

PEMBANGUNAN SISTEM KEHADIRAN PINTAR MENGGUNAKAN TEKNIK PENGECAMAN WAJAH BERASASKAN PEMBELAJARAN MENDALAM

Muhammad Muaz Husaini Rosli & Afzan Adam

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

Abstrak

Pengesahan kehadiran secara manual di setiap kelas memberikan beban kepada pengajar. Oleh itu, sistem pengurusan kehadiran automatik yang menggunakan pengenalan wajah telah diperkenalkan. Sistem ini mengatasi isu pelajar yang dianggap hadir tanpa fizikal berada di dalam kelas. Proses pengesahan wajah dan pengenalan pengguna adalah inti dalam sistem ini. Projek ini mengenalkan model pelaksanaan sistem pengurusan kehadiran automatik dengan menggunakan Rangkaian Konvolusi Neural (CNN) dan mencadangkan pendekatan pelaksanaannya. Sistem yang dicadangkan memanfaatkan Rangkaian Neural Siamese (SNN) untuk menganalisis ciri-ciri wajah, termasuk kawasan wajah, tekstur, warna, dan struktur wajah menggunakan FaceNet sebagai pengeluaran ciri. Tiga eksperimen telah dijalankan untuk menilai prestasi model pengenalan wajah menggunakan data dari pangkalan data UKM dan LFW. Eksperimen pertama melibatkan model dilatih dan diuji dengan data wajah UKM. Eksperimen kedua melibatkan model yang sama, tetapi dengan menggunakan data wajah LFW. Eksperimen ketiga menggabungkan kedua-dua pangkalan data, di mana model dilatih dengan data wajah LFW dan diuji dengan data wajah UKM. Prestasi eksperimen ketiga menunjukkan hasil rendah berbanding eksperimen pertama, menunjukkan keunggulan model dilatih dengan data UKM dalam mengenal wajah Asia. Kesemua data latihan merangkumi 80% daripada jumlah keseluruhan gambar dalam pangkalan data, manakala data ujian merangkumi 20% gambar selepas diaugmentasi.

Kata kunci: Pengesahan Kehadiran, Pengenalan Wajah, Sistem Pengurusan Kehadiran, Rangkaian Konvolusi Neural, Rangkaian Neural Siamese, FaceNet

Pengenalan

Isu utama kepada sistem pengurusan kehadiran yang lama merupakan penggunaannya yang agak memakan masa. Tandatanganan kehadiran memakan masa beberapa minit kerana keperluan untuk menulis nama penuh, nombor matrik dan juga tandatangan. Tidak dinafikan bahawa sekarang terdapat sistem kehadiran yang menggunakan QR Kod seperti di Universiti Kebangsaan Malaysia yang menggunakan aplikasi MyUKM: Smart Attendance sebagai sistem kehadiran mereka. Ketelusan data yang diperoleh juga antara isu utama (Mohamad 2019). Hal ini kerana data kehadiran seseorang pelajar itu boleh di buat secara tidak bertanggungjawab oleh pelajar lain, menyebabkan pelajar yang sebenarnya tidak hadir. Ini merupakan kekurangan pada Smart Attendance yang menggunakan QR Kod. Ini terjadi apabila terdapat beberapa pelajar yang tidak hadir sanggup memberikan kata laluannya kepada rakannya yang hadir bagi tujuan rekod kehadiran. Perkara ini berlaku di universiti lain yang menggunakan sistem pengurusan kehadiran dengan QR Kod (Masalha dan Hirzallah 2014).

Selain itu, pangkalan data wajah ialah elemen utama yang diperlukan apabila bekerja dengan mana-mana teknologi pengecaman wajah. Tanpanya, algoritma tidak dapat memastikan ketepatan apabila membandingkan atau mengecam imej yang tidak diketahui dengan imej yang telah dikenal pasti. Data asas ini penting untuk digunakan supaya algoritma dapat dilatih dan diuji. Terdapat banyak pangkalan data berbeza yang telah tersedia untuk digunakan secara percuma dalam talian. Namun, kekurangan utama penggunaan pangkalan data yang sedia ada ialah tiada kawalan ke atas spesifikasi data yang akan membawa kepada isu semasa ujian (Calistra 2016). Penggunaan pangkalan data wajah yang besar memainkan peranan penting dalam mencapai ketepatan pengecaman wajah yang tinggi (Schroff et al. 2015; Zhou et al. 2015). Namun, dalam membangunkan sistem pengecaman wajah kepada peranti mudah alih saiz pangkalan data yang terlalu besar akan menggunakan masa yang lama dalam proses pengecaman wajah dan menggunakan saiz memori yang besar kepada peranti mudah alih. Pangkalan data wajah yang sedia ada juga, sebahagian besar tidak merangkumi spesifikasi demografi penduduk di Malaysia yang berbilang bangsa.

Projek ini bertujuan untuk mencapai dua objektif utama. Pertama, objektifnya adalah untuk menganalisis dengan mendalam setiap ciri atau fitur yang terdapat pada wajah manusia. Analisis ini akan membantu kami memahami secara menyeluruh tentang bagaimana ciri-ciri wajah dapat diidentifikasi dan diukur secara objektif. Kedua, projek ini bertujuan untuk membangunkan model pengecaman wajah yang spesifik kepada konteks demografi Malaysia. Model ini akan dihasilkan dengan mengambil kira variasi wajah dalam kalangan populasi Malaysia, termasuk perbezaan dalam fitur wajah seperti warna kulit dan elemen demografi lain.

