บทความวิจัย

Computer Science Department
Faculty of Informatics, Mahasarakham University

โปรแกรมแปลงใบเสร็จรับเงินอัตโนมัติ
Automated Receipt Digitization

พิจิตรา เหราวรรณ์ , สุธาทินี สุมังคะเสริฐ, ระพีพร ช่ำชอง
สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
61011212053@msu.ac.th, 61011212071@msu.ac.th, rapeeporn.c@msu.ac.th

บทคัดย่อ

โครงงานนี้เป็นการวิจัยและพัฒนาการประมวลผลข้อมูลใบเสร็จที่เป็นข้อมูลภาพไปเป็นข้อมูลตัวอักษร เพื่อศึกษาและพัฒนาเทคนิคการตรวจจับข้อมูลและข้อความบนใบเสร็จ โดยในงานนี้จะใช้หลักการ Image Processing มาใช้ในขั้นตอน Pre-process เป็นขั้นตอนการเตรียมข้อมูลก่อนเข้าไปประมวลผลและทดสอบอัลกอริทึมที่ใช้ในการรู้จำตัวอักษร โดยใช้การประมวลผลภาพจากการทดลองเพื่อประเมินประสิทธิภาพของ Model สถาปัตยกรรม Convolutional Neural Network (CNN) ร่วมกับ Recurrent Neural Network (RNN) ในการรู้จำตัวอักษร ซึ่งเลือกใช้ Optimizers 3 ประเภท คือ RMSprop , Adam และ SGD ตามลำดับเพื่อหาผลการเรียนรู้ที่มีประสิทธิภาพที่ดีที่สุด จึงนำสถาปัตยกรรมดังกล่าวมาประยุกต์ใช้กับ Application และ Website เพื่อเป็นการเผยแพร่การใช้งานโปรแกรมประมวลผลการแปลงใบเสร็จรับเงินอัตโนมัติ (Automated Receipt Digitization) ต่อไป

คำสำคัญ: การรู้จำตัวอักษรจากรูปภาพ , Convolutional Neural Network , Recurrent Neural Network , Character Recognition

1.บทนำ

ในปัจจุบันมีเทคโนโลยีที่ถูกพัฒนาเพื่อสนองความต้องการของมนุษย์เพิ่มมากขึ้น เทคโนโลยีเหล่านั้นถูกพัฒนาขึ้นมาเพื่อจัดการกับปัญหาต่าง ๆ ที่พบเจอในชีวิต ตัวอย่างของปัญหาหนึ่งคือเรื่องของการจัดการกับข้อมูลด้านการเงิน เช่น ใบเสร็จรับเงินที่มาจากหลายที่ ทำให้องค์กรต่าง ๆ ต้องเสียทรัพยากรแรงงานและเงินจำนวนมากในการจัดการข้อมูลใบเสร็จเหล่านี้ลงในฐานข้อมูลระบบบัญชีในองค์กรของตนเอง นอกจากนั้นการใช้แรงงานที่เป็นมนุษย์นั้นอาจเกิดข้อผิดพลาดได้ในการนำข้อมูลเข้าจากเหตุผลบางประการเช่น การกรอกข้อมูลผิดพลาด หรือแม้แต่การทำงานที่ล่าช้าในการกรอกข้อมูล ดังนั้นถ้าหากมีการนำเอาเทคโนโลยีเข้ามาช่วยในการแปลงข้อมูลใบเสร็จที่เป็นรูปภาพไปเป็นข้อมูลตัวอักษรที่สามารถบันทึกลงฐานข้อมูลได้อย่างง่ายและรวดเร็ว จะช่วยให้องค์กรประหยัดทรัพยากรและเวลาในการจัดการข้อมูลใบเสร็จเหล่านี้ลงอย่างมาก

ด้วยเหตุผลดังกล่าวคณะผู้พัฒนาโครงงานจึงได้มีการพัฒนาการประมวลผลข้อมูลใบเสร็จที่เป็นข้อมูลภาพไปเป็นข้อมูลตัวอักษรแล้ว ซึ่งสามารถนำไปประยุกต์เพื่อบันทึกลงบนฐานข้อมูลได้แบบอัตโนมัติ โดยงานนี้มีลักษณะของการแปลงข้อมูลภาพไปเป็นตัวอักษรแบบอัตโนมัติ ซึ่งมีการดำเนินงาน 3

Computer Science Department
Mahasarakham University

ขั้นตอนหลัก ได้แก่ (1) การตรวจจับข้อความ (Text detection) (2) การรู้จำจำข้อความ (Text recognition) และ (3) การจำแนกข้อมูล (Information Extraction)

โครงการนี้จะช่วยลดระยะเวลาในการจัดการข้อมูลใบเสร็จและลดข้อผิดพลาดของข้อมูลในใบเสร็จที่อาจเกิดขึ้นได้โดยนำไปประยุกต์ใช้ในระบบบัญชีขององค์กร ซึ่งหากมีการนำโครงการนี้เข้าไปช่วยจะให้องค์กรประหยัดเวลาและต้นทุนในการจัดการข้อมูลใบเสร็จเหล่านี้มากยิ่งขึ้น

