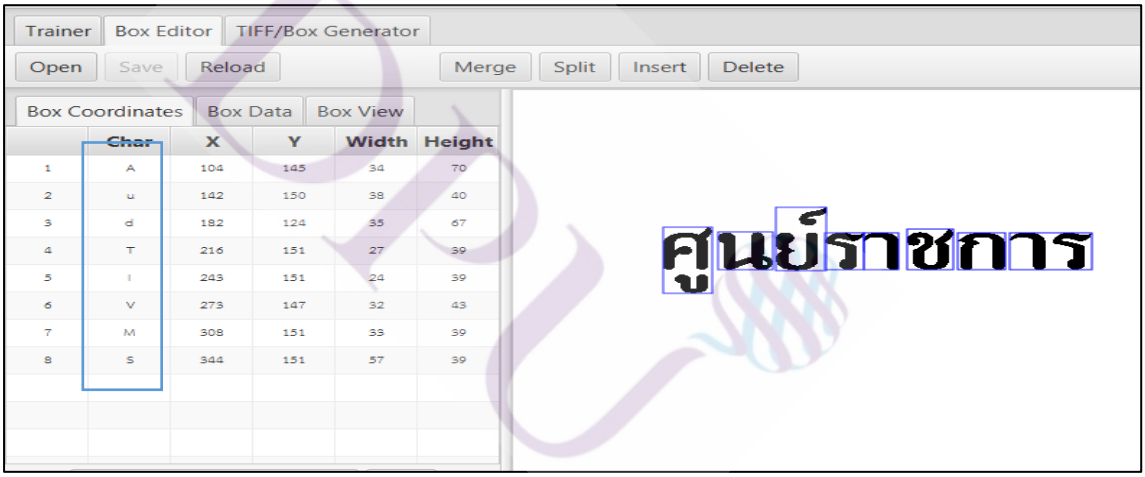


ภาพที่ 3.11 สร้าง box

4) การแก้ไขตัวละอักษร ด้านล่างเป็นภาพการรู้จำอักขระต้นฉบับ ในจอแสดงผลการรู้จำในคอลัมน์Charจะเห็นว่าผลการจดจำเป็นตัวอักษรภาษาอังกฤษซึ่งแตกต่างจากภาษาไทยทางด้านขวา ดังภาพที่ 3.12-3.13

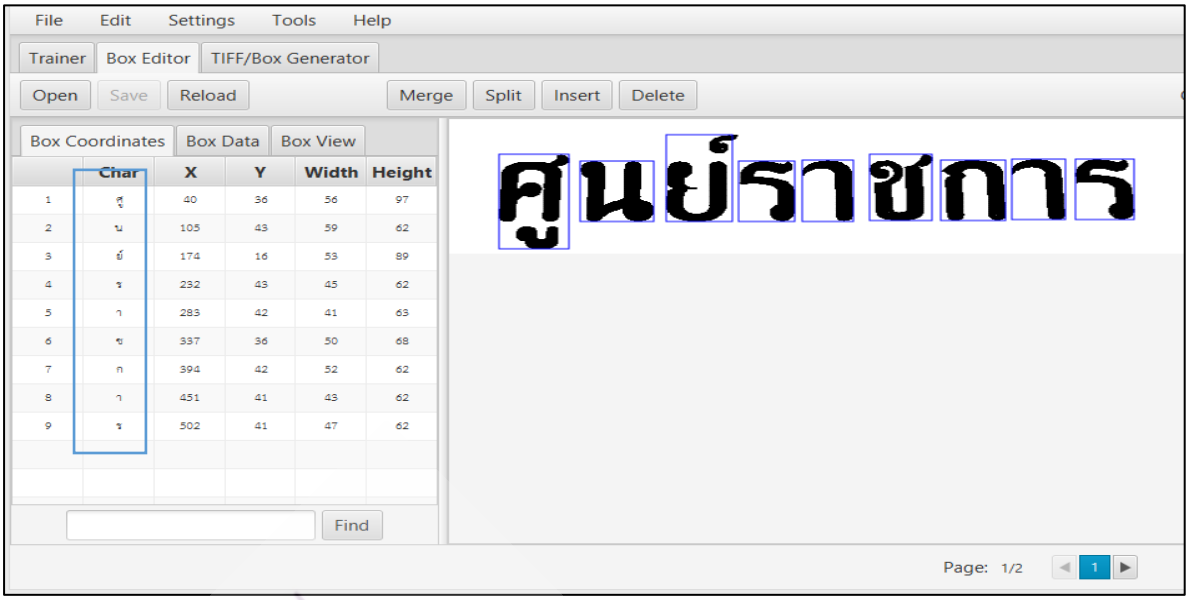


ภาพที่ 3.12 ภาพการรู้จำอักขระต้นฉบับจากป้าย

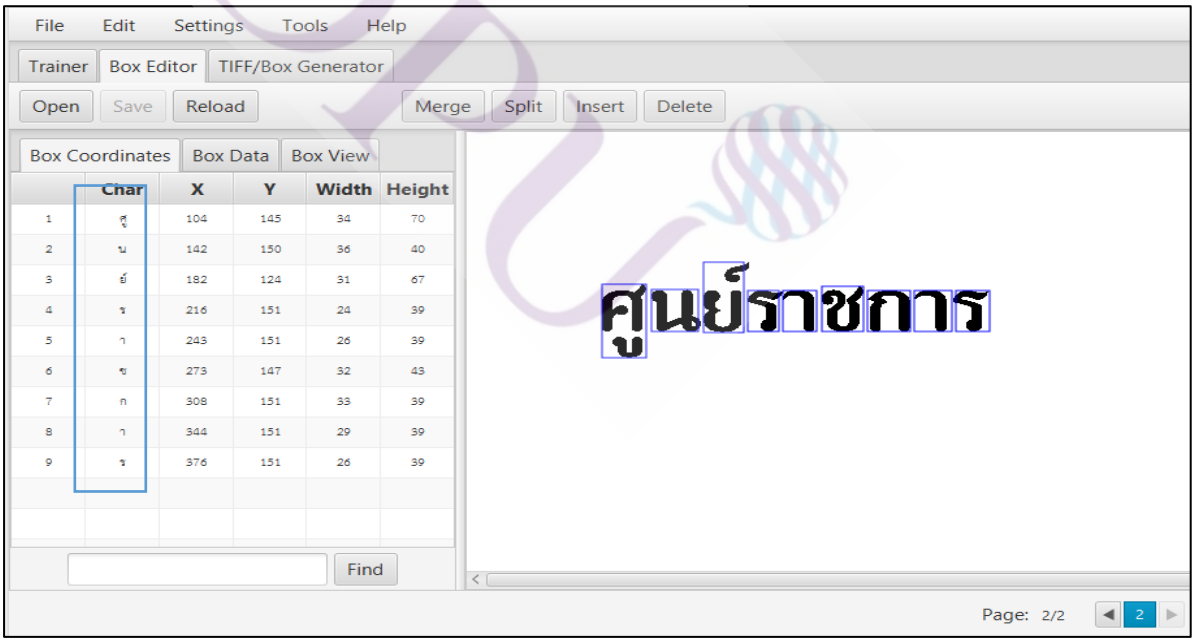


ภาพที่ 3.13 ภาพการรู้จำอักขระต้นฉบับจากการพิมพ์

ซึ่งข้อผิดพลาดในการรู้จำจะต้องแก้ไขโดยใช้มนุษย์ ในกระบวนการแก้ไขเราต้องแยกแต่ละอักษรภาษาไทยให้ถูกต้องด้วยตนเองและป้อนภาษาไทยที่ถูกต้องในคอลัมน์ Char ผลการแก้ไขดังภาพที่ 3.14-3.15