Projek ini juga memiliki skop yang terperinci untuk mencapai objektifnya. Pertama, dalam membina pangkalan data wajah, fokus akan diberikan kepada pelajar-pelajar dari Fakulti Teknologi dan Sains Maklumat (FTSM) yang mewakili tiga bangsa utama di Malaysia iaitu Melayu, Cina, dan India. Kedua, pengumpulan data wajah akan difokuskan terhadap kenalan peribadi yang dekat, termasuk rakan sekolah dan rakan dari latar belakang akademik yang serupa. Namun, terdapat beberapa kekangan yang mungkin mempengaruhi proses pembangunan. Pengumpulan data dalam skala besar mungkin terhad oleh faktor-faktor tertentu. Selain itu, kepakaran untuk melaksanakan projek dalam bidang ini masih dalam proses pemahaman, dan aktiviti pencarian sumber serta pembelajaran yang mendalam perlu dijalankan. Oleh itu, perancangan masa harus dilaksanakan dengan bijak untuk memastikan model pengecaman wajah yang sesuai dengan demografi Malaysia dapat dibangunkan dengan cekap dan berkualiti.

Selain itu, kepentingan yang besar dalam menangani masalah yang timbul dari sistem kehadiran yang sedia ada. Dengan penerapan sistem ini, isu berkaitan ketepatan data kehadiran dan potensi penipuan semasa pengumpulan data dapat diatasi. Sistem ini akan memberikan sokongan yang lebih baik kepada pensyarah dalam membuat keputusan mengenai pelajar yang kerap tidak hadir atau tidak hadir tanpa alasan yang jelas. Tambahan pula, pelajar juga akan mendapat faedah dengan kemampuan untuk menilai prestasi kehadiran mereka berdasarkan teknologi pengenalan wajah yang dibangunkan. Ini boleh meningkatkan motivasi mereka untuk menyimpan rekod kehadiran yang baik

dan mengurangkan risiko manipulasi ketidakhadiran. Sistem ini juga mempunyai potensi aplikasi yang luas dalam bidang keselamatan dan penyiasatan forensik, di mana ia boleh digunakan untuk mencocok dan membandingkan imej wajah individu dalam pelbagai konteks. Hasil projek ini bukan hanya memberikan penyelesaian praktikal dalam konteks pendidikan, tetapi juga memberi sumbangan penting dalam bidang keselamatan dan lain-lain.

Disamping itu, projek ini menerapkan pendekatan gabungan antara CRISP DM (Cross-Industry Standard Process for Data Mining) dan Agile dalam pembangunan sistem. Bahagian permodelan dalam kerangka kerja CRISP-DM akan diselarikan dengan kitaran pembangunan agile untuk mencapai hasil yang optimum. Pendekatan ini membolehkan proses merancang, menganalisis, dan memproses data menggunakan langkah-langkah yang terstruktur dalam CRISP-DM. Pada masa yang sama, prinsip-prinsip pembangunan agile digunakan untuk menyesuaikan dan meningkatkan model dan sistem secara berterusan berdasarkan keperluan semasa projek.

Bagi menjayakan projek ini, sistem pengesanan wajah yang merangkumi dua proses utama iaitu proses pengesanan lokasi wajah dalam imej atau video masa nyata dan aspek pengesanan identiti seseorang individu akan melibatkan penggunaan algoritma pengaturcaraan. Dengan penggunaan bahasa pengaturcaraan seperti Python, ia dapat membantu dalam melaksanakan dua proses ini kerana mempunyai jumlah pustaka yang meluas dan banyak yang boleh diimport dan diguna pakai bagi pembangunan projek ini. Selain daripada itu, Python juga dilengkapi dengan kod pratulis yang akan menjimatkan masa dalam penjanaan kod.

Dalam jadual 1.1 menunjukkan kajian-kajian tahun terkini mengenai pembangunan sistem kehadiran berdasarkan pengesanan wajah. Dapat diperhatikan bahawa terdapat tiga teknik pengesanan lokasi wajah yang kerap digunakan dalam kajian-kajian tahun terkini iaitu Viola-Jones, Haar Cascade dan juga Histogram of Oriented Gradients (HOG). Manakala algoritma bagi teknik pengesanan identiti yang digunakan untuk proses klasifikasi dapat diperhatikan masih lagi pelbagai dengan penggunaan CNN, SVM, LBPH dan sebagainya.

Jadual 1.1 Perbandingan algoritma pembelajaran mendalam bagi kajian-kajian tahun terkini (2018 – 2022)

Pengkaji dan Tahun	Sampel Data	Fitur Wajah	Teknik Pengesanan Identiti	Ketepatan	Kekangan
Bhattacharya et al.2018	Sampel Data Sendiri	Viola-Jones	CNN	Tidak jelas (disertakan dalam bentuk graf bar)	Tidak tepat sepenuhnya kerana masih gagal mengesan identiti pada situasi tertentu
Tej Chinimilli et al. 2020	60 imej untuk 18 individu berbeza	Haar Cascade	LBPH	77%	Sampel data yang kecil
Seal et al. 2020	Imej 8 individu berbeza	Viola-Jones dan HOG	Eigen Faces dan LBPH	94.81%	Tidak disertakan
Sanli dan Ilgen 2018	100 sampel data daripada 10 individu	HOG	LBPH	75%	Tidak disertakan
Pei et al. 2019	3563 sampel data daripada pengkalan data wajah VGG	Tidak disertakan	VGG-16	86.3%	Sampel data yang kecil
Indra et al. 2020	100 sampel data daripada 20 individu	Tidak disertakan	Bilateral Filter	81.25%	Tidak disertakan
Nurkhamid et al. 2021	150 sampel daripada 16 individu	HOG	CNN dan SVM	99.74%	Pengujian pada situasi dan keadaan yang berbeza tidak dilakukan
Shah et al. 2021	1200 sampel data daripada 6 individu	Haar Cascade	Haar Cascade	93.1%	Tidak disertakan

Laporan teknikal ini terdiri dari bahagian kajian kesusasteraan, metodologi kajian, keputusan dan perbincangan, kesimpulan, penghargaan, serta rujukan. Kajian kesusasteraan melibatkan analisis kajian terdahulu tentang pengecaman muka. Metodologi kajian menjelaskan kaedah yang digunakan.