2.ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

การตรวจจับข้อความบนฟอร์มใบเสร็จได้ใช้ทฤษฎีการประมวลผลภาพเบื้องต้น มาช่วยจัดการงานทางด้านการเตรียมข้อมูลก่อนเข้าไปเรียนรู้เพื่อให้การรู้จำตัวข้อความและการสกัดข้อมูลทำงานได้ดียิ่งขึ้น โดยใช้ทฤษฎีประมวลผลเบื้องต้นดังต่อไปนี้

หลังจากอ่านไฟล์ภาพเข้ามาแล้วจะนำภาพไปทำ Median Blur เพื่อจำกัดแสงเงาและสิ่งรบกวนต่างๆ ทำให้ภาพ Smoot ขึ้น จากนั้นนำภาพที่ได้ไปแปลงเป็นภาพระดับเทา แล้วนำไปแปลงเป็นภาพสองระดับด้วย Thresholding จากนั้นนำภาพขาวดำไปทำ Dilation แล้วจึงนำภาพนั้นไปหาพิกัดของบรรทัดและนำไป Crop เพื่อเป็น Input ของกระบวนการรู้จำตัวอักษรต่อไป

การรู้จำตัวอักษร จะใช้ Tesseract OCR ย่อมาจาก Optical Character Recognition ซึ่งเป็นกระบวนการของการแปลงสื่อสิ่งพิมพ์ เช่น กระดาษ นิตยสาร สัญญา หรือข้อมูลอะไรก็ตามที่อยู่ในรูปของเอกสารกระดาษ ให้กลายเป็นข้อความให้มีความฉลาดมากขึ้นกว่าการเป็นข้อความธรรมดา โครงสร้างของระบบ OCR ประกอบไปด้วยขั้นตอนการทำงานหลัก 2 ขั้นตอน ได้แก่

1.การประมวลผลขั้นต้น (Pre-process) เช่น การปรับแต่งข้อมูล (Normalization) การ กรองข้อมูลแทรกซ้อน (Noise Filtering) การตรวจจับวัตถุ (Object Detection) เป็นต้น

2.การรู้จำตัวอักษร (Character Recognition) เช่น วิธีทางโครงข่ายประสาทเทียม และ การเรียนรู้เชิงลึก เป็นต้น

การเรียนรู้เชิงลึก เป็นวิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาทเทียม (Neural Network) มา ซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งการเรียนรู้เชิงลึกที่จะกล่าวถึงในที่นี้ได้แก่ Convolutional Neural Network และ Recurrent Neural Network มีรายละเอียด ดังนี้

1. โครงข่ายประสาทเทียมแบบคอนโวลูชัน [3] (Convolutional Neural Network, CNN) สามารถประมวลผลกับรูปภาพได้มีประสิทธิผลมาก โดย ภายในระบบ CNN จะประกอบไปด้วย 3 Layer ได้แก่

- ชั้นคอนโวลูชัน (Convolution Layer) เป็น Layer ทำหน้าที่ในการสกัด Feature ออกมาจากรูปภาพ

- ชั้นพูล (Pooling layer) เป็นชั้นที่ทำหน้าที่ในการปรับขนาดและปริมาณของข้อมูล ตัวอย่าง (Sample)

- ชั้นฟูลลีคอนเนก (Fully connected layer) เป็น Layer ที่ประกอบด้วยระบบ Multilayer perceptron (MLP) ในการประมวลผลข้อมูลที่ได้มาจาก 2

Computer Science Department
Faculty of Informatics, Mahasarakham University

layer ก่อนหน้านี้เพื่อสังเคราะห์ และทำการแยกแยะรูปภาพออกเป็นหมวดหมู่

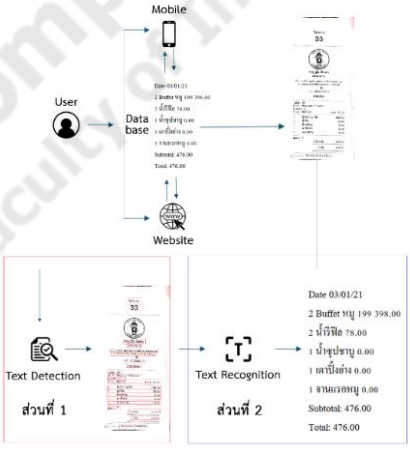
2. Recurrent Neural Network (RNN) [8] การทำงานของ RNN คือการเอาผลลัพธ์ที่ได้จากการคำนวณย้อนกลับมาใช้เป็นข้อมูลขาเข้าอีกครั้ง ซึ่งมีประโยชน์อย่างมากในข้อมูลที่มีความต่อเนื่อง เช่น ข้อมูลเสียง ข้อความ หรือแม้แต่รูปภาพเองก็ตาม

