

(Borang JKPTA FTSM UKM 3)



FAKULTI TEKNOLOGI DAN SAINS MAKLUMAT

BORANG PENYERAHAN LAPORAN ILMIAH

SEM 2 SESI 2020 / 2021

Bahagian A: Maklumat Diri Pelajar
Part A: Student's Details

No. Matrik (<i>Matric Number</i>)	A170950
Nama (<i>Name</i>)	Kh'ng Wei Yon
Program pengajian (<i>Programme</i>)	Ijazah Sarjana Muda Sains Komputer dengan Kepujian
No. Telefon (<i>Telephone Number</i>)	0164284287
Emel (<i>Email</i>)	a170950@siswa.ukm.edu.my

Tajuk Projek (*Project Title*):

ProReceipt: PENGECAMAN IMEJ RESIT MENGGUNAKAN TEKNIK OPTICAL CHARACTER RECOGNITION
(OCR) DAN CONNECTIONIST TEXT PROPOSAL NETWORK (CTPN)

Tandatangan (*Signature*): Kh'ng Wei Yon Tarikh (*Date*): 11/7/2021

Type text here

Bahagian B: Perakuan Penyelia
Part B: Supervisor's Approval

Saya peraku laporan ini telah disemak dan dibaiki, dan **menyokong / tidak menyokong*** penyerahan laporan ilmiah ini.

*I certify that this report has been reviewed and amended, and **approved / rejected*** the report submission.*

Tandatangan (*Signature*):  Tarikh (*Date*): 15/07/2021

Cap Rasmi :
(Official Stamp) 

**ProReceipt: PENGECAMAN IMEJ RESIT MENGGUNAKAN TEKNIK
OPTICAL CHARACTER RECOGNITION (OCR) DAN CONNECTIONIST
TEXT PROPOSAL NETWORK (CTPN)**

Kh'ng Wei Yon
Ts. Dr. Nor Samsiah Sani

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Dalam meniti arus kemodenan ini, resit merupakan suatu aspek dan bukti yang penting dalam kebanyakan jual beli seharian. Resit bukan sahaja penting untuk peniaga supaya mereka dapat memastikan urus rekod kewangan syarikat, tetapi juga terhadap pengguna yang hendak merekodkan perbelanjaan. Maka resit-resit ini, jika tidak disimpan dengan betul, tulisan atas resit akan dimakan usia atau mempunyai kualiti yang rendah menyebabkan kehilangan data yang penting untuk urusan lain. Oleh itu, tujuan utama projek ini dibangunkan adalah untuk membangunkan satu aplikasi mudah alih yang dapat mengecam imej resit, mengekstrak dan menganalisis maklumat yang terdapat dalam resit dengan menggunakan teknologi kecerdasan buatan (AI) iaitu pendekatan Pembelajaran Mendalam (*Deep Learning*) seperti *Connectionist Text Proposal Network* (CTPN). Dalam masa yang sama, teknik *Optical Character Recognition* (OCR) digunakan untuk pengecaman teks dan untuk menukar data gambar kepada teks bagi mendapatkan data dari gambar resit dengan cepat. OCR akan berfungsi untuk menukar teks yang dapat dari resit dan menukar teks tersebut menjadi data-data yang boleh simpan sepanjang masa. Kesimpulannya, dengan penggunaan aplikasi ini, ia mampu menukar dokumen menjadi data digital, menganalisis dan mengatur data pengguna (individu atau perniagaan bersaiz kecil hingga sederhana) dengan lebih bijak, cekap dan efektif.

1 PENGENALAN

Dalam era yang semakin maju ini, perniagaan memainkan peranan yang penting dalam pembangunan sosioekonomi masyarakat dan negara. Perniagaan adalah ekonomi kapitalis yang utama dan adalah sebuah bentuk organisasi yang menyediakan barangan atau perkhidmatan kepada pengguna. Walaupun dalam pelbagai tahap perniagaan, dari industri kecil sederhana sampai industri besar, resit memainkan peranan yang penting untuk merekodkan maklumat yang penting supaya rekod kewangan syarikat dapat diurus dengan mudah. Bukan itu sahaja, kebanyakan individu dalam era ini menyimpan resit untuk merekodkan perbelanjaan dan juga untuk menguruskan pelepasan cukai.