Keputusan dan perbincangan merujuk kepada hasil kajian. Kesimpulan merangkumi keseluruhan kajian. Penghargaan diberikan kepada sokongan yang diterima, dan rujukan merujuk kepada sumber yang digunakan.

Metodologi Kajian

Untuk memastikan kelancaran pelaksanaan projek, struktur pembangunan akan menjadi panduan projek sepanjang keseluruhan proses pelaksanaan. Struktur pembangunan ini adalah berdasarkan kitaran pembangunan CRISP-DM dan Agile seperti yang dinyatakan dalam bahagian Pengenalan. Proses pembangunan projek ini dimulakan dengan fasa spesifikasi keperluan, di mana analisis keperluan dalam domain permasalahan dilakukan. Fasa seterusnya adalah fasa reka bentuk, yang melibatkan pembangunan model pengecaman wajah dan pangkalan data wajah yang akan diimplementasikan dalam projek ini. Fasa pembangunan kemudian berfokus pada pembangunan prototaip antara wajah Sistem Kehadiran Pintar. Terakhir, terdapat fasa pengujian yang bertujuan untuk menganalisis fitur-fitur wajah antara pangkalan data wajah yang telah dibina dan yang telah ada, serta melakukan analisis prestasi model pengecaman wajah dalam keupayaannya untuk mengecam wajah pelajar dengan tepat.

Fasa Pengumpulan Data merupakan fasa di mana data imej wajah yang diperlukan diperolehi daripada subjek-subjek yang telah dipilih. Subjek terdiri daripada beberapa orang pelajar sukarela yang berbangsa antara Melayu, Cina atau India. Imej wajah diambil untuk setiap pelajar yang berbeza ekspresi iaitu biasa, tersenyum atau marah serta pandangan sisi dan jarak yang berbeza. Bagi langkah yang pertama, imej wajah akan diperolehi daripada imej sedia ada daripada laman sosial subjek. Namun jika jumlah imej wajah di laman sosial subjek tidak mencukupi atau tidak menepati ciri-ciri yang telah ditetapkan, imej wajah subjek akan ditangkap bagi memenuhi keperluan. Kemudian, imej wajah setiap subjek yang ditangkap disimpan di dalam fail dengan nama pelajar, dan semua imej adalah berformat JPEG sebagai pangkalan data UKM.

Seterusnya, untuk pangkalan data imej wajah yang kedua, Labeled Faces in the Wild (LFW) terkenal dengan koleksi imej wajah yang luas dan menampilkan pelbagai individu dari pelbagai sumber. Namun, kerana pangkalan data LFW terdiri daripada jumlah individu yang besar, ia mungkin tidak sepadan dengan ciri-ciri pangkalan data UKM yang terdiri daripada 30 pelajar dari FTSM. Untuk mengatasi perbezaan ini, satu proses pemilihan rawak telah dilaksanakan untuk memilih 30 individu daripada pangkalan data LFW. Proses pemilihan ini bertujuan untuk membentuk suatu subset yang mewakili ciri-ciri yang serupa dengan pangkalan data dalaman. Dengan memilih individu secara rawak, usaha dilakukan untuk memastikan jumlah yang adil dari jantina dan faktor lain yang terdapat dalam pangkalan data UKM.

Pemilihan individu secara rawak daripada pangkalan data LFW membantu menghubungkan perbezaan antara kedua-dua pangkalan data ini bagi membolehkan perbandingan yang lebih bermakna dan adil untuk model yang dilatih dengan pangkalan data LFW dan model yang dilatih dengan pangkalan data UKM. Pendekatan ini membolehkan proses menganalisis dan menilai prestasi model pengesanan wajah di kedua-dua pangkalan data dengan menilai sejauh mana model pengesanan wajah dapat belajar dengan baik di antara pangkalan data yang berbeza dengan ciri-ciri yang berbeza juga. Pemilihan dan persediaan yang teliti terhadap pangkalan data ini memainkan peranan penting dalam memastikan perbandingan yang bermakna untuk model pengesanan wajah.

Antara pembersihan data yang dilakukan adalah pengesanan wajah diaplikasikan untuk mencari dan mengekstrak koordinat wajah dalam imej. Kemudian, algoritma pengesanan titik rujukan wajah digunakan untuk mengenal pasti titik-titik utama pada wajah, seperti mata, hidung, dan mulut. Titik-titik rujukan ini kemudiannya dinormalisasi dengan mengira transformasi yang memetakan mereka ke set tetap titik rujukan, memastikan kedudukan dan skala yang konsisten. Bahagian wajah yang selaras kemudiannya diperolehi dari imej yang telah ditransformasikan berdasarkan koordinat kotak pembatas yang dinormalisasi. Tambahan lain, imej wajah yang diperolehi boleh diubah saiz kepada saiz tetap. Dengan mengaplikasikan teknik pra pemprosesan ini, imej-imej wajah disediakan

dalam cara yang mengurangkan variasi dalam kedudukan, rotasi, dan skala, dan seterusnya meningkatkan ketepatan algoritma pengenalan wajah.