3. Long Short-Term Memory (LSTM) [10] เป็นโครงข่ายประสาทเทียมที่เกิดซ้ำซึ่งแก้ปัญหาการไล่ระดับสีที่หายไป ของ RNN ลักษณะที่โดดเด่นของ LSTM คือ การที่สามารถเลือกได้ว่าข้อมูล ไหนควรที่จะจดจำข้อมูลไหนควรกำจัดทิ้ง

4. Gated Recurrent Units (GRU) [6] มีกลไกภายในที่ เรียกว่า Gate เพื่อควบคุมการไหลของข้อมูลเช่นการจดจำบริบทในช่วงเวลาต่างๆ ติดตามว่าข้อมูลใดใน อดีตสามารถเก็บสิ่งที่ลืมได้ ซึ่งพัฒนาต่อมาจาก LSTM เพื่อลดขั้นตอนการทำงานภายใน Gate และช่วย ให้การประมวลผลเร็วขึ้น

3.ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

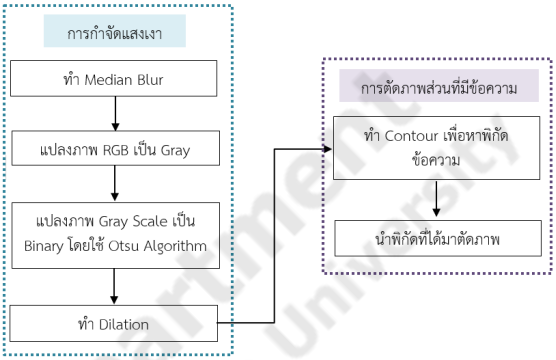
3.1 กรอบการดำเนินงาน



ภาพประกอบที่ 1 กรอบการดำเนินงานของระบบ

จากภาพประกอบที่ 3.1 จะเห็นว่าได้มีการแบ่งขั้นตอนการทำงานออกเป็น 2 ขั้นตอนหลัก ได้แก่

1. การตรวจจับข้อความ



ภาพประกอบที่ 2 ขั้นตอนการตรวจจับข้อความ

ขั้นตอนนี้คือขั้นตอนการเตรียมภาพก่อนนำไปรู้จำตัวอักษร โดยนำภาพที่เป็น Input ไปเข้าสู่กระบวนการ Pre-process เพื่อจำกัดแสงเงาและสิ่งรบกวนต่างๆบนรูปภาพที่อาจส่งผลต่อการรู้จำตัวอักษร โดยในขั้นตอนแรกของกระบวนการ Pre-process

2. การรู้จำตัวอักษร

ขั้นตอนนี้คือระบบจะทำการรู้จำตัวอักษร โดยภาพที่ถูกนำเข้าจะเข้าสู่กระบวนการทำ Pre-processing เพื่อเตรียมข้อมูลก่อนเข้าสู่ Model ที่เป็นกระบวนการในการรู้จำ แล้วจึงได้ผลลัพธ์ออกมา

3.2 การตรวจจับข้อความ

1. ภาพนำเข้า เป็นภาพใบเสร็จสี RGB ที่ได้มาจากการสแกน หรือการถ่ายรูปพื้นหลังของ ภาพไม่มีเส้นหรือบรรทัดลวดลายใดๆ

2. การเตรียมพร้อมภาพ (Pre-Processing เป็นขั้นตอนในการเตรียมภาพนำเข้าเพื่อให้พร้อมต่อการนำไปประมวลผลต่อไป โดยจะ ประกอบไปด้วยขั้นตอนในการ

Computer Science Department
Faculty of Informatics, Mahasarakham University

กำจัดแสงเงา การแปลงภาพให้เป็นไบนารี ขาว - ดำ และตัดรูปภาพให้มี เฉพาะตรงส่วนที่มีข้อความอยู่

3.3 ขั้นตอนการเรียกใช้งาน Tesseract OCR

วิธีการดำเนินการ การสกัดตัวหนังสือภาษาไทยและภาษาอังกฤษ โดยในขั้นตอนแรกนำ ภาพใบเสร็จเข้าไป จากนั้นเขียน Code เพื่อทดสอบการทำงานและตรวจสอบผลลัพธ์

Table: 20
Staff: Pimpavee Pimpavee
Guests: 2 ID: 5C5XQ
Date: 10/01/21 Time: 19:42

2	Buffet หมู 199	398.00
2	น้ำรีฟิล	78.00
1	น้ำซุปซานู	0.00
1	เตาปิ้งย่าง	0.00
1	จานแรกหมู	0.00

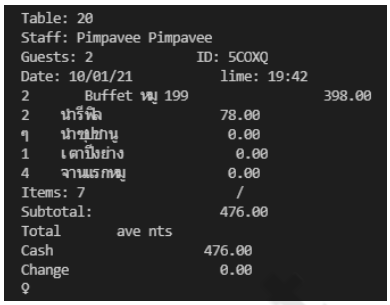
Items: 7
Subtotal: 476.00
Total: 476.00
Cash 476.00
Change 0.00
Powered by FoodStory