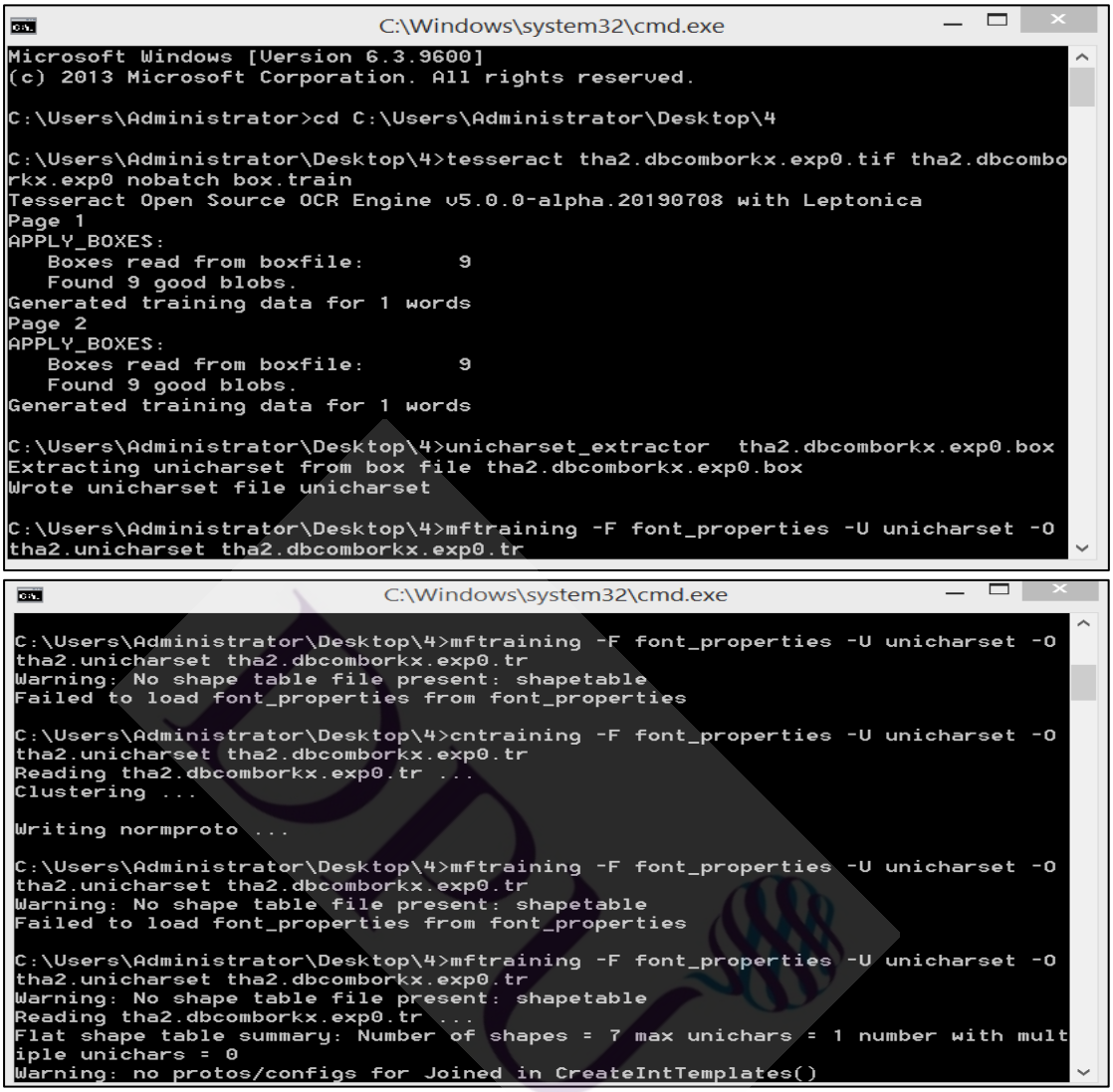


ภาพที่ 3.14 ภาพการรู้จำอักขระจากป้ายหลังการแก้ไข



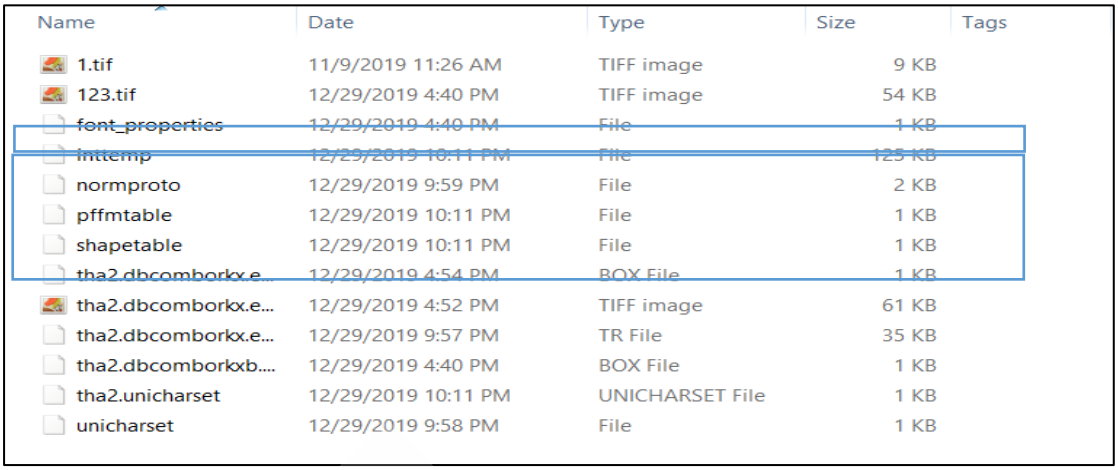
ภาพที่ 3.15 ภาพการรู้จำอักขระจากการพิมพ์หลังการแก้ไข

5) เริ่มการฝึกในขณะนี้ไฟล์ unicharset จะถูกสร้างขึ้นและไฟล์นี้จะใช้สำหรับการรวบรวมฟีเจอร์แบบอักษร ดังภาพที่ 3.16



ภาพที่ 3.16 เริ่มการฝึก unicharset

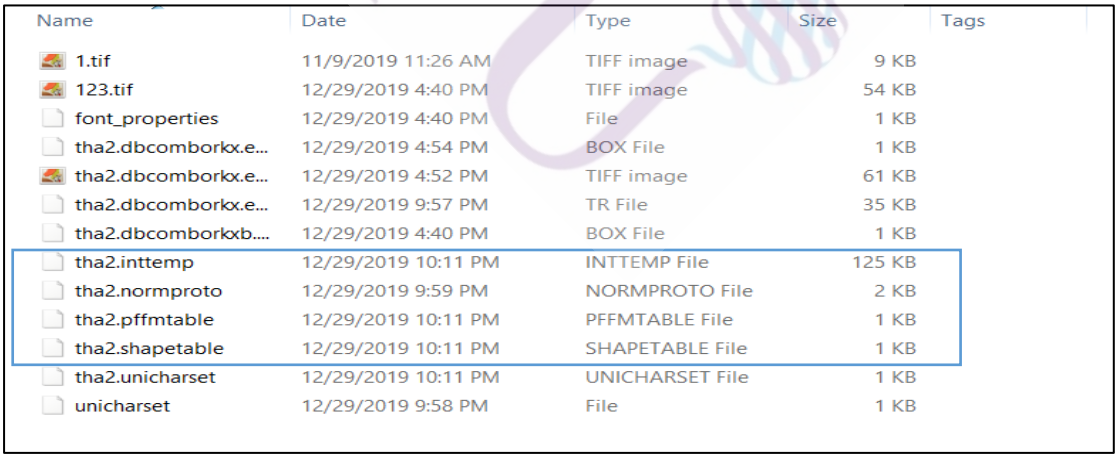
6) ทำFont characteristics กับ Aggregating font features :เนื่องจากมีฟอนต์ภาษาไทยจำนวนมากเราจึงกำหนดให้รู้จำภาษาไทยแบบฟอนต์มาตรฐานและรวม Font characteristics ดึ่นเข้าด้วยกัน จึงสร้างไฟล์ txt ก่อน ตั้งชื่อfont_properties แล้วใส่ normal 0 0 0 0ในไฟล์นี้ ใช้คำสั่ง mftrainig -F font_properties -U unicharset -O tha2.unicharset tha2.dbcomborkx.exp0.tr จะได้ไฟส์ unicharsetล่าสุดแล้ว ใช้คำสั่ง cntraining -F font_properties -U unicharset -O tha2.unicharset tha2.dbcomborkx.exp0.tr จะได้ไฟล์ shapetable ดังนั้นยังได้ไฟล์ inttemp pffmtable normproto ดังภาพที่ 3.17



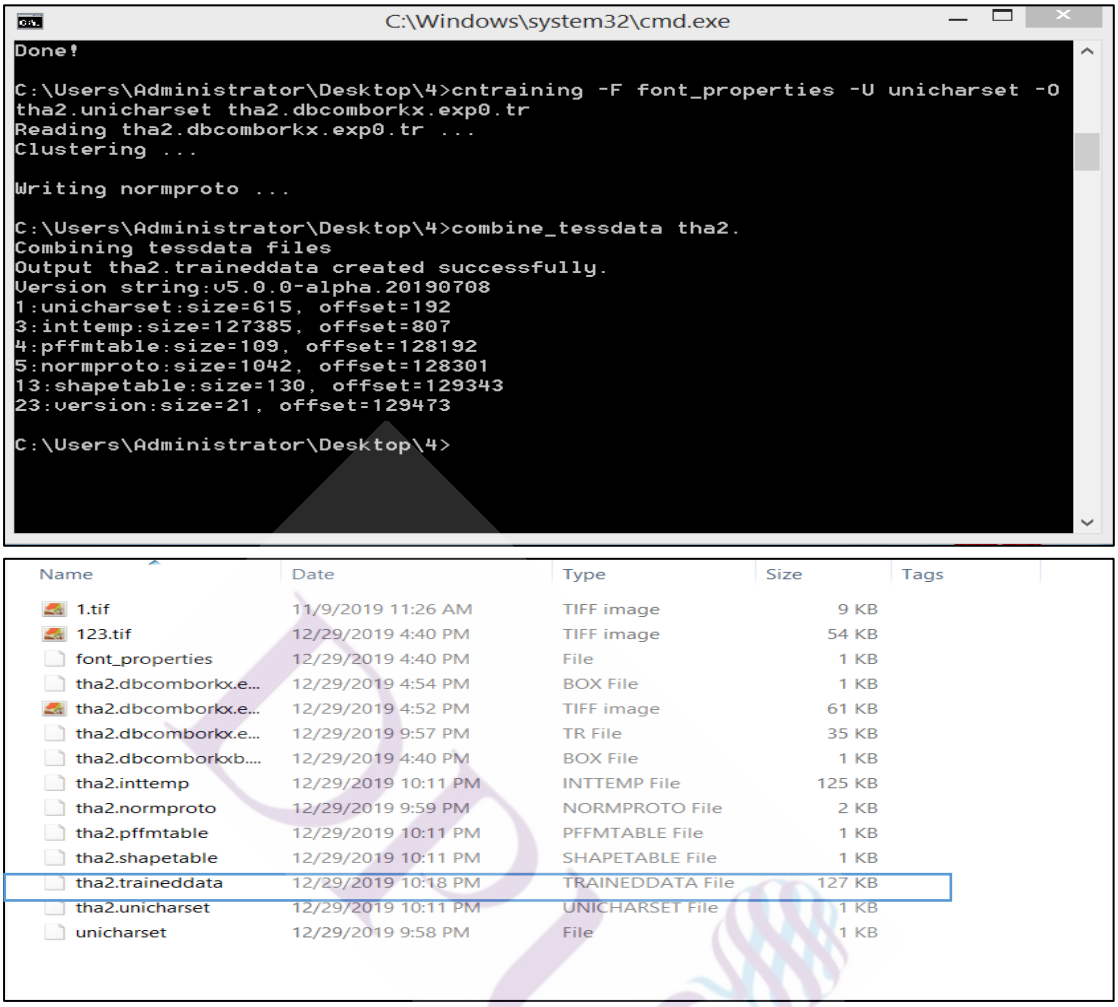
Name	Date	Type	Size	Tags
1.tif	11/9/2019 11:26 AM	TIFF image	9 KB	
123.tif	12/29/2019 4:40 PM	TIFF image	54 KB	
font_properties	12/29/2019 4:40 PM	File	1 KB	
inttemp	12/29/2019 10:11 PM	File	125 KB	
normproto	12/29/2019 9:59 PM	File	2 KB	
pffmtable	12/29/2019 10:11 PM	File	1 KB	
shapetable	12/29/2019 10:11 PM	File	1 KB	
tha2.dbcomborkx.e...	12/29/2019 4:54 PM	BOX File	1 KB	
tha2.dbcomborkx.e...	12/29/2019 4:52 PM	TIFF image	61 KB	
tha2.dbcomborkx.e...	12/29/2019 9:57 PM	TR File	35 KB	
tha2.dbcomborkxb....	12/29/2019 4:40 PM	BOX File	1 KB	
tha2.unicharset	12/29/2019 10:11 PM	UNICHARSET File	1 KB	
unicharset	12/29/2019 9:58 PM	File	1 KB	

ภาพที่ 3.17 ทำFont characteristics กับ Aggregating font features

7) สร้าง traineddataใหม่ขั้นตอนก่อนหน้าได้รับ 4 ไฟล์ก็คือ inttemp normproto pffmtable shapetable แก้ inttemp normproto pffmtable shapetable เป็นtha2.inttemptha2.normproto tha2.pffmtable tha2.shapetable และใช้คำสั่ง combine_tessdata tha2. เพื่อสร้างtraineddata ใหม่ ดังภาพที่ 3.18