Sejak akhir abad ke-19, Koperasi *National Cash Register* (NCR) bermula menggunakan gulungan kertas dan borang perakaunan untuk merekod detik-detik perniagaan. Lima tahun

kemudian, resit kertas berkembang dan digunakan dengan mesin daftar tunai. Tanpa kira apa jenis transaksi jualan yang berlaku, resit pembayaran akan diberikan apabila pengguna membeli barang atau perkhidmatan yang disediakan. Resit dikeluarkan sebagai bukti urusan jual beli tersebut wujud. Selain itu, kadangkala resit boleh digunakan sebagai bukti oleh pengguna untuk mendapat tuntutan jaminan jika barang atau perkhidmatan tidak memenuhi standard.

E-resit merupakan bukti jual beli tanpa menggunakan kertas resit, tetapi melalui e-mel untuk memberitahu pembeli lulusan jual beli. Sebagai pembeli, e-resit mudah disimpan dan dicari semasa diperlukan. Untuk peniaga, e-resit merupakan sebuah aset yang penting untuk membina senarai pembeli dan menjana data pembeli dengan kemudahan.

Walaupun e-resit banyak mendatangkan faedah kepada peniaga dan pengguna. Namun, di Malaysia, kertas resit masih banyak diguna dalam urusan jual beli seharian kecuali mungkin dalam jual beli atas talian. Oleh itu, matlamat kajian ini adalah untuk membangunkan sebuah aplikasi pengecaman imej resit untuk menyimpan data resit dan juga merekodkan gambar resit untuk rujukan masa hadapan. Melalui aplikasi ini, pengguna dan peniaga dapat menjalankan urusan seharian dengan lebih mudah.

2 PENYATAAN MASALAH

Dalam era yang modenisasi ini, sebanyak 90 peratus resit kertas dibuang ke dalam tong sampah dan tidak disimpan. Manakala resit yang disimpan juga tidak simpan dengan betul, oleh sebab reka bentuk resit yang tidak mudah disimpan. Tulisan atas resit akan dimakan usia atau mempunyai kualiti yang rendah menyebabkan kehilangan data yang penting untuk urusan seharian lain. (Sarthak 2018). Selain itu, kegunaan e-resit juga jarang disaksikan dalam jual beli seharian di negara kita. Oleh itu, masalah akan timbul apabila peniaga atau pengguna memerlukan resit tersebut sebagai bukti pembelian untuk tuntutan jaminan, pelepasan cukai atau pengurusan kewangan peribadi dan pelaporan perbelanjaan perniagaan.

Resit kertas mudah hilang kerana ia susah untuk disimpan atau difailkan. Kebanyakan resit kertas disediakan dengan kertas yang berkualiti rendah, dakwat yang dicetak juga mudah dimakan usia. Ini berlaku kerana resit kertas direka untuk merendahkan kos peniaga sebanyak mungkin. Selain itu, resit kertas juga mudah terlupa dan tercuci dengan baju kerana saiz resit kertas yang kecil dan ringan.

Walaupun, kertas resit susah untuk disimpan dan juga membahayakan alam sekitar serta kesihatan kita. Kertas resit masih banyak digunakan di Malaysia dalam urusan atau jual beli seharian. Oleh itu, kajian mengubahsuai aplikasi pengecaman imej resit ini masih relevan di kalangan kita. Walaupun e-resit boleh menggantikan resit kertas, tetapi kebanyakan orang berasa tidak selesa untuk berkongsi e-mel atau nombor telefon di kaunter dengan idea ketakutan digodam atau kehilangan data kebersendirian.

3 OBJEKTIF KAJIAN

Sebagaimana yang dinyatakan di atas, resit mengandungi data yang penting dan boleh digunakan untuk analisis masa depan serta kegunaan persendirian. Oleh itu, untuk seorang peniaga, resit dapat mengesan jualan dan juga menyenaraikan senarai pembelian oleh pelanggan yang dapat membantu dalam perniagaan pada masa depan. Di samping itu, untuk kegunaan persendirian, kita dapat mendapat tuntutan jaminan, pelepasan cukai dan juga menyediakan kiraan perbelanjaan. Oleh itu, kajian ini mempunyai keupayaan untuk menukar resit dari format kertas ke format digital dan menyimpan data-data yang penting untuk urusan-urusan lain. Bukan sahaja data yang ada pada resit perlu disimpan dalam format digital, tetapi juga gambar resit perlu disimpan sebagai bukti untuk masa hadapan.