ภาพประกอบที่ 3 ตัวอย่างภาพใบเสร็จ

```
#Text from Image
data_Img = pytesseract.image_to_string(cropX, lang = 'Tha+Eng')
print(data_Img)
```

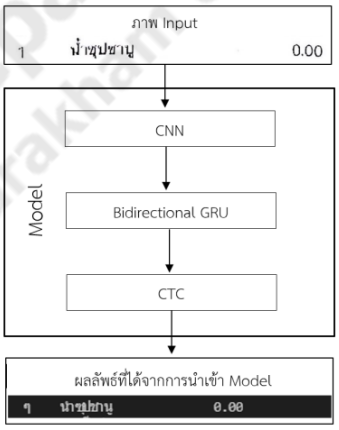
ภาพประกอบที่ 4 ตัวอย่าง Code และ การสกัดตัวอักษรภาษาไทยและภาษาอังกฤษ

โดยใช้ pytesseract สกัดข้อความจากภาพ ซึ่ง parameter lang= ‘tha+eng’ หมายถึง การใช้ Training Dataset ในการสกัดตัวอักษรไทยและอังกฤษจากภาพ จะพบว่าสามารถสกัดข้อความออกมาได้ถูกต้องบางคำ


```
Table: 20
Staff: Pimpavee Pimpavee
Guests: 2              ID: 5COXQ
Date: 10/01/21          lime: 19:42
2      Buffet หมู 199                398.00
2    น้ำรีฟิล                78.00
ๆ    น้ำซุปซานู               0.00
1    เตาปิ้งย่าง              0.00
4    จานแรกหมู               0.00
Items: 7                      /
Subtotal:                   476.00
Total       ave nts
Cash                      476.00
Change                      0.00
Q
```

ภาพประกอบที่ 5 ผลลัพธ์หลังจากทดลองใช้งาน Tesseract

3.4 การรู้จำข้อความ (Text Recognition)



ภาพประกอบที่ 6 ขั้นตอนการทำงานของกระบวนการรู้จำตัวอักษร

เมื่อผ่านกระบวนการ Pre-Processing แล้ว ผลลัพธ์ที่ได้จากขั้นตอนดังกล่าวจะถูกส่งเข้าสู่กระบวนการรู้จำตัวอักษร

3.5 การวัดประสิทธิภาพ

การประเมินความถูกต้องของการรู้จำตัวอักษร โดยใช้ Levenshtein Edit Distance [11] ซึ่งเป็นขั้นตอนการวัดเพื่อหาค่าความแตกต่างของข้อความทั้งสองชุดระหว่างชุดแรกที่เป็นผลเฉลย และชุดที่สองที่เป็นชุดผลการทำนาย โดยจะนับจำนวนครั้งของการแทรก การตัด และการแทนที่

3.6 ชุดข้อมูลที่ใช้ในการทดลอง

ชุดข้อมูลที่ใช้ในการสร้างโมเดลรู้จำ

Computer Science Department Faculty of Informatics, Mahasarakham University

500 ภาพจะแบ่งออกเป็นชุดสอน 60% ชุดวาลิเดท 20% และชุดทดสอบ 20% ซึ่งเป็นใบเสร็จของร้านชาบู อู๊ด เป็นต่อ และนำมาตัดเป็นข้อมูล 1 บรรทัด

ภาพประกอบที่ 7 ตัวอย่างข้อมูลใบเสร็จของร้านชา บู อู๊ดเป็นต่อ (ชุดข้อมูล 1 ภาพ)

ภาพประกอบที่ 8 ตัวอย่างชุดข้อมูลใบเสร็จของร้านชาบู อู๊ด เป็นต่อ (ชุดข้อมูล 1 บรรทัด)

ภาพประกอบที่ 9 ตัวอย่างชุดข้อมูลใบเสร็จของร้านชาบู อู๊ด เป็นต่อ (ชุดข้อมูล 1 คำ)

4. ผลการทดลอง

4.1 ผลการทดลองขั้นตอนการ Pre-process

ในขั้นตอนนี้ จะแสดงผลการทดลองการเตรียมภาพเพื่อนำไปประมวลผล โดยประกอบไปด้วย 5 ขั้นตอน