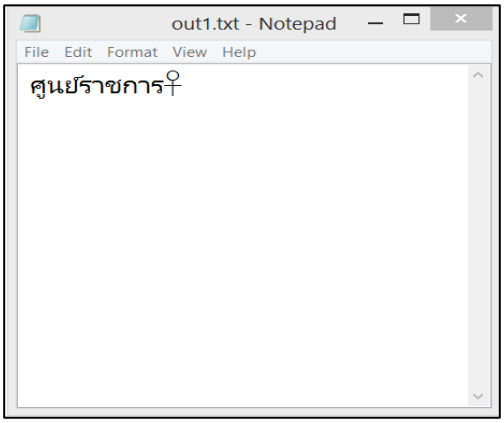


Name	Date	Type	Size	Tags
1.tif	11/9/2019 11:26 AM	TIFF image	9 KB	
123.tif	12/29/2019 4:40 PM	TIFF image	54 KB	
font_properties	12/29/2019 4:40 PM	File	1 KB	
tha2.dbcomborkx.e...	12/29/2019 4:54 PM	BOX File	1 KB	
tha2.dbcomborkx.e...	12/29/2019 4:52 PM	TIFF image	61 KB	
tha2.dbcomborkx.e...	12/29/2019 9:57 PM	TR File	35 KB	
tha2.dbcomborkxb....	12/29/2019 4:40 PM	BOX File	1 KB	
tha2.inttemp	12/29/2019 10:11 PM	INTTEMP File	125 KB	
tha2.normproto	12/29/2019 9:59 PM	NORMPROTO File	2 KB	
tha2.pffmtable	12/29/2019 10:11 PM	PFFMTABLE File	1 KB	
tha2.shapetable	12/29/2019 10:11 PM	SHAPETABLE File	1 KB	
tha2.unicharset	12/29/2019 10:11 PM	UNICHARSET File	1 KB	
unicharset	12/29/2019 9:58 PM	File	1 KB	



ภาพที่ 3.18 สร้าง traineddata ใหม่

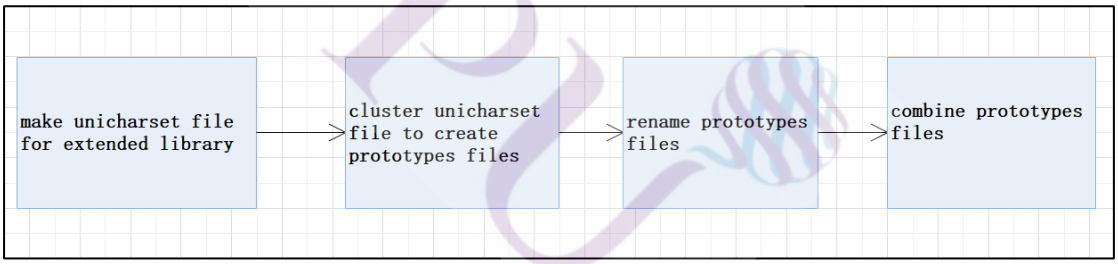
8) ผลทดสอบใหม่ดังรูปที่3.19 มันเป็นไปได้ที่จะรวมวิธีการฝึกอบรมผ่านโปรแกรมการฝึกอบรม tesseract และ jTessBoxEditor แก้ไขตัวอักษรซอฟต์แวร์ ใช้วิธีการที่สร้างขึ้นข้างต้นสามารถจดจำภาษาไทยในภาพที่ 3.2 เท่านั้นเพื่อให้บรรลุผลการจดจำ OCR ไทยจำนวนมาก บทความนี้ใช้วิธีการฝึกอบรมข้างต้นเพื่อฝึกอบรมภาษาไทยจำนวนมาก ในที่สุดสร้าง Library ฝึกอบรมภาษาไทย



ภาพที่ 3.19 ผลทดสอบใหม่

3.3.2 การปรับแต่งเมื่อการรู้จำผิดพลาด

ไลบรารีที่ผ่านการฝึกอบรมสามารถแก้ไขได้อักขระบางตัวที่ไม่ถูกต้อง แต่ปัญหาก็คือการรู้จำยังมีข้อผิดพลาดอยู่ ซึ่งเราต้องเพิ่มการฝึกอบรมอักขระ ดังภาพที่ 3.20



ภาพที่ 3.20 กระบวนการการปรับแต่งเมื่อการรู้จำผิดพลาด

ที่มา: อธิบายภาพ 3.20

ขั้นตอนที่ 1 make unicharset file for extended library

สร้างไฟล์ font_properties เพื่อรวมไลบรารีขนาดใหญ่

ใช้คำสั่ง:mftraining -F font_properties -U unicharset -O tha.unicharset tha.font.exp0.tr

ขั้นตอนที่ 2 cluster unicharset file to create prototypes files

ใช้คำสั่ง cntraining -F font_properties -U unicharset -O tha.unicharset tha.font.exp0.tr

จะได้ไฟล์ unicharset, inttemp, normproto, pfftable

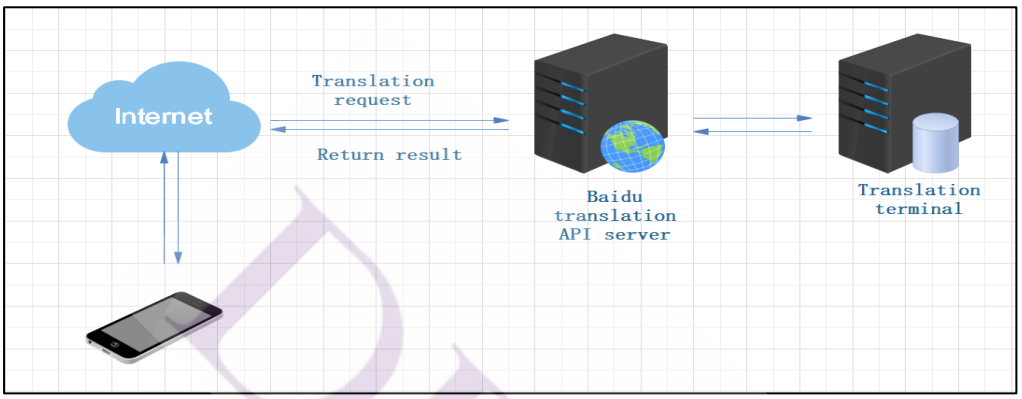
ขั้นตอนที่ 3 rename prototypes files

ไฟล์ normproto; pffmtable; inttemp; shapetable แก้เป็น tha.normproto; tha.pffmtable; tha.inttemp; tha.shapetable

ขั้นตอนที่ 4 combine prototypes files

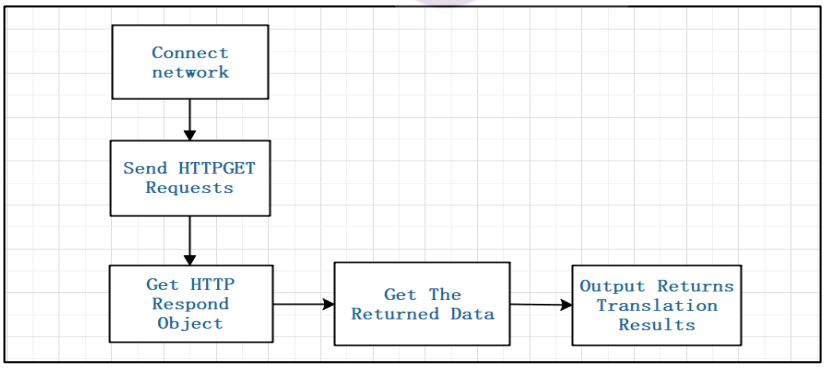
ใช้คำสั่ง combine_tessdata tha. จะได้ไลบรารีใหญ่

3.4 baidu Translation ดังภาพที่ 3.21



ภาพที่ 3.21 Baidu access process

วิธีใช้ baidu Translation API ดังภาพที่ 3.22



ภาพที่ 3.22 Translation process

ที่มา: ขั้นตอนที่ 1 : connect network

แปลภาษาต้องใช้ server จึงต้องต่อnetwork ซึ่ง code ดังภาพที่ 3.23

```
String requestUrl = getUrlWithQueryString(host, params);
URL uri = new URL(requestUrl);
HttpURLConnection conn = (HttpURLConnection) uri.openConnection();
```

ภาพที่ 3.23 code ต่อ network

ขั้นตอนที่ 2 : Send http get request

การแปล baidu สามารถใช้วิธี get ของ HTTP เพื่อทำการขอร้อง (ใช้ http://api.fanyi.baidu.com/api/translate ไปขอ) ซึ่ง code ดังภาพที่ 3.24

```
private static final String TRANS_API_HOST = "http://api.fanyi.baidu.com/api/translate";
public String getTransResult(String query, String from, String to) {
    Map<String, String> params = buildParams(query, from, to);
    return HttpGet.get(TRANS_API_HOST, params);
}
```

ภาพที่ 3.24 code ทำการขอร้อง

ยกตัวอย่าง : แปลภาษาไทย(สวัสดีค่ะ)เป็นอังกฤษ

q=สวัสดีค่ะ

from=th

to=en

appid=2015063000000001

salt=1435660288 (random code)

Platform-assigned keys: 12345678

Generate sign :

Step1.Stitching strings 1: splice

appid=2015063000000001+q=สวัสดีค่ะ+salt=1435660288+key=12345678

Get string1 : “2015063000000001สวัสดีค่ะ143566028812345678”

Step2. Calculate signature : (MD5 encryption for string 1)

sign=md5(2015063000000001สวัสดีค่ะ1435660288123456 78),get sign=e3813ffed6da733d38b71793d653231f

Splicing complete request:

http://api.fanyi.baidu.com/api/trans/vip/translate?q=สวัสดีค่ะ&from=th&to=en&appid=2015063000000001&salt=1435660288&sign=e3813ffed6da733d38b71793d653231f

ตารางที่ 3.2 data type

Field name	type	not null	description
from	TEXT	YES	Translation source language
to	TEXT	YES	Target language
Client_id	Appid	YES	A authorized key,registered by developers in Baidu Developer Center
q	TEXT	YES	Request translated text, The field must be UTF--8 encoded,urlencode encoding required
src	TEXT	YES	Source language
dst	TEXT	YES	Target language

ขั้นตอนที่3: Get HTTP Respond Object

การแปล baidu ข้อมูลที่ส่งกลับมาคือแพ็คเกจ json ซึ่ง code ดังภาพที่ 3.25

```
String query = srcText.getText().toString().trim();
String tranString = api.getTransResult(query, "th", "en");
```

ภาพที่ 3.25 code แพ็คเกจ json

ที่มา: เช่น {"from":"th","to":"en","trans_result":[{"src":"สวัสดีค่ะ ","dst":"hello"}]

ขั้นตอนที่4: Get The Returned Data ซึ่ง code ดังภาพที่ 3.26

```
String result = api.getTransResult(query, "th", "en");
String translation = jsonParse.parse(result);
resultText.setText(translation);
```

