

ALPHAOCR: PENGEKSTRAKAN MAKLUMAT PADA RESIT BERASASKAN MODEL PEMBELAJARAN MENDALAM PICK DAN PENGECAMAN AKSARA OPTIK (OCR)

Muhammad Haikal Iman Bin Osman^{1*},

Norsamsiah Sani²

¹*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia*

²*Center for Artificial Intelligence Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia*

Abstrak

Resit merupakan suatu dokumen yang berperanan sebagai bukti keatas setiap transaksi pembelian yang sah. Resit menyimpan maklumat seperti tarikh transaksi, barangan yang dibeli, harga barangan, nama penjual, dan nama pembeli. Kesemua maklumat ini penting bagi tujuan penyimpanan rekod dan penjejakan perbelanjaan yang bukan sahaja mustahak bagi seseorang individu, malah syarikat bagi tujuan dokumentasi, belanjawan, dan audit. Namun begitu, dokumentasi resit bukanlah mudah kerana proses ini melibatkan jumlah resit yang banyak di samping kondisi resit yang semakin kabur hari demi hari. Tujuan utama projek ini dibangunkan adalah untuk menghasilkan satu aplikasi berasaskan web yang berkebolehan mengekstrak maklumat pembelian pada resit dengan menggunakan model pembelajaran mendalam *Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks (PICK)* dan Pengecaman Aksara Optik (OCR). OCR berupaya mengecam aksara pada imej resit yang kemudiannya boleh disimpan dan digunakan bagi tujuan analisis. Pembangunan aplikasi berasaskan web AlphaOCR memperolehi maklum balas yang positif daripada pengguna dalam ujian kebolehgunaan di samping berupaya meningkatkan nilai *Mean Entity F1-Score (mEF)* model pembelajaran mendalam PICK daripada 83.32% ke 83.88%. Model OCR yang sesuai untuk digunakan bagi mencapai matlamat kajian ini adalah EasyOCR dengan nilai

Character Error Rate (CER) yang setara dengan TesseractOCR iaitu 92.04%. Ini menjadikan ketepatan model AlphaOCR secara keseluruhannya adalah 87.96%. Dengan penggunaan aplikasi berasaskan web ini, kesemua maklumat penting pada resit pembelian boleh ditukarkan menjadi digital sekaligus memudahkan proses dokumentasi, analisis, dan penjejakan perbelanjaan pada sesuatu resit.

Kata kunci: Pengekstrakan Maklumat, Pembelajaran Mendalam, Pengecaman Aksara Optik

Pengenalan

Keberadaan teknologi serba canggih pada hari ini telah dimanfaatkan oleh pelbagai lapisan individu, masyarakat, dan organisasi bagi meningkatkan produktiviti dalam aktiviti seharian terutamanya dengan kewujudan teknologi baru seperti AI dalam pengecaman dan pengekstrakan teks. Menurut sebuah artikel daripada IBM Cloud Education (2022), pengecaman dan pengekstrakan teks pada dokumen dapat dilakukan dengan menggunakan sebuah teknik pemprosesan imej yang dikenali sebagai OCR ke atas dokumen yang diimbas atau gambar dokumen tersebut.

Mendokumentasikan resit bukanlah suatu perkara yang mudah kerana proses ini lazimnya dilakukan secara manual dengan cara menyalin maklumat pembelian ke dalam buku atau komputer. Proses ini amat sulit untuk dilakukan kerana melibatkan jumlah resit yang cukup banyak serta bakal menjadi kabur, rapuh, dan berkedut akibat dimakan usia. Aktiviti penjejakan perbelanjaan bagi individu dan perniagaan turut terjejas terutamanya apabila ingin meneliti pembelian yang dilakukan pada masa lalu seperti bagi tujuan audit dan siasatan mahkamah akibat daripada ketiadaan rekod simpanan digital dan aplikasi berasaskan web yang mampu digunakan di semua platform sistem operasi untuk mengekstrak dan menyimpan maklumat pembelian pada resit.

Selain itu, kekurangan set data terbuka berkenaan maklumat pembelian di peringkat nasional telah merencatkan aktiviti perlombongan data bagi tujuan kajian kepenggunaan yang bukan sahaja mampu dimanfaatkan oleh sesebuah perniagaan, malah berupaya membantu kerajaan menentukan

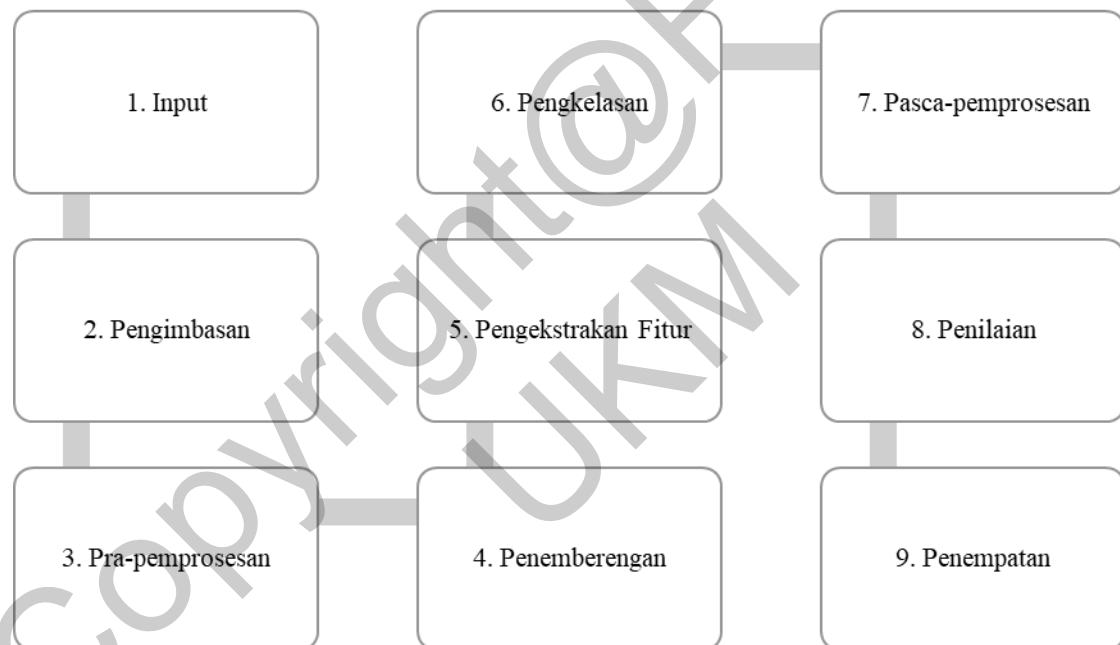
dasar-dasar ekonomi dan kepenggunaan yang baharu seperti penetapan harga siling barangan berpermintaan tinggi berdasarkan set data pembelian yang diekstrak daripada resit.

Bagi mengecilkan jurang ini, pembangunan aplikasi berasaskan web AlphaOCR merupakan suatu ciptaan yang perlu bagi membolehkan pengguna mengekstrak maklumat pada resit pembelian, mendapatkan laporan ringkas berkenaan maklumat pembelian, dan memuat turun laporan tersebut bagi tujuan dokumentasi. Dalam mencapai hasil pembangunan ini, dua objektif utama pembangunan perlulah dicapai iaitu membangunkan dan mengenal pasti model pembelajaran mendalam terbaik dan OCR terbaik untuk mengekstrak maklumat dari resit, dan membangunkan sebuah aplikasi berasaskan web yang berkebolehan mengekstrak maklumat dengan menggunakan model pembelajaran mendalam dan OCR terbaik serta menghasilkan set data baharu berkenaan maklumat pembelian bagi tujuan kajian pada masa akan datang.

