# gGHOSTFACE: MODEL PENGECAMAN EKSPRESI MUKA RINGAN DAN TEGUH

TANG JIA HUI

KOK VEN JYN

*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia*



**ABSTRAK**

Ekspresi muka merupakan suatu tindak balas fenomenologi yang ditunjukkan oleh manusia dalam corak kompleks yang signifikan dalam komunikasi bukan lisan. Walaupun wujudnya banyak rangkaian *state-of-the-art* pengecaman ekspresi muka (FER), penyelidikan FER masih merupakan kajian yang berterusan dalam bidang visi komputer kerana fokus utama kajian-kajian FER kini adalah untuk mengatasi pengklasifikasian ekspresi muka yang serupa dengan senario kehidupan sebenar di mana wujudnya kepelbagaian yang dinamik dalam keadaan cahaya, kehadiran oklusi, operasi pasca pemprosesan dan lain-lain dalam pada imej yang merosot ketepatan pengklasifikasian. Walaupun penambahan komponen auksiliari berjaya mengatasi masalah ini, ia gagal mengambil kira aspek kos komputasi yang dikenakan oleh rangkaian serta peningkatan kebergantungannya terhadap GPU yang canggih pada fasa pembelajaran dan peramalan. Kegagalan untuk memenuhi keperluan perkakasan oleh model FER menyebabkan peningkatan masa pembelajaran model dan ketidakcekapan rangkaian untuk mengklasifikasikan ekspresi muka pada masa nyata. Dengan ini, model FER lambat mengklasifikasikan ekspresi muka untuk menjadi sepadan dengan perubahan ekspresi muka manusia yang cepat dan halus dalam kehidupan sebenar. Selain itu, tumpuan untuk meningkatkan ketepatan pengklasifikasian rangkaian state-of-the-art FER adalah ketepatan daripada hasil pengklasifikasian rangkaian terhadap set data imej ekspresi muka yang serupa dengan senario kehidupan sebenar tanpa menyedari masalah kewujudan contoh musuh dan kesalahan pelabelan dalam set data FER. Ia diperkenalkan untuk mengelirukan pembelajaran rangkaian tentang ciri-ciri muka yang ditunjukkan dengan betul berdasarkan ekspresi muka yang spesifik, menyebabkan penghafazan label yang rosak oleh model FER yang akhirnya merosotkan ketepatan pengklasifikasiannya. Seterusnya, kajian FER yang melibatkan data sensitif, iaitu imej ekspresi muka mempunyai kerentaan yang tinggi terhadap serangan musuh seperti serangan inferens keahlian (MIA) yang menyerang model FER disebabkan oleh kebolehan rangkaian untuk menghafaz imej ekspresi muka. MIA lalu akan mengenal pasti jikalau suatu contoh imej ekspresi muka merupakan ahli kepada set data pembelajarannya. Pengabaian isu ini meningkatkan risiko akses tanpa bertauliah terhadap imej-imej tersebut bagi perlaksanaan aktiviti tidak beretika melalui MIA. Projek ini mencadangkan model FER – GhostFace yang hanya menggunakan depthwise convolution dan triplet attention supaya ia berupaya untuk mengklasifikasi ekspresi muka manusia kehidupan sebenar dengan kos komputasi yang rendah. Demi mengatasi kekurangan keteguhan terhadap contoh musuh, pelabelan salah dan MIA, teknik pembesaran data nyah-hafazan (de-memorization) yang cekap telah digunakan, iaitu RandAug dan MixUp supaya model FER tidak menghafaz sifat ciri-ciri muka yang salah terhadap suatu ekspresi muka tertentu dan menghafaz imej ekspresi muka dalam set data pembelajarannya secara kritikal demi mengurangkan kejayaan MIA. Daripada eksperimen yang dijalankan, GhostFace berjaya menggunakan FLOPs yang lebih sedikit dalam pengklasifikasian ekspresi muka set data RAF-DB dimana imej-imejnya menyerupai ekspresi muka dalam kehidupan seharian manusia. Dengan teknik pembesaran data nyah-hafazan, keteguhan dan ketepatan pengklasifikasian GhostFace terhadap imej ekspresi muka kehidupan seharian yang rumit telah meningkat walaupun imej tersebut disalah labelkan. Penyiasatan projek ini mendapati hubungan berkadar langsung antara penghafazan ekspresi muka oleh model FER dengan kejayaan MIA.