1.ภาพ input

2.การทำ Median Blur

3.แปลงภาพ RGB เป็น Gray Scale

4.แปลงภาพ Gray Scale เป็น Binary โดยใช้ Otsu Algorithm

5.ทำ Dilation

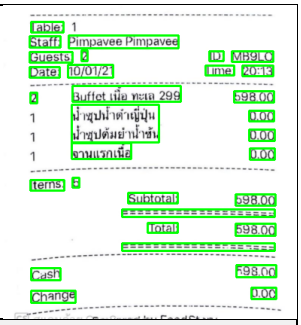
4.2 การตัดภาพส่วนที่มีข้อความ

ในขั้นตอนนี้จะทำการตัดภาพให้เหลือเฉพาะส่วนที่มีข้อความอยู่เท่านั้น โดยทำการหาพิกัดเริ่มต้นของภาพ ความกว้าง และความสูงของภาพตรงส่วนที่มีข้อความปรากฏ แล้ว crop ตัดรูปภาพนั้น ตามพิกัดที่ได้มา โดยจะยกตัวอย่างผลลัพธ์ดังตารางต่อไปนี้

ตารางที่ 1 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 1

ภาพที่ 1	
ทำ Contour เพื่อหาพิกัดข้อความเพื่อ crop ภาพ	

Computer Science Department
Faculty of Informatics, Mahasarakham University

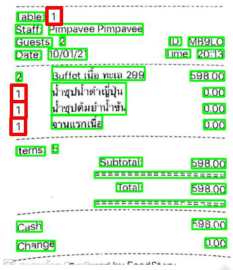
นำพิกัดความสูงของรูปภาพมาหารสอง เพื่อทำการ crop ภาพ	

รายละเอียด			
การตรวจจับข้อความทั้งหมด	ข้อความรู้จำถูกต้อง	การตรวจจับข้อความผิดพลาด	ตรวจจับข้อความไม่เจอ
36	30	1	4

- การตรวจจับข้อความผิดพลาด 1 ข้อความ เป็นเพราะตัวอักขระที่อยู่ชิดกันจึงทำให้เกิดการตรวจจับข้อความผิดพลาด

ภาพประกอบที่ 10 การตรวจจับข้อความผิดพลาด ภาพที่ 1

- ตรวจจับข้อความไม่ได้ ได้แก่ ส่วนของเลข 1 จำนวน 4 ข้อความ เป็นเพราะเลข 1 เป็นตัวอักษรที่มีขนาดเล็กเกินไป จึงทำให้ตรวจจับข้อความไม่ได้ ภาพประกอบที่ 4.2 การตรวจจับข้อความไม่ได้ ภาพที่ 1



ภาพประกอบที่ 11 การตรวจจับข้อความไม่ได้ ภาพที่ 1

4.3 การเปรียบเทียบประสิทธิภาพการรู้จำโดยการปรับค่า Optimizers

4.3.1 การประเมินประสิทธิภาพ Model

ในการแปลงใบเสร็จรับเงินติโนมัตินี้ ได้ใช้สถาปัตยกรรม CNN ร่วมกับ RNN โดยได้ทำการเลือกใช้ Optimizers 3 ประเภท ได้แก่ RMSprop , Adam และ SGD เพื่อหาผลการเรียนรู้ที่มีประสิทธิภาพที่ดีที่สุด โดยในการทดลองนี้ ได้ใช้ข้อมูลในการเรียนรู้ทั้งหมด 17,028 ภาพ และทำการแบ่งชุดข้อมูลออกเป็น 3 ชุดได้แก่ ชุดเรียนรู้ 60% ชุดวาลิเดชัน 20% และชุดทดสอบ 20% ใช้จำนวนรอบ (epoch) ในการรู้จำทั้งหมด 100 รอบ

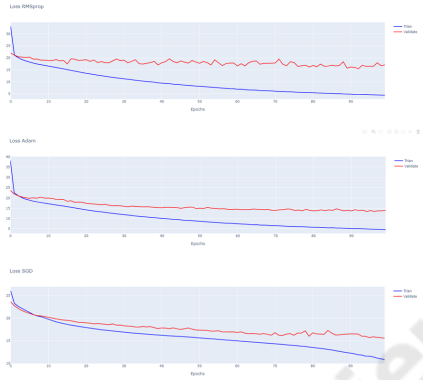
ตารางที่ 2 รูปแบบ Model ที่ทำการทดลองทั้ง 2 Model

Model
Input Image 150 x 32
CNN 2 Layer (32,64)
RNN 2 Layer (128,64)
SoftMax Activation ReLU Optimizer
CTC
Output

Computer Science Department Faculty of Informatics, Mahasarakham University

ตารางที่ 3 ผลการประเมินประสิทธิภาพค่า CER ที่ได้จากการเรียนรู้ Optimizer ทั้ง 3 ประเภท

Optimizer	CER (%)
RMSprop	20.3907
Adam	20.0639
SGD	21.0043

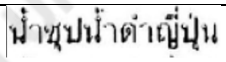


ภาพประกอบที่ 12 กราฟแสดงค่า Loss ที่ได้จากการเรียนรู้ด้วย Optimizers 3 ประเภท ได้แก่ RMSprop , Adam และ SGD ตามลำดับ