Kajian ini terhad kepada resit pembelian di Malaysia yang bercetak dan tidak menggunakan teknik pemprosesan bahasa tabii (NLP) yang kompleks kepada teks yang diekstrak. Sebagai contoh, sekiranya kedai A merujuk “KitKat” dengan kata kunci “KK” pada resit manakala kedai B merujuknya dengan kata kunci “KitK”, kajian ini tidak akan menggabungkan kedua-dua kata kunci tersebut menjadi satu objek yang sama iaitu “KitKat”. Sebaliknya, akan menganggap kedua-dua kata kunci tersebut sebagai dua objek berbeza. Dari segi aplikasi berasaskan web pula, pengguna tidak berupaya memanggil semula maklumat yang diekstrak sebaik sahaja aplikasi ditutup.

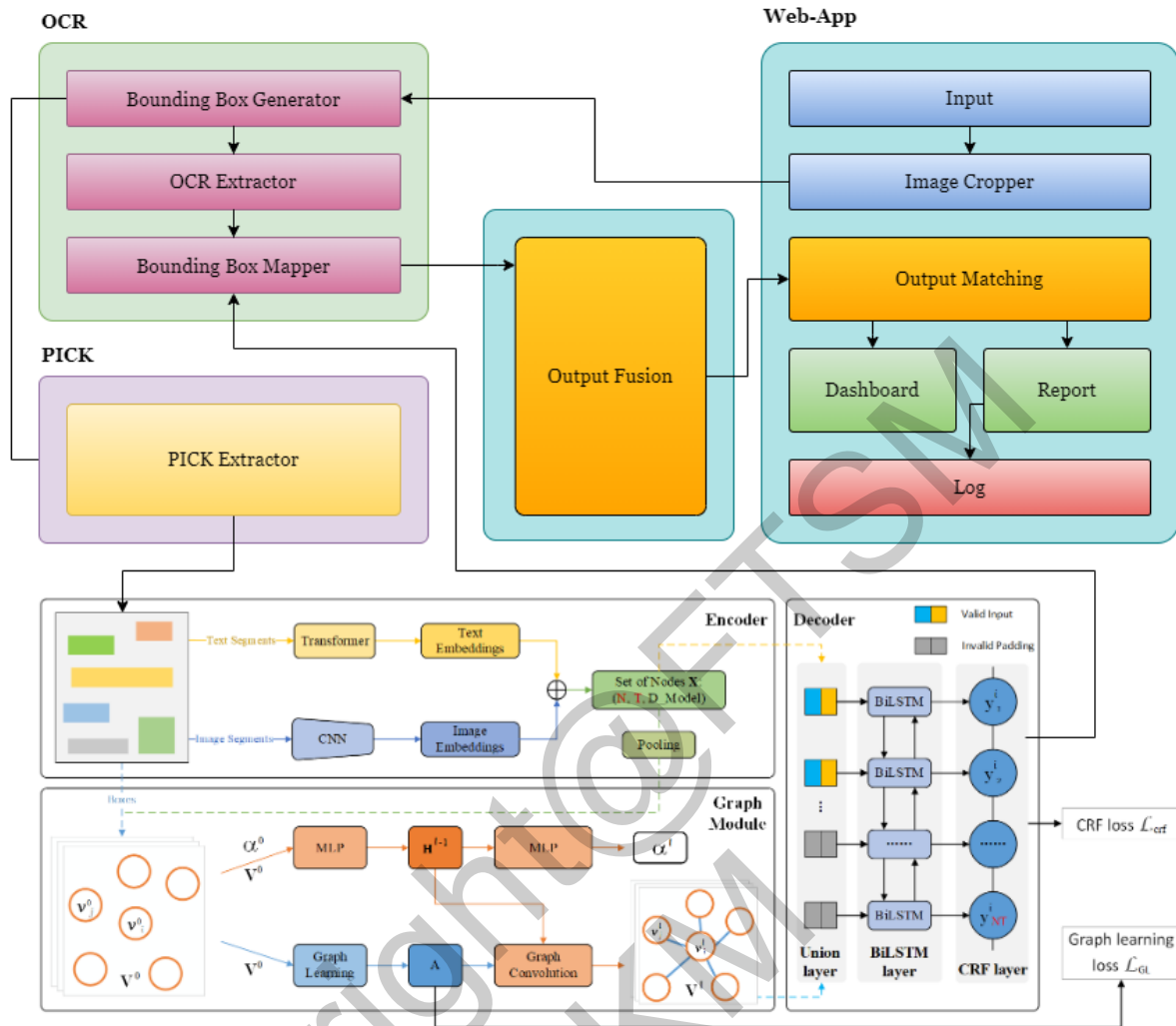
Antara kekangan untuk membangunkan projek ini termasuklah kekurangan set data lokal untuk disuap kepada mesin sewaktu proses pemodelan. Kebanyakan sumber atas talian hanya membekalkan koleksi resit dari luar negara sedangkan skop projek ini hanyalah terhad kepada Malaysia. Kekangan lain adalah keupayaan memahami domain perniagaan seperti keperluan resit sewaktu pengauditan dan siasatan mahkamah, dan kemahiran mengekstrak maklumat dengan menggunakan OCR.

Metodologi kajian yang digunakan dalam pembangunan projek ini adalah metodologi pengecaman corak seperti yang dicadangkan oleh Arindam Chaudhuri et. al (2017) dan Shubham Singh (2023) dalam sebuah artikel yang ditulisnya dengan tajuk *Pattern Recognition – Phases and Activities*. Menurut Shubham Singh (2023), pendekatan yang digunakan dalam membangunkan sistem pengecaman corak dapat dibahagikan kepada kira-kira lapan fasa iaitu input, pengimbasan, penemberengan, pengekstrakan fitur, pengkelasan, pasca-pemprosesan, penilaian, dan penempatan. Berikut merupakan rajah bagi metodologi kajian pengecaman corak yang akan digunakan bagi membangunkan model pembelajaran mendalam PICK berserta OCR yang digunakan oleh aplikasi berasaskan web AlphaOCR.



Rajah Error! No text of specified style in document..1 Metodologi pengecaman corak

Sumber: Shubham Singh 2023



Rajah 1.2 Struktur menyeluruh pembangunan model AlphaOCR

Seperti yang ditunjukkan dalam rajah di atas, struktur menyeluruh pembangunan model AlphaOCR terdiri daripada tiga komponen utama iaitu komponen model pembelajaran mendalam PICK terbaik, model OCR terbaik, dan aplikasi berasaskan web.

Resit secara umumnya memiliki tiga bahagian utama iaitu kepala, badan, dan kaki. AlphaOCR akan membahagikan imej input yang diterima kepada dua iaitu imej asal, dan imej badan resit. Imej asal akan dihantar ke modul pemangkasan imej (*image cropper*) bagi menghasilkan imej badan resit. Kedua-dua imej ini kemudiannya akan dihantar ke modul penghasilan kotak sempadan (*bounding box generator*) bagi menentukan kawasan perbatasan bagi setiap aksara yang terdapat pada resit. Modul pengestrakan OCR (*OCR extractor*) akan menjalankan pengestrakan ke atas imej badan resit bagi mengekstrak maklumat berkenaan produk dan harga produk manakala imej asal resit

pula akan dihantar ke modul pengekstrakan PICK (*PICK extractor*) bagi mengekstrak nama syarikat, alamat syarikat, tarikh, dan jumlah harga yang dibayar.

Kesemua hasil pengekstrakan ini akan diteliti oleh modul pemetaan kotak sempadan atau bounding box mapper bagi memetakan kawasan yang diekstrak pada resit secara visual dan dalam masa yang sama, hasil pengekstrakan daripada kedua-dua model ini digabungkan menjadi satu melalui modul output fusion. Sistem akan memaparkan hasil pengekstrakan yang sudah pun dilakurkan kepada pengguna untuk memperoleh maklum balas pengguna melalui modul padanan output atau output matching yang memerlukan pengguna membetulkan atau mengelaskan teks yang diekstrak kepada entiti yang berikaitan iaitu sama ada nama syarikat, alamat syarikat, tarikh, jumlah harga yang dibayar, produk, atau harga produk.

