

BLAZEX: MODEL PENGECEMAN EKSPRESI MUKA BERASASKAN PEMBELAJARAN MENDALAM

Firdanish Rafique Mohd Hamid & Nor Samsiah Sani

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

Abstrak

Projek ini bertujuan untuk membangunkan sistem Pengecaman Ekspresi Wajah (FER) menggunakan teknologi pembelajaran mendalam dan set data FER2013. Proses pembangunan melibatkan pemilihan model pembelajaran mendalam terbaik untuk FER dan pemrosesan data yang canggih untuk meningkatkan kualiti set data. Model yang dibangunkan dioptimalkan agar mencapai prestasi maksimum, dengan ketepatan keseluruhan mencapai 95.5% semasa diuji dengan set data FER2013. Namun, ketika diuji dengan set data KDEF yang berbeza, prestasi menurun menjadi 47.9%. Model FER yang telah dioptimalkan akan diintegrasikan ke dalam aplikasi dunia nyata dan disepadukan dalam sistem aplikasi web BlazeX untuk memberi pengecaman dan analisis masa nyata. Sistem ini menawarkan potensi besar untuk pelbagai aplikasi seperti pengecaman ekspresi dalam gambar dan video, analisis sentimen, penilaian respons pengguna dan sistem kecerdasan buatan. Dengan teknologi FER ini, organisasi dapat membuat keputusan berasaskan data dan memperoleh analisis mendalam tentang ekspresi manusia dalam pelbagai konteks kehidupan dunia nyata.

Kata kunci: Pengecaman Ekspresi Muka, Pembelajaran Mendalam, Convolutional Neural Network, Aplikasi Web

Pengenalan

Dalam bahagian Pengenalan bagi laporan teknikal ini, beberapa elemen penting yang akan diterangkan secara ringkas dan komprehensif seperti latar belakang, objektif, skop, justifikasi dan kepentingan, metodologi dan organisasi laporan.

Dalam beberapa tahun kebelakangan ini, bidang *computer vision* mengalami kemajuan yang pesat terutama dalam bidang *facial expression recognition* (FER). FER telah menjadi topik yang menarik dalam penyiasatan kerana pelbagai aplikasinya dalam dunia nyata, termasuk dalam sektor

keselamatan, hiburan, pendidikan dan perubatan (Gan & Zhang 2022). Ekspresi muka (Gambarajah 1) berfungsi sebagai saluran komunikasi bukan lisan yang penting, menyampaikan maklumat berharga tentang emosi, niat dan reaksi seseorang. Memahami dan mengenali ekspresi muka adalah kunci untuk meningkatkan interaksi manusia-teknologi, meningkatkan pengalaman pengguna dan meningkatkan kualiti perkhidmatan.



Gambarajah **Error! No text of specified style in document.** 7 ekspresi muka universal yang ditunjukkan oleh manusia iaitu kegembiraan, kesedihan, kemarahan, ketakutan, kejutan, kejijikan dan neutral

Dalam dunia sebenar, permintaan untuk teknologi FER meningkat dengan pesat. FER memainkan peranan penting dalam membantu organisasi memahami kepuasan pelanggan dan menilai pengalaman pengguna mereka. Dengan mengenali ekspresi muka, organisasi boleh mengumpulkan cerapan berharga tentang perasaan dan kepuasan pelanggan terhadap produk dan perkhidmatan yang mereka tawarkan.

Oleh itu, projek ini bertujuan untuk membangunkan teknologi FER berdasarkan model *deep learning* (DL) yang relevan dan boleh digunakan dalam dunia sebenar. Teknologi ini akan menganalisis dan mengklasifikasikan ekspresi muka dalam interaksi masa nyata dengan sistem, serta memberikan maklum balas segera tentang tahap kepuasan pengguna. Dengan menggunakan teknologi FER yang canggih, organisasi boleh memperoleh cerapan berharga tentang kepuasan pengguna dan mengambil langkah proaktif untuk meningkatkan kualiti perkhidmatan dan produk mereka.

Isu dan masalah yang ingin diselesaikan oleh projek ini adalah untuk membangunkan model FER yang tepat dan boleh dipercayai untuk mengenali ekspresi muka dengan betul dalam situasi dunia sebenar. Model FER yang canggih ini dijangka dapat meningkatkan pengalaman pengguna,

meningkatkan kualiti perkhidmatan dan membantu organisasi membuat keputusan berdasarkan pemahaman mendalam tentang perasaan dan kepuasan pelanggan.

Projek ini mempunyai kaitan dan kepentingan yang tinggi kerana keupayaannya untuk mengenali dan memahami ekspresi muka adalah aspek penting dalam pembangunan produk dan perkhidmatan di dunia sebenar. Dengan memahami ekspresi muka pelanggan, organisasi boleh menyesuaikan strategi pemasaran, meningkatkan antara muka pengguna dan menyediakan perkhidmatan yang lebih cekap.

Pembangunan teknologi FER yang sesuai dan cekap adalah penting dalam meningkatkan pengalaman pengguna dan memberi organisasi kelebihan daya saing dalam dunia yang serba pantas hari ini. Projek ini memberikan sumbangan berharga untuk menangani cabaran dalam pengecaman ekspresi muka dan memberi peluang untuk mempersembahkan teknologi inovatif yang berguna dalam pelbagai aplikasi dunia sebenar.

Objektif utama projek ini adalah untuk membangunkan dan mengenalpasti model FER berdasarkan DL yang dapat mengesan dan mengelaskan ekspresi muka dengan tepat. Projek ini bertujuan untuk meningkatkan prestasi model FER melalui penalaan hiperparameter semasa proses latihan model. Matlamatnya adalah untuk mengenalpasti nilai hiperparameter yang optimum bagi mencapai prestasi model yang lebih baik daripada sebelumnya. Selain itu, projek ini juga bertujuan untuk menilai prestasi model dengan menguji model pada set data yang berbeza. Tujuannya adalah untuk menilai keupayaan generalisasi model, iaitu kemampuannya untuk mengenali ekspresi wajah dengan tepat dalam situasi dunia nyata yang berbeza. Selain itu, penilaian ini juga akan menilai kesesuaian model FER yang dibangunkan untuk aplikasi dunia nyata, terutamanya dalam meningkatkan pengalaman pengguna dan kualiti perkhidmatan. Dengan mencapai kesemua objektif ini, projek ini diharapkan dapat menyumbang secara signifikan dalam pembangunan model FER yang tepat, mantap dan dapat digunakan secara meluas dalam pelbagai aplikasi dunia nyata. Keberhasilan

projek ini dapat membawa manfaat besar dalam meningkatkan interaksi manusia dengan teknologi, pengalaman pelanggan dan kepuasan pengguna secara keseluruhan.