Manakala bagi augmentasi data yang pertama adalah kecerahan rawak dan pelarasan kontras. Dengan mengubah secara rawak tahap kecerahan dan kontras imej, model menjadi lebih mahir dalam mengendalikan variasi dalam keadaan pencahayaan. Selain itu, penukaran mendatar digunakan untuk mencipta variasi sudut pandangan. Dengan mengimejkan imej wajah secara mendatar, model mendapat pendedahan kepada orientasi yang berbeza dan akan meningkatkan kebolehan model untuk mengecam dan mengelaskan wajah tanpa mengira arah. Antara lain, teknik augmentasi yang digunakan ialah manipulasi kualiti JPEG dan pelarasan tepu untuk mengubah suai keamatan warna dalam imej. Dengan mengubah tahap ketepuan secara rawak, model lebih mampu mengendalikan variasi dalam penampilan warna, menjadikannya lebih mudah disesuaikan dengan tona warna dan keadaan pencahayaan yang pelbagai.

Fasa analisis data terbahagi kepada dua peringkat pembelajaran mesin iaitu pengesanan wajah dan pengesanan identiti. Bagi peringkat pertama iaitu pengesanan wajah, model Rangkaian Neural Konvolusi berbilang tugas (MTCNN) akan di terapkan. Model MTCNN ialah pendiskrip ciri yang sering digunakan untuk mengekstrak ciri daripada data imej. Ia digunakan secara meluas dalam tugas penglihatan komputer untuk pengesanan objek. Dalam model MTCNN, bingkai imej dipaparkan dengan menggabungkan ciri dan kemudian ia dihantar ke rangkaian Konvolusi Rangkaian Neural (CNN) untuk meramalkan lokasi wajah di sekeliling kotak sempadan yang perlu dipaparkan.

Selepas itu, algoritma yang telah dipilih iaitu Face Net akan diuji prestasinya. Face Net telah dipilih sebagai salah satu algoritma bagi pengekstrakan ciri kerana prestasinya yang tinggi dan masa pengekstrakan yang singkat daripada model CNN yang lain. Dalam model FaceNet, fungsi Triplet Loss dipilih untuk digunakan. Triplet Loss dilakukan daripada padanan antara imej wajah dan contoh positif serta negatif. Umumnya, pengekstrakan ciri menggunakan model FaceNet, seperti yang dilihat dalam Rajah 3.4, ia melibatkan proses input imej ke dalam seni bina pembelajaran mendalam, yang

kemudiannya dinormalisasikan kepada L2, menghasilkan ciri wajah yang dilatih menggunakan fungsi Triplet Loss.

Dalam projek ini, fungsi Triplet Loss telah digunakan untuk meningkatkan prestasi model pengecaman wajah. Fungsi Triplet Loss membolehkan model mempelajari daripada lapisan pembedaan yang berkesan dengan menekankan hubungan jarak antara sampel utama, positif dan negatif. Dengan meminimumkan kehilangan tiga kali ganda semasa latihan, model akan dapat menangkap persamaan dan ketidaksamaan antara wajah. Pelaksanaan fungsi Triplet Loss ini meningkatkan ketepatan dan keteguhan sistem pengecaman wajah, menghasilkan keupayaan pengecaman yang lebih baik dan pengenalan individu yang lebih dipercayai.

Tiga eksperimen dijalankan untuk menilai prestasi model pengecaman wajah yang telah ditingkatkan menggunakan fungsi Triplet Loss. Eksperimen pertama melibatkan pengujian model yang telah dilatih dan diuji menggunakan pangkalan data wajah UKM. Eksperimen kedua, melibatkan pengujian model yang sama, tetapi kali ini model ini dilatih dan diuji menggunakan pangkalan data wajah LFW. Manakala eksperimen ketiga adalah gabungan antara kedua-dua pangkalan data. Model telah dilatih menggunakan data wajah dari pangkalan data LFW dan kemudian diuji menggunakan data wajah dari pangkalan data UKM. Pendekatan ini membolehkan penilaian sejauh mana model yang telah dilatih menggunakan pangkalan data asing mampu mengenal pasti individu dari pangkalan data lain.

Pada peringkat seterusnya, dalam setiap eksperimen prestasi model pengecaman wajah, data latihan dibagi menggunakan teknik pembahagian data dengan nisbah 80:20. Teknik pembahagian data ini adalah cara umum dalam membahagi data, di mana 80% daripada keseluruhan data wajah digunakan sebagai data latihan untuk melatih model. Tujuan dari penggunaan sebahagian besar data untuk latihan adalah untuk membolehkan model memahami ciri-ciri wajah dengan lebih baik agar dapat meningkatkan kemampuan model. Sementara itu, 20% digunakan sebagai data ujian untuk menguji prestasi model. Data ujian ini terdiri dari wajah-wajah yang tidak pernah dilihat oleh model

sebelumnya, sehingga menguji sejauh mana model dapat mengenali wajah-wajah yang tidak dikenali sebelumnya dan menilai keupayaannya dalam mengatasi situasi yang belum pernah dihadapi sebelumnya.