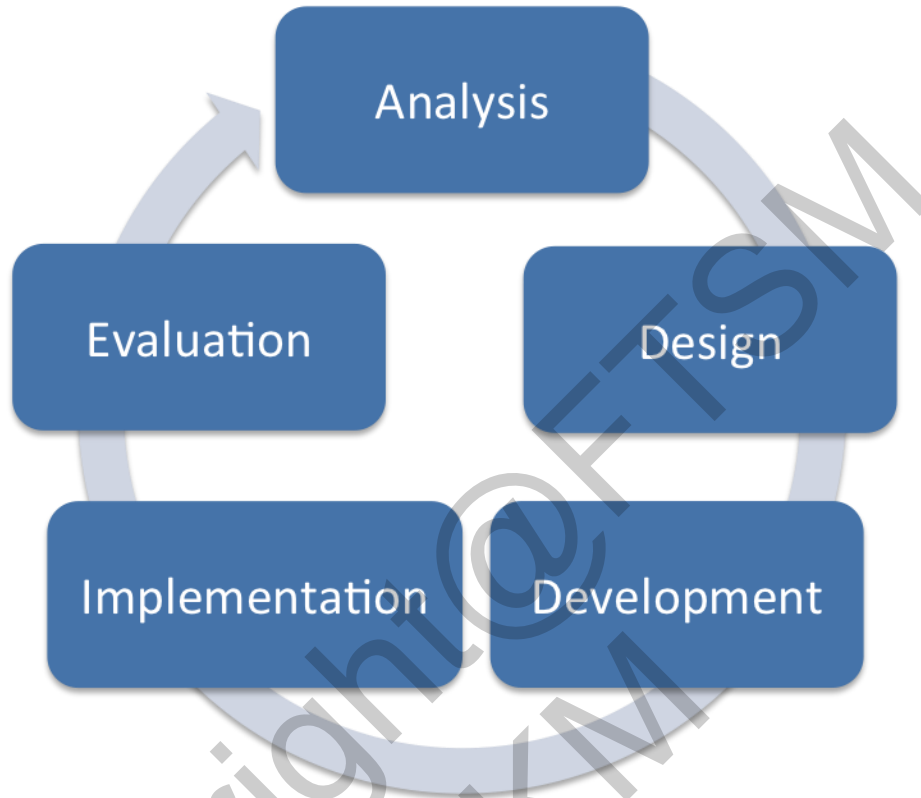
Berikut merupakan objektif-objektif yang ingin dicapai melalui projek ini:

- a. Membangunkan teknik pembelajaran mendalam CTPN yang dapat meningkatkan prestasi pengecaman imej.
- b. Mengekstrak maklumat pembelian produk, jumlah pembayaran dan data peniaga di dalam resit kepada digital data, dengan mengimplimentasikan CTPN yang dibangunkan mendalam dan Pengecaman Aksara Optik (OCR)
- c. Membangunkan aplikasi digital resit yang boleh memuat naik, mengekstrak dan menyimpan data ke dalam pangkalan data.

4 METOD KAJIAN

Dalam projek ini, kaedah Kitar Hayat Pembangunan Sistem (SDLC) yang akan digunakan adalah model ADDIE (Analyze, Design, Development, Implementation, and Evaluate) untuk

mengubahsuaikan algoritma dan membina aplikasi pengecaman imej resit yang sempurna. Model ini merupakan model yang lebih interaktif serta dinamik dan mengandungi lima fasa iaitu fasa analisis, fasa reka bentuk, fasa pembangunan, fasa implementasi dan fasa penilaian. (Serhat Kurt,2018)



4.1 Fasa Analisis

Fasa ini merupakan asas yang pertama dan terpenting. Fasa ini merupakan fasa untuk menetapkan tujuan. Dalam fasa ini, skop dan masalah telah dikenalpasti. Selain itu, dalam fasa ini, kajian dan artikel orang lain telah dianalisis untuk menyempurna kajian ini dan menyempurnakan algoritma OCR yang sedia ada. Dalam fasa ini, keperluan seperti gambar resit kertas perlu disimpan juga ditentukan supaya dapat diimplementasikan di fasa akan datang.

4.2 Fasa Reka Bentuk

Dalam fasa ini, alat dan set data yang sesuai dicari dan dicuba untuk meningkatkan prestasi algoritma dan aplikasi yang akan diubahsuai dalam fasa lain. Pengkaji akan memastikan sumber data dan kajian lain merupakan sumber yang sesuai untuk kajian ini. Setiap perincian perlu diperhatikan dalam fasa ini supaya aplikasi yang dibangunkan dapat memenuhi keperluan pengguna dan mudah diguna. Pembangunan pembelajaran mendalam dijalankan dalam fasa ini, untuk mereka bentuk algoritma yang paling sesuai untuk aplikasi.

4.3 Fasa Pembangunan

Fasa pembangunan akan memberi lebih banyak tumpuan dalam pembangunan algoritma mengikut objektif kajian yang ditetapkan supaya algoritma serta aplikasi memenuhi keperluan reka bentuk dan menyelesaikan masalah yang dinyatakan. Dalam fasa ini, pembangunan algoritma dan pembangunan aplikasi pengecaman imej resit dijalankan.