Projek ini bertujuan untuk membina dan menggunakan teknologi FER yang direka khusus untuk aplikasi dunia nyata. Komponen utama produk akhir termasuk pembangunan model FER berdasarkan DL, penyepaduan kamera web untuk menangkap imej muka masa nyata sebagai input dan pengecaman ekspresi muka masa nyata menggunakan kamera web, aplikasi web untuk menjalankan model DL dengan input antara muka pengguna juga akan dibangunkan. Aplikasi web ini membolehkan pengguna mendapatkan maklum balas segera tentang ekspresi muka mereka dalam interaksi masa nyata. Ini akan memperluas fungsi model FER dan meningkatkan perkaitannya dalam menangani keperluan dunia sebenar. Model FER akan difokuskan pada mengenali dan mengklasifikasikan ekspresi wajah yang ditunjukkan dengan tepat dan model ini akan dilatih menggunakan set data imej muka pelbagai yang telah diberi label ekspresi yang sepadan.

Salah satu kekangan utama dalam pelaksanaan projek ini adalah masalah masa dan sumber. Pembangunan teknologi FER berdasarkan DL memerlukan masa dan sumber yang besar untuk menjalankan pelbagai latihan dan penalaan model. Keperluan ini mungkin menyebabkan keterbatasan dalam pelaksanaan dan kemajuan projek jika sumber yang mencukupi tidak tersedia. Selain itu, penggunaan perkakasan yang memadai juga menjadi kekangan. Latihan model DL memerlukan akses kepada sumber pengiraan yang canggih seperti GPU dan CPU. Keterhadan akses kepada perkakasan yang mencukupi boleh mempengaruhi skop projek dan kerumitan model yang dapat dibangunkan. Aspek privasi data juga merupakan kekangan penting yang perlu diambil kira. Penggunaan set data dunia nyata yang mengandungi imej muka individu mungkin menimbulkan kebimbangan privasi, terutama jika imej-imej tersebut mengandungi maklumat peribadi yang sensitif. Oleh itu, untuk memastikan kepatuhan undang-undang dan peraturan privasi data, projek ini hanya akan menggunakan set data yang telah dibuat tersedia untuk umum dan tidak melanggar privasi data individu.

Projek ini menghadapi beberapa kekangan yang perlu diatasi untuk mencapai tujuan dan hasil yang diinginkan. Usaha yang berterusan dan strategi yang sesuai akan diperlukan untuk memastikan kejayaan projek dan pembangunan teknologi FER yang cekap dan selamat untuk aplikasi dunia nyata.

Projek ini dilaksanakan dengan tujuan untuk mengembangkan dan mengoptimumkan model FER berasaskan DL untuk aplikasi dunia sebenar yang berfokus pada penggunaan masa nyata. Justifikasi dan kepentingan projek ini dapat diuraikan dalam beberapa aspek.

Pertama, dalam bidang ilmu FER, projek ini memberi sumbangan penting dengan mengintegrasikan teknologi terkini DL, khususnya CNN. Penyelidikan dalam bidang ini terus berkembang pesat dan memainkan peranan penting dalam *computer vision* dan pembelajaran mesin. Dengan meningkatnya keperluan akan aplikasi yang memahami ekspresi dan emosi manusia, model FER yang tepat dan cekap sangat diperlukan untuk menjawab cabaran dalam bidang ini.

Kedua, dalam aplikasi dunia sebenar, kesan dan manfaat projek ini adalah jelas dalam meningkatkan interaksi manusia-komputer dan pengalaman pengguna secara keseluruhan. Model FER yang dibangunkan dalam projek ini akan memungkinkan sistem untuk mengenali dan menginterpretasi ekspresi muka manusia secara real time. Dengan demikian, sistem yang menggunakan model ini dapat memberikan respons dan tindakan yang lebih tepat dalam situasi yang memerlukan pemahaman emosi manusia.

Ketiga, manfaat projek ini adalah bagi industri yang memerlukan analisis dan pengecaman ekspresi muka dalam penggunaan masa nyata. Misalnya, dalam sektor peruncitan, sistem FER ini dapat membantu memahami reaksi pelanggan terhadap produk atau perkhidmatan tertentu. Begitu juga, dalam sektor pengangkutan awam, sistem FER yang berfokus pada penggunaan masa nyata dapat membantu mengenal pasti ketidakpuasan hati pelanggan dan mengambil tindakan segera untuk meningkatkan kualiti perkhidmatan.

Keempat, penggunaan model FER dalam aplikasi dunia sebenar juga membuka potensi penggunaan luas dalam bidang sosial dan psikologi. Ia dapat digunakan untuk memantau respons

emosi dalam konteks sosial, membantu dalam terapi psikologi dan memahami reaksi manusia dalam interaksi antara manusia.

Oleh yang demikian, projek ini memberikan justifikasi yang kukuh dan kepentingan yang jelas dalam mengembangkan model FER berasaskan DL untuk aplikasi dunia sebenar yang berfokus pada penggunaan masa nyata. Dengan mengintegrasikan teknologi terkini dan memanfaatkan potensi model FER dalam meningkatkan interaksi dan pemahaman manusia, projek ini diharapkan akan memberi sumbangan yang berharga kepada bidang ilmu FER dan memberikan manfaat nyata dalam berbagai industri dan aplikasi dunia sebenar.

Dalam memahami kepentingan FER, kajian literatur yang telah dilaksanakan menekankan bagaimana teknologi ini telah memberikan analisis berharga tentang tahap kepuasan pelanggan dalam pelbagai industri. Kepuasan pelanggan adalah konsep pelbagai dimensi yang dipengaruhi oleh kualiti perkhidmatan, kemudahan, kebolehcapaian dan keupayaan. Oleh itu, memahami ekspresi muka pelanggan memberikan pandangan berharga tentang keadaan emosi mereka dan tahap kepuasan keseluruhan.

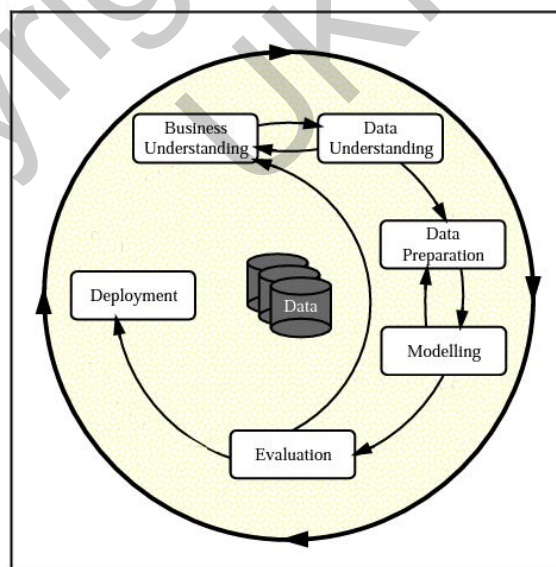
Pengaplikasian FER dalam industri dunia sebenar telah membuka peluang baru untuk meningkatkan kualiti perkhidmatan dan strategi pemasaran. Penyedia perkhidmatan awam dapat mengenal pasti ketidakpuasan hati pelanggan dalam masa nyata dan mengambil tindakan pembetulan yang segera. Demikian juga, dalam sektor peruncitan, FER memainkan peranan penting dalam memahami pilihan pelanggan dan menyediakan pengalaman yang disesuaikan untuk meningkatkan kepuasan.