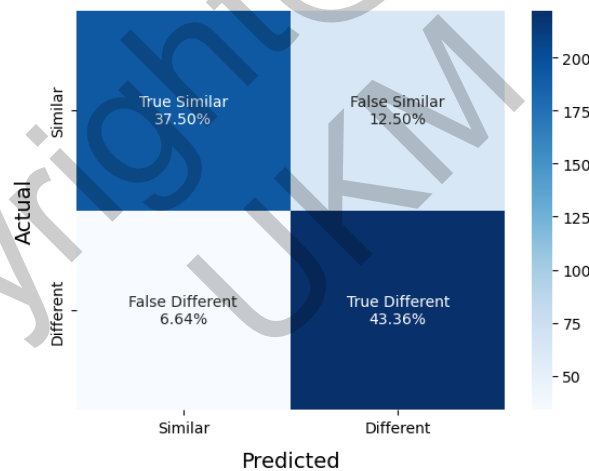
Seterusnya, kaedah penilaian sistem pengecaman yang pertama ialah matriks kekeliruan iaitu matriks yang meringkaskan prestasi model pengelasan. Ia menunjukkan bilangan positif benar (TP), negatif benar (TN), positif palsu (FP) dan negatif palsu (FN) untuk setiap kelas. Daripada matriks kekeliruan, beberapa metrik prestasi boleh dikira. Matriks kekeliruan berguna dalam menilai prestasi model pada setiap kelas dan mengesan set data tidak seimbang. Daripada matriks kekeliruan 'Accuracy' dan 'Precision' juga boleh didapatkan. Ia biasanya digunakan dalam kajian penyelidikan, seperti karya Bao et al. (2021), yang menggunakannya untuk menilai prestasi model pembelajaran mesin untuk mendiagnosis kanser payudara daripada imej mamografi.

Selain itu, ketepatan merupakan salah satu faktor penilaian yang akan mengukur kadar positif benar antara jumlah ramalan positif. Ketepatan berguna apabila nilai positif palsu adalah tinggi. Skor ketepatan yang tinggi menunjukkan bahawa model membuat sedikit ramalan positif palsu. Ketepatan digunakan secara meluas dalam kajian penyelidikan, seperti karya Wang et al. (2021), yang menggunakannya untuk menilai prestasi model pembelajaran mesin untuk mendiagnosis COVID-19 daripada imej CT dada. Selain itu adalah 'recall' yang mengukur perkadaran positif benar antara jumlah bilangan sampel positif sebenar. Ingat kembali berguna apabila kos negatif palsu tinggi. Skor ingatan yang tinggi menunjukkan bahawa model membuat sedikit ramalan negatif palsu. Recall digunakan secara meluas dalam kajian penyelidikan, seperti karya Chen et al. (2021), yang menggunakannya untuk menilai prestasi model pembelajaran mesin untuk meramalkan retinopati diabetik daripada imej retina.

Keputusan dan Perbincangan

Eksperimen pertama merangkumi pengujian model yang dilatih dan diuji dengan pangkalan data wajah UKM. Model yang dilatih menunjukkan ketepatan yang baik dalam mengklasifikasikan imej-imej wajah. Berdasarkan matriks kekeliruan daripada Rajah 1.1, terdapat informasi sebagai berikut:

- *True Similar* (TS) sebesar 37.50%. Ini menunjukkan bahwa model dengan tepat mengecam imej-imej wajah yang sebenarnya positif.
- *False Similar* (FS) sebesar 12.50%. Ini mengindikasikan bahwa model cenderung memberikan prediksi positif yang salah pada imej-imej wajah yang sebenarnya negatif.
- *False Different* (FD) sebesar 6.64%. Terdapat imej-imej wajah yang sebenarnya positif tetapi tidak dicam sebagai positif oleh model pengecaman wajah.
- *True Different* (TD) sebesar 43.36%. Model dengan tepat mengklasifikasikan imej-imej wajah yang sebenarnya negatif sebagai negatif.



Rajah 1.1 Matriks Kekeliruan bagi latihan model pengecaman wajah dengan pangkalan data UKM

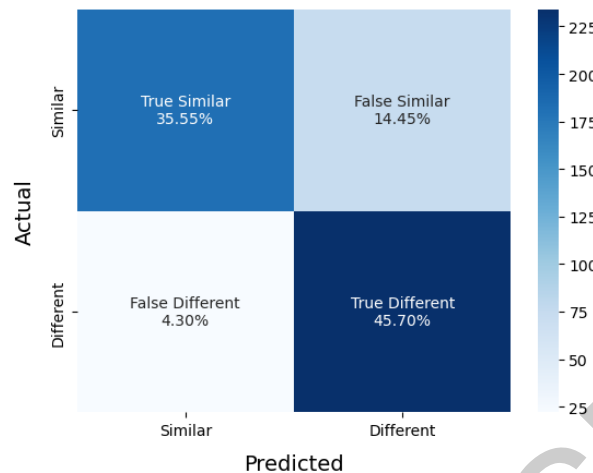
Kesimpulannya, model yang dilatih memiliki ketepatan yang agak tinggi dalam mengklasifikasikan imej-imej wajah sebagai positif atau negatif. Namun, terdapat kesalahan dalam prediksi, seperti *False Similar* dan *False Different*, yang perlu diperbaiki atau ditingkatkan. Selain itu, matrik-matrik prestasi yang lain seperti ketepatan, kejituan dan ingatan semula diperolehi bagi mendapatkan imejan yang lebih lengkap tentang prestasi model. Jadual 1.2 menunjukkan penilaian prestasi model dengan data latihan daripada pangkalan data UKM yang dibangunkan.