ภาพที่ 3.26 code Get The Returned Data

ขั้นตอนที่5: ผลการแปล ซึ่ง code ดังภาพที่ 3.27

```
String result = api.getTransResult(query, "th", "en");
String translation = jsonParse.parse(result);
resultText.setText(translation);
```

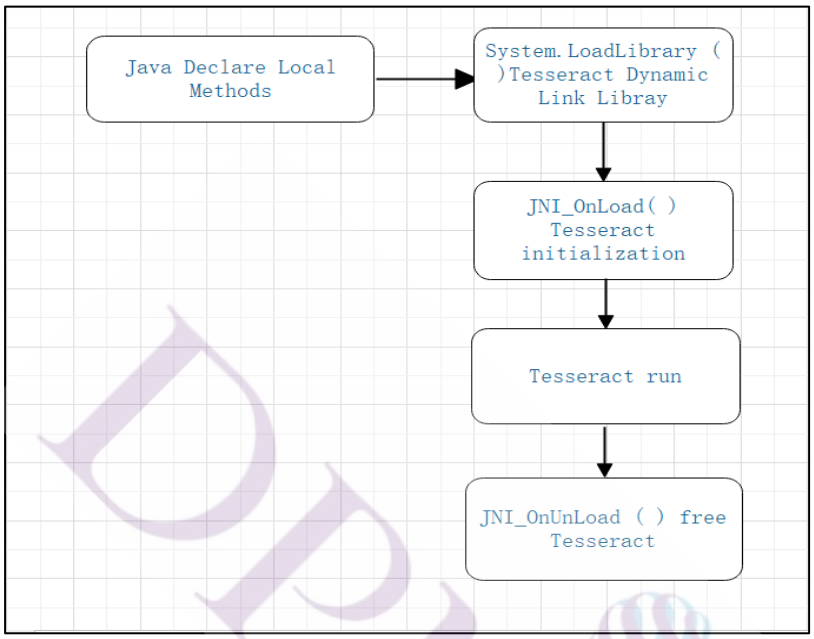
ภาพที่ 3.27 code ผลการแปล

3.5 อินเตอร์เฟซ JNI

JNI (Java Native Interface) คือ วิธีการในจาวาที่ทำให้เราสามารถเรียกใช้งาน native code (ส่วนของโปรแกรมที่เขียนขึ้นโดยใช้ภาษา C หรือ C++) จากโปรแกรมจาวาที่เราเขียนขึ้นได้ หรือในทางกลับกันก็ได้ คือเรียกใช้ส่วนของโปรแกรมที่เขียนขึ้นด้วยภาษาจาวาจากโปรแกรมที่เขียนขึ้น ด้วยภาษา C หรือ C++ จะเห็นว่า JNI จะเป็น interface แบบสองทางที่ทำให้ส่วนของ

โปรแกรมจาวาและส่วนของโปรแกรมที่เขียนด้วย C หรือ C++ ทำงานร่วมกันได้ภายใต้ process เดียวกัน (ทำงานแบบเป็นโปรแกรมเดียวกัน ไม่ใช่เป็นแบบสองโปรแกรมทำงานร่วมกัน)

เนื่องจากใช้ Tesseract ต้องผ่านอินเทอร์เฟซ JNI ดังนั้นเราต้องเชื่อมต่อระหว่างโปรแกรมJAVA และ JNI เพื่อ transfer Tesseract ซึ่งกระบวนการระบบดังภาพที่ 3.28



ภาพที่ 3.28 กระบวนการระบบอินเตอร์เฟซ JNI

ที่มา: อธิบายภาพ 3.29

ขั้นตอนที่ 1 : Java declare local methods.ใช้งานcode native เพื่อบอกว่าวิธีนี้ใช้ภาษาอื่น ๆ ไม่ใช้ภาษา Java ดังภาพที่ 3.29

```
public native void nativeSetImage(
        byte[] imagedata, int width, int height, int bpp, int bpl);
```

ภาพที่ 3.29 code Java declare local methods

ขั้นตอนที่ 2 : สร้างเป็นไลบรารี(Dynamic link library) และใช้ JNI-OnLoad เพื่อ Tesseract initialization ดังภาพที่ 3.30

```
static {
   System.loadLibrary("lept");
   System.loadLibrary("tess");
   nativeClassInit();
}
```

ภาพที่ 3.30 code Dynamic link library

ขั้นตอนที่ 3 : run Tesseract ดังภาพที่ 3.31

```
Bitmap bitmap = BitmapFactory.decodeFile(filePath);
TessBaseAPI baseApi = new TessBaseAPI();
baseApi.setImage(bitmap);
String text = baseApi.getUTF8Text();
```

ภาพที่ 3.31 code run Tesseract

ขั้นตอนที่ 4 :เมื่อเครื่องเสมือน(Virtual Machine) free จะไปหาฟังก์ชัน JNI_OnUnload ( ) เพื่อ delete ดังภาพที่ 3.32

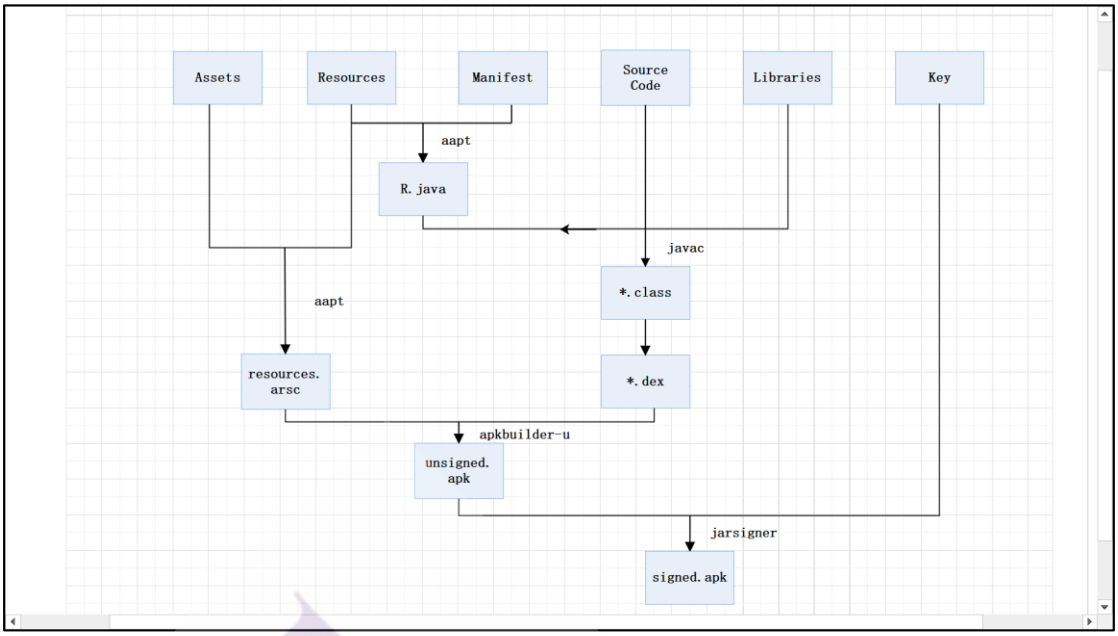
```
public native void Clear();
public native void End();
```

```
baseApi.clear();
baseApi.end();
```

ภาพที่ 3.32 code delete

3.6 OCR ระบบบรรจุภัณฑ์

หลังจากเขียนโปรแกรมระบบการรู้จำเสร็จแล้วเราต้องใช้codeแอปพลิเคชันบรรจุภัณฑ์เป็นแพ็คเกจการติดตั้งแอปพลิเคชันซึ่งก็คือไฟล์ apk กระบวนการบรรจุไฟล์ ดังภาพที่ 3.33



ภาพที่ 3.33 กระบวนการ OCR ระบบบรรจุภัณฑ์

ที่มา: อธิบายภาพ 3.34

ขั้นตอนที่ 1: ใช้เครื่องมือ aapt(Android Asset Package Tool) เพื่อทำแพ็คเกจไฟล์ Assets ไฟล์Resources ไฟล์Manifestเป็นแพ็คเกจบีบอัดresources.arsc จากนั้นสร้าง ID ที่สอดคล้องกันในไฟล์ R.java

ขั้นตอนที่ 2: คอมไพล์ source codeคือใช้ java compilerคอมไพล์ต้นฉบับ java ทั้งหมดจะได้ *.class

ขั้นตอนที่ 3: เครื่องเสมือนdalvikเปลี่ยน *.class เป็น *.dex

ขั้นตอนที่ 4: ใช้เครื่องมือapkbuilderเปลี่ยนแพ็คเกจบีบอัดresources.arsc ไพล์ *.dex เป็นไพล์ unsigned.apk

ขั้นตอนที่ 5: ใช้ไพล์ Android.keystoreเซ็นชื่อไฟล์apk

ขั้นตอนที่ 6: ใช้เครื่องมือjarsignerเปลี่ยนไพล์ unsigned.apk กับ key เป็นไพล์ signed.apk OCR ระบบบรรจุภัณฑ์ใช้งานจริงดังภาพที่ 3.34

ภาพที่ 3.34 OCR ระบบบรรจุภัณฑ์ใช้งานจริง

3.7 สรุป

ในบทนี้จะกล่าวถึงการออกแบบระบบแอพมือถือ ระบบแอพมือถือประกอบด้วย 3 ส่วนมี Photograph ,Translation ,Vocabulary ส่วนโมดูล Photograph ใช้เครื่องมือ Tesseract open source เพื่อการฝึกอปรมและการรู้จำภาพ ส่วนโมดูล Translation ใช้ Protocol HTTP กับ Baidu Translation เพื่อแปลภาษาไทยเป็นภาษาจีนกับภาษาอังกฤษ ส่วนโมดูล Vocabulary โดยใช้ระบบ Android เพื่อ save ภาษาไทย การออกแบบระบบแอพมือถือเสร็จแล้วต้องใช้เครื่องมือ Android Studio เพื่อบรรจุในตัวติดตั้งแอพพลิเคชัน Android ให้ผู้ใช้ใช้งานได้

DPU

บทที่ 4
ผลการทดสอบระบบ

เนื้อหาในบทนี้จะกล่าวถึงการผลทดสอบการใช้งานของโปรแกรมที่ผู้วิจัยได้ออกแบบแนวคิดโปรแกรมนี้ขึ้นมาและประสานงานในการพัฒนาโปรแกรม การประเมิลผลการใช้งานและการปรับแต่งเมื่อการรู้จำผิดพลาด โดยทำการทดสอบระบบที่ทำขึ้น ในการทดสอบกำหนดให้มีการใช้งานจริง โดยการนำโปรแกรมที่พัฒนาขึ้นแล้วมาติดตั้งบนเครื่องโทรศัพท์มือถือ เพื่อที่จะใช้งานได้จริง