Dari kajian literatur, ditekankan bahawa model DL, terutamanya CNN, telah menunjukkan hasil yang menjanjikan dalam tugas FER. Model ini mampu mempelajari perwakilan hierarki dari imej muka dan mengekstrak ciri diskriminatif untuk mengenali dan mentafsir ekspresi muka dengan tepat. Hasil ini memberikan keyakinan bahawa penggunaan model DL adalah pilihan utama dalam pembangunan model FER yang berkesan.

Dalam perbandingan dan perbezaan antara beberapa kajian, diketahui bahawa penggunaan model DL, terutamanya CNN, memberikan ketepatan yang tinggi dalam pengecaman dan klasifikasi ekspresi muka. Walaupun terdapat beberapa pendekatan alternatif seperti pengelasan Haar dan teknik hibrid, CNN tetap menjadi pilihan utama dalam membina model FER yang tepat dan boleh dipercayai.

Dengan memahami kajian sedia ada dan implikasinya, projek ini dapat menggunakan pengetahuan latar belakang ini untuk menghasilkan model FER yang mantap dan berkesan dalam memahami ekspresi muka pelanggan dalam aplikasi berasaskan pelanggan. Penggunaan teknik DL, khususnya CNN, diharapkan dapat memberikan ketepatan yang tinggi dalam pengecaman ekspresi muka dan membantu meningkatkan pengalaman pelanggan dalam pelbagai sektor dunia sebenar.

Metodologi yang digunakan dalam projek ini ialah Cross-Industry Standard Process for Data Mining (CRISP-DM) (rujuk Gambarajah 2), yang menyediakan pendekatan berstruktur dan berulang untuk membangunkan model FER. CRISP-DM memastikan kemajuan sistematik daripada memahami keperluan projek kepada penggunaan akhir model FER.



Gambarajah 2 Seni bina model CRISP-DM

Sumber: Wirth & Hipp 2000

Fasa pertama dalam metodologi ialah Pemahaman Perniagaan, di mana keperluan dan objektif projek ditentukan untuk aplikasi dunia sebenar. Langkah seterusnya ialah Pemahaman Data, di mana data tentang ekspresi muka dikumpul dan dimanipulasi. Set data yang sesuai telah dipilih dan teknik pra-pemprosesan digunakan untuk memastikan kualiti optimum untuk latihan dan penilaian model FER.

Fasa Penyediaan Data melibatkan pembersihan, mengubah dan memperhalusi data untuk menyediakan set data yang sesuai untuk melatih model FER. Teknik pra-pemprosesan, seperti *cropping*, *smoothing*, *resizing* dan *rotation* digunakan untuk meningkatkan kualiti set data.

Pada peringkat Permodelan, model FER dibangunkan dan disempurnakan. Teknik Pembelajaran Dalam khususnya CNN digunakan untuk latihan model menggunakan set data yang telah disediakan. Pelbagai seni bina, algoritma pengoptimuman dan penalaan hiperparameter diterokai untuk mencari konfigurasi yang paling berkesan dalam FER.

Fasa Penilaian menilai model FER yang dibangunkan dengan teliti menggunakan metrik pemarkahan dan set data yang sesuai. Prestasi model diukur dari segi ketepatan, ingatan dan skor-F1 untuk mendapatkan gambaran tentang keberkesanan, keteguhan dan keupayaan model untuk digunakan dalam pelbagai situasi.

Selepas model FER memenuhi kriteria prestasi yang diinginkan, ia akan digunakan dan disepadukan ke dalam aplikasi web, menggunakan kamera web sebagai sumber input. Antara muka pengguna akan dibuat untuk menyediakan pengecaman ekspresi muka masa nyata dan memberikan maklum balas.

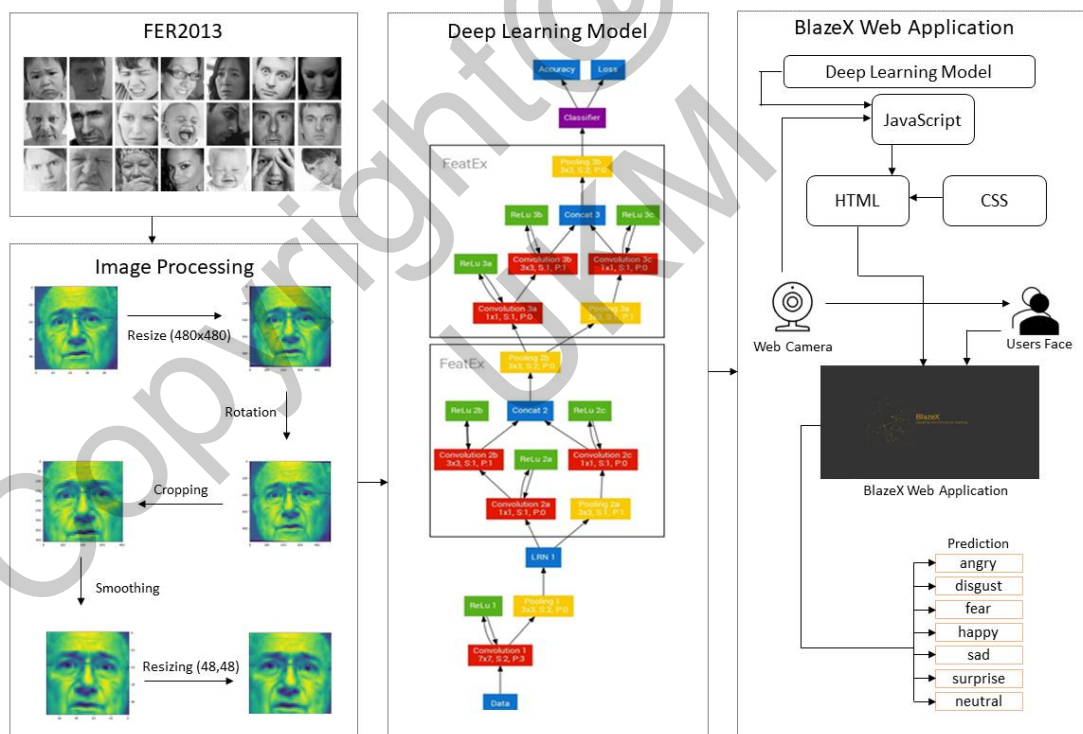
CRISP-DM dipilih untuk kesesuaiannya dalam projek ini kerana fleksibilitinya dalam membenarkan binaan berulang, pelarasan dan peningkatan pada setiap peringkat berdasarkan hasil penilaian sepanjang projek. Metodologi ini menggalakkan pendekatan berstruktur yang memastikan pemahaman menyeluruh tentang matlamat perniagaan, keperluan data dan penilaian prestasi model.