Jadual 1.1 Penilaian prestasi latihan model pengecaman wajah dengan pangkalan data UKM

NILAI LATIHAN	
Ketepatan	0.8086
Kejituan	0.7519
Ingatan Semula	0.9741

Eksperimen kedua merangkumi pengujian model yang dilatih dan diuji dengan pangkalan data wajah LFW. Prestasi model, seperti yang dipaparkan dalam Rajah 1.2 daripada matriks kekeliruan, menunjukkan nilai-nilai berikut:

- *True Similar (TS)* sebesar 37.50%: Model dengan tepat mengidentifikasi 37.50% imej positif secara benar. Ini adalah imej wajah yang sebenarnya positif dan diklasifikasikan dengan benar sebagai positif oleh model.
- *False Similar (FS)* sebesar 12.50%: Model salah mengklasifikasikan 12.50% imej negatif sebagai positif. Ini adalah imej wajah yang sebenarnya negatif tetapi secara keliru diidentifikasi sebagai positif oleh model.
- *False Different (FD)* sebesar 6.64%: Model salah dalam mengklasifikasikan 6.64% imej positif sebagai negatif. Ini adalah imej wajah yang sebenarnya positif tetapi secara keliru diidentifikasi sebagai negatif oleh model.
- *True Different (TD)* sebesar 43.36%: Model dengan tepat mengidentifikasi 43.36% imej negatif secara benar. Ini adalah imej wajah yang sebenarnya negatif dan diklasifikasikan dengan benar sebagai negatif oleh model.



Rajah 1.2 Matriks Kekeliruan bagi latihan model pengecaman wajah dengan pangkalan data LFW

Dari nilai-nilai tersebut, dapat disimpulkan bahwa prestasi model adalah baik tetapi masih ada ruang untuk perbaikan. Ketepatan model dapat dihitung dengan menjumlahkan jumlah insiden yang diklasifikasikan dengan benar (TS dan TD) dan membaginya dengan keseluruhan jumlah insiden. Kejituan dapat dihitung dengan membagi TS dengan jumlah TS dan FS, sedangkan ingatan semula dapat dihitung dengan membagi TS dengan jumlah TS dan FD.

Jadual 1.3 Penilaian prestasi latihan model pengecaman wajah dengan pangkalan data LFW

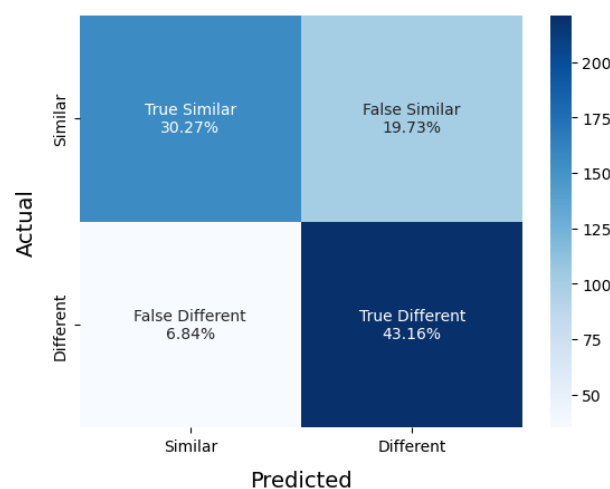
NILAI LATIHAN	
Ketepatan	0.8125
Kejituan	0.8222
Ingatan Semula	1.0000

Nilai ketepatan sebesar 81.25% yang menunjukkan bahwa model memiliki tingkat ketepatan yang baik dalam ramalan kelas secara keseluruhan. Selain itu, nilai kejituan sebesar 82.22% menunjukkan bahawa sebahagian besar jangkaan positif yang dilakukan oleh model adalah benar. Selanjutnya, nilai ingatan semula sebesar 100.00% menunjukkan bahawa model mampu mengidentifikasi semua contoh positif dengan benar. Dalam kesimpulan ini, dapat disimpulkan bahawa model berhasil memberikan hasil yang baik dengan tahap ketepatan yang tinggi, serta mampu

mengidentifikasi semua contoh positif. Hal ini menunjukkan bahwa model ini efektif dalam mengenali wajah yang benar

Eksperimen ketiga merangkumi pengujian model yang dilatih dengan pangkalan data wajah LFW dan diuji dengan pangkalan data wajah UKM. Rajah 1.3 menunjukkan matriks kekeliruan prestasi model sebagai berikut:

- *True Similar* (TS) sebesar 30.27%: Model dengan tepat mengklasifikasikan 30.27% imej positif secara benar. Ini adalah imej wajah yang sebenarnya positif dan berhasil diidentifikasi dengan benar oleh model.
- *False Similar* (FS) sebesar 19.73%: Model secara salah mengklasifikasikan 19.73% imej negatif sebagai positif. Ini adalah imej wajah yang sebenarnya negatif tetapi salah diklasifikasikan sebagai positif oleh model.
- *False Different* (FD) sebesar 6.84%: Model secara salah mengklasifikasikan 6.84% imej positif sebagai negatif. Ini adalah imej wajah yang sebenarnya positif tetapi salah diklasifikasikan sebagai negatif oleh model.
- *True Different* (TD) sebesar 43.16%: Model dengan tepat mengklasifikasikan 43.16% imej negatif secara benar. Ini adalah imej wajah yang sebenarnya negatif dan berhasil diidentifikasi dengan benar sebagai negatif oleh model.



Rajah 1.3 Matriks Kekeliruan pengujian dengan data latihan UKM

Dari matriks kekeliruan tersebut, dapat disimpulkan bahawa model memiliki ketepatan baik dalam mengklasifikasikan insiden dengan ketepatan sebesar 73.43%. Namun, terdapat penurunan ketepatan yang signifikan jika dibandingkan dengan model yang dilatih menggunakan pangkalan data UKM seperti yang ditunjukkan didalam Jadual 1.4.