## 1 PENGENALAN

Emosi merupakan keadaan fisiologi dan suatu tindak balas fenomenologi yang dimanifestasikan oleh manusia dalam pelbagai corak yang rumit dan ketara seperti perubahan

*Rajah 1.1: Tujuh ekspresi muka universal yang ditunjukkan oleh manusia, iaitu (a) Kegembiraan, (b) Kesedihan, (c) Kemarahan, (d) Ketakutan, (e) Kejutan, (f) Kejijikan, dan (g) Neutral (Ekman, P. 2009.)*

ekspresi muka yang jelas dan lambat ataupun yang sebaliknya (Kret. 2015). Perubahan ini adalah bersifat naluri dan bergantung kepada pengalaman dan peristiwa yang dialami oleh individu tersebut. Terdapat tujuh emosi ekspresi muka universal yang ditunjukkan oleh manusia iaitu kegembiraan, kesedihan, kemarahan, ketakutan, kejijikan, kejutan dan neutral (Rajah 1.1). Ia adalah penting bukan sahaja sebagai mod komunikasi tambahan bagi manusia yang berkesan, tetapi kajian yang berkaitan juga menyumbang kepada bidang psikologi (Sheaffer, B.L et al. 2009), automobil (Katsis, C.D. et al. 2015) dan bioperubatan (Haber, N. et al. 2020) yang berkongsi matlamat yang sama untuk meningkatkan taraf kehidupan manusia.

Walaupun manusia dikurniakan dengan keupayaan untuk mengenal pasti emosi melalui ekspresi muka dengan serta-merta, keperluan untuk mengautomasikan proses pengecaman ekspresi muka (FER) meningkat demi mengurangkan kebergantungan kepada ahli pakar untuk menganalisis emosi ekspresi muka yang ditunjukkan. Pendekatan pertama pengautomasian FER bermula dengan Sistem Pengekodan Tindakan Muka (FACS) (Martinez, B. et al. 2016) (Pantic, M. et al. 2005) tetapi beralih kepada kaedah pembelajaran mesin tradisional (ML) kerana FACS adalah tenaga buruh dan masa intensif, dimana ia bergantung kepada subjektiviti penilai (annotators) yang berpengalaman dalam penandaan emosi ekspresi muka secara manual. Kaedah ML dalam FER terdiri daripada tiga proses utama, iaitu pengesanan muka, pengekstrakkan ciri muka seperti dengan Local Binary Pattern (LBP), Gabor Filter dan diakhiri dengan pengklasifikasian ekspresi muka dengan model pembelajaran mesin seperti Support Vector Machine (SVM), K-Nearest Neighbors (KNN) dan lain-lain (Dino, H.I. & Abdulrazzaq, M.B. 2019) (Noor, et al. 2020). Walaupun pendekatan ML berjaya mengklasifikasikan ekspresi muka dengan baik dalam kos komputasi yang rendah, tetapi ia gagal mengklasifikasikan imej ekspresi muka yang lebih mencabar, iaitu imej ekspresi muka yang menyerupai kehidupan seharian manusia dimana muka subjek adalah tertutup ataupun menjadi tidak jelas disebabkan oleh kewujudan oklusi, keadaan pencahayaan dan posisi kepala, tetapi ia gagal mengklasifikasikan imej ekspresi muka yang lebih mencabar, iaitu imej ekspresi muka yang menyerupai kehidupan seharian manusia dimana muka subjek adalah tertutup ataupun menjadi tidak jelas disebabkan oleh kewujudan oklusi, keadaan pencahayaan dan posisi kepala yang berbeza serta operasi pasca pemprosesan. Seterusnya, pendekatan ML melibatkan proses

percubaan-dan-ralat untuk menghasilkan standard ataupun urutan proses bagi teknik pengesanan muka dan penggabungan teknik pengekstrakan ciri untuk menghasilkan ketepatan pengklasifikasian ekspresi muka yang tinggi. Dengan ini, sifat teknik ML yang tenaga buruh dan masa intensif serta ketidakupayaannya untuk mengklasifikasikan ekspresi muka yang lebih mencabar, teknik pembelajaran mendalam (DL) yang menggunakan rangkaian saraf berasaskan fungsi sistem saraf dan otak manusia dalam pemprosesan dan analisis maklumat adalah diperkenalkan untuk mengatasi kekangan ini.