Menggunakan metodologi CRISP-DM, proyek ini memastikan proses binaan yang sistematis dan teratur, penggunaan sumber yang cekap dan penyampaian teknologi FER berkualiti tinggi.

Seterusnya, dalam laporan teknikal ini akan diterangkan diikuti dengan bahagian Metodologi Kajian, Keputusan dan Perbincangan dan Kesimpulan.

Metodologi Kajian

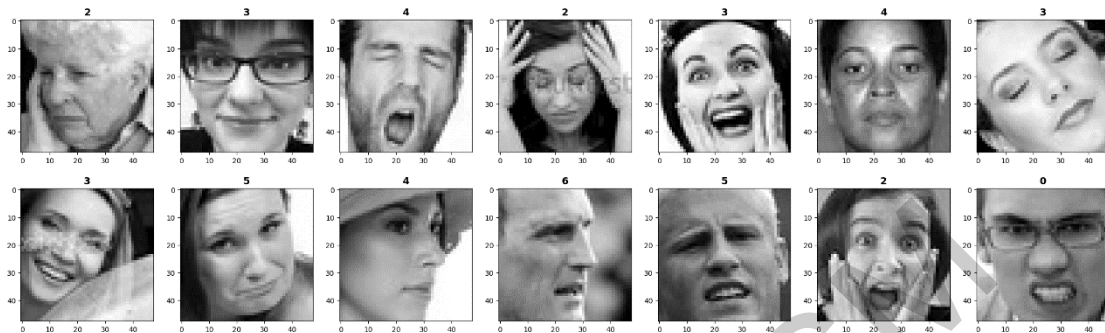
Proses pembangunan merupakan bahagian yang kritikal dalam projek ini, di mana kesemua komponen aplikasi dihasilkan dan diintegrasikan dengan lancar. Proses pembangunan terdiri daripada beberapa langkah penting, termasuk pemilihan set data, pemrosesan data, pembangunan model pengecaman dan pembangunan web aplikasi berintegrasi model pengecaman. Gambarajah 3 menunjukkan carta alir bagi keseluruhan sistem aplikasi Blaze X.



Gambarajah 3 Carta Alir Keseluruhan Sistem Aplikasi BlazeX

Set data yang dipilih untuk tujuan latihan model DL adalah FER2013, iaitu set data yang mengandungi imej wajah manusia dengan ekspresi yang berbeza. FER2013 adalah satu set data awam

yang digunakan secara meluas dalam pembangunan model pengecaman ekspresi berdasarkan DL (Goodfellow et al. 2013).



Gambarajah **Error! No text of specified style in document.** Contoh-contoh imej ekspresi muka dalam set data FER2013 yang terdiri daripada 7 kategori ekspresi berbeza (marah, jijik, takut, gembira, sedih, terkejut dan neutral)

FER2013 mengandungi lebih daripada 35,000 imej wajah dalam 7 kategori ekspresi berbeza, iaitu marah, jijik, takut, gembira, sedih, terkejut dan neutral (rujuk Gambarajah 4). Set data ini adalah unik dan mencerminkan pelbagai keadaan ekspresi yang berbeza, menjadikannya ideal untuk latihan dan pengujian model FER.

Set data FER2013 terdiri daripada imej wajah dengan pelbagai dimensi dan resolusi. Untuk memastikan konsistensi dan ketepatan dalam pembangunan model, semua imej wajah dalam set data telah diambil alih dan diuruskan dengan baik. Sebagai persediaan awal, set data ini telah dipecahkan kepada dua bahagian utama, iaitu set latihan dan set ujian. Set latihan digunakan untuk melatih model pengecaman ekspresi, manakala set ujian digunakan untuk menguji prestasi model yang dicadangkan. FER2013 adalah satu sumber yang berkualiti untuk latihan dan pengujian model FER. Dengan menggunakan set data ini, diharapkan model sistem FER yang dicadangkan dapat berfungsi dengan baik dan memberikan hasil yang tepat dan boleh dipercayai.

Pengujian model adalah langkah kritikal dalam menilai prestasi dan keberkesanan model DL yang telah dibangunkan. Set data yang dipilih bagi tujuan pengujian model FER adalah FER2013 dan KDEF (KAROLINSKA Directed Emotional Faces). KDEF adalah set data ekspresi wajah yang mempunyai variasi ekspresi dan wajah dari pelbagai subjek. Set data ini direka untuk menyediakan

imej-imej wajah dengan ekspresi muka yang berbeza seperti gembira, sedih, marah, takut, terkejut dan neutral.

Set data FER2013 telah dipisahkan menjadi dua subset, iaitu set data latihan dan set data pengujian. Set data latihan digunakan untuk melatih model, manakala set data pengujian digunakan untuk menguji prestasi model pada data yang tidak pernah dilihat oleh model semasa proses latihan. Di samping itu, set data KDEF dibiarkan sebagai satu set data pengujian yang terpisah sepenuhnya untuk menilai prestasi model pada set data yang berbeza dan tidak dikenalpasti oleh model semasa proses latihan. Dengan memisahkan set data dengan cara ini, kita dapat menyediakan penilaian yang objektif terhadap keupayaan model dalam mengenali ekspresi wajah di kedua-dua set data FER2013 dan KDEF.

Kedua-dua set data ini juga telah melalui proses pemprosesan sebelum digunakan untuk pembangunan model. Pemprosesan set data merupakan langkah penting dalam pembangunan model FER yang berkesan. Proses ini melibatkan beberapa langkah untuk menyesuaikan dan menyediakan gambar wajah sebelum digunakan untuk melatih dan menguji model FER.

Langkah pertama dalam pemprosesan adalah mengubah saiz imej wajah asal kepada saiz yang tetap, iaitu 480x480 piksel. Ini dilakukan dengan teknik downsizing menggunakan kaedah penyusutan *bilinear*. Teknik ini penting untuk memastikan setiap imej mempunyai saiz seragam yang sesuai dengan input model FER. Selanjutnya, imej diubah kepada saiz 48x48 piksel sebelum dimasukkan ke dalam model.

Langkah kedua adalah melakukan pengecaman wajah menggunakan algoritma Viola-Jones berdasarkan analisis ciri Haar. Algoritma ini membolehkan pengesanan lokasi dan saiz wajah dengan cepat dan tepat. Pengecaman ini memastikan hanya imej wajah yang digunakan untuk langkah berikutnya dalam pemprosesan.

Langkah seterusnya adalah menetapkan sudut rotasi yang diperlukan untuk menyelaraskan imej wajah. Algoritma ini mencari kedudukan mata dalam imej dan mengukur sudut rotasi yang

diperlukan untuk membentangkan imej mengikut kedudukan mata. Proses ini memastikan imej wajah menjadi seimbang dan tidak condong sebelum dihantar ke model FER.