4.1 สภาพแวดล้อมในการทดสอบ

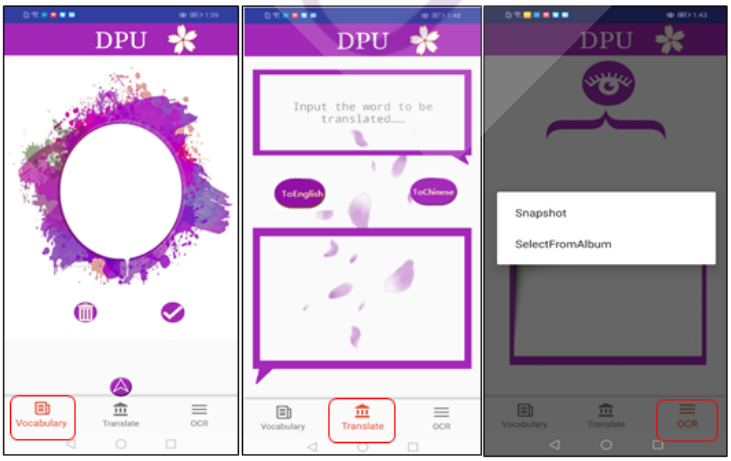
ผู้วิจัยได้ทำการทดสอบโปรแกรมดังกล่าวบนเครื่องโทรศัพท์มือถือที่ติดตั้งระบบปฏิบัติการแอนดรอยด์เวอร์ชั่น 9 รุ่น Huawei

ภาพที่ 4.1 โทรศัพท์มือถือแอนดรอยด์ รุ่น Huawei 9 ที่ใช้ในการทดสอบโปรแกรม

4.2 หน้าจอแสดงผลการทำงานของโปรแกรม

ภาพที่ 4.2 แสดงThaiOCR APPบนเครื่องโทรศัพท์มือถือ

4.3 ThaiOCR มีปุ่มฟังก์ชั่นคือ Vocabulary, Translate, OCR



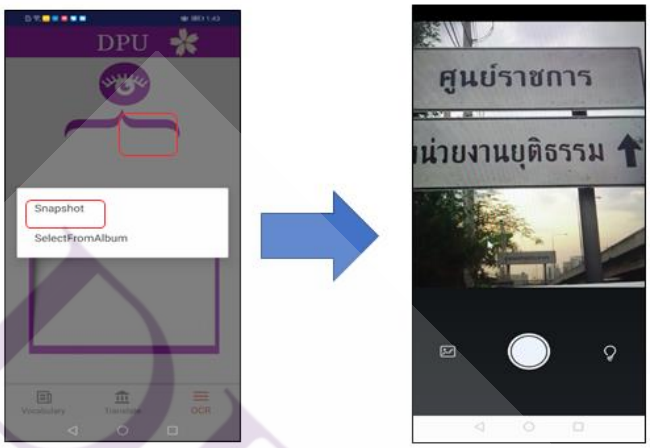
ภาพที่ 4.3 หน้าหลักของโปรแกรม

4.4 ผลการทดสอบ

ผลการทดสอบของการใช้งานโปรแกรม เพื่อให้เป็นทำงานได้จริง

1) ผลการทดสอบ ปุ่มฟังก์ชั่นของ OCR

1.1 ถ่ายภาพจาก Snapshot

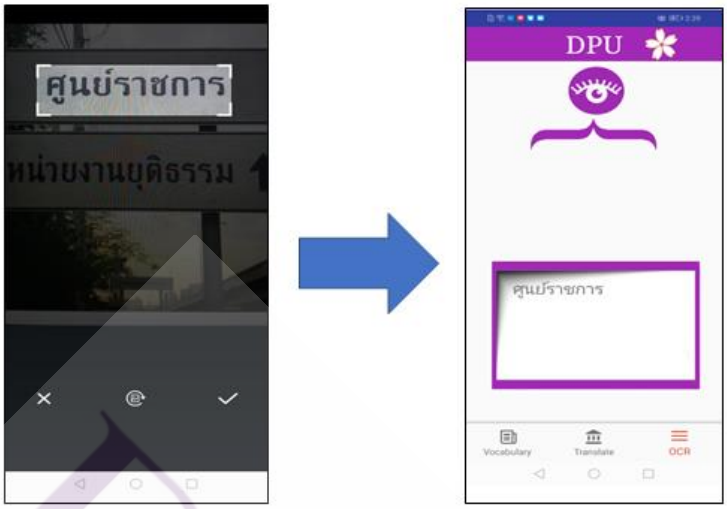


ภาพที่ 4.4 แสดงการถ่ายรูปจาก Snapshot

1.2 ถ่ายภาพจาก Select From Album

ภาพที่ 4.5 แสดงการถ่ายรูปจาก Select From Album

1.3 ตัดภาพ



ภาพที่ 4.6 แสดงการตัดภาพ

2) ผลการทดสอบ Translate

ภาพที่ 4.7 แสดงภาษาไทยแปลเป็นภาษาอังกฤษกับจีน

3) ผลการทดสอบ Vocabulary

ภาพที่ 4.8 แสดงการ Vocabulary

4) ผลการทดสอบฟังก์ชั่นการรับรู้ของ OCR

ความแม่นยำของระบบ OCR มาจาก อัตราการรู้จำที่ถูกต้อง อัตราการรู้จำที่ข้อผิดพลาด ซึ่งอัตราความถูกต้อง=จำนวนรู้จำที่ถูกต้อง/จำนวนตัวอักษรทั้งหมด ดังนั้นผู้วิจัยนำไปทำการทดสอบกับป้ายทั้งหมด14 ภาพป้าย ดังต่อไปนี้

ภาพที่ 4.9 ตัวอย่างของป้ายถนน

ภาพที่ 4.10 ป้ายดั้งเดิมที่ใช้ในการทดสอบ

5) การทดสอบประสิทธิภาพด้านความแม่นยำของป้ายดั้งเดิม

ตารางที่ 4.1 ป้ายดั้งเดิมที่ยังไม่ได้เพิ่มความแม่นยำ

ป้ายดั้งเดิม	จำนวนตัวอักษร	จำนวนอักษรที่ถูกต้อง	อัตราความถูกต้อง	จำนวนตัวอักษรที่ข้อผิดพลาด	อัตราข้อผิดพลาด
ป้ายที่ 1	13	7	54%	6	46%
ป้ายที่ 2	23	11	48%	12	52%
ป้ายที่ 3	8	5	62%	3	38%
ป้ายที่ 4	29	27	93%	2	7%
ป้ายที่ 5	9	7	78%	2	22%
ป้ายที่ 6	11	10	91%	1	9%
ป้ายที่ 7	33	24	73%	9	27%
ป้ายที่ 8	8	6	75%	2	25%
ป้ายที่ 9	16	10	63%	6	38%
ป้ายที่ 10	18	15	83%	3	50%
ป้ายที่ 11	6	3	50%	3	50%
ป้ายที่ 12	11	10	91%	1	9%
ป้ายที่ 13	11	5	45%	6	55%
ป้ายที่ 14	11	8	73%	3	27%
ทั้งหมด	207	148	71.5%	59	28.5%

ตารางที่ 4.2 ป้ายดั้งเดิมที่เพิ่มความแม่นยำ

ป้ายดั้งเดิม	จำนวน ตัวอักษร	จำนวนอักษร ที่ถูกต้อง	อัตรา ความถูกต้อง	จำนวน ตัวอักษรที่ ข้อผิดพลาด	อัตรา ข้อผิดพลาด
ป้ายที่ 1	13	13	100%	0	0%
ป้ายที่ 2	23	23	100%	0	0%
ป้ายที่ 3	9	9	100%	0	0%
ป้ายที่ 4	29	28	97%	1	3%
ป้ายที่ 5	9	9	100%	0	0%
ป้ายที่ 6	11	11	100%	0	0%
ป้ายที่ 7	33	30	91%	3	9%
ป้ายที่ 8	8	8	100%	0	0%
ป้ายที่ 9	16	16	100%	0	0%
ป้ายที่ 10	18	17	94%	1	6%
ป้ายที่ 11	6	5	83%	1	17%
ป้ายที่ 12	11	10	91%	1	9%
ป้ายที่ 13	11	11	100%	0	0%
ป้ายที่ 14	11	11	100%	0	0%
ทั้งหมด	207	200	97%	7	3%

อธิบายตารางที่ 4.2 ในการทดลองครั้งนี้ใช้คำที่เป็นป้ายที่ภาษาไทยมาทำการทดสอบทั้งหมด 207 คำ ซึ่งคำที่ประมวลผลลัพธ์ถูกต้อง 97% ป้ายส่วนใหญ่ความแม่นยำสูงแต่ป้ายที่ 7 กับ ป้ายที่ 11 ความแม่นยำแค่ 91% กับ 83%เพราะว่าป้ายที่มีตัวอักขระลักษณะคล้ายคลึงกัน

ภาพที่ 4.11 ป้ายเอียงที่ใช้ในการทดสอบ

6) การทดสอบประสิทธิภาพด้านความแม่นยำของป้ายเอียง

ตารางที่ 4.3 ป้ายเอียงที่ไม่ได้เพิ่มความแม่นยำ

ป้ายเอียง	จำนวนตัวอักษร	จำนวนอักษรที่ถูกต้อง	อัตราความถูกต้อง	จำนวนตัวอักษรที่ข้อผิดพลาด	อัตราข้อผิดพลาด
ป้ายที่ 1	13	7	54%	6	46%
ป้ายที่ 2	23	9	39%	14	61%
ป้ายที่ 3	8	5	63%	3	37%
ป้ายที่ 4	29	25	86%	4	14%
ป้ายที่ 5	9	6	67%	3	33%
ป้ายที่ 6	11	10	91%	1	9%
ป้ายที่ 7	33	20	61%	13	39%
ป้ายที่ 8	8	7	87%	1	13%
ป้ายที่ 9	16	15	94%	1	6%
ป้ายที่ 10	18	15	83%	3	17%
ป้ายที่ 11	6	4	67%	2	33%
ป้ายที่ 12	11	9	82%	2	18%
ป้ายที่ 13	11	7	64%	4	36%
ป้ายที่ 14	11	8	73%	3	27%
ทั้งหมด	207	147	71%	60	29%