DL dalam kajian FER semakin meluas kerana tugasan pengesanan muka, pengekstrakan ciri-ciri muka serta pengklasifikasian adalah digabungkan sebagai satu penyelesaian yang komprehensif dilaksanakan oleh model DL dengan campur tangan manusia yang minimum. Seterusnya, penambahan pelbagai komponen bantuan pada tulang belakang model FER meningkatkan keupayaan model untuk mengklasifikasikan imej ekspresi muka yang menyerupai senario kehidupan sebenar manusia. Akan tetapi, model FER berasaskan DL dan komponen tambahan yang diperkenalkan telah dilaksanakan tanpa pertimbangan keringanan model dimana kos komputasi tinggi telah diperkenalkan, peningkatkan dalam kerumitan seni bina model FER serta kebergantungannya untuk menggunakan GPU yang lebih canggih semasa pembelajaran dan pengklasifikasian ekspresi muka. Selain itu, model berasaskan DL mempunyai kemampuan yang tinggi untuk mempelajari corak set data melalui berjuta-juta parameter yang akhirnya meningkatkan risikonya terhadap penghafazan sifat ciriciri muka yang tidak munasabah bagi sesuatu kelas ekspresi muka akibat daripada pelabelan salah dan contoh musuh yang diperkenalkan untuk mengelirukan pembelajaran model. Ini memberi pengertian palsu terhadap generalisasi dan kesahihan hasil peramalan model FER terhadap imej ekspresi muka di luar set data pembelajarannya, akibat daripada kekurangan keteguhannya terhadap masalah pelabelan salah dan contoh musuh (Rabin, M.R.I, et al, 2021). Walaupun terdapat kajian yang sedia ada untuk mengatasi kekurangan keringanan dan keteguhan terhadap masalah pelabelan salah dan contoh musuh seperti (Laha Ale et al, 2019) (Ferro-P'erez, R. et al, 2020) (Zhao, ZengQun. et alt. 2021) (Li, Yong. et al. 2019) (Hui, Ding. et al. 2020), komponen tambahan telah diperkenalkan secara berlebihan pada tulang belakang model FER dan meningkatkan kerumitan seni binanya serta kos komputasinya masih boleh ditambahbaikkan lagi demi mempercepatkan proses pengklasifikasian ekspresi muka pada masa nyata. Selain itu, kajian FER yang sedia ada melibatkan penggunaan mercu tanda muka pada imej ekspresi muka dan teknik pra-latihan model dengan set data imej ekspresi muka yang besar dan mahal untuk diperolehi sebagai pendekatan untuk mencapai ketepatan pengklasifikasian imej ekspresi muka yang baik.

Selain itu, imej ekspresi muka sebagai data biometrik sensitif juga merupakan suatu kebimbangan dalam FER (Vemou. & Horvath, 2021) disebabkan oleh risiko penyalahgunaan dan akses tanpa kebenaran terhadap imej muka yang tinggi (Martinez-Martin, 2019) dan penekanan terhadap isu perlindungan data privasi muka yang tidak komprehensif, kabur (Nor, Tasrib, Francis, Hesham, & Othman, 2021) dan perlindungan privasi muka yang masih bergantung kepada cara privasinya dicabulkan. Serangan inferens keahlian (MIA) merupakan serangan terhadap model ML dan DL untuk meramalkan dan membandingkan persamaan antara imej ekspresi muka seorang individu yang dikehendaki (image of interest) dan koleksi imej ekspresi muka dalam pangkalan data imej ekspresi muka, bersama dengan data tambahan (quasi-identifiers) seperti umur dan jantina. Sekiranya persamaan antara kedua-dua imej adalah tinggi, maka imej individu yang dikehendaki merupakan ahli kepada set data pembelajaran yang digunakan oleh model FER. Dengan ini, imej ekspresi muka dan quasi-identifiers individu tersebut adalah dikenal pastikan dan sekiranya maklumat ini didedahkan kepada pihak ketiga tanpa kebenaran, hal ini menyebabkan kebocoran data peribadi individu yang terlibat serta meningkatkan risiko data mereka disalahgunakan. Kejayaan MIA didorong oleh keupayaan model FER berasaskan rangkaian saraf untuk menghafaz setiap titik data ataupun label dalam set data pembelajaran (Carlini, N., 2021). Walaupun terdapat kajian yang sedia ada mencadangkan teknik yang berlainan seperti (Mahawaga Arachchige, P. et al. 2020) dan (Singh, Fan, & Kankanhalli, 2021) untuk mengurangkan masalah pelanggaran privasi, akan tetapi perdagangan antara kualiti penggunaan model dan privasi yang dicapai adalah besar, menyebabkan model yang dibangunkan dengan pemeliharaan privasi yang amat baik berkemungkinan mempunyai penggunaan yang lemah (Leino, K. et al, 2020) ataupun kaedah yang digunakan masih mempunyai kerentanan yang tinggi terhadap MIA, iaitu serangan yang biasa dikenakan terhadap model ML dan DL.