Imej muka akan dipotong mengikut kedudukan mata dan mengekalkan kawasan yang relevan sahaja bagi pengecaman ekspresi muka. Pemotongan ini mengoptimumkan saiz imej dan memfokuskan imej pada bahagian muka yang penting untuk analisis ekspresi muka.

Sebelum akhirnya menghantar imej wajah ke dalam model FER, langkah terakhir pemrosesan adalah melaksanakan proses penapisan *bilateral* pada imej muka. Penapisan *bilateral* membantu mengurangkan kebisingan dalam imej dan menonjolkan ciri-ciri muka dengan lebih jelas. Proses ini meningkatkan kualiti imej dan memberikan hasil yang lebih tajam dan jelas untuk analisis.

Bagi model FER berasaskan DL dibangunkan pula adalah model CNN yang direka khusus untuk FER. Model ini berfungsi untuk mengklasifikasikan gambar *grayscale* berukuran 48x48 yang mewakili ekspresi wajah ke dalam tujuh kategori ekspresi: marah, jijik, takut, gembira, sedih, terkejut, atau neutral. Struktur model ini terdiri daripada empat lapisan utama.

Lapisan pertama adalah lapisan konvolusi (conv2) dengan 64 penapis berukuran 7x7 dan langkah (2, 2). Selepas konvolusi, fungsi aktivasi ReLU digunakan untuk memperkenalkan *non-lineariti* ke dalam model. Hasil dari lapisan konvolusi ini kemudian melewati lapisan *max pooling* (max_pool2) dengan saiz pengagihan 3x3 dan langkah (2, 2) untuk mengurangkan dimensi ruang. *Batch normalization* digunakan untuk normalisasi pengaktifan lapisan sebelumnya, membantu dalam menstabilkan dan mempercepatkan proses latihan.

Lapisan kedua terdiri daripada tiga sub-lapisan. Sub-lapisan pertama adalah konvolusi 1x1 (conv4) dengan 96 penapis dan aktivasi ReLU. Sub-lapisan kedua adalah lapisan *max pooling* (max_pool3) dengan saiz pengagihan 3x3 dan langkah (1, 1). Sub-lapisan ketiga adalah konvolusi 3x3 (conv5) dengan 208 penapis dan aktivasi ReLU. Satu lagi konvolusi 1x1 (conv6) dengan 64 penapis dan aktivasi ReLU dikenakan kepada output lapisan *max pooling* sebelumnya. Output dari

konvolusi 3x3 dan konvolusi 1x1 kemudian disambungkan bersama-sama mengikut paksi saluran, menghasilkan peta ciri dengan kedalaman yang lebih tinggi (*chunk_1*).

Lapisan ketiga serupa dengan lapisan kedua, tetapi dengan saiz dan jumlah penapis yang berbeza. Sub-lapisan pertama adalah konvolusi 1x1 (*conv7*) dengan 96 penapis dan aktivasi ReLU. Sub-lapisan kedua adalah lapisan *max pooling* (*max_pool4*) dengan saiz pengagihan 3x3 dan langkah (1, 1). Sub-lapisan ketiga adalah konvolusi 3x3 (*conv8*) dengan 208 penapis dan aktivasi ReLU. Konvolusi 1x1 lain (*conv9*) dengan 64 penapis dan aktivasi ReLU dikenakan kepada output lapisan *max pooling* kedua. Output dari konvolusi 3x3 dan konvolusi 1x1 kemudian disambungkan bersama-sama mengikut paksi saluran, menghasilkan peta ciri yang lain dengan kedalaman yang lebih tinggi (*chunk_2*).

Lapisan keempat dan terakhir ialah lapisan output. Peta ciri dari *chunk_2* dikembangkan menjadi vektor dan lapisan dropout dengan kadar 0.5 dikenakan untuk mencegah *overfitting*. Akhirnya, lapisan *dense* dengan 7 unit dan aktivasi *softmax* digunakan untuk menghasilkan kebarangkalian bagi setiap kelas ekspresi.

Model ini direka untuk mengekstrak dan mempelajari ciri-ciri tahap tinggi dari ekspresi wajah melalui lapisan konvolusi bertumpuk. Lapisan *max pooling* membantu mengurangkan dimensi ruang, manakala penyambungan peta ciri daripada laluan konvolusi yang berbeza menyediakan representasi yang berbeza bagi input. Lapisan *dropout* membantu dalam pelarasan untuk mencegah *overfitting* dan aktivasi *softmax* dalam lapisan output memastikan model menghasilkan pengagihan kebarangkalian untuk ketujuh kelas ekspresi. Model ini dilatih menggunakan *optimizer* Adam dengan *categorical cross-entropy loss* dan dinilai berdasarkan ketepatan semasa proses latihan.

Dalam pembentukan model DL, pemilihan hiperparameter adalah langkah kritikal untuk mencapai prestasi yang optimum. Beberapa hiperparameter yang dipilih termasuk batch size dan jenis *optimizer*. Berdasarkan perbandingan prestasi antara *optimizer* Adam, Adamax, Nadam, RMSprop dan SGD, *optimizer* 'Adam' dipilih kerana memberikan ketepatan yang optimum pada data latihan

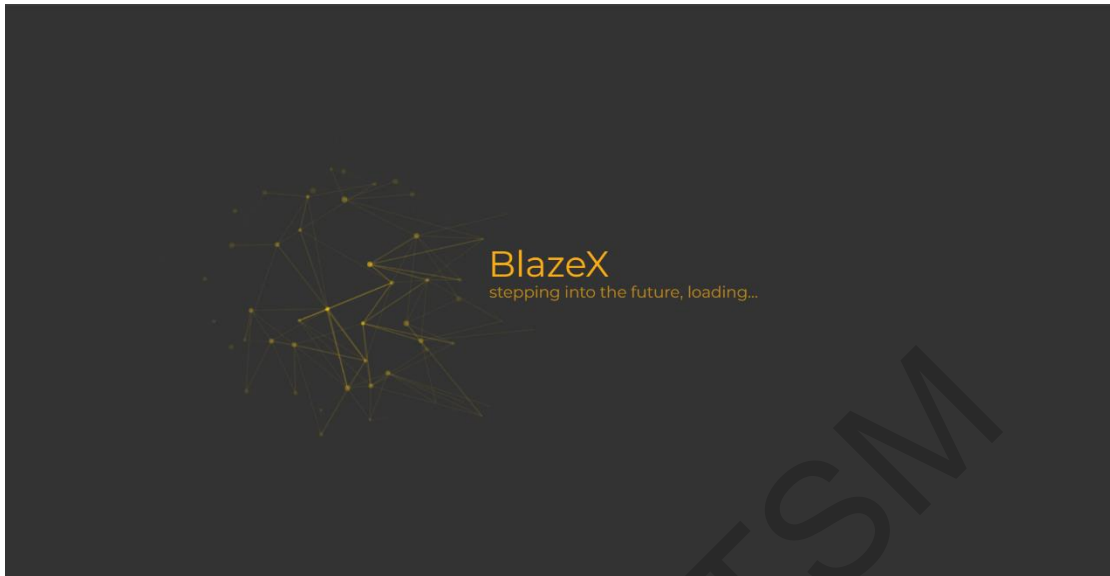
dan pengesahan serta prestasi kebolehgnerelasian yang baik. Saiz *batch* juga telah diuji pada dua kadar yang berbeza, nilai 32 dan 64. Saiz *batch* 64 dipilih oleh kerana mempunyai nilai ketepatan yang lebih optimum.