4.4 Fasa Implementasi

Fasa ini dalam modul bertujuan untuk menambah baik algoritma dan mengaplikasikan algoritma yang sedia diukur dalam aplikasi *Android* untuk diuji dengan set data latihan supaya dapat produk akhiran yang sempurna.

4.5 Fasa Penilaian

Fasa terakhir untuk model ini, merupakan fasa untuk menguji produk akhiran. Dalam fasa ini, set data yang baru telah dicari akan digunakan untuk menilai produk akhiran. Dalam fasa ini, maklum balas perlu disimpan untuk meningkatkan prestasi aplikasi. Selain itu, pilihan untuk mengubahsuai algoritma atau tidak, jika ubahsuai diperlukan model ADDIE boleh ulangi semula

dari fasa pertama.

Copyright@FTSM
UKM

5 HASIL KAJIAN

Dalam proses pembangunan sistem ini, peringkat pengaturcaraan dan rekabentuk akan membahagikan sistem ini ke dalam beberapa modul yang lebih terperinci supaya penulisan kod aturcara dapat dilaksanakan dengan lebih baik. Sistem ini, mempunyai dua hasil iaitu ketepatan model CTPN dan hasilan aplikasi yang dibuat dengan Kivy.

5.1 Modul CTPN

Bahagian belakang sistem dibangunkan dengan menggunakan bahasa pengaturcaraan Python. Beberapa langkah telah disediakan dalam bahagian belakang sistem.

i) Model CTPN dilatih

Kod sumber Model CTPN yang dapat dari *eragonruan-github* telah diubahsuai untuk melatih model CTPN supaya kotak pengikat imej akan dikenal pasti. Model ini dilatih dengan menggunakan data imej IDCAR 2013.

ii) Model CTPN diguna

Selepas latihan model, model ini akan digunakan dalam pengecam kotak pengikat imej. “*OpenCV*” dan “*sys*” merupakan modul yang digunakan untuk mendapat imej dan digunakan untuk mendapat kotak pengikat. Untuk mendapat kotak pengikat, saiz imej mungkin perlu diubah kerana saiz imej yang terlalu besar akan mempengaruhi keputusan. Oleh itu, fungsi untuk ubah saiz imej mengikut saiz imej telah disediakan. Selepas itu, kotak pengikat untuk setiap imej telah dapat dengan penggunaan model yang dilatih. Tetapi susunan untuk kotak pengikat tidak mengikut dari atas ke bawah. Oleh itu, fungsi untuk menyusun kotak pengikat dalam susunan dari atas ke bawah telah disediakan. Selain dari itu, kotak pengikat yang berada dalam barisan yang sama juga digabung menjadi satu kotak pengikat yang besar. Fungsi ini diadakan adalah untuk mendapat maklumat dari setiap barisan dalam susunan yang betul. Akhirnya “*tesseract*” akan disambungkan dengan menggunakan “*pytesseract*” untuk mengekstrak maklumat dari setiap kotak pengikat yang diproses dalam langkah sebelumnya. Maklumat yang disimpan akan dicam dengan ekspresi biasa (*regular expression*) dan menyimpan dalam format JSON.

Tesseract adalah Mesin Pengenalan teks sumber terbuka (OCR), tersedia di bawah lesen Apache 2.0. Ia boleh digunakan secara langsung, atau menggunakan API untuk mengekstrak teks bercetak dari gambar. Ia menyokong pelbagai bahasa.

iii) Model CTPN diuji

Dalam pengujian ini beberapa aspek semasa melatih model telah ditunjukkan. CTPN Model yang telah dilatih untuk mendapat kotak pengikat (bounding box) telah dilatih dengan beberapa set data.

Untuk model CTPN saya telah menggunakan set data idcar 2013.

Net	Dataset	Recall	Precision	Hmean
Origin CTPN	ICDAR13 training data	73.72%	92.77%	82.15%

Oleh sebab melatih model CTPN akan mengambil masa yang agak lama, model ini telah dipilih sebagai model untuk mendapat kotak pengikat kerana mendapat ketepatan yang agak tinggi. Selepas itu, sistem akan mengambil model ini untuk mendapat kotak pengikat dan mengekstrak maklumat dari setiap kotak pengikat melalui OCR. f-score akan dikira dengan menggunakan formula ini.