## 2 PENYATAAN MASALAH

Kepentingan FER dalam kehidupan manusia telah memotivasikan automasi teknologi FER supaya emosi manusia yang ditunjukkan melalui ekspresi muka dapat dikenal pastikan dengan cepat dan tepat. Pendekatan awal pengautomasian FACS dan ML telah digantikan dengan pendekatan DL supaya pengesanan muka, pengekstrakan ciri-ciri muka dan pengklasifikasian imej ekspresi muka yang mencabar tetapi menyerupai senario kehidupan sebenar manusia di mana wujudnya oklusi, keadaan pencahayaan, posisi kepala yang berlainan dan lain-lain boleh dilaksanakan melalui satu model tanpa campur tangan manusia secara berlebihan. Ini direalisasikan oleh pendekatan DL melalui penambahan komponen tambahan ataupun

rangkaian kepada model tulang belakang FER tetapi tanpa pertimbangan terhadap kos komputasi yang diperkenalkan, kecanggihan GPU yang diperlukan demi mencapai proses pembelajaran yang cepat dan merealisasi pengklasifikasian pada masa nyata. Hal ini menyebabkan pengklasifikasian ekspresi muka pada masa nyata susah direalisasikan dan hasil peramalan ekspresi muka adalah tidak sepadan dengan perubahan ekspresi muka manusia yang cepat dan halus dalam kehidupan sebenar. Selain itu, pembelajaran rangkaian saraf melibatkan jumlah parameter yang tinggi dan kemampuan penghafazan datanya adalah baik, menyebabkan kerentanannya untuk menghafaz sifat ciri-ciri muka yang tidak sepadan dengan emosi ekspresi muka yang ditunjukkan oleh subjek imej adalah tinggi. Ini adalah akibat daripada pelabelan salah imej ekspresi muka dan contoh musuh yang diperkenalkan dengan sengaja dalam set data pembelajaran FER untuk mengelirukan pembelajaran model FER terhadap ciri-ciri muka yang munasabah bagi suatu kelas ekspresi muka. Dengan ini, ketepatan pengklasifikasian model FER terhadap ekspresi muka di luar set data pembelajaran akan merosotkan. Seterusnya, penghafazan imej ekspresi muka oleh model FER demi mencapai ketepatan pengklasifikasian yang tinggi telah meningkatkan kerentanannya terhadap MIA yang membina semula imej ekspresi muka yang digunakan dalam pembelajaran model. Ini meningkatkan risiko pelanggaran privasi imej ekspresi muka yang digunakan dalam pembelajaran model.