Hasil pengekstrakan yang sudah dikemas kini akan dipaparkan dalam bentuk papan analisis dengan keupayaan memuat turun maklumat yang diekstrak tadi. Kesemua maklumat ini akan disimpan dalam satu fail log yang boleh diakses oleh admin melalui bahagian belakang aplikasi.

Sorotan Susastera

Pada tahun 2019, suatu kajian oleh Anh Le Duc, Dung Van Pham, dan Tuan Anh Nguyen telah dijalankan bagi tujuan pengesanan dan pengecaman teks pada resit dengan tajuk kajian *Deep Learning Approach for Receipt Recognition*. Metode kajian ini terbahagi kepada dua: pengesanan teks dan pengecaman teks. Pengesanan teks dilakukan menggunakan algoritma CTPN dengan keupayaan mengekstrak fitur melalui VGG-16, BiLSTM, dan regresi kotak sempadan atau bounding box regression manakala algoritma AED dengan keupayaan mengekstrak fitur melalui DenseNet, dan pengesanan teks melalui OCR. CTPN dipilih dalam kajian tersebut bagi tujuan pengesanan teks kerana berupaya mengesan teks secara horizontal dengan baik manakal AED pula digunakan bagi tujuan pengecaman teks kerana terkenal dengan kebolehnya mengecam ekspresi matematik, dan aksara Vietnam yang bertulis tangan. Kajian ini dilakukan dengan menggunakan set data SROIE2019

yang dibekalkan oleh ICDAR. Anh Le Duc dan rakan-rakannya membahagikan set data SROIE2019 dengan nisbah 80:10:10 bagi menghasilkan set latihan (500), ujian (63), dan pengesanan (63). CTPN, CTPN+pra-pemrosesan, dan CTPN+pra-pemrosesan+pengesanan teks melalui OCR adalah tiga variasi algoritma CTPN yang dihasilkan dalam kajian tersebut. Ketiga-tiga algoritma ini diuji sebanyak dua kali: tanpa AED bagi tujuan pengesanan teks, dan kemudiannya dengan AED bagi tujuan pengesanan dan pengecaman teks secara serentak. Hasil kajian mendapati algoritma CTPN+pra-pemrosesan+pengesanan teks melalui OCR bersama AED telah menunjukkan prestasi terbaik dengan memperoleh skor F1, Precision, dan Recall masing-masing sebanyak 71.9%, 72.5%, dan 71.3%. Walaupun algoritma CTPN+pra-pemrosesan menunjukkan peratusan Recall yang lebih baik (72.3%), namun secara keseluruhannya algoritma CTPN+pra-pemrosesan+pengesanan teks melalui OCR berjaya mencapai peratusan yang lebih tinggi dalam metrik penilaian yang lain. Kekurangan kajian ini adalah tidak melibatkan pengesanan dan pengecaman teks bertulis tangan.

Vedant Kumar bersama-sama ahli pasukan pengkajinya telah menjalankan sebuah kajian pada tahun 2020 bertajuk *Extraction of Information from Bill Receipts Using Optical Character Recognition* bagi tujuan pengaplikasian teknologi OCR untuk mengekstrak maklumat daripada gambar resit dan bil, dan menyimpannya dalam bentuk yang boleh difahami oleh mesin. Empat metode yang digunakan adalah termasuk pembuangan bayang dan tera air (*watermark*), pembahagian kepada dua invois panjang (*bifurcation of long invoices*) bagi invois dengan panjang melebihi 26cm dengan mengambil kira 55% bagi bahagian invois terawal menjadi segmen pertama dan 50% bagi bahagian invois terakhir menjadi segmen kedua, pengecaman dan pengekstrakan teks, dan yang terakhir adalah mendapatkan maklumat relevan. Algoritma yang digunakan adalah OpenCV bagi pemrosesan imej manakala Tesseract OCR bagi pengekstrakan teks. OpenCV dipilih kerana kemasyhurannya dalam pemrosesan imej data manakalah Tesseract OCR digunakan kerana berupaya mengesan teks dalam pelbagai bahasa dengan kebolehan menghasilkan output teks mengikut format atau kedudukan yang serupa dengan imej input. Set data yang digunakan adalah

tidak dinyatakan kerana pengkaji tersebut hanya melakukan demonstrasi pengaplikasian teknologi OCR bagi mengekstrak maklumat pada bil resit. Walau bagaimanapun, para pengkaji ada memaklumkan bahawa aplikasi yang dihasilkan telah diuji dengan 25 keping bil resit. Metrik penilaian yang digunakan adalah *accuracy* dan PSNR. Nilai *accuracy* yang dicapai adalah setinggi 97% pada bil resit bersaiz kecil, dan 83% pada bil resit bersaiz besar manakala nilai PSNR pula adalah di antara 38.13-42.57db. Nilai PSNR yang ditunjukkan tidaklah begitu mustahak kerana nilai PSNR adalah spesifik terhadap setiap satu imej input sahaja yang mana dalam kes ini nilai yang ditunjukkan adalah bagi dua keping resit yang digunakan dalam kajian tersebut. Kelebihan utama hasil kajian ini adalah berupaya mengekstrak teks walaupun dengan kehadiran tera air dan bayang pada bil resit manakala kekurangannya adalah hanya berupaya menghasilkan output optimum pada imej input dengan teks bercetak (*printed text*).

CUTIE: Learning to Understand Documents with Convolutional Universal Text Information Extractor adalah suatu kajian yang dilakukan oleh Xiaohui, Endi Niu, Zhuo Wu, dan Xiaoguang Wang pada tahun 2019 bagi menghasilkan OCR resit dan pengekstrakan maklumat penting. Metode yang digunakan adalah pemetaan kedudukan grid atau grid positional mapping bagi menghasilkan grid input data untuk kegunaan algoritma CNN. Jenis algoritma CNN yang digunakan dalam kajian tersebut adalah CUTIE-A dan CUTIE-B. CUTIE-A merupakan model CNN berkapasiti tinggi dengan keupayaan menggabung fitur dari pelbagai resolusi manakala CUTIE-B merupakan model CNN dengan *atrous convolution* dan *ASPP* yang berupaya meningkatkan saiz pandangan dan mengesan konteks pelbagai skala. Kajian tersebut juga menggunakan algoritma CloudScan dan BERT bagi tujuan perbandingan dengan dua algoritma yang diperkenalkan iaitu CUTIE-A dan CUTIE-B. Set data yang digunakan adalah SROIE2019 yang dibekalkan oleh ICDAR dengan nisbah set latihan dan ujian masing-masing sebanyak 75:25 daripada jumlah set data yang dikhususkan bagi tujuan ini iaitu sebanyak 1000. Dengan menggunakan metrik penilaian AP dan softAP, CUTIE-A memperoleh keputusan bagi AP sebanyak 90.8% (resit teks), 77.7% (resit makanan dan hiburan), dan 69.5% (resit

hotel) manakala untuk softAP pula adalah sebanyak 97.2% (resit teksi), 91.4% (resit makanan dan hiburan), dan 87.8% (resit hotel). Bagi CUTIE-B pula, AP yang diperoleh adalah 94.0% (resit teksi), 81.5% (resit makanan dan hiburan), dan 74.6% (resit hotel) manakala softAP pula adalah sebanyak 97.3% (resit teksi), 89.7% (resit makanan dan hiburan), dan 87.0% (resit hotel). Kedua-dua algoritma CUTIE ini berjaya mengatasi algoritma CloudScan dengan keputusan AP masing-masing sebanyak 82.0%, 64.0%, dan 60.0% bagi resit teksi, resit makanan dan hiburan, dan resit hotel. Algoritma BERT pula memperoleh peratusan AP sebanyak 88.1% (resit teksi), 80.1% (resit makanan dan hiburan), dan 71.7% (resit hotel). Keputusan bagi softAP untuk CloudScan dan BERT adalah tidak dinyatakan. Kekurangan kajian tersebut adalah tidak melibatkan langkah pra-pemprosesan yang menghubungkan teks-teks menjadi entiti bermakna (meaningful entitie) untuk dimanfaatkan sewaktu proses pemetaan kedudukan grid. Kajian tersebut hanya menukar teks menjadi token dengan menggunakan algoritma greedy longest-match-first yang sedia ada di dalam gedung kamus BERT.