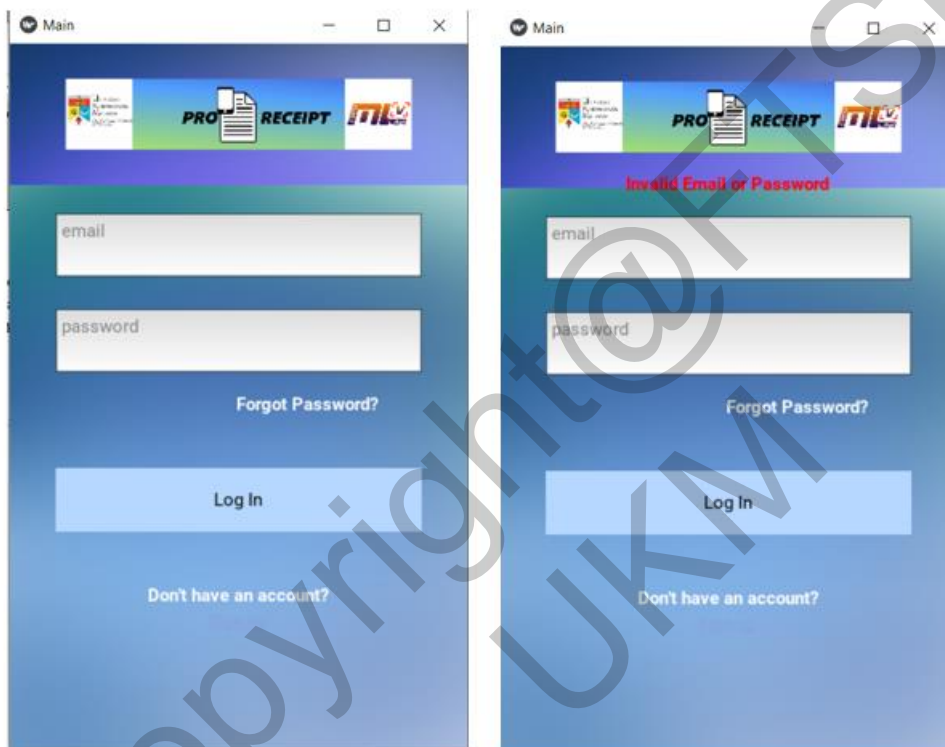
$$F1 = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})$$

f-score yang dikira merupakan 82.16%.

5.2 Pengujian Aplikasi

Setelah proses implementasi selesai, proses pengujian sistem bagi aplikasi ini akan dijalankan. Ia bagi memastikan aplikasi ini bebas dari ralaat serta dapat memenuhi keperluan pengguna. Setiap fungsi dalam aplikasi ini akan diuji bagi memastikan ia berfungsi dengan baik. Fasa pengujian sistem ini terbahagi kepada pengujian antaramuka dan pengujian integrasi.

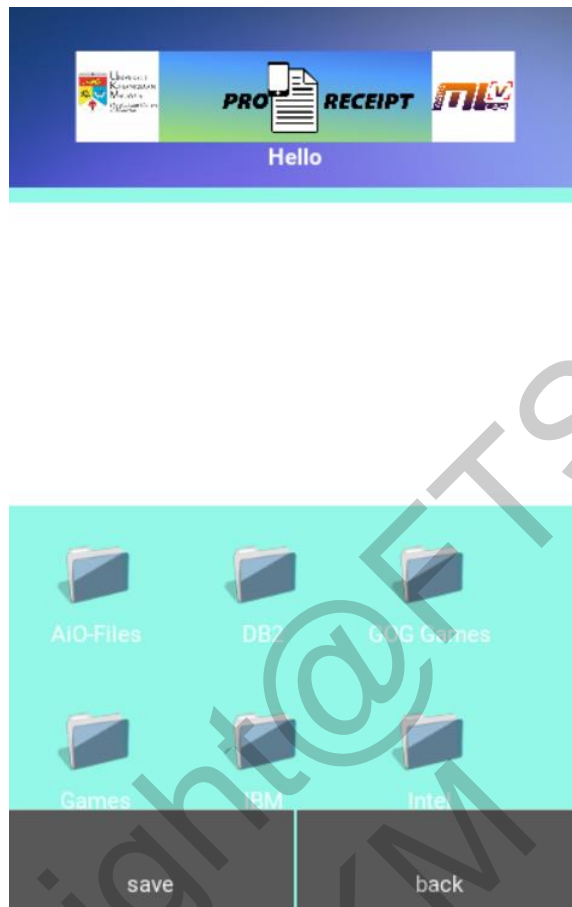
i. Antaramuka Log Masuk Aplikasi



Rajah 5.1 Antara Muka Log Masuk Akaun

Rajah 5.1 menunjukkan antara muka log-masuk ke dalam sistem. Pengguna yang sudah berjaya mendaftar akaun untuk aplikasi ini dan dapat memasukkan e-mel dan kata laluan yang betul akan terus masuk ke dalam aplikasi. Sekiranya, jika e-mel dan kata laluan yang dimasukkan tidak memadan dengan maklumat dalam *firebase*, mesej ralat “*invalid Email or Password*” akan dipaparkan supaya pengguna boleh menukarkan e-mel dan kata laluan yang lain. Pengguna perlu menekan butang “Log In” selepas memasukkan perincian yang betul.