Jadual 1.2 Penilaian prestasi latihan model pengecaman wajah dengan data latihan UKM

NILAI LATIHAN	
Ketepatan	0.7344
Kejituan	0.6863
Ingatan Semula	0.8633

Model pra-latih, terutamanya yang dilatih dengan pangkalan data yang besar dan pelbagai, dapat menangkap pola dan fitur umum yang berkaitan di pelbagai domain. Walau bagaimanapun, ia mungkin mempunyai kelemahan apabila melibatkan ciri-ciri tertentu yang berbeza secara ketara antara set data pra-latih dan set data yang diguna pakai untuk fasa pengujian. Dalam kes pengecaman wajah, kepelbagaian fitur wajah boleh menjadi salah satu aspek yang mempengaruhi prestasi model. Jika model pra-latih tidak diperkenalkan kepada pelbagai fitur wajah semasa latihan, ia mungkin tidak mempelajari ciri-ciri dan variasi yang halus khusus kepada sesuatu demografi bagi kumpulan yang khusus. Akibatnya, prestasi model dalam mengenal pasti wajah pelajar tidak setepat berbanding wajah semasa latihan model. Untuk menangani isu ini, penggunaan set data yang mempunyai fitur-fitur wajah yang mewakili populasi atau demografi subjek sasaran adalah relevan.

Pembangunan model pengecaman wajah ini merupakan satu pencapaian yang penting kerana ia membentuk asas kepada sistem yang kukuh dan boleh dipercayai. Dengan memanfaatkan pangkalan data UKM, ia terbukti berkesan dalam mengenal pasti wajah pelajar seperti yang dibuktikan di bahagian atas, projek ini memastikan bahawa sistem boleh dipercayai dan diguna pakai dalam mengecam wajah pelajar dengan tepat.

Selain itu, penambahbaikan adalah perkara penting yang perlu diambil kira untuk memastikan sistem yang dibangunkan dan mampu untuk mengecam wajah dengan sempurna. Antara penambahbaikan yang boleh dilakukan ke atas Sistem Kehadiran Pintar adalah untuk membangunkan aplikasi yang sebenar, penting untuk melakukan penelitian mendalam tentang keperluan pengguna serta melakukan pengujian aplikasi secara menyeluruh. Penelitian keperluan pengguna akan memberikan pemahaman yang lebih baik tentang keperluan dan harapan pengguna, sehingga aplikasi dapat direka dengan tepat untuk memenuhi keperluan mereka. Selain itu, pengujian aplikasi yang melibatkan ujian fungsi, prestasi, kualiti, dan keselamatan akan memastikan aplikasi berfungsi dengan baik dan dapat digunakan.

Kesimpulan

Model yang diujikan menunjukkan ketepatan yang baik dalam mengklasifikasikan insiden dengan nilai ketepatan sebesar 73.43%. Namun, terdapat penurunan yang signifikan dibandingkan dengan model yang dilatih dengan data latihan UKM (seperti yang terlihat dalam Jadual 1.4). Model pra-latih memiliki kemampuan untuk menangkap pola dan fitur umum di berbagai domain, tetapi mungkin memiliki kelemahan ketika berhadapan dengan ciri-ciri yang berbeda antara data pra-latih dan pengujian. Khususnya dalam kasus pengecaman wajah, variasi fitur wajah menjadi faktor kunci yang mempengaruhi kinerja model. Jika model tidak diperkenalkan dengan berbagai fitur wajah selama pelatihan, ia mungkin tidak dapat memahami variasi halus yang khusus bagi kelompok tertentu dalam populasi. Oleh karena itu, untuk mengatasi masalah ini, dianjurkan untuk menggunakan dataset yang mencakup variasi fitur wajah yang relevan dengan populasi subjek yang dituju.

Sistem Kehadiran Pintar yang dibangunkan juga berjaya mencapai objektif pembangunan sistem dan berjaya menyelesaikan masalah yang dinyatakan dalam bahagian pengenalan dokumen ini. Walaupun sistem ini masih berada dalam peringkat prototaip, ia telah berjaya mengaplikasikan sebuah model pengecaman wajah yang menggunakan pangkalan data wajah yang baru yang

dinamakan UKM. Model ini telah menunjukkan tahap ketepatan dan prestasi yang tinggi dalam mengenali wajah pelajar.

Walaupun Sistem Kehadiran Pintar ini menawarkan kelebihan yang signifikan, terdapat beberapa kekurangan yang perlu diberi perhatian. Pertama, perlu diingatkan bahawa sistem ini masih berada dalam fasa prototaip. Oleh itu, terdapat kemungkinan adanya kekurangan dalam fungsi dan prestasi sistem. Pengguna perlu memahami bahawa dalam peringkat ini, sistem mungkin belum sempurna dan masih memerlukan penyempurnaan sebelum diimplementasikan secara meluas. Selain itu, kekurangan lain yang perlu diambil perhatian adalah kesalahan dalam mengecam wajah pelajar oleh model pengecaman wajah. Walaupun model ini telah mencapai tahap ketepatan yang tinggi, masih terdapat kemungkinan adanya kesalahan dalam mengecam wajah pelajar. Hal ini mungkin disebabkan oleh variasi penampilan wajah, pencahayaan yang tidak ideal, atau faktor-faktor lain yang mempengaruhi kualiti imej sewaktu perolehan data input oleh Sistem Kehadiran Pintar. Penting untuk diakui bahawa Sistem Kehadiran Pintar yang menggunakan pengecaman wajah masih tidak sempurna dan memerlukan penambahbaikan untuk mengatasi kekurangan-keurangan tersebut.