จากกราฟผลการเรียนรู้ข้างต้น จะเห็นว่าการเรียนรู้โดยใช้ Optimizer RMSprop และ Adam พบว่าค่า Loss มีการลดลงอย่างรวดเร็ว ส่วน Optimizer SGD นั้นพบว่าค่า Loss มีอัตราการลดลงค่อนข้างช้ากว่า 2 วิธีข้างต้น ซึ่งหากต้องการให้ค่า Loss ลดลงอาจจะต้องเพิ่มจำนวนรอบในการเรียนรู้มากขึ้น ซึ่งจะต้องใช้เวลาในการเรียนรู้มากขึ้น นอกจากนี้จะเห็นว่าการเลือกใช้ SDG อาจก่อให้เกิด Under-fitting (ชุดวาลิเดท มีค่า Loss ต่ำกว่า ชุดเรียนรู้) หากจำนวนรอบน้อยเกินไป

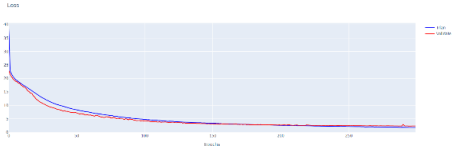
จากผลการเรียนรู้ด้วย Optimizer ทั้ง 3 ประเภท จากการเรียนรู้เป็นจำนวน 100 รอบ พบว่า Adam นั้นมีค่า Character Error Rate (CER) ต่ำสุด และ SDG มีค่าสูงสุด โดยการทดลองดังกล่าว สามารถแสดงภาพตัวอย่างที่ได้จากการทำนายผลลัพธ์โดยยกตัวอย่าง ได้ดังตารางต่อไปนี้

ตารางที่ 4 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 1

ภาพต้นฉบับ	
ผลเฉลย	น้ำซุปน้ำดำญี่ปุ่น
Optimizer RMSprop	น้ำซุปดิป่น
Optimizer Adam	น้ำซุปมำน
Optimizer SGD	นำซปย

4.4 ผลลัพธ์ประสิทธิภาพการรู้จำจาก Model Optimizer Adam

จากการทดลองที่ 4.3 พบว่า Optimizer Adam ให้ประสิทธิภาพที่ดีที่สุดจึงเลือกใช้วิธีดังกล่าว และได้ใช้ข้อมูลในการเรียนรู้ทั้งหมด 20,611 ด้วยชุดข้อมูลสำหรับเรียนรู้ 80% ชุดวาลิเดชัน 10% และชุดทดสอบ 10% จากข้อมูลทั้งหมด แสดงผลการฝึกทั้ง 300 รอบดังตารางต่อไปนี้



ภาพประกอบที่ 13 ภาพกราฟแสดงค่า Loss ที่ได้จาก Model

Computer Science Department Faculty of Informatics, Mahasarakham University

จากผลการเรียนรู้ในการรู้จำนี้ จึงนำไปใช้ในการประเมินประสิทธิภาพ แสดงผลลัพธ์ดังตารางต่อไปนี้

ตารางที่ 5 ตัวอย่างผลลัพธ์การทำนาย ภาพที่ 1

ใบเสร็จที่ 1	การตรวจจับข้อความ	การตรวจจับความถูกต้อง		ข้อความรู้จำถูกต้อง	
	จำนวนข้อความทั้งหมด	จำนวนข้อความ	ความถูกต้อง (%)	จำนวนข้อความ	ความถูกต้อง (%)
	29	26	89.65	9	31.03
ผลเฉลย	Table: 38 Staff: Pimpavee Pimpavee Guests: 2 Date: 10/01/21 Time 19:48 2 Buffet เนื้อ ทะเล 299 598.00 น้ำซุปน้ำต๋ำญี่ปุ่น 0.00 เตาปิ้งย่าง 0.00 จานแรกเนื้อ 0.00 Items: 5 Subtotal: 598.00 Total: 598.00				
ผลทำนายถูกต้อง	['Guests:'] ['2'] ['Date:'] ['598.00'] ['0.00'] ['0.00'] ['0.00'] ['Items:'] ['598.00']				

ตารางที่ 6 ตัวอย่างผลลัพธ์การทำนาย ภาพที่ 2

ใบเสร็จที่ 2	การตรวจจับข้อความ	การตรวจจับความถูกต้อง		ข้อความรู้จำถูกต้อง	
	จำนวนข้อความทั้งหมด	จำนวนข้อความ	ความถูกต้อง (%)	จำนวนข้อความ	ความถูกต้อง (%)
	34	31	91.18	9	26.47
ผลเฉลย	Table: 33 Staff: Pimpavee Pimpavee Guests: 2 ID: 6QG0X Date: 10/01/21 Time 20:17 2 Buffet เนื้อ ทะเล 299 598.00 น้ำซุปชาบู 0.00 เตาปิ้งย่าง 0.00 จานแรกเนื้อ 0.00 Items: 5 Subtotal: 598.00 Total: 598.00				
ผลทำนายถูกต้อง	['Guests:'] ['2'] ['Date:'] ['2'] ['0.00'] ['0.00'] ['0.00'] ['Items:'] ['598.00']				