ii. Antaramuka mengambil gambar resit kertas untuk pengestrakan maklumat



Rajah 5.2 Pengujian mengambil gambar resit kertas untuk pengestrakan maklumat

Rajah 5.2 adalah antara muka untuk penangkapan resit. Antara muka ini akan mengambil gambar resit melalui kamera mobil. Selain itu, butang “show” adalah untuk menunjukkan maklumat resit selepas penangkapan resit. Butang “save” adalah untuk memberi pengguna mengesahkan penyimpanan maklumat dalam pangkalan data yang disediakan.

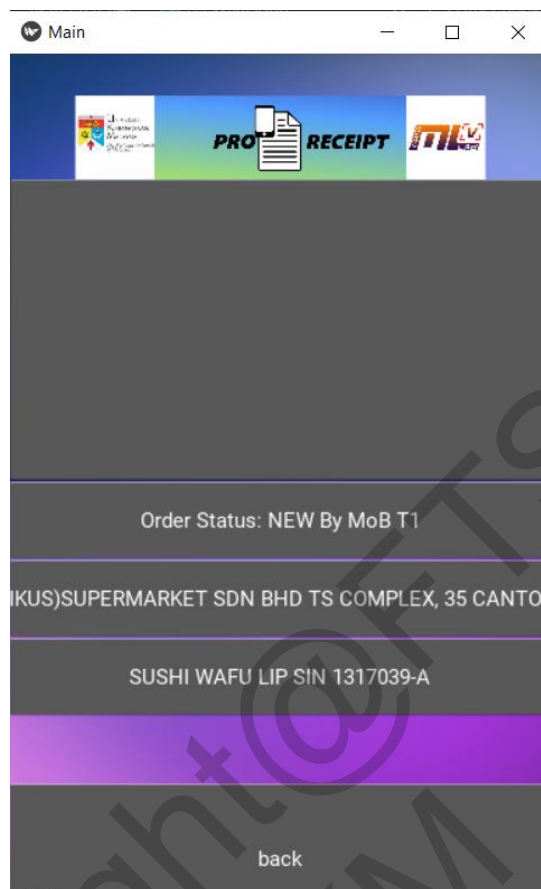
iii. Antaramuka Paparan Maklumat Resit

Label	Value
Date	null
Store Name	"R BARU(PULAU TIKUS)SUPERMARKET SDN BHD TS COMPLEX, 35 CANTONMENT RD, 10350"
Item	["R BARU(PULAU TIKUS)SUPERMARKET SDN BHD", "Item Code Description Amt (RM)", "16300651 FARM FRESH MILK 2L", "Cash : RM 50,00", "Change ; RM: 10.40"]
Total	"Total : RM 39.60"

back

Rajah 5.3 Antara Muka Pemaparan Maklumat Resit


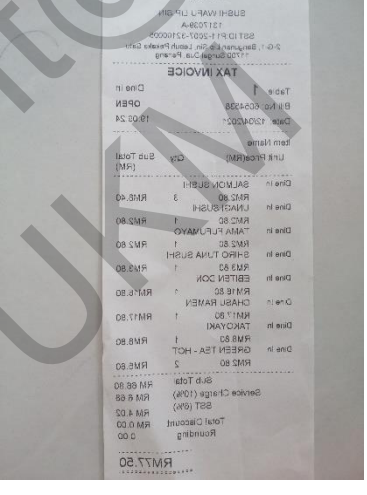
Rajah 5.3 adalah antara muka paparan maklumat resit. Antara muka ini akan menunjukkan “date”, “bil no”, “store name”, “item”, “total” sebagai maklumat resit yang telah diekstrak.

iv. Antaramuka senarai Resit dari pangkalan data

Rajah 5.4 Antara Muka Senarai Resit

Rajah 5.4 antara muka ini, pengguna boleh melihat senarai butang yang mengandungi nama penjual resit. Butang yang ditekan akan menunjukkan antara muka pemaparan maklumat resit.