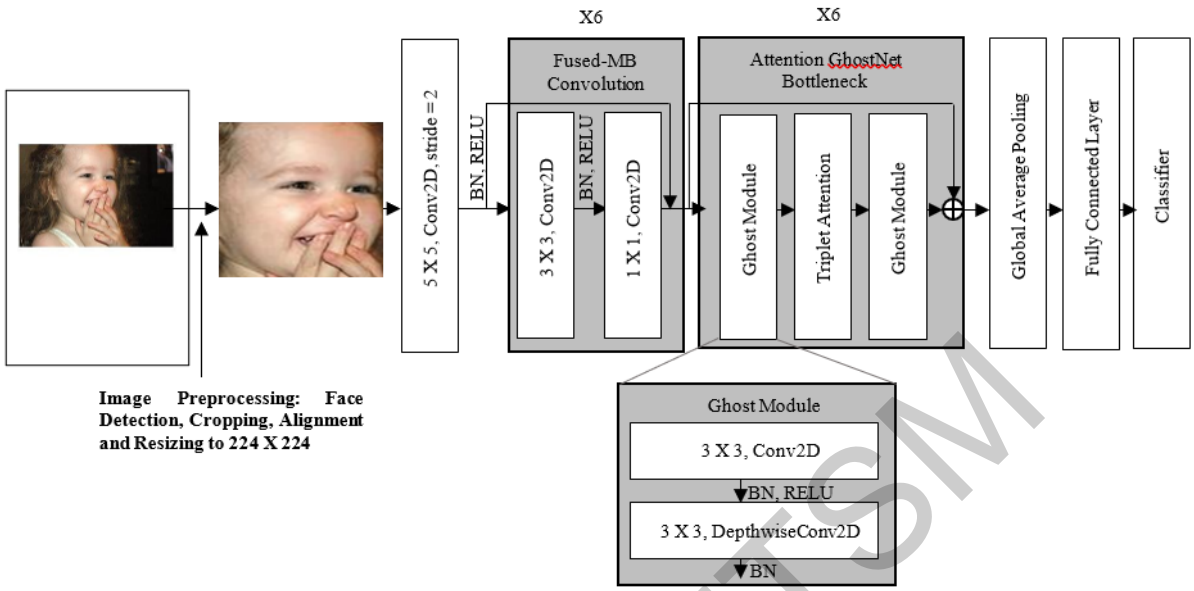
## 3 METHOD KAJIAN

Objektif projek ini adalah untuk membangunkan model pengecaman ekspresi muka yang ringan dan teguh dalam mengklasifikasikan ekspresi muka kepada salah satu daripada tujuh ekspresi universal iaitu gembira, sedih, marah, jijik, takut, terkejut dan neutral. Sub-objektif dan sumbangan projek ini adalah:-

a. Mencadangkan model FER ringan yang menggunakan mekanisme perhatian merentas dimensi untuk mencapai ketepatan yang baik terhadap pengklasifikasian ekspresi muka yang membayangkan senario kehidupan sebenar.
b. Meningkatkan keteguhan model FER yang dicadangankan terhadap contoh musuh dan masalah pelabelan salah dengan menggunakan teknik pembesaran data nyah-hafazan RandAug dan MixUp.
c. Menyiasat hubungan antara kekuatan penghafazan model FER dan keupayaannya untuk memelihara privasi imej ekspresi muka yang digunakan dalam pembelajaran model daripada MIA.

# 4 METHOD KAJIAN

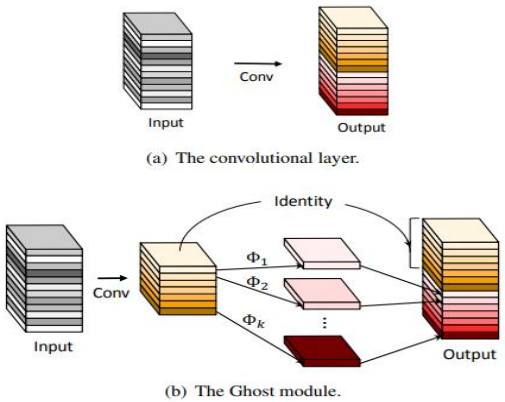
Penunjukkan dan perubahan ekspresi muka manusia dalam kehidupan sebenar adalah cepat, halus dan wujudnya cabaran seperti oklusi, pencahayaan dan posisi kepala yang berlainan serta operasi pasca pemprosesan yang menutupkan ataupun merosotkan kejelasan bahagian muka yang penting dalam pengklasifikasian ekspresi muka yang baik. Walaupun penambahan komponen bantuan mampu mengatasi cabaran-cabaran ini, ia meningkatkan kos komputasi dan permintaan model FER terhadap GPU canggih dalam pembelajaran dan pengklasifikasian ekspresi muka. Sekiranya permintaan ini tidak dipenuhi, pengklasifikasikan ekspresi muka menjadi lambat dan tidak sepadan dengan perubahan ekspresi muka manusia yang cepat dan halus seperti dalam kehidupan sebenar dan pengklasifikasian ekspresi muka tidak boleh diimplementasikan pada masa nyata. Selain itu, kemampuan model FER berasaskan rangkaian saraf adalah tinggi dalam penghafazkan sifat ciri-ciri muka yang penting untuk mengklasifikasikan ekspresi muka dengan baik tetapi kewujudan contoh musuh dan pelabelan salah menyebabkannya untuk menghafaz ciri-ciri muka yang salah bagi suatu kelas emosi ekspresi muka serta menyebabkan MIA berjaya mengenal pasti imej ekspresi muka yang digunakan dalam pembelajaran model. Ini masing-masing menyebabkan prestasi model dalam mengklasifikasikan ekspresi muka di luar set data pembelajarannya untuk merosot dan meningkatkan risiko pelanggaran privasi dan penyalahgunaan imej ekspresi muka yang digunakan dalam pembelajaran model. Oleh itu, model FER baru GhostFace (Rajah 4.1) telah dicadangkan dan ia terdiri daripada dua blok utama yang menyumbang kepada keringanan model, iaitu Fused-MB Convolution (Tan, Mingxing.et al. 2021) dan Attention GhostNet Bottleneck (Misra, D., et al, 2021) (Han, Kai. et al. 2020). Fused-MB Convolution membantu GhostFace untuk memanfaatkan penggunaan pemecut mudah alih moden manakala Attention GhostNet Bottleneck menggunakan kos komputasi depthwise convolution yang lebih rendah, konsep kewujudan kelewahan ciri-ciri muka yang diekstrakkan dalam peta ciri untuk menjana lebih banyak peta ciri (ghost features) lagi dan konsep perhatian interaksi silang dimensi untuk meningkatkan pembelajaran dan pemfokusan GhostFace terhadap ciri-ciri muka yang intrinsik dalam pengklasifikasian ekspresi muka yang baik di bawah kos komputasi yang lebih rendah. Bagi mengatasi masalah contoh musuh dan pelabelan salah dalam set data FER, teknik pembesaran data nyah-hafazan MixUp (Zhang, Hongyi. et al. 2018) dan RandAug (Cubuk, E.D. et al. 2020) telah digunakan untuk meningkatkan kepelbagaian yang rawak dalam suatu imej ekspresi muka supaya model FER tidak menghafaz kepada sifat ciri-ciri muka spesifik yang salah bagi suatu kelas ekspresi muka dan akhirnya meningkatkan ketepatan model dalam