Teknik K-Fold Cross Validation juga digunakan untuk meningkatkan ketepatan generalisasi dan mencegah *overfitting* pada model. Dalam contoh ini, model dilatih menggunakan 10 *fold cross-validation*, di mana data dibahagikan kepada 10 bahagian yang sama besar. Setiap iterasi, salah satu bahagian digunakan sebagai data pengujian, manakala bahagian-bahagian lain digunakan sebagai data latihan. Hasil akhir model diperoleh melalui purata daripada keseluruhan iterasi ini.

Untuk menilai prestasi model pula, laporan klasifikasi digunakan. Laporan klasifikasi adalah ringkasan statistik yang menyediakan maklumat tentang ketepatan model dalam mengenali dan mengklasifikasikan setiap ekspresi dengan benar. Laporan ini memberikan maklumat tentang metrik seperti ketepatan, kecermatan, skor-F1 dan jumlah sampel untuk setiap kategori ekspresi.

Untuk mendapatkan pandangan lebih mendalam tentang keputusan klasifikasi model untuk setiap kategori ekspresi, satu laporan klasifikasi akan dihasilkan. Ketepatan mewakili nisbah ramalan *true positive* berbanding semua ramalan positif, manakala *recall* mengira nisbah ramalan *true positive* berbanding semua sampel positif sebenar. Skor-F1 menggabungkan ketepatan dan ketepatan pemulihan menjadi satu metrik, memberikan penilaian seimbang terhadap prestasi model untuk setiap kelas ekspresi.

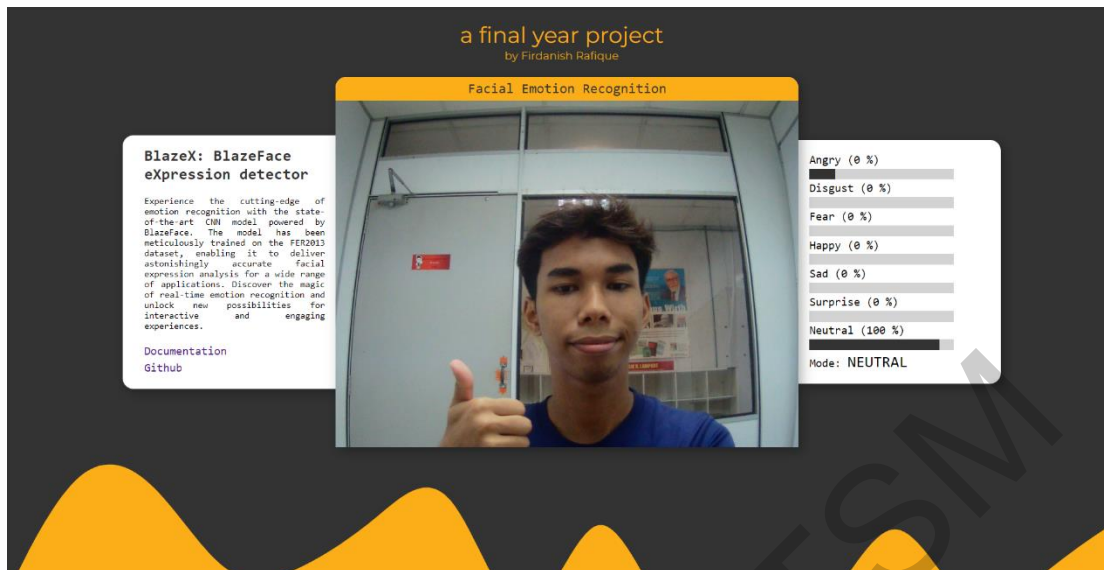
Bagi pembangunan aplikasi web berintegrasi model pembelajaran mendalam, ia dibahagikan kepada dua komponen penting iaitu pembangunan *front-end* dan *back-end*. Pembangunan *front-end* berfokus pada pengaturcaraan elemen visual laman web yang akan berinteraksi dengan pengguna. Bahagian ini bertanggungjawab untuk memaparkan antara muka yang interaktif dan menarik bagi pengguna. Gambarajah 5 hingga Gambarajah 7 menunjukkan *user interface* (UI) pengguna yang telah dibangunkan.



Gambarajah 5 UI Halaman Pramuat



Gambarajah 6 UI Halaman Utama



Gambarajah 7 UI Halaman Utama setelah diberi akses kepada kamera

Sementara itu, pembangunan *back-end* memberi tumpuan kepada bahagian laman web yang tidak dapat dilihat oleh pengguna. Di dalam projek ini, bahagian belakang web ini berfungsi sebagai tempat untuk melakukan pengiraan logik yang dijalankan oleh model DL. Bahagian ini menangani pemrosesan dan analisis data, termasuk pengecaman ekspresi muka oleh model FER.

Untuk mencapai objektif ini, projek ini menggunakan *front-end* sistem terdiri daripada modul HTML, CSS dan Javascript untuk membina antara muka pengguna yang menarik dan responsif. Manakala bahagian *back-end* terdiri daripada modul Python untuk mengendalikan logik perhitungan dan pemrosesan data oleh model DL.

Keputusan dan Perbincangan

Untuk set data FER2013, model FER memperlihatkan prestasi yang baik dengan ketepatan keseluruhan sebanyak 95.5%. Model ini berjaya mengenal pasti pelbagai ekspresi dengan baik, khususnya dalam mengklasifikasikan ekspresi “Happy” dengan ketepatan 96.4% dan skor-F1 97.3%. Hasil pengujian menunjukkan model ini juga berjaya dengan baik dalam mengenali ekspresi “Fear”, “Neutral” dan “Surprise” dengan ketepatan di atas 95.0%.

Namun, dalam pengujian menggunakan set data KDEF, prestasi model menurun dengan ketepatan keseluruhan sebanyak 47.9%. Model ini menghadapi cabaran terutamanya dalam mengklasifikasikan ekspresi “Disgust” dan “Sad”, yang memperlihatkan ketepatan yang rendah dengan masing-masing nilai 0.714 dan 0.314. Keputusan ini menunjukkan bahawa model mungkin menghadapi kesukaran dalam mengenal pasti ekspresi wajah yang jarang atau tidak seimbang dalam set data KDEF. Jadual 1 dan Jadual 2 menunjukkan laporan klasifikasi pengujian model pada set data FER2013 dan KDEF.