ตารางที่ 4.4 ป้ายเอียงที่เพิ่มความแม่นยำ

ป้าย เอียง	จำนวน ตัวอักษร	จำนวนอักษร ที่ถูกต้อง	อัตรา ความถูกต้อง	จำนวน ตัวอักษรที่ ข้อผิดพลาด	อัตรา ข้อผิดพลาด
ป้ายที่ 1	13	12	92%	1	8%
ป้ายที่ 2	23	23	100%	0	0%
ป้ายที่ 3	8	8	100%	0	0%
ป้ายที่ 4	29	26	90%	3	10%
ป้ายที่ 5	9	9	100%	0	0%
ป้ายที่ 6	11	11	100%	0	0%
ป้ายที่ 7	33	28	85%	5	15%
ป้ายที่ 8	8	8	100%	0	0%
ป้ายที่ 9	16	16	100%	0	0%
ป้ายที่ 10	18	16	89%	2	11%
ป้ายที่ 11	6	5	83%	1	17%
ป้ายที่ 12	11	10	91%	1	9%
ป้ายที่ 13	11	9	82%	2	18%
ป้ายที่ 14	11	11	100%	0	0%
ทั้งหมด	207	18	93%	15	7%

อธิบายตารางที่ 4.4 ในการทดลองครั้งนี้ใช้คำที่เป็นป้ายที่ภาษาไทยมาทำการทดสอบทั้งหมด 207 คำ ซึ่งคำที่ประมวลผลลัพธ์ถูกต้อง 93% ป้ายส่วนใหญ่ความแม่นยำสูงแต่ป้ายที่ 11 ความแม่นยำแค่ 82% เพราะว่าป้ายที่มีตัวอักขระลักษณะคล้ายคลึงกันและป้ายเอียงถ่ายไม่ชัด

ภาพที่ 4.12 ป้ายแสงน้อยที่ใช้ในการทดสอบ

7) การทดสอบประสิทธิภาพด้านความแม่นยำของป้ายแสงน้อย

ตารางที่ 4.5 ป้ายแสงน้อยที่ยังไม่ได้เพิ่มความแม่นยำ

ป้ายแสงน้อย	จำนวนตัวอักษร	จำนวนอักษรที่ถูกต้อง	อัตราความถูกต้อง	จำนวนตัวอักษรที่ข้อผิดพลาด	อัตราข้อผิดพลาด
ป้ายที่ 1	13	6	46%	7	54%
ป้ายที่ 2	23	15	65%	8	35%
ป้ายที่ 3	8	5	62%	3	38%
ป้ายที่ 4	29	21	72%	8	28%
ป้ายที่ 5	9	6	67%	3	33%
ป้ายที่ 6	11	10	91%	1	9%
ป้ายที่ 7	33	19	58%	14	42%
ป้ายที่ 8	8	7	87%	1	13%
ป้ายที่ 9	16	15	94%	1	6%
ป้ายที่ 10	18	17	94%	1	6%
ป้ายที่ 11	6	5	83%	1	17%
ป้ายที่ 12	11	9	82%	2	18%
ป้ายที่ 13	11	5	45%	6	55%
ป้ายที่ 14	11	8	73%	3	27%
ทั้งหมด	207	148	71%	59	29%

ตารางที่ 4.6 ป้ายแสงน้อยที่เพิ่มความแม่นยำ

ป้ายแสงน้อย	จำนวนตัวอักษร	จำนวนอักษรที่ถูกต้อง	อัตราความถูกต้อง	จำนวนตัวอักษรที่ข้อผิดพลาด	อัตราข้อผิดพลาด
ป้ายที่ 1	13	12	92%	1	80%
ป้ายที่ 2	23	20	87%	3	13%
ป้ายที่ 3	8	7	87%	1	13%
ป้ายที่ 4	29	28	97%	1	3%
ป้ายที่ 5	9	9	100%	0	89%
ป้ายที่ 6	11	11	100%	0	0%
ป้ายที่ 7	33	27	82%	6	18%
ป้ายที่ 8	8	8	100%	0	0%
ป้ายที่ 9	16	16	100%	0	0%
ป้ายที่ 10	18	17	94%	1	6%
ป้ายที่ 11	6	5	83%	1	17%
ป้ายที่ 12	11	10	91%	1	9%
ป้ายที่ 13	11	10	91%	1	9%
ป้ายที่ 14	11	11	100%	0	0%
ทั้งหมด	207	191	92%	16	8%

อธิบายตารางที่ 4.6 ในการทดลองครั้งนี้ใช้คำที่เป็นป้ายที่ภาษาไทยมาทำการทดสอบทั้งหมด 207คำซึ่งคำที่ประมวลผลลัพธ์ถูกต้อง 92% ป้ายส่วนใหญ่ความแม่นยำสูงแต่ป้ายที่ 7 กับ ป้ายที่ 11ความแม่นยำแค่82%กับ83%เพราะว่าป้ายที่มีตัวอักขระลักษณะคล้ายคลึงกันและแสงน้อยถ่ายไม่ชัด

(1)บางนาดาวคะนอง

(2)เดอะแพลนท์ แจ้งวัฒนะ ราชพฤกษ์

(3)แยกเกษตร

(4)กระทรวงสาธารณสุข สำนักงาน กพ . (นนทบุรี)

(5)ศูนย์ราชการ

(6)สะพานพระราม

(7)โรงเรียนอนุบาลจุติพร เมืองทองธานี โครงการ

(8)ธาราพาร์ค

(9)เดอะแพลนท์ ปากเกร็ด

(10)สำนักงานกองทุนยุติธรรม

(11)แอสปาย

(12)โพลิแทนคอนโด

(13)ศูนย์ราชการ 2 กม.

(14)ศูนย์ OTOP รามอินทรา

ภาพที่ 4.13 ป้ายตัวพิมพ์ที่ใช้ในการทดสอบ

8) การทดสอบประสิทธิภาพด้านความแม่นยำของป้ายตัวพิมพ์

ตารางที่ 4.7 ป้ายตัวพิมพ์ที่ยังไม่ได้เพิ่มความแม่นยำ

ป้ายตัวพิมพ์	จำนวนตัวอักษร	จำนวนอักษรที่ถูกต้อง	อัตราความถูกต้อง	จำนวนตัวอักษรที่ข้อผิดพลาด	อัตราข้อผิดพลาด
ป้ายที่ 1	13	13	100%	0	0%
ป้ายที่ 2	23	18	78%	5	22%
ป้ายที่ 3	9	9	100%	0	0%
ป้ายที่ 4	29	26	90%	3	10%
ป้ายที่ 5	9	7	78%	2	22%
ป้ายที่ 6	11	9	82%	2	18%
ป้ายที่ 7	33	29	88%	4	12%
ป้ายที่ 8	8	5	62%	3	38%
ป้ายที่ 9	16	15	94%	1	60%
ป้ายที่ 10	18	17	94%	1	6%
ป้ายที่ 11	6	6	100%	0	0%
ป้ายที่ 12	11	9	82%	2	18%
ป้ายที่ 13	11	8	73%	3	27%
ป้ายที่ 14	11	6	55%	5	45%
รวม	207	176	85%	31	15%

ตารางที่ 4.8 ป้ายตัวพิมพ์ที่เพิ่มความแม่นยำ

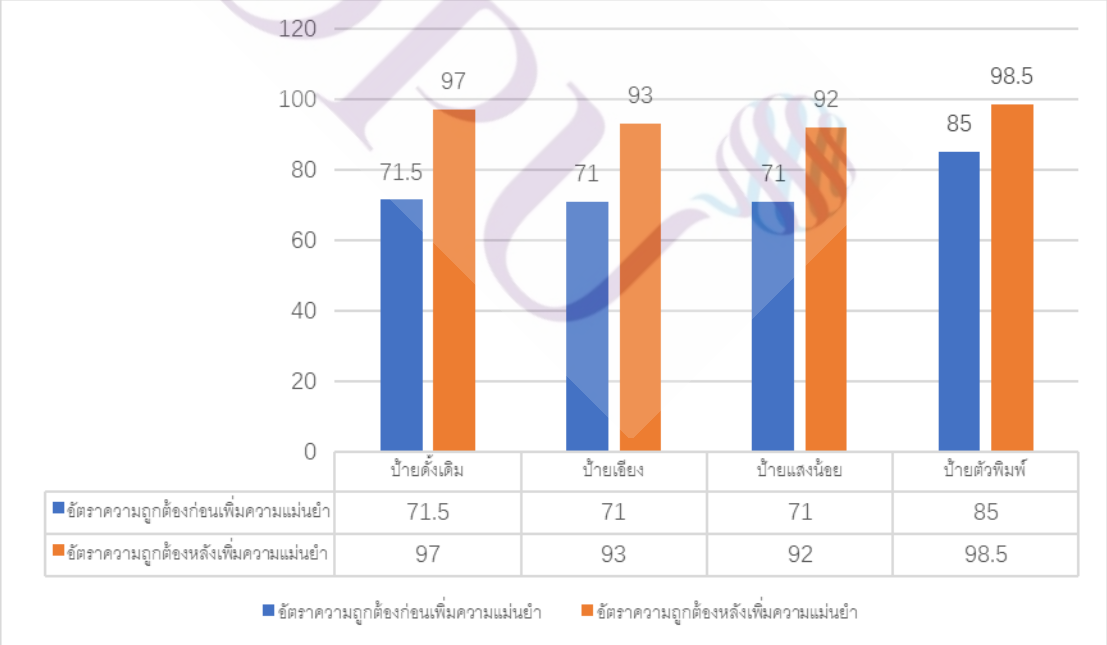
ป้ายตัวพิมพ์	จำนวน ตัวอักษร	จำนวนอักษร ที่ถูกต้อง	อัตรา ความถูกต้อง	จำนวน ตัวอักษรที่ ข้อผิดพลาด	อัตรา ข้อผิดพลาด
ป้ายที่ 1	13	13	100%	0	0%
ป้ายที่ 2	23	23	100%	0	0%
ป้ายที่ 3	9	9	100%	0	0%
ป้ายที่ 4	29	29	100%	0	0%
ป้ายที่ 5	9	9	100%	0	0%
ป้ายที่ 6	11	11	100%	0	0%
ป้ายที่ 7	33	31	94%	2	6%
ป้ายที่ 8	8	8	100%	0	0%
ป้ายที่ 9	16	16	100%	0	0%
ป้ายที่ 10	18	18	100%	0	0%
ป้ายที่ 11	6	6	100%	0	0%
ป้ายที่ 12	11	11	100%	0	0%
ป้ายที่ 13	11	0	100%	11	0%
ป้ายที่ 14	11	10	91%	1	9%
รวม	207	204	98.5%	3	1.5%

อธิบายตารางที่ 4.8 ในการทดลองครั้งนี้ใช้คำที่เป็นป้ายที่ภาษาไทยมาทำการทดสอบทั้งหมด 207 คำ ซึ่งคำที่ประมวลผลลัพธ์ถูกต้อง 98.5% ป้ายตัวพิมพ์ส่วนใหญ่ความแม่นยำสูงแต่ป้ายที่ 7 กับ ป้ายตัวพิมพ์ที่ 14ความแม่นยำแค่94%กับ91%เพราะว่าป้ายที่มีตัวอักขระลักษณะคล้ายคลึง