Kesimpulannya, projek ini berjaya mencapai dua objektif utama sebagaimana yang dinyatakan dalam bahagian Pengenalan dokumen ini. Analisis fitur-fitur muka telah berhasil dilakukan untuk memperkuat keperluan adanya pangkalan data wajah yang baru. Di samping itu, model pengecaman muka yang telah dibangunkan juga telah berjaya mengecam muka sesuai untuk kegunaan demografi Malaysia seperti yang dinyatakan dalam bahagian Keputusan dan Perbincangan dokumen ini.

Penghargaan

Dengan segala kerendahan hati, terlebih dahulu saya ingin bersyukur kepada Ilahi atas limpah kurnia dan rahmat-Nya, kerana berhasil menyiapkan Projek Tahun Akhir ini dengan penuh kejayaan dalam masa yang telah ditetapkan. Saya ingin mengucapkan setinggi-tinggi penghargaan

kepada Ts. Dr. Afzan Adam, yang bertanggungjawab sebagai penyelia Projek Tahun Akhir saya. Terima kasih atas segala bimbingan, tunjuk ajar, dan cadangan yang diberikan sepanjang pelaksanaan projek ini.

Saya juga ingin menghargai jasa Ts. Dr. Nor Samsiah Sani, selaku Ketua Program Kecerdasan Buatan, kerana sentiasa memberi peringatan tentang tarikh penghantaran projek dan format penulisan yang betul. Penghargaan setinggi-tinggi juga kepada semua pensyarah yang telah berusaha memberikan kursus dan bimbingan sepanjang pelaksanaan kajian ini.

Tidak lupa juga ucapan rasa terima kasih kepada ibu bapa saya, Encik Rosli dan Puan Normah, atas sokongan yang diberikan dari segi semangat, moral, dan kewangan. Terakhir, jutaan terima kasih saya tujukan kepada rakan-rakan seperjuangan saya yang dengan sukacitanya menjadi subjek kajian ini, serta menyediakan data-data berharga untuk projek ini.

RUJUKAN

- Bhattacharya, S., Nainala, G. S., Das, P., & Routray, A. (2018). Smart Attendance Monitoring System (SAMS): A Face Recognition Based Attendance System for Classroom Environment. *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*. <https://doi.org/10.1109/icalt.2018.00090>
- Calistra, C. (2016, July 12). Pros and Cons of Private and Public Face Recognition Databases [Online forum post]. LinkedIn. <https://www.linkedin.com/pulse/pros-cons-private-public-face-recognition-databases-cole-calistra>.
- Indra, E., Yasir, M., Andrian, A., Sitanggang, D., Sihombing, O., Tamba, S. P., & Sagala, E. (2020). Design and Implementation of Student Attendance System Based on Face Recognition by Haar-Like Features Methods. *2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT)*. <https://doi.org/10.1109/mecnit48290.2020.9166595>

- Masalha, F., & Hirzallah, N. (2014). A Students Attendance System Using QR Code. *International Journal of Advanced Computer Science and Applications*, 5(3).
<https://doi.org/10.14569/ijacsa.2014.050310>.
- Mohamed, B. K. P., & Raghu, C. V. (2012). Fingerprint attendance system for classroom needs. *2012 Annual IEEE India Conference (INDICON)*.
<https://doi.org/10.1109/indcon.2012.6420657>.
- Nurkhamid, Setialana, P., Jati, H., Wardani, R., Indrihapsari, Y., & Norwawi, N. M. (2021). Intelligent Attendance System with Face Recognition using the Deep Convolutional Neural Network Method. *Journal of Physics: Conference Series*, 1737(1), 012031.
<https://doi.org/10.1088/1742-6596/1737/1/012031>.
- Pei, Z., Xu, H., Zhang, Y., Guo, M., & Yang, Y. H. (2019). Face Recognition via Deep Learning Using Data Augmentation Based on Orthogonal Experiments. *Electronics*, 8(10), 1088. <https://doi.org/10.3390/electronics8101088>
- Sanli, O., & Ilgen, B. (2018). Face Detection and Recognition for Automatic Attendance System. *Advances in Intelligent Systems and Computing*, 237–245.
https://doi.org/10.1007/978-3-030-01054-6_17.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr.2015.7298682>.
- Seal, S., Sen, A., Mukerjee, R., & Das, A. K. (2020). An approach towards development of automated attendance system using face detection and recognition. *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. <https://doi.org/10.1109/iemcon51383.2020.9284817>

Shah, K., Bhandare, D., & Bhirud, S. (2020). Face Recognition-Based Automated Attendance System. *Advances in Intelligent Systems and Computing*, 945–952.

https://doi.org/10.1007/978-981-15-5113-0_79

Tej Chinimilli, B., T., A., Kotturi, A., Reddy Kaipu, V., & Varma Mandapati, J. (2020). Face Recognition based Attendance System using Haar Cascade and Local Binary Pattern

Histogram Algorithm. *2020 4th International Conference on Trends in Electronics and*

Informatics (ICOEI)(48184). <https://doi.org/10.1109/icoei48184.2020.9143046>

Muhammad Muaz Husaini bin Rosli (A182954)

Ts. Dr. Afzan Adam

Fakulti Teknologi & Sains Maklumat,

Universiti Kebangsaan Malaysia

Copyright@FTSM
UKM