v. Pengujian ketepatan OCR dalam aplikasi

Pengujian	Hasil	Resit	Gagal/Lulus
Resit1	<p>Sushi Zento Japanese Restaurant Oishii Avenue Sdn Bhd</p> <p>[Date: 31/03/2021 Time, 8:11:51']</p> <p>[Table: A3-Dine In', 'RM 2.50 4 10%/1.00', 'RM 2.60 2 10%/0.52 RM4.68', 'RM 5.60 1 10%/0.56 RM5.04', 'RM 22.00 1 10%/2.20 RM19.80', 'RM 25.80 1 104/258 RM23.22', 'RM 24.00 1 104/240 RM21.60', '02 10%/0.36—_RM3.24', 'Discount: RMO.00']</p> <p>Total Amount: RMB6.58</p>		Gagal
Resit2	<p>SUSHI WAFU LIP SIN 1317039-A</p> <p>[Date: 12/04/2021 19:09:24']</p> <p>[Table: 1 Dine In g', '(RM)', 'Dine In SALMON SUSHI', 'RM2.80 5 RMB8.40', 'Dine in UNAGI SUSHI', 'RM2 80 1 RM2.80', 'Dine in TAMA FUFUMAYO', 'RM2.80 1 RM2.80', 'Dine In SHIRO TUNA SUSHI', 'RM3.80 1 RM3.80', 'Dinein EBITEN DON', 'RM16.80 41 RM16.80', 'Dine in GHASU RAMEN', 'RM17.80 t RM17.80', 'Dine in = TAKOYAKI', 'RMB.80 1 RM8.80', 'RM2 80 2', 'RM 66.60']</p> <p>Unit Price(RM) Qty Sub Total</p>		Lulus

Resit3

es Texas Chicken (M) Sdn
Bhd
[Date: 24/04/202']
[Lot 1.02 & 1.03, Ground',
"Annexed Building 1,
Sunshine S', 'ad No 1, Jalan
Mayar', '41950 Bayan Ba',
'SST ID: Wi0-180']



Gagal kerana tidak
dapat kenal item

Resit4

[]
[R BARU(PULAU
TIKUS)SUPERMARKET
SDN BHD', 'Item Code
Description Amt (RM)',
'16300651 FARM FRESH
MILK 2L', 'Cash : RM
50,00', 'Change ; RM:
10.40']
Total : RM 39.60
{store: 'R BARU(PULAU
TIKUS)SUPERMARKET
SDN BHD TS COMPLEX,
35 CANTONMENT RD,
10350', 'date': [], 'items': ['R
BARU(PULAU
TIKUS)SUPERMARKET
SDN BHD', 'Item Code
Description Amt (RM)',
'16300651 FARM FRESH
MILK 2L', 'Cash : RM
50,00', 'Change ; RM:
10.40'], 'total': 'Total : RM
39.60'}



Lulus

6 KESIMPULAN

Projek ini menumpukan penggunaan teknologi Pembelajaran Mendalam dan Optical Character Recognition (OCR) dalam pembelajaran berdasarkan platform Windows. Tujuan pembangunan aplikasi ini adalah untuk menguruskan resit secara sistematik dengan mengekstrak maklumat penting darinya. Melalui aplikasi ini, pengguna diharapkan dapat menguruskan resit dengan mudah.

Antara kelebihan aplikasi ProReceipt yang dibangunkan termasuklah:

- a. Antara muka dan sistem aplikasi yang mudah diguna oleh pengguna
- b. Aplikasi yang dibina boleh diguna dalam telefon pintar dan juga dalam komputer.
- c. Teknologi OCR dan CTPN yang dapat mengecam maklumat imej resit

Antara kelemahan aplikasi ProReceipt yang dibangunkan termasuklah:

- a. Ketepatan dan kecekapan OCR yang diancam kerana kualiti imej resit
- b. Fungsi aplikasi yang terhad.

Antara kekangan yang akan dihadapi semasa melaksanakan kajian ini termasuklah:

- a. kekurangan pengetahuan dalam algoritma OCR.
- b. data resit kertas yang tidak sempurna akan menyebabkan aplikasi menyimpan data digital yang tidak sempurna contohnya: resit hanya menunjukkan kod produk tanpa nama produk
- c. kekurangan pengalaman dalam mencari set data yang sesuai untuk pembelajaran mendalam.

kekurangan teknik dalam mengimplementasikan model dalam aplikasi mobil

Akibat kekurangan pengalaman dalam pembinaan model dan juga aplikasi, pelbagai tutorial dan latihan dijadikan rujukan bagi menyempurna projek ini.