9) สรุปผลการทดลองทั้งหมด

ตารางที่ 4.9 ผลการเปรียบเทียบความแม่นยำ

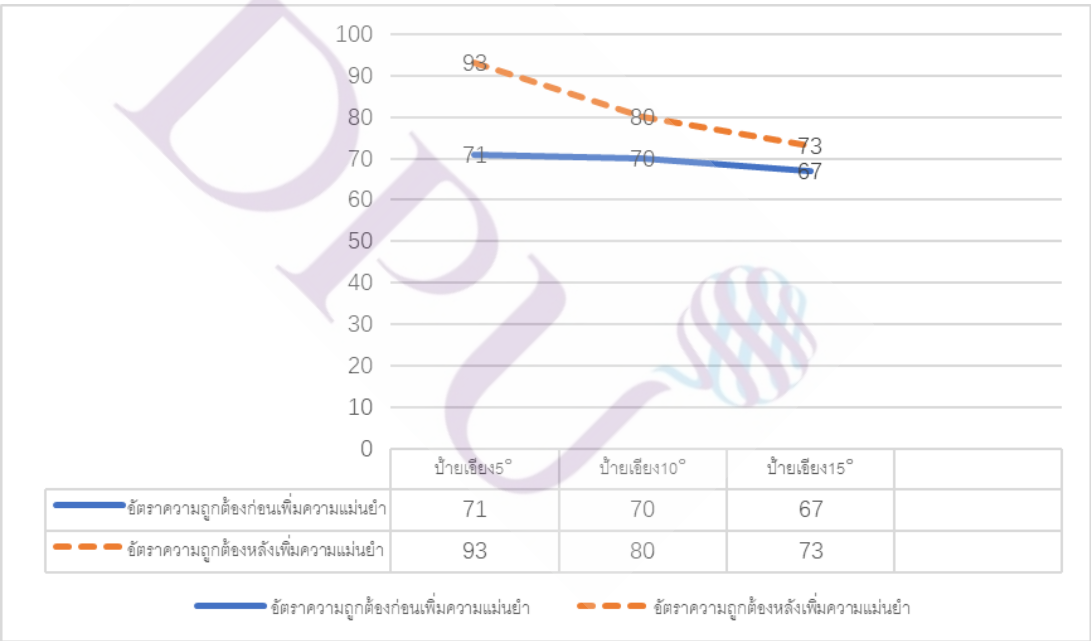
	จำนวนตัวอักษร	อัตราความถูกต้องของยังไม่ได้เพิ่มความแม่นยำ	อัตราความถูกต้องของเพิ่มความแม่นยำ
ป้ายดั้งเดิม	207	71.5%	97%
ป้ายเอียง	207	71%	93%
ป้ายแสงน้อย	207	71%	92%
ป้ายพิมพ์	207	85%	98.5%



ภาพที่ 4.14 ผลการเปรียบเทียบความแม่นยำ

จากผลการทดลองสรุปได้ว่าเมื่อยังไม่ได้เพิ่มความแม่นยำ ป้ายดังเดิม ป้ายเอียง ป้ายแสงน้อย พิมพ์ มีค่าความถูกต้องเฉลี่ยเป็น 71.5%, 71% ,71% ,85% เพราะยังไม่มีการปรับให้เหมาะสมสำหรับฟอนต์ภาษาไทยและยังไม่มีห้องสมุดฝึกอบรมที่พร้อมใช้งานได้ซึ่งความถูกต้องเฉลี่ยค่อนข้างต่ำ เมื่อนำเอาเทคนิคการแปลงค่าสี RGB เป็น Gray Scale , การกรองค่ามัธยฐานเสร็จ ,การแปลงภาพระดับเทาให้เป็นภาพขาว -ดำ ฯที่นำเสนอมาใช้ร่วมกันกับOCRสามารถเพิ่มความถูกต้องของการรู้จำตัวอักขระด้วยแสงสูงขึ้น เมื่อตัวอักษรยิ่งมีความชัดเจน ตรง แสงปกติ ส่งผลให้ผลลัพธ์ที่ได้จากกระบบรู้จำตัวอักษรมีประสิทธิภาพมากยิ่งขึ้น ซึ่งมีค่าความถูกต้องเฉลี่ยเป็น 97%, 93%, 92%, 98.5%.

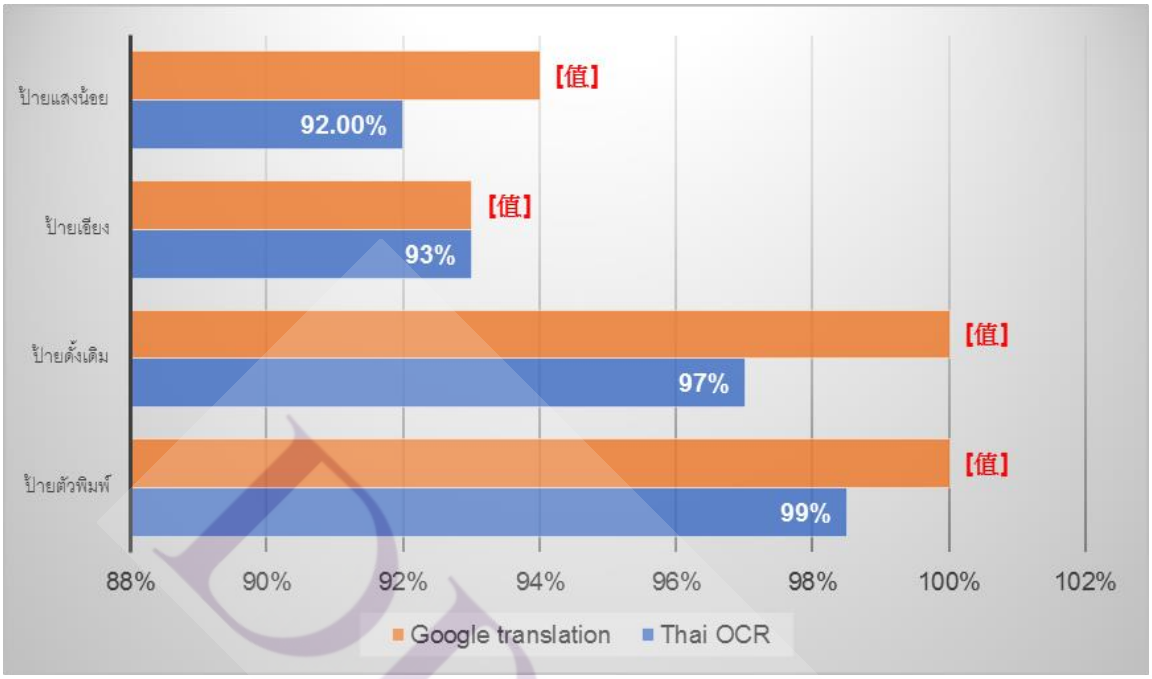
10) สรุปผลการทดลองป้ายเอียง



ภาพที่ 4.15 ผลการทดลองป้ายเอียง

จากภาพจะเห็นได้ป้ายที่มีความเอียงน้อยได้แก่ 5 องศา เมื่อมีการเพิ่มความแม่นยำเข้าไปทำให้ผลการทดลองมีอัตราความแม่นยำที่สูงมากกว่าป้ายที่มีความเอียง 10 องศา และ 15 องศา อย่างไรก็ตามเมื่อเพิ่มความแม่นยำเข้าไปจะเห็นได้ชัดเจนว่าผลการทดลองมีความแม่นยำมากขึ้นในทุกระดับของความเอียง

11) ผลการเปรียบเทียบความแม่นยำของGoogle translation



ภาพที่ 4.16 ผลการเปรียบเทียบความแม่นยำของ Google translation

จากภาพเมื่อทำการเปรียบเทียบความแม่นยำระหว่าง google translation กับ Thai OCR จะเห็น google translation มีความแม่นยำกว่าแต่ไม่มากนัก โดยในการแปลภาพพิมพ์ google translation มีความแม่นยำ 100% ส่วน Thai OCR แม่นยำ 99% ป้ายดั้งเดิม google translation แม่นยำ 100% ส่วน Thai OCR แม่นยำ 97% ป้ายเอียงมีความแม่นยำเท่ากัน คือ 93% ส่วนป้ายแสงน้อย google translation แม่นยำ 94% ส่วน Thai OCR แม่นยำ 92%

4.5 สรุป

ผลการทดสอบการใช้งานของโปรแกรมที่ผู้วิจัยได้ออกแบบแนวคิดโปรแกรมนี้ขึ้นมาและประสานงานในการพัฒนาโปรแกรมการประเมิลผลการใช้งานและการปรับแต่งเมื่อการรู้จำผิดพลาด โดยทำการทดสอบระบบที่ทำขึ้น โดยการใช้งานจริงและนำโปรแกรมที่พัฒนาขึ้นมาติดตั้งบนเครื่องโทรศัพท์มือถือ ระบบปฏิบัติการแอนดรอยด์เวอร์ชั่น 9 รุ่น Huawei เพื่อที่จะใช้งานได้จริง โดยการทดสอบนั้นได้ทำการทดสอบกับตัวอักษรพิมพ์ และตัวอักษรที่เป็นป้าย ป้ายแสงน้อย และป้ายเอียง ผลการทดสอบพบว่า

การทดสอบก่อนการเพิ่มความแม่นยำของข้อความที่พิมพ์อยู่ที่ 85% หลังจากที่เพิ่มความแม่นยำแล้วเพิ่มขึ้นเป็น 98.5% ส่วนป้ายดังเดิมก่อนการเพิ่มความแม่นยำอยู่ที่ 71.5% หลังจากที่เพิ่มความแม่นยำแล้ว 97% สำหรับป้ายแสงน้อยพบว่าก่อนการเพิ่มความแม่นยำอยู่ที่ 71% หลังการเพิ่มความแม่นยำแล้ว 92% นอกจากนี้เมื่อนำโปรแกรมไปทดสอบกับป้ายที่มีลักษณะเอียงพบว่าก่อนเพิ่มความแม่นยำอยู่ที่ 71% และหลังเพิ่มความแม่นยำอยู่ที่ 93% ซึ่งถือว่ามีประสิทธิภาพสูง

บทที่ 5
บทสรุป และข้อเสนอแนะ

ในบทนี้จะเป็นการอภิปรายเพื่อสรุปผลที่ได้จากการทดสอบงานวิจัยในครั้งนี้ รวมทั้งจะกล่าวถึงที่มาของงานวิจัย การพัฒนาโปรแกรม ผลที่ได้จากทดสอบโปรแกรม และข้อเสนอแนะสำหรับแนวทางในการพัฒนางานวิจัยนี้ต่อไป เพื่อแก้ไขข้อบกพร่องของโปรแกรมให้มีประสิทธิภาพมากขึ้น