ตารางที่ 7 ตัวอย่างผลลัพธ์การทำนาย ภาพที่ 3

ใบเสร็จที่ 3	การตรวจจับข้อความ	การตรวจจับความถูกต้อง		ข้อความรู้จำถูกต้อง	
	จำนวนข้อความทั้งหมด	จำนวนข้อความ	ความถูกต้อง (%)	จำนวนข้อความ	ความถูกต้อง (%)
	30	26	86.67	9	30
ผลเฉลย	Table: Staff: Pimpavee Pimpavee Guests: 2 ID: MB9LO Date: 10/01/21 Time 20:13 2 Buffet เนื้อ ทะเล 299 598.00 0.00 0.00 จานแรกเนื้อ 0.00 Items: 5 Subtotal: 598.00 Total: 598.00				
ผลทำนายถูกต้อง	['Guests:'] ['2'] ['Date:'] ['0.00'] ['0.00'] ['0.00'] ['Items:'] ['598.00'] ['598.00']				

4.4 เปรียบเทียบประสิทธิภาพการรู้จำใบเสร็จจาก Tesseract

ในการทดลองนี้ได้ทำการเปรียบเทียบการรู้จำใบเสร็จจากวิธีที่พัฒนากับการใช้ Tesseract โดยได้ผลการทดลอง ดังนี้

ตารางที่ 8 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 1

ภาพต้นฉบับ	Table:	CER (%)
ผลเฉลย	Table:	
Model Optimizer Adam	Table:	100
Tesseract	lable:	84

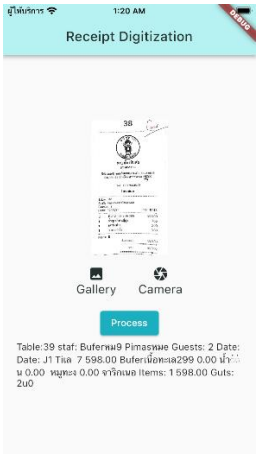
ตารางที่ 9 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 2

ภาพต้นฉบับ	Staff:	CER (%)
ผลเฉลย	Staff:	
Model Optimizer Adam	Staff:	100
Tesseract	Staff:	0

จากการทดลองดังกล่าว ได้ข้อสรุปว่า Model Optimizers Adam มีประสิทธิภาพในการประมวลผลที่ดีกว่าการใช้ Tesseract โดยหากใช้ Tesseract มีค่า CER อยู่ที่ 58.9411% ในขณะที่การใช้ Model Optimizers Adam มีค่า CER อยู่ที่ 17.6620%

Computer Science Department Faculty of Informatics, Mahasarakham University

4.5 ตัวอย่างการทำนายผลบน Mobile Application



ภาพประกอบที่ 14 ตัวอย่างการทำนายผลบน Mobile Application

4.6 ตัวอย่างการทำนายผลบน Website

ภาพประกอบที่ 15 ตัวอย่างการทำนายผลบน Website

4.7 สรุปและวิเคราะห์ผลการทดลอง

จากการทดลองในการหาตัว Optimizers ที่ดีที่สุดโดยทำการทดลองจากภาพจำนวน 17,028 ภาพ พบว่า ตัว Optimizers Adam มีค่า CER คิดเป็น 20.0639 % ซึ่งต่ำที่สุด เมื่อเทียบกับ ตัว Optimizers อื่น การเปรียบเทียบประสิทธิภาพการรู้จำโดยการปรับค่า Optimizers จึงทำให้สรุปได้ว่า Optimizer Adam มีประสิทธิภาพที่ดีกว่า Optimizer อื่น

ในส่วนของการเปรียบเทียบประสิทธิภาพการรู้จำใบเสร็จพบว่า การใช้ Model Optimizer Adam จะทำให้ได้ผลลัพธ์ดีกว่าการใช้ Tesseract ซึ่งมีค่า CER อยู่ที่ 58.9411% ดังนั้น ในงานวิจัยนี้จึงเลือกใช้ Optimizer Adam และทำการทดลองกับชุดข้อมูล 500 ภาพ ตัดเป็นข้อความได้ทั้งหมด 20,611 ภาพ ใช้ชุดข้อมูลสำหรับเรียนรู้ 80% ชุดวาลิเดชัน 10% และชุดทดสอบ 10% ฝึกทั้ง 300 รอบ โดยมีค่า CER คิดเป็น 17.6620% มาประยุกต์ใช้กับ Application และ Website เพื่อเป็นการเผยแพร่การใช้งานโปรแกรมการประมวลผลการแปลงใบเสร็จรับเงินเป็นตัวพิมพ์อัตโนมัติ (Automated Receipt Digitization) ต่อไป