Pada tahun 2020, sebuah kajian bertajuk *An End to End Method for Taxi Receipt Automatic Recognition Based on Neural Network* telah dijalankan oleh pasukan pengkaji yang terdiri daripada Weiliang Liu, Xueguang Yuan, Yangan Zhang, Mengya Liu, Zhenyu Xiao, dan Jianlan Wu. Objektif kajian tersebut adalah untuk menghasilkan model pembelajaran mendalam dengan keupayaan mengecam maklumat pada resit teksi tanpa menggunakan pemprosesan imej tradisional dan OCR yang biasa. Tiga metode kajian tersebut adalah termasuk penentuan kawasan informatif, pemangkasan imej resit teksi, dan pengecaman maklumat pada imej. Algoritma SSD dengan VGG-16 digunakan bagi proses penentuan kawasan informatif pada resit teksi manakal algoritma CNN dengan GRU direka khas bagi tujuan pengecaman aksara pada resit teksi. Proses pemangkasan imej resit teksi dilakukan bagi tujuan membahagikan resit teksi mengikut bahagian-bahagian maklumatnya yang tersendiri. Dalam kes ini, pengkaji tersebut membahagikan resit teksi kepada empat belas bahagian kerana set data yang digunakan merupakan resit teksi yang selaras dari segi formatnya atau dalam bahasa yang mudah, set datanya adalah terhad kepada satu domain sahaja iaitu resit teksi dari

syarikat yang sama di Beijing. Set data yang digunakan pula terdiri daripada 1800 imej resit teks yang kemudiannya dibahagi kepada lima lipatan (*fold*) bagi tujuan latihan dan penilaian rentas (*cross validation*). Weiliang Liu et al ada melakukan penambahbaikan pada set data yang digunakan untuk mengatasi keadaan resit yang kurang baik seperti gelap akibat kurang pencahayaan, dan kedudukan yang tidak selari dengan kamera. Keputusan yang direkodkan adalah dengan membandingkan *accuracy* sistem pengesanan maklumat pada resit teks yang lain iaitu *NN and OCR* (90.62%), *Template Match and OCR* (95.92%), *Image Processing and NN* (95.75%), dan *NN and ML* (99.2% bagi aktiviti pengesanan aksara sahaja) manakala algoritma ciptaan para pengkaji tersebut iaitu kombinasi antara SSD dan CNN-GRU memperoleh *accuracy* sebanyak 94.36%. Kekurangan kajian ini adalah hanya terbatas kepada pengesanan maklumat pada resit dan bukannya pengekstrakan maklumat yang ada padanya di samping keadah yang digunakan adalah khusus buat resit teks dengan format yang digunakan di Beijing.

Joel Odd dan Emil Theologou telah melakukan sebuah kajian yang bertajuk *Utilize OCR Text To Extract Receipt Data and Classify Receipts With Common Machine Learning Algorithms* pada tahun 2018 bagi menguji sama ada teknik OCR yang ada pada pembelajaran mesin berupaya mengelaskan resit dan mengekstrak maklumat khusus titik data tertentu. Kajian tersebut menggunakan metode yang lazim iaitu dengan mengumpul set data resit fizikal, menukarkan resit fizikal menjadi digital, latihan dan ujian algoritma pembelajaran mesin, dan mereka bentuk algoritma khusus untuk mengekstrak maklumat tertentu yang diinginkan pada resit. Dua belas algoritma lazim yang digunakan oleh para pengkaji, tetapi hanya empat algoritma yang dipilih bagi tujuan perbandingan prestasi iaitu LinearSVC, MLP, Naïve Bayes, dan *DummyClassifier*. *DummyClassifier* merupakan sebuah algoritma yang direka hanya untuk melakukan perkara asas dalam pengelasan data dan tidaklah optimum untuk menyelesaikan masalah-masalah tertentu. Bagi algoritma OCR pula, pengkaji tersebut telah memilih *Google Drive REST API* berbanding *Azure Computer Vision API* setelah mendapati bahawa *REST API* berjaya memperoleh *accuracy* tertinggi pada ketiga-tiga

penilaian ringkas berbeza yang dijalankan iaitu sebanyak 87.16% (ujian 1), 84.12% (ujian 2), dan 88.79% (ujian 3) berbanding *Azure Computer Vision API* yang memperoleh keputusan *accuracy* sebanyak 53.99% (ujian 1), 55.80% (ujian 2), dan 52.22% (ujian 3). Set data yang digunakan adalah terdiri daripada 556 resit dengan lima kategori iaitu kedai barangan dapur, kedai makan, kedai runcit, dan lain-lain. Kajian tersebut menggunakan metrik penilaian *accuracy* bagi membandingkan prestasi keempat-empat algoritma yang digunakan. Nilai *accuracy* yang diperoleh adalah seperti berikut: LinearSVC (94.07%), MLP (93.53%), Naïve Bayes (68.52%), dan *DummyClassifier* (28.06%). Hasil kajian tersebut mendapati bahawa algoritma terbaik iaitu LinearSVC berjaya memperoleh *accuracy* 72.32% bagi pengekstrakan titik data khusus iaitu harga manakala 50.89% bagi pengekstrakan tarikh. Kelebihan kajian ini adalah mampu mengekstrak titik data khusus seperti harga dan tarikh berbanding mengekstrak keseluruhan maklumat daripada resit. Walau bagaimanapun, prestasi *accuracy* yang ditunjukkan masih lemah dan suatu metode kajian yang lebih kompleks perlu diperkenalkan bagi mengoptimumkan output yang dihasilkan.

Melalui kajian yang dilakukan oleh Peng Zhang, Yunlu Xu, Zhazhang Cheng, Shiliang Pu, Jing Lu, Liang Qiao, Yi Niu, dan Fei Wu dengan tajuk *TRIE: End-to-End Text Reading and Information Extraction for Document Understanding* mendapati bahawa model yang diperkenalkan iaitu *TRIE* berjaya mencapai keputusan *accuracy* tertinggi sebanyak 93.26% berbanding *GCN* yang hanya pada 92.60% dalam pengekstrakan maklumat pada resit. Set data yang digunakan adalah *SROIE2019*, 5000 resit teks, dan 2475 resume dalam bahasa Cina. Algoritma *TRIE* merupakan kombinasi antara *FPN* dan *ResNet-50*.

Wenwen Yu, Ning Lu, Xianbiao Qi, Ping Gong, dan Rong Xiao mengusulkan sebuah model pembelajaran mendalam yang diberi nama *PICK* melalui tajuk kajian mereka *PICK: Processing Key Information Extraction from Documents using Improved Graph Learning-Convolutional Networks* pada tahun 2020. Tujuan utama kajian tersebut adalah bagi menghasilkan sebuah model dengan keupayaan mengekstrak maklumat penting (*KIE*) yang masih menjadi cabaran utama dalam aspek

pengekstrakan maklumat pada dokumen bertulis mahupun bercetak. PICK merupakan sebuah model ensemble yang menggabungkan algoritma CNN, GLCN, BiLSTM, transformer dan CRF. Algoritma PICK diuji dengan menggunakan set data SROIE2019, 2,630 invois perubatan, dan 2,000 tiket tren. Dua model yang menjadi saingan PICK dalam kajian tersebut adalah model asas (*baseline*) dihasilkan dengan menggunakan gabungan algoritma BiLSTM dan CRF, dan model LayoutLM. Ketiga-tiganya diuji dengan metrik penilaian mEF. Keputusan yang diperoleh bagi mEF untuk PICK adalah 96.1% (SROIE2019), 87.0% (resit perubatan), dan 98.6% (resit tren). Nilai mEF bagi model asas pula adalah sebanyak 72.3% (resit perubatan), 85.4% (resit tren), dan SROIE2019 tidak dinyatakan manakala keputusan mEF bagi algoritma LayoutLM hanya dinyatakan untuk set data SROIE2019 iaitu 95.2%.