5.1 สรุปผลการวิจัย

จากปัญหาทั้งหมดที่ผู้วิจัยได้นำเสนอในบทที่ 1 ส่วนของบทนำจะเห็นได้ว่าการเดินทางท่องเที่ยวในประเทศไทยและอาจจะส่งผลต่อการกลับมาเที่ยวซ้ำที่เมืองไทยอีกครั้งของนักท่องเที่ยว ซึ่งเป็นผลสืบเนื่องมาจากการไม่รู้ภาษาไทย การออกแบบและประยุกต์ใช้งานโปรแกรมประยุกต์เพื่อให้นักท่องเที่ยวชาวจีนได้ใช้แอพลิเคชั่นดังกล่าวในการแปลภาษาไทยได้ถูกพัฒนาขึ้นมาเพื่อที่จะแก้ปัญหา ดังกล่าว ระบบนี้พัฒนาบนโปรแกรมแอพมือถือใช้ Tesseract OCR เพิ่มความแม่นยำตัวอักษรไทย ดังนั้นโปรแกรมประยุกต์เพื่อให้นักท่องเที่ยวชาวจีนได้ใช้แอพลิเคชั่นดังกล่าวในการแปลภาษา เป็นประโยชน์ต่อผู้ใช้

ซึ่งผลการพัฒนาโปรแกรมพบว่าการทดสอบก่อนการเพิ่มความแม่นยำของข้อความที่พิมพ์อยู่ที่ 85% หลังจากที่เพิ่มความแม่นยำแล้วเพิ่มขึ้นเป็น 98.5% ส่วนป้ายดังเดิมก่อนการเพิ่มความแม่นยำอยู่ที่ 71.5% หลังจากที่เพิ่มความแม่นยำแล้ว 97% สำหรับป้ายแสงน้อยพบว่าก่อนการเพิ่มความแม่นยำอยู่ที่ 71% หลังการเพิ่มความแม่นยำแล้ว 92% นอกจากนี้เมื่อนำโปรแกรมไปทดสอบกับป้ายที่มีลักษณะเอียงพบว่าก่อนเพิ่มความแม่นยำอยู่ที่ 71% และหลังเพิ่มความแม่นยำอยู่ที่ 93% ซึ่งถือว่ามีประสิทธิภาพสูง

5.2 อภิปรายผลการวิจัย

การพัฒนาโปรแกรมพบว่าการทดสอบก่อนการเพิ่มความแม่นยำของข้อความที่พิมพ์มีความแม่นยำน้อยโดยหลังจากที่เพิ่มความแม่นยำแล้วความถูกต้องเพิ่มขึ้นมาก ในส่วนของป้ายก่อนการเพิ่มความแม่นยำมีค่าต่ำแต่หลังจากที่เพิ่มความแม่นยำแล้วมีความถูกต้องสูง สำหรับป้าย

แสงน้อยพบว่าก่อนการเพิ่มความแม่นยำมีความถูกต้องต่ำ และหลังการเพิ่มความแม่นยำแล้วมีความถูกต้องเพิ่มขึ้น นอกจากนี้เมื่อนำโปรแกรมไปทดสอบกับป้ายที่มีลักษณะเอียงพบว่าก่อนเพิ่มความแม่นยำมีความถูกต้องต่ำเมื่อเพิ่มความแม่นยำอยู่ความถูกต้องสูงขึ้นมาก ซึ่งถือว่าการพัฒนาโปรแกรมได้อย่างมีประสิทธิภาพสูง สอดคล้องกับ Martinez Carballido (2011) ที่เสนอวิธีการรู้จำเฉพาะตัวเลขโดยการใช้แม่แบบ (Template) ของตัวเลขขนาด 5x7 พิกเซล แล้วนำไปเปรียบเทียบกับแม่แบบวิธีนี้มีความถูกต้องในการรู้จำตัวเลขร้อยละ 97.3 สอดคล้องกับ Abdurrahman (2012) ได้เสนอวิธีการ OCR Tesseract คือการจับภาพบนสมาร์ทโฟน ซึ่งประยุกต์ใช้กับกล้องโทรศัพท์มือถือผลลัพธ์ที่ได้มีความถูกต้องร้อยละ 70

5.3 ข้อจำกัดและแนวทางแก้ไขของงานวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อทดสอบผลกระทบของการรู้จำภาษาไทย โดยการถ่ายภาพป้ายจากปกติ ป้ายเอียง ป้ายแสงน้อย พิมพ์ แสดงผลการทดลอง ป้ายปกติ พิมพ์ ความแม่นยำสูง แต่ป้ายเอียง ป้ายแสงน้อย ความแม่นยำไม่ค่อยสูงซึ่งงานวิจัยนี้ปรับปรุงฟังก์ชันการประมวลผลล่วงหน้าเพื่อเพิ่มความแม่นยำ

5.4 ข้อเสนอแนะและแนวทางการพัฒนาในอนาคต

จากผลสรุปในงานวิจัยนี้ ยังมีข้อจำกัดบางประการซึ่งอาจจะทำให้การรู้จำภาพป้ายยังไม่ถูกต้องเท่าที่ควร โดยข้อจำกัดดังกล่าวนั้นสามารถที่จะนำมาเพิ่มเป็นประเด็นสำหรับวิจัยครั้งต่อไปได้ดังนี้

1) สำหรับในงานวิจัยนี้ ป้ายที่จัดเก็บมาจากป้ายถนนเท่านั้น โดยหากนำมาใช้ในด้านอื่น ๆ ผลที่ได้รับอาจจะไม่ถูกต้องเท่าที่ต้องการจึงมีข้อเสนอแนะสำหรับการวิจัยครั้ง ต่อ ๆไปคือ การเพิ่มขอบเขตของจำนวนภาพป้าย เพื่อที่จะเพิ่มขอบเขตของการรู้จำภาพให้กว้างมากขึ้น

2) นอกจากจะขยายขอบเขตของจำนวนภาพป้ายแล้ว เพื่อที่จะพัฒนาความถูกต้องในการรู้จำภาพป้ายยังสามารถทำได้โดยการเพิ่มจำนวน font ที่นำมาสร้าง Libraryในด้านเดิมในงานวิจัยครั้งต่อไป เพื่อให้การรู้จำภาพป้ายมีความถูกต้องมากขึ้น

3) ในงานวิจัยครั้งนี้ การออกแบบระบบแอพมือถือมีแค่ ฟังก์ชัน Photograph,Translate (Via Baidu),Vocabularyเท่านั้น ซึ่งจะยังไม่มีฟังก์ชันอื่นๆ ทำให้ผู้ใช้ไม่สะดวกสบาย ดังนั้นในงานวิจัยครั้งต่อไปมีการกำหนดให้ระบบมี ฟังก์ชันมากขึ้น

4) ในงานวิจัยครั้งนี้ ระบบแค่ใช้ได้โทรศัพท์มือถือแอนดรอยด์เวอร์ iOSยังใช้ไม่ได้ ดังนั้นในงานวิจัยครั้งต่อไปจะพัฒนาระบบ iOS เพื่อให้ผู้ใช้ใช้ได้สะดวกสบาย

บรรณานุกรม

บรรณานุกรม

สารสนเทศจากสื่ออิเล็กทรอนิกส์

[1]Peris, Álvaro, Cebrián, Luis, Casacuberta F . Online Learning for Neural Machine Translation Post-editing[J]. 2017.

[2]Wang Y L , Xu Z J , Li S S . Based on HTTP Protocol Android Phones to Access the Web Server[J]. Applied Mechanics and Materials, 2014, 543-547:3441-3444.

[3]“Thai OCR : Thai Optical Character Recognition”, [ออนไลน์] เข้าถึง ได้จาก : http://thaiocr.phaisarn.com/introduce.html(วันที่สืบค้นข้อมูล : 8 กันยายน 2553).

[4]วันทนีย์ พันธชาติ ,ผศ.พุทธชาติ โปธิบาล , อรอินทรา ภู่ประเสริฐ, น้ า หนึ่ง มิตรสมาน, ปิยดา สบายใจ, สรินยา ชมพูบุตร, “อุปกรณ์สิ่งอ านวย ความสะดวกส าหรับเด็กที่มีปัญหาการเรียนรู้ ในต่างประเทศ”, [ออนไลน์] เข้าถึงได้จาก: http://www.nectec.or.th/index.php?option=com_content&view=article&id=441:2010-03-04-09-13-08&catid=40:technology-news&Itemid=165 (วันที่สืบค้นข้อมูล : 10 กันยายน 2553

บทความ

[5]Absar J. Programming for the Android Dalvik Virtual Machine[M]. 2017.

[6]Driss S B , Soua M , Kachouri R , et al. A comparison study between MLP and Convolutional Neural Network models for character recognition[C] SPIE Commercial + Scientific Sensing and Imaging. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2017.

[7]Hao Rao, Yuman Lin, Xiaofeng Chen. Implementation of English-Chinese Translation Based on Android OCR[C] 2015.

[8]He S , Hu X . Chinese Character Recognition in Natural Scenes[C] International Symposium on Computational Intelligence & Design. IEEE, 2017.

บรรณานุกรม (ต่อ)

[9]Kim Y J , Cho S J , Kim K J , et al. Benchmarking Java application using JNI and native C application on Android[C] Control, Automation and Systems (ICCAS), 2012 12th International Conference on. IEEE, 2012.

[10]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.

[11]Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study

[12]尹磊. 基于Android平台文字识别应用的设计与实现[D].北京邮电大学,2016.

[13]Elie Krevat and Elliot Cuzzillo., “Improving Off-line Handwritten Character Recognition with Hidden Markov Models”.

[14]Wojciech Bieniecki, Szymon Grabowski and Wojciech Rozenberg, “Image Preprocessing for Improving OCR Accuracy”, MEMSTECH2007, 2007.

ภาษาไทย

[15]ณัฐธิดา ลีสม , โอฬาริก สุรินต๊ะ, “การตัดตัวอักษรลายมือเขียน ภาษาไทยออกจากเอกสารภาพเชิงดิจิทัล”, สาขาวิชาระบบสารสนเทศเพื่อ การจัดการ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม.

[16]นราพงษ์ อภิรัตน์วรากุล, นเรศ เกื้อปัญญากูล, “การศึกษาหลักการ การจดจ าตัวอักษรไทย”, ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น, 2548.

[17]]พงศธรครเกษมและฉัตรชัย ศุภพิทักษ์สกุล,การจำแนกประเภทและรู้จำป้ายทะเบียนรถยนต์ในประเทศไทยโดยใช้เทคนิคเคเนียเรสเนเบอร์, Walailak Procedia,2018

ประวัติผู้เขียน

ชื่อ-นามสกุล	Zhou Hongying
ประวัติการศึกษา	หลักสูตรบริหารธุรกิจบัณฑิต
	สาขาธุรกิจระหว่างประเทศ
	คณะบริหารธุรกิจ มหาวิทยาลัยธุรกิจบัณฑิตย์

DPU