5. สรุปผลและอภิปรายผล

โครงงานปริญญานิพนธ์ฉบับนี้นำเสนอการประมวลผลการแปลงใบเสร็จรับเงินเป็นตัวพิมพ์อัตโนมัติ (Automated Receipt Digitization) โดยใช้การประมวลผลภาพเบื้องต้นและการเรียนรู้เชิงลึกมาใช้งานนี้ ซึ่งสามารถใช้งานการแปลงใบเสร็จรับเงินได้ผ่าน Mobile Application และ Website Application

จากการทดลองวัดประสิทธิภาพการรู้จำตัวอักษรจากฐานข้อมูลรูปภาพรวมกันทั้งสิ้น 20,611 ภาพ โดยใช้ Optimizer Adam

Computer Science Department
Faculty of Informatics, Mahasarakham University

ซึ่งมีประสิทธิภาพในการทำนายผลสูงสุด โดยมีค่า Character Error Rate (CER) ต่ำสุดอยู่ที่ 17.6620%

5.1 ปัญหา และอุปสรรคในการดำเนินงาน

5.1.1 เนื่องจาก Algorithm ที่ใช้ในการรู้จำมีความซับซ้อนสูงมากจำเป็นต้องใช้ทรัพยากรในการประมวลผลสูงและใช้ระยะเวลานาน

5.1.2 ชุดข้อมูลเป็นภาพที่ถ่ายหรือสแกนจากมือถือ มีสัญญาณรบกวนไม่ว่าจะเป็นแสง ขนาดของภาพ คุณภาพของรูปถ่าย ผลลัพธ์ที่ได้จากการแปลงใบเสร็จจึงมีความคลาดเคลื่อน

5.1.3 Algorithm การแปลงใบเสร็จรับเงินเป็นตัวพิมพ์อัตโนมัติจำเป็นต้องมีการทำนายผลเฉลยด้วยตัวเองจึงใช้ระยะเวลานานในการสร้างผลเฉลย

5.2 ข้อเสนอแนะ

5.2.1 ควรพัฒนาเพิ่มเติมในการนำเข้าข้อมูลใบเสร็จให้หลากหลายรูปแบบ

5.2.2 ควรปรับปรุงการกำจัดสัญญาณรบกวน แสงเงา และการปรับปรุงคุณภาพของรูปภาพเพื่อให้ได้คุณภาพที่ดีขึ้น

เอกสารอ้างอิง

1. วิณัฏฐา แสงสุข และ ฐิติพร ลิ้มแหลมทอง, ความรู้เบื้องต้นเกี่ยวกับกฎหมายทั่วไป. กรุงเทพมหานคร: สำนักพิมพ์มหาวิทยาลัยรามคำแหง, 2555.
2. ธีระ สิงหพันธุ์, กฎหมายอาญาภาค 1. กรุงเทพมหานคร : สำนักพิมพ์มหาวิทยาลัยรามคำแหง, 2556.
3. R. Hermawan, *Natural language processing with python*, June 2009., vol. 1, no. 1. United States of America: O'Reilly Media, 2011.
4. R. Zhao and K. Mao, "Fuzzy bag-of-words model for document representation," *IEEE Trans. fuzzy Syst.*, vol. 26, no. 2, pp. 794–804, 2017.
5. N. Durrani and S. Hussain, "Urdu word segmentation," *NAACL HLT 2010 - Hum. Lang. Technol. 2010 Annu. Conf. North Am. Chapter Assoc. Comput. Linguist. Proc. Main Conf.*, no. June, pp. 528–536, 2010.
6. K. Maher and M. S. Joshi, "Effectiveness of Different Similarity Measures for Text Classification and Clustering," *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 4, pp. 1715–1720, 2016.
7. R. Ghawi and J. Pfeffer, "Efficient Hyperparameter Tuning with Grid Search for Text Categorization using kNN Approach with BM25 Similarity," *Open Comput. Sci.*, vol. 9, no. 1, pp. 160–180, 2019, doi: 10.1515/comp-2019-0011.
8. K. W. Boyack *et al.*, "Clustering more than two million biomedical publications: Comparing the accuracies of nine text-based similarity approaches," *PLoS One*, vol. 6, no. 3, 2011, doi:

Computer Science Department Faculty of Informatics, Mahasarakham University

10.1371/journal.pone.0018029.

9. E. Haddi, X. Liu, and Y. Shi, “The Role of Text Pre-processing in Sentiment Analysis,” *Procedia Comput. Sci.*, vol. 17, pp. 26–32, 2013, doi: 10.1016/j.procs.2013.05.005.
10. C. Grover, B. Hachey, I. Hughson, and C. Korycinski, “Automatic summarisation of legal documents,” *Proc. Int. Conf. Artif. Intell. Law*, pp. 243–251, 2003, doi: 10.1145/1047788.1047839.
11. B. Hachey and C. Grover, “Sentence Classification Experiments for Legal Text Summarisation,” *Leg. Knowl. Inf. Syst. Jurix 2004, Seventeenth Annu. Conf.*, no. May, pp. 29–38, 2004.

Computer Science Department
Faculty of Informatics, Mahasarakham University