Graph Convolution for Multimodal Information Extraction from Visually Rich Documents

merupakan sebuah kajian menarik yang dilakukan oleh Xiaojing Liu et al. pada tahun 2019 bagi menghasilkan sebuah model dengan keupayaan mengekstrak maklumat daripada dokumen dari perspektif tulisan dan visual tanpa bergantung kepada pemrosesan bahasa tabii (NLP) semata-mata. Kebiasaannya, para pengkaji dalam kajian yang lain hanya mengambil kira maksud tersurat tulisan-tulisan yang ada pada resit atau dokumen. Namun bagi Xiaojing Liu dan rakan-rakan, visual seperti kedudukan setiap teks pada dokumen tersebut juga memainkan peranan penting dalam proses mengekstrak maklumat daripadanya. VATI dengan 3,000 data dalam bahasa Cina dan IPR dengan 1,500 data dalam bahasa Inggeris adalah dua set data yang digunakan dalam kajian tersebut. Bagi menguji prestasi algoritma yang diusulkan iaitu model BiLSTM+CRF+GCN, dua model asas yang diberi nama Baseline 1, dan Baseline 2 dicipta dengan menggunakan model BiLSTM+CRF. Baseline 1 diuji dengan menggunakan teks secara individu manakala Baseline 2 diuji dengan menggunakan teks yang telah dicantumkan (*concatenated*). Keputusan skor F1 bagi Baseline 1, Baseline 2, dan BiLSTM+CRF+GCN bagi set data VATI masing-masing adalah 74.5%, 85.4%, dan 87.3% manakala bagi set data IPR pula masing-masing sebanyak 74.7%, 82.0%, dan 83.6%. Kelebihan teknik ini

berbanding teknik pengkaji lain adalah dengan mengambil kira aspek visual yang ditunjukkan di dalam sesuatu dokumen.

Pada tahun 2020, Bodhisattwa Prasad Majumder dan rakan-rakan pengkajinya telah menemui bahawa prestasi pengekstrakan maklumat pada resi dan invois mencapai tahap terbaik melalui model Representation Learning berbanding model BoW dan MLP dengan nilai F1 skor sebanyak 83.9% bagi set data Invoices1 dan Invoices2 manakala 83.3% bagi set data SROIE2019. Kajian tersebut telah menggunakan data sebanyak 14,237 bagi Invoices1, 595 data bagi Invoices2, dan kira-kira 626 data bagi SROIE2019 bagi tujuan pengekstrakan maklumat daripada dokumen dengan tajuk kajian *Representation Learning for Information Extraction from Form-like Documents*.

Cheng Jian Lin dan rakan-rakannya mengusulkan model YOLOv4-s dalam kajian mereka pada tahun 2022 yang bertajuk *Automatic Receipt Recognition System Based on Artificial Intelligence Technology* bagi menghasilkan sebuah sistem pengecaman resit automatic. Antara algoritma lain yang dibandingkan dengan YOLOv4-s adalah YOLOv3, YOLOv4, dan YOLOv5 dengan menggunakan set data resit bercetak dan bertulis tagan masing-masing sebanyak 300 dan 100. Metrik penilaian Accuracy yang digunakan mendapati bahawa YOLOv4-s mencapai nilai Accuracy tertinggi iaitu 99.39% berbanding YOLOv3 (71.49%), YOLOv4 (78.82%) dan YOLOv5 (92.27%). Kelebihan yang ditunjukkan oleh model YOLOv4-s adalah keupayaannya untuk mengurangkan process pengecilan saiz imej input atau downsampling yang mana boleh mempengaruhi pengecaman objek kecil jika saiz imej input dkecilkan. Walaupun model YOLOv4-s digunakan dalam kajian tersebut, namun model CNN turut menjadi asas dalam kajian yang berperanan sebagai pengecam aksara bagi set data yang digunakan.

Pada tahun 2022, Vedhaviyas et. al telah memperkenalkan kaedah pengecaman nombor plat baharu dengan menggunakan tiga modul utama iaitu modul pemerolehan imej, pengecaman nombor plat, dan pengecaman aksara. Bagi tujuan pengecaman aksara, pengkaji menggunakan dua buah model OCR berbeza iaitu EasyOCR dan TesseractOCR. Kedua-dua model ini menggunakan 1200

imej set data daripada GitHub dengan 800 imej digunakan sewaktu latihan dan 400 lagi digunakan bagi tujuan ujian. Dengan menggunakan metrik penilaian CER, EasyOCR berjaya memperoleh nilai CER tertinggi iaitu sebanyak 95.00% berbanding TesseractOCR (88.00%). Menurut Vedhaviyas et. al, CER adalah metrik penilaian ketepatan yang digunakan bagi mengukur keupayaan sesebuah model OCR.

Dari segi perbandingan aplikasi, pengkaji telah membandingkan aplikasi berasaskan web yang akan dibangunkan iaitu AlphaOCR dengan tiga aplikasi lain iaitu Easy Expense, Smart Receipts, dan Foreceipt dari segi fitur utama keempat-empat daripadanya. Kesemua aplikasi berupaya mengecam kewujudan resit dan kemudiannya mengekstrak maklumat pada resit tersebut. Walau bagaimanapun, hanya AlphaOCR sahaja yang berkebolehan mengekstrak maklumat pembelian secara terperinci seperti perkhidmatan atau barang apakah yang dibeli beserta harga setiap satu daripadanya.

Metodologi Kajian

Metodologi kajian yang digunakan dalam pembangunan projek ini adalah metodologi pengecaman corak seperti yang dicadangkan oleh Arindam Chaudhuri et. al (2017) dan Shubham Singh (2023) seperti yang ditunjukkan dalam rajah berikut.



Rajah Error! No text of specified style in document..2 Metodologi pengecaman corak

Sumber: Shubham Singh 2023

Set data yang digunakan bagi melatih model PICK berserta OCR bagi AlphaOCR adalah set data SROIE2019 oleh ICDAR. Set data yang digunakan adalah imej mengenai pelbagai jenis resit pembelian bercetak di Malaysia dalam format 'jpeg' atau 'jpg'. Kesemua imej ini adalah resit pembelian dengan bentangan (*layout*) yang berbeza. Sebagai tambahan, pengkaji menggunakan set data resit pembelian di Malaysia yang dikumpul sendiri bagi menambah dan memvariasikan lagi set data sedia ada dengan format yang sama iaitu 'jpeg' atau 'jpg'. Rajah di bawah menunjukkan sebahagian kecil imej resit pembelian yang digunakan.

Input yang diperlukan adalah berasal daripada gambar atau imbasan resit pembelian Malaysia yang dicetak dan bukan ditulis tangan yang akan dihantar kepada model pembelajaran mendalam PICK berserta OCR melalui aplikasi berasaskan web AlphaOCR.

Sewaktu di fasa pengimbasan, AlphaOCR akan mengimbas dokumen asal menjadi imej digital dengan menggunakan mekanisma pengangkutan dan peranti pengesan yang menukarkan keamatan cahaya menjadi garisan-garisan kelabu (*grey levels*). AlphaOCR akan menukarkan imej pelbagai tahap (*multilevel image*) menjadi dwi-tahap (*bi-level*) yang hanya terdiri daripada warna hitam dan putih bagi menjimatkan ruang memori komputer. Proses ini dikenali sebagai pengambangan atau thresholding dan merupakan proses yang penting kerana output pengecaman adalah bergantung sepenuhnya kepada kualiti imej dwi-tahap yang dihasilkan dalam proses pengimbasan aksara ini. Secara ringkasnya, nilai ambang (*threshold*) akan menentukan sama ada sesuatu garis kelabu (*grey levels*) yang diimbas itu adalah hitam ataupun putih. Sekiranya nilai garis kelabu yang diimbas berada di bawah nilai ambang, maka garis kelabu tersebut akan dianggap sebagai hitam dan begitulah sebaliknya sekiranya nilai garis kelabu tersebut berada di atas nilai ambang yang menjadikannya putih (Chaudhuri 2017).