Jadual 1 Laporan klasifikasi pada set data FER2013 yang menunjukkan metrik *precision*, *recall*, *F1-score*, *support*, *accuracy*, *macro average* dan *weighted average* pada label ekspresi (“0”: marah, “1”: jijik, “2”: takut, “3”: gembira, “4”: sedih, “5”: terkejut, “6”: neutral)

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
0	0.949	0.951	0.950	490
1	1.000	0.945	0.972	55
2	0.937	0.928	0.932	528
3	0.964	0.981	0.973	879
4	0.946	0.944	0.945	594
5	0.983	0.954	0.968	416
6	0.942	0.963	0.953	626
Accuracy			0.955	3588
Macro Average	0.962	0.952	0.957	3588
Weighted Average	0.956	0.956	0.956	3588

Jadual 2 Laporan klasifikasi pada set data KDEF yang menunjukkan metrik *precision*, *recall*, *F1-score*, *support*, *accuracy*, *macro average* dan *weighted average* pada label ekspresi (“0”: marah, “1”: jijik, “2”: takut, “3”: gembira, “4”: sedih, “5”: terkejut, “6”: neutral)

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
0	0.365	0.503	0.423	161
1	0.714	0.031	0.060	160
2	0.361	0.366	0.364	164
3	0.694	0.950	0.802	160
4	0.314	0.510	0.388	157
5	0.776	0.519	0.622	160
6	0.521	0.475	0.497	158
<i>Accuracy</i>			0.479	1120
<i>Macro Average</i>	0.535	0.479	0.451	1120
Weighted Average	0.535	0.479	0.451	1120

Penurunan prestasi model FER semasa diuji dengan set data KDEF boleh disebabkan oleh beberapa faktor. Salah satu kemungkinan adalah perbezaan ciri-ciri antara set data FER2013 dan KDEF, seperti saiz imej, kualiti dan variasi dalam ekspresi wajah. Oleh itu, model yang dioptimumkan untuk set data FER2013 mungkin tidak sesuai sepenuhnya untuk set data KDEF kerana perbezaan ini.

Selain itu, kekurangan data dalam set data KDEF juga boleh menjadi masalah. Jika set data KDEF mempunyai taburan yang berbeza daripada set data FER2013, model mungkin menghadapi kesulitan dalam mengenali dan mengklasifikasikan ekspresi dengan betul. Jika jumlah imej ekspresi dalam set data KDEF terhad, model mungkin tidak dapat mempelajari dengan baik dan menghadapi kesukaran dalam membuat ramalan yang tepat.

Faktor lain yang mungkin menyumbang kepada penurunan prestasi adalah *overfitting*. Model FER yang telah dilatih dengan set data FER2013 mungkin menjadi terlalu terbiasa dengan data latihan dan tidak dapat berfungsi dengan baik untuk data ujian yang berbeza, seperti set data KDEF. Akibatnya, model kehilangan keupayaan pengelasan dan prestasinya menurun.

Variabiliti dalam data juga boleh menyumbang kepada masalah ini. Set data KDEF mungkin mempunyai variasi yang lebih tinggi dalam ekspresi wajah dan keadaan pencahayaan berbanding set data FER2013. Variabiliti ini boleh menyulitkan model dalam mengenali ekspresi dengan betul apabila diuji dengan set data KDEF.

Untuk meningkatkan prestasi model pada set data KDEF, beberapa langkah boleh diambil. Ini termasuk penyesuaian model untuk mengatasi perbezaan ciri-ciri set data, pengoptimuman hiperparameter untuk memperbaiki kesesuaian model, penambahbaikan data untuk meningkatkan kepelbagaian set data, atau penggunaan teknik pembelajaran pindahan (Ramalingam & Garzia 2018) untuk menggunakan pengetahuan daripada model yang sudah dilatih pada set data lain yang serupa dengan KDEF. Analisis lanjut seperti pemeriksaan ciri-ciri set data juga penting untuk mengenal pasti penyelesaian yang lebih sesuai untuk meningkatkan prestasi model pada set data KDEF. Keputusan ini menunjukkan bahawa model mungkin menghadapi kesukaran dalam mengenal pasti ekspresi wajah yang jarang atau tidak seimbang dalam set data KDEF.

Secara keseluruhan, model FER menunjukkan prestasi yang memuaskan pada set data FER2013, tetapi perlu peningkatan untuk prestasi lebih baik pada set data KDEF. Dalam usaha untuk meningkatkan prestasi model, cadangan untuk meningkatkan keberkesanan dan ketepatan termasuklah meningkatkan data latihan dengan set data yang lebih pelbagai dan melaksanakan teknik penyesuaian model yang lebih efisien. Dengan langkah-langkah penambahbaikan ini, diharapkan model ini dapat mencapai prestasi yang lebih baik dan lebih sesuai untuk aplikasi FER pada masa hadapan.

Kesimpulannya, sistem FER yang diintegrasikan dengan model DL telah menunjukkan prestasi yang baik dalam mengenali pelbagai ekspresi muka. Walaupun terdapat beberapa halangan yang dihadapi, usaha ini membuka potensi besar untuk penggunaan teknologi FER dalam pelbagai aplikasi dan perkhidmatan merentas sektor. Dengan beberapa penambahbaikan yang dicadangkan,

sistem ini berpotensi untuk menjadi lebih tepat, berkesan dan memberikan faedah yang ketara dalam bidang keselamatan, analisis ekspresi dan interaksi manusia-komputer yang lebih inovatif dan cekap.

Kesimpulan

Hasil daripada bahagian Keputusan dan Perbincangan menunjukkan bahawa sistem FER yang diintegrasikan dengan model DL mencapai prestasi yang baik dalam mengenali pelbagai ekspresi muka, terutamanya "Happy" dengan ketepatan dan skor-F1 yang tinggi. Walau bagaimanapun, apabila diuji dengan set data KDEF, prestasi model menurun dengan ketara, terutamanya dalam mengklasifikasikan ekspresi "Disgust" dan "Sad".

Beberapa faktor menyumbang kepada prestasi yang lebih rendah pada set data KDEF, termasuk perbezaan dalam ciri imej dan taburan data antara set data FER2013 dan KDEF, serta kemungkinan *overfitting*. Variabiliti dalam data dan kekurangan data dalam set data KDEF juga menyulitkan prestasi model.

Dalam bahagian Pengenalan, objektif utama projek telah dinyatakan sebagai pembangunan teknologi FER berdasarkan model DL untuk mengenali pelbagai ekspresi muka dengan tepat. Oleh itu, dapat dinyatakan bahawa objektif ini telah berhasil dicapai dengan prestasi yang baik pada set data FER2013. Model FER telah berjaya mengenali dan mengklasifikasikan ekspresi muka dengan ketepatan keseluruhan sebanyak 95.5%, dengan hasil yang terbaik untuk ekspresi "Happy" dengan ketepatan 96.4% dan skor-F1 97.3%. Walau bagaimanapun, prestasi model menurun apabila diuji dengan set data KDEF, menunjukkan bahawa terdapat cabaran dalam menghadapi ekspresi muka yang jarang dalam set data tersebut. Meskipun demikian, objektif keseluruhan projek ini dapat dianggap berjaya berdasarkan pencapaian prestasi model yang memuaskan pada set data FER2013.