*Rajah 4.1: Seni bina model pengecaman ekspresi muka yang akan mengklasifikasikan imej ekspresi muka kepada satu daripada tujuh ekspresi muka universal, iaitu kegembiraan, kesedihan, kemarahan, neutral, kejijikan, kejutan dan ketakutan.*

mengklasifikasikan ekspresi muka di luar set data pembelajarannya. Disebabkan kejayaan MIA adalah berdasarkan keupayaan model FER untuk menghafaz imej ekspresi muka dalam fasa pembelajarannya (Carlini, N., 2021), penyiasatan terhadap hubungan antara kekuatan penghafazan imej ekspresi muka dan tahap pemeliharaan privasi imej ekspresi muka yang digunakan dalam pembelajaran model daripada MIA melibatkan keadaan jikalau teknik pembesaran data nyah-hafazan telah digunakan dalam pembelajaran model. Selepas itu, MIA adalah diperkenalkan dengan sengaja terhadap model FER selepas fasa pembelajarannya terhadap ekspresi muka demi mempelajari kejayaannya untuk menentukan imej ekspresi muka yang digunakan dalam pembelajaran model.

Proses pengklasifikasian ekspresi muka GhostFace bermula dengan prapemprosesan imej ekspresi muka yang menyerupai senario kehidupan sebenar manusia, iaitu pengesanan dan pemangkasan muka oleh RetinaFace untuk memastikan muka yang menghadap ataupun tidak menghadap ke hadapan (non-frontal) berjaya dikesan dan dikeratkan. Pemangkasan muka membolehkan GhostFace untuk berfokus lagi kepada bahagian muka yang signifikan dalam pengklasifikasian imej ekspresi muka yang ditunjukkan. Seterusnya, imej tersebut adalah disaiz semulakan kepada dimensi 224x224x3 untuk mengatasi variasi dalam saiz imej ekspresi muka akibat daripada proses pengesanan dan pemangkasan muka. Imej prapemprosesan ini berakhir dengan imej normalisasi di mana berlakunya transformasi terhadap taburan nilai piksel supaya ia mengikut Gaussian piawai, di mana nilai-nilai piksel mempunyai min 0 dan




*Rajah 4.2: Perbandingan operasi penjanaan peta ciri oleh lapisan konvolusi (a) biasa bersama dengan modul Ghost (b).*
Sumber Rujukan: (Han, Kai. et al. 2020)

sisihan piawai 1. Seterusnya, imej ekspresi muka yang diproses akan digunakan sebagai input kepada GhostFace yang terdiri daripada blok *Fused-MB Convolution* dan *Attention GhostNet Bottleneck* untuk mengekstrak dan berfokus kepada ciri-ciri muka yang signifikan dalam pengklasifikasian ekspresi muka di bawah kos komputasi yang rendah. Akhirnya, peta ciri adalah diratakan dengan lapisan Global Average Pooling sebelum diproseskan oleh lapisan *Fully-Connected* demi mengklasifikasikan imej ekspresi muka kepada salah sata daripada tujuh emosi universal. Untuk meningkatkan keteguhan GhostFace terhadap masalah pelabelan salah, contoh musuh dan MIA, teknik pembesaran data nyah-hafazan adalah disertakan dalam proses imej prapemprosesan.