7 RUJUKAN

- Adrian Sarno, 2020. *Information Extraction from Receipts with Graph Convolutional Networks.* [Online] Available at <https://nanonets.com/blog/information-extraction-graph-convolutional-networks/> Accessed on 27/10/2020
- Anh Duc Le, Dung Van Pham, Tuan Anh Nguyen, 2019. *Deep Learning Approach for Receipt Recognition.*[Online] Available at https://www.researchgate.net/publication/333505533_Deep_Learning_Approach_for_Receipt_Recognition
- Anh Duc Le, Dung Van Pham, Tuan Anh Nguyen. *Deep Learning Approach for Receipt Recognition* <https://arxiv.org/ftp/arxiv/papers/1905/1905.12817.pdf>
- Ankur choudhary, *User Requirement Specification* <https://medium.com/@ankur1857/user-requirement-specification-7dc1830abbef>
- Anupama Ray, Sai Rajeswar, Santanu Chaudhury, 2015. *Text recognition using deep BLSTM networks.* [Online] Available at https://www.researchgate.net/publication/274255467_Text_recognition_using_deep_BLSTM_networks
- ConferenceCall, 2018: *What is an e-receipt? And should your business use them?* [Online] Available at <https://www.conferencecall.co.uk/blog/what-is-an-e-receipt/>
- Design Specification , <http://www.ofnisystems.com/services/validation/design-specification/>
- Erez Hasson, *Black Box Testing* <https://www.imperva.com/learn/application-security/black-box-testing/>
- Filip Zelic & Anuj Sable, *A comprehensive guide to OCR with Tesseract, OpenCV and Python.* <https://nanonets.com/blog/ocr-with-tesseract/>
- Gilab David Maayan,2020. *Optical Character Recognition Using Deep Learning Techniques.* [Online] Available at <https://heartbeat.fritz.ai/optical-character-recognition-using-deep-learning-techniques-1376605b022a>
- Joel Odd, Emil Theologou, 2018. *Utilize OCR text to extract receipt data and classify receipts with common Machine Learning algorithms.* [Online] Available at: <https://www.diva-portal.org/smash/get/diva2:1215460/FULLTEXT01.pdf>
- Joseph Wenhao Mao, 2017 *Self-Learning Receipt Optical Character Recognition Engine.* US9824270 Available at: <https://patentimages.storage.googleapis.com/c7/20/43/8f22533597cb4a/US9824270.pdf>
- Neil Kokemuller, 2017. *What is the Purpose of a Receipt?* [Online] Available at <https://bizfluent.com/info-8386381-returns-allowances.html> Accessed on 1/11/2020
- Özal YILDIRIM, 2018. *A NOVEL WAVELET SEQUENCES BASED ON DEEP BIDIRECTIONAL LSTM NETWORK MODEL FOR ECG SIGNAL CLASSIFICATION.*

Available at https://www.researchgate.net/figure/Basic-structure-of-the-BLSTM-network-The-LSTM-nets-at-the-bottom-indicate-the-forward_fig3_324056729

Rafi Ullah, Ali Sohani, Faraz Ali, Athaul Rai, 2017. *OCR Engine to extract Food-items and Prices from Receipt Images via Pattern matching and heuristics approach*. Conference: *International Conference on Computing and Related Technologies 2017*. Project: *Machine Learning*. Available online: https://www.researchgate.net/publication/322222389_OCR_Engine_to_extract_Food-items_and_Prices_from_Receipt_Images_via_Pattern_matching_and_heuristics_approach

Sarthak Moghe, 2018: *Everyone wants Digital Receipts, So Why Is the Retail Industry Not Adopting It?* [Online] Available at <https://www.conferencecall.co.uk/blog/what-is-an-e-receipt/>

Serhat Kurt, 2018: *ADDIE Model Instructional Design* [Online] Available at <https://educationaltechnology.net/the-addie-model-instructional-design/>

Spently, 2013: *Using the Past to Improve the Future: A Brief History of the Receipt* [Online], Available at <https://blog.spently.com/using-the-past-to-improve-the-future-a-brief-history-of-the-receipt-93bc5c6e796a>

William Todt. *Receipt manager with OCR support written in dart*. <https://medium.com/swlh/fuzzy-receipt-parser-and-manager-cb614e4aa6a>

Xiaohui Zhao, Zhuo Wu, Xiaoguang Wang. *CUTIE: Learning to Understand Documents with Convolutional Universal Text Information Extractor* https://www.researchgate.net/publication/332109910_CUTIE_Learning_to_Understand_Documents_with_Convolutional_Universal_Text_Information_Extractor