

INTRODUCTION

In this project we were asked to choose any one of the tasks between task2 and task3 for the – “ICDAR 2019 Robust Reading Challenge on Scanned Receipts OCR and Information Extraction”.

For optical character recognition purposes till now the general model used is Connectionist Text Proposal Network (CTPN) for text recognition on the image and For text line recognition, Convolutional Recurrent Neural Network and Attention-based Encoder-Decoder (AED) have been achieved state-of-the-art for many problems such as Scene Text Recognition, handwritten recognition, etc.

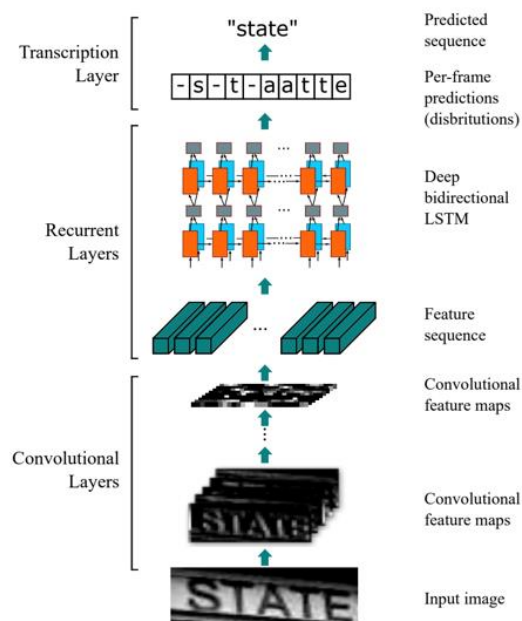
Here I chose task2 and the statement for task 2 (**Scanned Receipt OCR**) is to accurately recognize the text in a receipt image. No localisation information is provided, or is required.

METHODOLOGY

Models

Here I had applied a combination of CNN and RNN. Since this is an image processing task so CNN will play a major role in extracting image features that contain text and through RNN we will be able to find information out of the text within that image.

Here I applied a modified CRNN in this task. CRNN is a conventional scene text recognition method including convolutional layers, bidirectional LSTM layers, and a transcription layer in sequence.



In scanned receipts each text usually contains several words. Here I add the blank space between words to the alphabet for LSTM prediction and thus improve the network from single word recognition to multiple words recognition. Moreover, we double the input image width to tackle the overlap problem of long texts after max-pooling and stack one more LSTM, enhancing the accuracy per character in the training set from 62% to 83%.

Procedure

Here I had performed task2 by taking bounding box from task1 and leveraging it for task2 since we are not required to perform task1. I also taken test data in account which is provided by ICDAR 2019. I had also tried for training the model but unfortunately I was ran out of time since I had performed both the task due to some unfortunate consequences, So I had taken pretrained weights for the model in account. I had used Pytorch framework for this project

I had used google colab for the coding purpose and the flow is-

- Firstly there are some utility functions which are required for handling some distortations in the images, making the dataset ready for prediction purpose, etc.
- Then we have Bidirectional LSTM and CRNN as our models for predicting our result.
- Then we have lmdbDataset class for making data set ready for data loader,resizeNormalize Class for normalizing and resizing the image since CRNN accepts a fixed size at last we have RandomSequenceSampler class for randomly organizing the dataset So we will not be biased in our experiment.
- Then at last we are loading the images from the directory and predicting the text with help of bounding boxes and pre-trained model.

I had also tried for training the model but the problem with that was some assertion errors which I couldn't get to found out and I also tried for different things such as I had tried to extract all the bounding box by myself but my model which was CTPN doesn't able to figure it out probably due to the wrong tuning, I had also tried to visualize the results for task2 and this time got success.

RESULTS

- We are accurately able to recognize the text in the invoice as scene in the figure below.
- Example result is given below and further results can be found here <https://drive.google.com/drive/folders/1LavzikM-I35hN4RFQ9o-sA7oU-tY9hYC?usp=sharing> –

total sales (inclusive gst) m 98.3
28-01-18
49 speed mart s/b (519537-x)

change km
(8-74
ss7s lba & pervim5 sds worc nm989s
.00
.01
tax(rm)
rounding rm 98.90
taman berkeley
gst id. no : 000181747712
@ 11 x rm 8.99
invoice no : 18291/102/t0380
5.61 vy
1313-subang bestari
s = 6%
thank you. please come again
credit rm 98.90
0s
gst susmarv aoun:(rm)
93.28
4 /4/4:
lgt p.i. 2811 jalan angsa
06:04pm
41150 klang selangor
keep the invoice #cr applicable reworks
rounding adjustment rm
569004

- Task2 is completed Now.

3 18.01 0.17

SYARIKAT PERNIAGAAN GIN KEE
 (811109-A)
 NO 290, JALAN AIR PANAS,
 SETAPAK,
 53200, KUALA LUMPUR.
 TEL : 03-40210276
 GST ID : 000750673920
SIMPLIFIED TAX INVOICE

CASH

Doc No. CS00012518 Date: 03/01/2018
 Cashier USER Time: 13:04:00
 Salesperson Ref.:

Item	Qty	S/Price	Amount	Tax
1539	1	90.10	90.10	SR
5LT BIG TREE HIGH GLOSS PAINT				
1611	1	4.24	4.24	SR
THINNER BOTTLE				
Total Qty.	2		94.34	
Total Sales (Excluding GST)			89.00	
Discount			0.00	
Total GST			5.34	
Rounding			0.00	
Total Sales (Inclusive of GST)			94.34	
CASH			94.34	
Change			0.00	

GST SUMMARY			
Tax Code	%	Amt (RM)	Tax (RM)
SR	6	89.00	5.34
Total		89.00	5.34

GOODS SOLD ARE NOT RETURNABLE. THANK YOU.

3 18.01 0.17

SYARIKAT PERNIAGAAN GIN KEE
 (811109-A)
 NO 290, JALAN AIR PANAS,
 SETAPAK,
 53200, KUALA LUMPUR.
 TEL : 03-40210276
 GST ID : 000750673920
SIMPLIFIED TAX INVOICE

CASH

Doc No. CS00012518 Date: 03/01/2018
 Cashier USER Time: 13:04:00
 Salesperson Ref.:

Item	Qty	S/Price	Amount	Tax
1539	1	90.10	90.10	SR
5LT BIG TREE HIGH GLOSS PAINT				
1611	1	4.24	4.24	SR
THINNER BOTTLE				
Total Qty.	2		94.34	
Total Sales (Excluding GST)			89.00	
Discount			0.00	
Total GST			5.34	
Rounding			0.00	
Total Sales (Inclusive of GST)			94.34	
CASH			94.34	
Change			0.00	

GST SUMMARY			
Tax Code	%	Amt (RM)	Tax (RM)
SR	6	89.00	5.34
Total		89.00	5.34

GOODS SOLD ARE NOT RETURNABLE. THANK YOU.

FUTURE WORK

- In this I had not applied training which I will figure out as soon as possible and then I will perform training with the pytorch framework itself.
- Task-3 can also easily performed after this as only we have design algorithm to extract key information out of it and then we will save it in json format

GOOGLE DRIVE LINK

All the result and intermediate datas are present in the drive below.

<https://drive.google.com/drive/folders/1Xs5SaiEZndU2P4nccxZGB2JjtVSTF1sa?usp=sharing>

REFERENCES

- <https://arxiv.org/ftp/arxiv/papers/1905/1905.12817.pdf>
- <https://github.com/zzzDavid/ICDAR-2019-SROIE>
- <https://github.com/Michael-Xiu/ICDAR-SROIE>