Imej digital yang terhasil di fasa pengimbasan akan cenderung untuk mengandungi hingar (Chaudhuri 2017). Melalui aktiviti pra-pemprosesan, hingar dapat dihapuskan dengan menggunakan dua teknik iaitu pengisian (*filling*) dan penipisan (*thinning*). AlphaOCR akan melakukan proses pengisian bagi menghapuskan ruangan, dan kecacatan kecil dalam aksara yang telah melalui pendigitalan manakala penipisan pula mengurangkan lebar garisan-garisan pada aksara. Normalisasi adalah suatu proses dalam pra-pemprosesan yang dijalankan bagi memperoleh aksara dengan saiz, condong, dan putaran yang seragam. Empat metode lazim dalam aktiviti normalisasi adalah normalisasi pencong dan pengestrakan asas (*skew normalization and baseline extraction*), normalisasi sendeng (*slant normalization*), normalisasi saiz (*size normalization*), dan pelicinan kontur (*contour smoothing*). Mampatan merupakan aktiviti pra-pemprosesan yang penting bagi mendapatkan saiz imej input yang optimum sekaligus mengurangkan penggunaan storan melalui teknik pengambangan (*thresholding*) dan penjarangan (*thinning*). Teknik pengambangan digunakan bagi menukarkan imej input yang berwarna atau dalam skala kelabu (*gray scale*) menjadi imej dedua (*binary image*) yang hanya terdiri daripada warna hitam atau putih sahaja. Arindam Chaudhuri et. al (2017) menjelaskan bahawa dua kategori terpenting bagi pengambangan adalah viz sejagat (*viz global*) dan viz tempatan (*viz local*). Viz sejagat menetapkan suatu nilai ambangan (*threshold*) yang terpakai terhadap keseluruhan imej berdasarkan anggaran yang dibuat daripada aras latar belakang (*background level*) dan keamatan rajah galur (*histogram intensity*) imej tersebut. Viz tempatan pula menggunakan nilai ambangan (*threshold*) yang berbeza bagi setiap piksel (*pixel*) berpandukan maklumat kawasan setempat (*local area information*). Teknik penjarangan mengekstrak maklumat mengenai rupa bentuk aksara sekaligus mengecilkan saiz imej input. Teknik ini pada kebiasannya digunakan bagi menyeragamkan teks bertulis kerana setiap penulis memiliki gaya penulisan yang berbeza serta mengakibatkan ketebalan garisan yang dihasilkan adalah pelbagai. Teknik penjarangan tidak dilakukan pada teks bercetak kerana ketebalan garisan yang terhasil sentiasa seragam (Reddy 2019).

Penemberengan merupakan proses membahagikan imej aksara kepada subkomponen-subkomponen yang berkaitan. Menurut Susmith Reddy (2019), penemberengan adalah mustahak bagi meningkatkan kadar pengecaman aksara oleh sesebuah sistem OCR. Walaupun pelbagai teknik penemberengan aksara sudah pun diperkenalkan, namun Susmith Reddy (2019) menyatakan bahawa masalah berkenaan penemberengan aksara berangkai masih belum dapat diselesaikan. Terdapat tiga teknik penemberengan aksara iaitu penemberengan tersurat, penemberengan tersirat, dan pendekatan holistik.

Pengekstrakan fitur akan mendapatkan ciri-ciri penting sesuatu aksara yang terdapat pada resit pembelian Malaysia. Teknik-teknik bagi pengekstrakan fitur dapat dibahagikan kepada tiga iaitu taburan titik (*distribution of points*), transformasi dan pengembangan bersiri (*transformation and series expansions*), dan analisis berstruktur (*structural analysis*). Kumpulan fitur yang berbeza akan dinilai berdasarkan kepekaan terhadap hingar (*noise sensitivity*), canggaan (*deformation*), dan kemudahan pelaksanaan dan guna (*ease of implementation and use*) (Chaudhuri 2017). Pengekstrakan fitur juga menjalankan satu tugas penting iaitu pengkelasan. Pengkelasan adalah proses mengenal pasti setiap aksara dan menempatkannya ke dalam kelas aksara yang betul. Dua kategori pengkelasan adalah teori keputusan dan metode berstruktur.

Sistem OCR menggunakan secara meluas metodologi pengecaman corak yang menempatkan sempel yang tidak diketahui kepada suatu kelas yang telah dipratakri. Empat pendekatan umum bagi pengecaman corak adalah padanan templat (*template matching*), teknik statistik (*statistical techniques*), teknik berstruktur (*structural techniques*), dan ANN. Dalam kajian ini, pendekatan yang akan digunakan adalah dengan menggunakan model pembelajaran mendalam PICK berserta OCR. Fasa ini akan berlangsung dengan membangun dan menguji model pembelajaran mendalam PICK berserta OCR yang terbaik untuk tujuan pengekstrakan maklumat pada resit. Aktiviti-aktiviti lain yang berlangsung di fasa ini adalah menghasilkan ujian terhadap reka bentuk algoritma dengan cara membahagikan set data kepada data latihan dan data ujian, dan yang terakhir adalah menilai prestasi

model melalui kefahaman berkenaan domain, tahap pencapaian objektif, dan reka bentuk ujian. Fasa ini perlu diulang sehingga model yang optimum berjaya dihasilkan sebelum meneruskan ke fasa seterusnya (Hotz 2022).

Pasca-pemprosesan adalah mustahak bagi menghasilkan output yang diinginkan. Antara aktiviti lazim bagi pasca-pemprosesan adalah pengkelasan dan pengesanan ralat dan pembetulan. Aktiviti pasca-pemprosesan bagi pembangunan projek ini amat sukar kerana setiap penjual menggunakan istilah atau nama yang berbeza walaupun barang jualan antara kesemua kedai tersebut adalah sama. Namun begitu, skop kajian adalah terhad kepada resit Malaysia dan NLP berskala besar bagi mengelaskan produk yang sama antara kedai yang berbeza adalah diluar skop projek ini.

Bagi fasa penilaian, metrik penilaian yang digunakan untuk menilai prestasi model pembelajaran mendalam PICK terbaik adalah *Mean Entity F1-Score (mEF)*. Metrik penilaian yang digunakan bagi menilai model OCR terbaik pula adalah *Character Error Rate (CER)* manakala bagi ujian kebolegunaan pula digunakan bagi menilai prestasi aplikasi berasaskan web yang dibangunkan.

Output yang dijana adalah hasil aksara yang diesktrak daripada resit pembelian Malaysia dengan betul mengikut atribut yang telah ditetapkan. Model pengecaman corak yang berasaskan model pembelajaran mendalam PICK berserta OCR ini akan ditempatkan bagi kegunaan umum dalam bentuk sebuah aplikasi berasaskan web di suatu server yang boleh dicapai semua pengguna bagi tujuan pengekstrakan maklumat pada resit sesuai dengan objektif kajian yang telah dinyatakan.

Keputusan dan Perbincangan

Aktiviti pengujian bagi AlphaOCR dilakukan dengan menggunakan tiga teknik pengujian yang berbeza. Jadual 1.1 di bawah menunjukkan ringkasan bagi ketiga-tiga teknik atau pelan pengujian yang dijalankan.

Jadual 1.1 Pelan Pengujian Yang Dijalankan

Komponen	Jenis Pengujian	Model Terlibat
Model Pembelajaran Mendalam	Menilai prestasi dua buah model PICK melalui beberapa metrik penilaian	PICK-1 PICK-2
Model OCR	Menilai prestasi tiga buah model OCR melalui ketepatan pengekstrakan dua puluh frasa dengan betul	TesseractOCR KerasOCR EasyOCR
Aplikasi Berasaskan Web	Menjalankan ujian kebolegunaan terhadap aplikasi yang dibangunkan	AlphaOCR

Melalui ujian yang dijalankan ke atas dua buah model pembelajaran mendalam PICK iaitu PICK-1 dan PICK-2, kita dapati bahawa kedua-dua model berupaya mengekstrak maklumat daripada resit dengan baik dengan nilai mEF yang setara bagi PICK-1 dan PICK-2 iaitu masing-masing sebanyak 83.32% dan 83.88%. Walau bagaimanapun, model PICK-2 menunjukkan prestasi yang lebih baik berbanding PICK-1. Jadual 1.2 dan Jadual 1.3 menunjukkan keputusan terperinci yang diperoleh melalui ujian yang dijalankan ke atas model pembelajaran mendalam PICK.