Hasil kajian ini memberikan impak dan implikasi yang penting kepada bidang ilmu dan industri yang berkaitan dengan teknologi FER. Keberkesanan model FER berdasarkan DL yang telah dibangunkan dapat memberikan manfaat yang besar dalam pelbagai bidang seperti sistem

keselamatan, analisis ekspresi dan interaksi manusia-komputer. Teknologi FER yang tepat dan berkesan mempunyai potensi untuk digunakan dalam sistem keselamatan awam, seperti pemantauan dan pengawasan di tempat awam untuk mengenal pasti potensi ancaman. Selain itu, teknologi ini juga dapat digunakan dalam analisis ekspresi manusia untuk mengkaji reaksi pengguna terhadap produk atau perkhidmatan, yang dapat membantu meningkatkan strategi pemasaran dan pengalaman pengguna secara keseluruhan.

Walaupun sistem FER ini telah mencapai objektif utamanya, terdapat beberapa kekangan yang perlu diambil kira. Pertama, meskipun model DL berhasil mencapai tingkat yang baik dalam FER, terdapat situasi di mana model mungkin menghadapi kesulitan dalam mengenali ekspresi dengan tepat, terutamanya ketika wajah terlalu samar atau kurang cahaya. Kedua, aplikasi ini menggunakan model yang kompleks untuk pengesanan ekspresi wajah dan ini mungkin menyebabkan masa respons yang agak lambat, terutamanya apabila dijalankan pada peranti dengan sumber daya yang terhad. Ketiga, prestasi model sangat bergantung pada kualiti dan jumlah data latihan. Kekangan dalam memperoleh set data yang pelbagai dan mencukupi mungkin mempengaruhi keseluruhan keberkesanan model.

Untuk meningkatkan lagi sistem FER ini, beberapa cadangan penambahbaikan adalah seperti berikut. Pertama, kajian lanjut boleh dijalankan untuk menambah baik model DL dengan menggunakan set data yang lebih pelbagai dan mencukupi. Selain itu, teknik peningkatan seperti augmentasi data boleh digunakan untuk meningkatkan prestasi model. Kedua, teknik penjajaran wajah yang lebih cekap boleh diterapkan untuk memastikan wajah dapat dikesan dan dikenal pasti dengan lebih tepat, terutamanya dalam keadaan wajah yang kabur atau kurang cahaya. Ketiga, pengoptimuman model dan penggunaan teknik pemangkasan model *pruning* boleh membantu meningkatkan kelajuan aplikasi tanpa mengorbankan ketepatan pengesanan. Keempat, menggunakan teknik pembelajaran pindahan untuk meningkat prestasi model pada set data KDEF. Iaitu, menggunakan model yang sudah dilatih pada set data FER2013 sebagai model asas, kemudian

lanjutkan latihan pada set data KDEF untuk mengambil manfaat dari pengetahuan yang ada. Kelima, mengkaji dengan lebih teliti ciri-ciri set data KDEF dan bandingkan dengan FER2013. Dengan memahami perbezaan di antara kedua-dua set data ini, prestasi FER pada set data KDEF dan memperbaikinya. Dengan mengambil langkah-langkah ini dan melakukan analisis yang mendalam, ia boleh meningkatkan prestasi model pada set data KDEF dan meningkatkan pengecaman ekspresi yang lebih baik pada data ujian. Terakhir, aplikasi ini boleh ditingkatkan dengan memperkenalkan versi mudah alih, membolehkan pengguna mengakses FER dengan mudah melalui peranti mudah alih.

Secara keseluruhannya, kajian ini berjaya membina dan menggunakan system aplikasi FER berasaskan model DL yang berjaya mengecam dan mengklasifikasikan pelbagai ekspresi muka dengan berkesan. Walaupun menghadapi beberapa cabaran, hasil kajian ini menunjukkan potensi besar untuk memanfaatkan teknologi FER dalam pelbagai aplikasi dunia sebenar dan sektor industri. Dengan cadangan penambahbaikan, sistem ini berpotensi untuk meningkatkan ketepatan, keberkesanan dan kebolehgunaannya pada masa hadapan. Hasil penyelidikan ini menawarkan faedah yang ketara dalam bidang keselamatan, analisis ekspresi dan interaksi manusia-komputer, yang membawa kepada inovasi yang lebih cemerlang dan berkesan dalam dunia teknologi.

Penghargaan

Saya ingin merakamkan penghargaan atas peluang yang diberi bagi mengusahakan Projek Tahun Akhir untuk ijazah saya. Projek ini merupakan peluang yang menarik untuk saya menggunakan pengetahuan dan kemahiran yang telah saya perolehi sepanjang pengajian saya.

Saya ingin berterima kasih atas bimbingan dan sokongan yang diberikan oleh penyelia saya, Dr. Nor Samsiah binti Sani dan juga bekas penyelia saya, Dr. Kok Ven Jyn sepanjang proses pengkajian. Kepakaran dan pengalaman mereka sangat berharga dalam membantu saya memperhalusi soalan dan objektif penyelidikan saya dan untuk membangunkan cadangan yang jelas dan tersusun dengan baik.

Saya juga ingin berterima kasih atas sumber dan sokongan yang diberikan oleh institusi saya, Universiti Kebangsaan Malaysia terutamanya kepada pensyarah dan rakan-rakan saya daripada Fakulti Teknologi dan Sains Maklumat

Akhir sekali, saya sangat menghargai peluang untuk mengusahakan projek ini dan saya sentiasa menantikan cabaran dan peluang pada masa yang akan datang.

RUJUKAN

- Gan, B., Zhang, C. (2022). Target Detection and Network Optimization: Deep Learning In Face Expression Feature Recognition. *Journal of Sensors*, (2022), 1-10.
<https://doi.org/10.1155/2022/5423959>
- Wirth, R. and Hipp, J. 2000. April. CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (Vol. 1, pp. 29-39)*.
- Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.H. and Zhou, Y. 2013. Challenges in representation learning: A report on three machine learning contests. In *Neural Information Processing: 20th International Conference, ICONIP 2013, Daegu, Korea, November 3-7, 2013. Proceedings, Part III 20 (pp. 117-124)*. Springer berlin heidelberg.
- Ramalingam, S. and Garzia, F. 2018, October. Facial expression recognition using transfer learning. In *2018 International Carnahan Conference on Security Technology (ICCST) (pp. 1-5)*. IEEE.

Firdanish Rafique bin Mohd Hamid (A182072)
Ts. Dr. Nor Samsiah Sani
Fakulti Teknologi & Sains Maklumat,
Universiti Kebangsaan Malaysia.