Jadual 1.2 Keputusan latihan model PICK-1

Entiti	mEP	mER	mEF	mEA	GL Loss	CRF Loss	Total Loss
<i>address</i>	0.8782 61	0.8058 51	0.8404 99	0.8058 51			
<i>total</i>	0.7788 78	0.7737 70	0.7763 16	0.7737 70			
<i>date</i>	0.9683 54	0.9807 69	0.9745 22	0.9807 69			
<i>company</i>	0.7272 73	0.7826 09	0.7539 27	0.7826 09			
Keseluruhan	0.8441 99	0.8223 90	0.8331 52	0.8223 90	0.0035 44	6.5727 36	6.9271 27

Jadual 1.3 Keputusan latihan model PICK-2

Entiti	mEP	mER	mEF	mEA	GL Loss	CRF Loss	Total Loss
<i>address</i>	0.8575 27	0.8484 04	0.8529 41	0.8484 04			
<i>total</i>	0.7500 00	0.8152 17	0.7812 50	0.8152 17			
<i>date</i>	0.9012 35	0.9358 97	0.9182 39	0.9358 97			

<i>company</i>	0.8150 68	0.7803 28	0.7973 20	0.7803 28			
Keseluruhan	0.8401 73	0.8374 60	0.8388 14	0.8374 60	0.0035 61	5.4333 38	5.7893 97

Bagi model OCR pula, TesseractOCR mencapai keputusan ketepatan pengekstrakan frasa purata yang hampir sama dengan EasyOCR iaitu masing-masing sebanyak 0.9266 dan 0.9204. Walaupun TesseractOCR mencapai keputusan tertinggi dalam ketepatannya mengekstrak aksara, namun TesseractOCR memisahkan frasa menjadi perkataan-perkataan yang berbeza berbanding dengan EasyOCR yang mengambil kira frasa dan semantik perkataan dalam pengekstrakannya. Dengan mengambil kira aspek ini, model OCR yang terbaik dan sesuai mengikut kegunaan projek ini adalah EasyOCR kerana memiliki ketepatan pengekstrakan yang setara dengan TesseractOCR tetapi dalam masa yang sama berkebolehan dalam mengambil kira makna atau semantik setiap frasa yang diekstrak. Berikut merupakan keputusan pengujian keseluruhan bagi tiga model OCR yang diuji.

Jadual 1.4 Keputusan pengujian keseluruhan model OCR

Model	CER
TesseractOCR	0.9266 3108 92
KerasOCR	0.7572 3917 43
EasyOCR	0.9203 9880 84

Pengujian aplikasi berasaskan web adalah berpandukan teknik ujian kebolegunaan terhadap aplikasi yang dibangunkan. Mengikut sebuah kajian yang dijalankan oleh Jakob Nielsen (2012), keputusan ujian kebolegunaan yang terbaik boleh diperolehi dengan cara menguji sesebuah aplikasi kepada tidak lebih daripada lima orang pengguna bagi setiap ujikaji kecil yang ingin dijalankan. Dengan mengambil kira panduan ini, pengujian aplikasi berasaskan web dijalankan kepada lima belas orang pengguna melalui tiga iterasi atau kumpulan berbeza dan keputusan maklum balas yang diperolehi adalah seperti dalam jadual berikut.

Jadual 1.5 Keputusan pengujian kebolegunaan aplikasi

Soalan	Keputusan	Analisis
<i>What are you thinking as you view the web-based AI?</i>	<p><i>For me personally everyone can access web based ai because they only need a web browser to access it compared to application based AI.</i></p> <p><i>Minimal and interesting user interface.</i></p> <p><i>I think of it as a regular website with functioning features (extract information from receipts).</i></p> <p><i>I think is was attractive.</i></p> <p><i>First impression was very great, i love the web design, it is very aesthetic and colors were on point.</i></p>	<p>Sebanyak tiga daripada lima orang responden menyentuh aspek kosmetik aplikasi yang tampak menarik. Ini menunjukkan bahawa aspek kosmetik sesuatu aplikasi memainkan peranan penting dalam mengekalkan minat pengguna untuk terus menggunakan aplikasi.</p>
<i>How was the experience of using the product to complete this task?</i>	<p><i>Curious because I kinda explore and expose myself to a new thing, the website interface is simple which i can easily using it without need to really think about how to using the website.</i></p> <p><i>Simple and easy to use.</i></p> <p><i>The product experience was fun and easy, despite encountering errors for a few times.</i></p> <p><i>It is cool.</i></p> <p><i>It was easy and simple to use, I believe many people can benefit from this project when it comes to documentation.</i></p>	<p>Sebanyak empat daripada lima orang responden berpendapat bahawa aplikasi ini mudah untuk digunakan.</p>
<i>What are your thoughts on the instructions provided?</i>	<p><i>The instruction is direct-foward and easily to understand.</i></p> <p><i>The instruction is very clear.</i></p> <p><i>The instructions were simple and direct.</i></p> <p><i>It was easy to understand.</i></p> <p><i>It was clear and precise.</i></p>	<p>Kesemua lima orang responden bersetuju bahawa arahan yang dipaparkan dalam aplikasi sangat jelas dan mudah difahami.</p>
<i>How easy or difficult was it to navigate?</i>	<p><i>It is easy to navigate.</i></p> <p><i>Easy to navigate.</i></p> <p><i>The navigation was fairly easy.</i></p>	<p>Kesemua lima orang responden bersetuju bahawa aplikasi ini mudah untuk diterokai dan digunakan.</p>

	<p><i>Easy.</i></p> <p><i>It is easy to navigate as there's a flow for the navigation process.</i></p>	
<i>What are your thoughts on the design and layout?</i>	<p><i>My time almost 15-17 minutes because of the technical error and I read and do the instruction 1 by 1</i></p> <p><i>Design is very minimal, and the layout is perfectly arranged</i></p> <p><i>The aesthetics, design and layout of the website is minimalistic and well done.</i></p> <p><i>Design and layout are arranged to make it easier to navigate.</i></p> <p><i>Design and layout is very good. It is well structured, simply to say, it is pleasant to see web design.</i></p>	<p>Sebanyak empat daripada lima orang responden berpendapat bahawa reka bentuk antara muka aplikasi adalah kemas dan cantik.</p>
<i>How long did it take the user to complete this task?</i>	<p><i>Estimated around 5-10 minutes.</i></p> <p><i>Around 5 minutes.</i></p> <p><i>Approximately 20 minutes to complete the task for one image.</i></p> <p><i>It is not take too long and not too fast. Just nice!</i></p> <p><i>It didn't take long, it took less than 5 minutes for generating the output for 1 receipt.</i></p>	<p>Sebanyak empat daripada lima orang responden berpendapat bahawa mereka mengambil masa kurang dari 6 minit untuk menyiapkan tugas.</p>
<i>How would you describe your overall experience with the product?</i>	<p><i>My overall experience, the product is easy to access and it is easy to use which we can freely adjust box of the information that we want to extract and this product I think have a bright future ahead</i></p> <p><i>Very satisfy with the user experience</i></p> <p><i>Overall experience was refreshing and interesting.</i></p> <p><i>It is the greatest experience.</i></p> <p><i>It was a refreshing experience, it made user imagine how beneficial it is if it were to be implemented.</i></p>	<p>Kesemua lima orang responden berpendapat bahawa penggunaan aplikasi ini adalah sesuatu yang baharu dan mudah digunakan.</p>
<i>What did you like the most about using this product?</i>	<p><i>The adjustment of the box to extract the information and the ability to read the receipt.</i></p>	<p>Sebanyak tiga orang responden paling menyukai keupayaan aplikasi mengekstrak maklumat</p>

	<p><i>Simple and clean layout.</i></p> <p><i>The information extraction AI achieved its function and purpose to a degree, showing potential for future development. The extraction process is also transparent and allows real-time updates.</i></p> <p><i>The process of extracting the information.</i></p> <p><i>The convenience usability. I believe even elderly able to use this product just through the guide in the web page.</i></p>	<p>pada resit.</p> <p>Seorang responden paling menyukai susun atur bentangan aplikasi</p> <p>Seorang responden paling menyukai kondisi aplikasi yang mesra pengguna dan mudah digunakan oleh orang yang lebih tua</p>
<p><i>What did you like the least?</i></p>	<p><i>The error when I try to get all of the information of the receipt.</i></p> <p><i>Some text boxes are too small to read, but still readable.</i></p> <p><i>None to be honest, only the few errors and technical bugs but it is normal and totally understandable.</i></p> <p><i>Nothing.</i></p> <p><i>The loading time.</i></p>	<p>Sebanyak dua orang responden tidak menghadapi apa-apa masalah.</p> <p>Sebanyak dua orang responden menghadapi ralat sewaktu pengekstrakan.</p> <p>Seorang responden tidak menyukai saiz ruangan teks yang kecil.</p>
<p><i>What, if anything, surprised you about the experience?</i></p>	<p><i>The experience using new developed ai technology.</i></p> <p><i>The AI that can read any type of font provided is cool.</i></p> <p><i>The information extraction was accurate to a degree. Alphaocr also had nailed the user interface and experience (UI/UX).</i></p> <p><i>I was speechless I guess.</i></p> <p><i>The modeling structure for this product. I didn't expect it to implement different deep learning models.</i></p>	<p>Kesemua lima orang responden menyukai keupayaan AI mengekstrak maklumat pada resit.</p>
<p><i>What, if anything, caused you frustration?</i></p>	<p><i>The error.</i></p> <p><i>The error that happened when i selected the whole receipt, not only the body part.</i></p> <p><i>The few errors which prohibited me to obtain the output (image 1).</i></p>	<p>Sebanyak tiga orang responden paling tertekan ketika menghadapi ralat.</p>

	<i>I will consider that is not my sustenance.</i>	
	<i>The loading time.</i>	
<i>On a scale from 1 to 5 (1=not at all likely, 5=very likely), how likely are you to recommend this product to a friend?</i>	4, likely. 4 out of 5. 5. 5. 5.	Kesemua lima orang responden memberikan mahu memperkenalkan produk ini kepada kawan.
<i>How frequently would you use this product?</i> [Option 1:] Never [Option 2:] Very Rarely (once per month) [Option 3:] Rarely (2-3 times/month) [Option 4:] Occasionally (2-3 times/week) [Option 5:] Frequently (1-2 times/day) [Option 6:] Very Frequently (3+ times/day)	Frequently. Occasionally. 3. Never. I have never seen before. 4.	Setiap responden memiliki matlamat berbeza.
<i>Please paste the txt files' output generated.</i>	<pre>{'date': '12-01-19', 'company': None, 'address': 'Kamasan PERINDUSTRIAN ', 'product': 'AIR PRESSUPE SPRAYER SX-575-1 1.5L', 'price': '3.48', 'total': '33.f0'} {'date': '10/08/2017', 'company': 'TRI SHAAS SDN BHD', 'address': None, 'product': 'ien 0 EYLOPE ISINXTOIN TQ-S1015', 'price': '2.60', 'total': None} {'date': '12-01-19', 'company': None, 'address': 'tan w', 'product': 'chopping BOARO 35.5x25.5CM 803m#', 'price': '1 X 19.00', 'total': '33.f0'} {'date': '12-01-19', 'company': None, 'address': 'tan w', 'product': 'AIR PRESSUPE SPRAYER SX-575-1 1.5L', 'price': 'X 19.00', 'total': '33.f0'} {'date': '10/08/2017', 'company': 'TRI SHAAS SDN BHD', 'address': None, 'product': 'GS1Ju: 000429165592', 'price': '2.50', 'total': None}</pre>	Kesemua lima orang responden berjaya mengekstrak maklumat pada resit.

Kesimpulan

Secara keseluruhannya, pembangunan aplikasi berasaskan web AlphaOCR memperolehi maklum balas yang positif daripada pengguna di samping berupaya meningkatkan nilai mEF model pembelajaran mendalam PICK daripada 83.32% ke 83.88%. Model OCR yang sesuai untuk digunakan bagi mencapai matlamat kajian ini adalah EasyOCR dengan nilai CER yang setara dengan TesseractOCR iaitu 92.04%. Ini menjadikan ketepatan model AlphaOCR secara keseluruhannya adalah 87.96%.

Kajian ini telah berjaya mencapai objektif pembangunannya iaitu membangunkan dan mengenal pasti model pembelajaran mendalam terbaik dan OCR terbaik untuk mengekstrak maklumat dari resit dan membangunkan sebuah aplikasi berasaskan web yang berkebolehan mengekstrak maklumat dengan menggunakan model pembelajaran mendalam dan OCR terbaik menghasilkan set data baharu berkenaan maklumat pembelian bagi tujuan kajian pada masa akan datang.

Walau bagaimanapun, kajian ini boleh ditambah baik pada masa akan datang dengan menambah entiti yang boleh diekstrak pada resit seperti kuantiti dan kod diskaun. Selain itu, teknik NLP juga boleh digunakan bagi meningkatkan semantik pengekstrakan teks seperti menggabungkan singkatan pada dua resit berbeza yang memiliki makna yang sama.

Penghargaan

Segala pujian bagi Allah atas kurniaan ilhamNya di sepanjang berlangsungnya perancangan dan pembangunan bagi projek ini.

Terima kasih yang tidak terhingga kepada penyelia projek tahun akhir saya, Ts. Dr. Nor Samsiah Sani atas dorongan dan panduan bermakna yang telah diberikan selama setahun kajian ini dilakukan. Juga kepada para pensyarah dan pihak Fakulti Teknologi & Sains Maklumat yang telah banyak

memberikan tunjuk ajar melalui sesi taklimat projek tahun akhir mingguan serta bahan-bahan rujukan berkaitan.

Terima kasih yang tidak terkata kepada ibu bapa, keluarga, dan rakan-rakan seperjuangan atas doa, sokongan dan tunjuk ajar yang telah diberikan kepada saya yang bergelar insan. Akhir bicara, terima kasih dan tahniah kepada diri saya sendiri kerana telah berjaya melengkapkan projek tahun akhir dengan penuh dedikasi walau pelbagai kekangan dan halangan yang muncul sepanjang kajian ini dijalankan. Anda memang hebat dan kacak.

RUJUKAN

Chaudhuri, A., Mandaviya, K., Badelia, P., Ghosh, S.K. (2017). Optical Character Recognition Systems for Different Languages with Soft Computing. Warsaw: Springer.

Hotz, N. (2022). What is CRISP DM? <https://www.datascience-pm.com/crisp-dm-2/> [16 November 2022]

IBM Cloud Education. (2022). What Is Optical Character Recognition (OCR)? <https://www.ibm.com/cloud/blog/optical-character-recognition> [16 November 2022]

Nielsen, J. (2012). Usability 101: Introduction To Usability. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>

Reddy, S. (2019). Segmentation In OCR. <https://towardsdatascience.com/segmentation-in-ocr-10de176cf373> [16 November 2022]

Singh, S. (2023). Pattern Recognition – Phases and Activities. <https://www.geeksforgeeks.org/pattern-recognition-phases-and-activities/> [1 Mac 2023]

Muhammad Haikal Iman Bin Osman (A179376)
Ts. Dr. Nor Samsiah Sani
Fakulti Teknologi & Sains Maklumat,
Universiti Kebangsaan Malaysia