## 4.1 BLOK FUSED-MV CONVOLUTION

Blok Fused-MB Convolution (*Rajah 4.2*) yang terdiri daripada lapisan konvolusi 3X3, lapisan Batch Normalization (BN) dan pengaktifan ReLu serta diakhiri dengan lapisan konvolusi 1X1 diperkenalkan pada awal GhostFace demi mengatasi masalah pembelajaran model yang lambat akibat daripada penggunaan lapisan *depthwise convolution* pada awal model klasifikasi ini (Tan, Mingxing.et al. 2021). Walaupun depthwise convolution mampu mengurangkan kos komputasi dari segi FLOPS dan jumlah parameter GhostFace, pengurangan tersebut tidak membayangkan kepada kecekapan inferens oleh pemecut mudah alih moden dan pengurangan terhadap masa pembelajaran model untuk mengklasifikasikan imej ekspresi muka. Dengan ini, proses konvolusi *depthwise convolution* KxK akan digantikan dengan konvolusi biasa KxK, di mana K adalah saiz kernel bagi lapisan konvolusi tanpa memperkenalkan kos komputasi dari segi FLOPs dan jumlah parameter secara berlebihan. Konvolusi 3X3 akan mengekstrakkan corak ciri-ciri muka daripada imej ekspresi muka ataupun peta ciri dan konvolusi 1X1 (*pointwise convolution*) pada akhir blok ini akan menggabungkan ciri-ciri yang diekstrakkan

daripada lapisan konvolusi 3x3 dan sebagai pengurangan dimensi sebelum bermulanya proses konvolusi 3x3 yang baru. Lapisan BN diperkenalkan demi mengurangkan peralihan kovariat dalaman yang berlaku disebabkan oleh pengkemaskinian parameter rangkaian yang mengubahkan pengedaran aktivasi rangkaian semasa pembelajaran model FER justeru mengurangkan masa pembelajaran model.

### 4.2 ATTENTION GHOSTNET BOTTLENECK

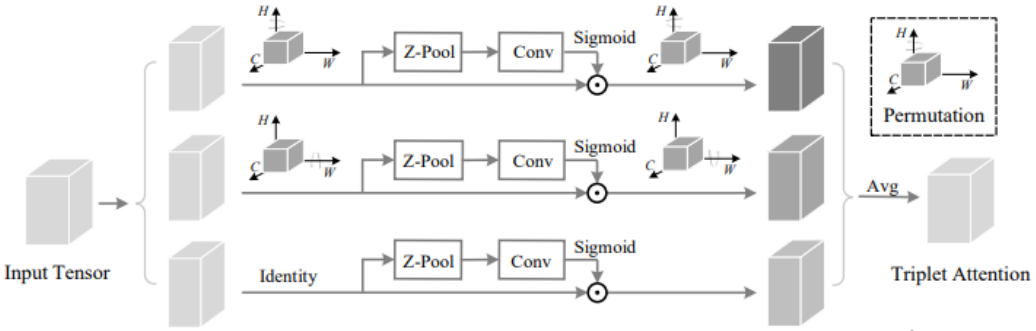
#### MODUL GHOST

Modul Ghost (Rajah 4.2) diperkenalkan untuk memanfaatkan kewujudan kelewahan ciri-ciri muka yang diekstrakkan dalam peta ciri supaya model cadangan FER yang dibangunkan mempunyai kefahaman yang komprehensif berkaitan dengan input wajah, tanpa memperkenalkan kos komputasi yang tinggi seperti konvolusi konvensional dan pada masa yang sama, mengekalkan prestasi pengklasifikasian ekspresi muka GhostFace pada tahap yang baik. Modul Ghost adalah terdiri daripada lapisan konvolusi biasa yang bertanggungjawab untuk menjanakan peta ciri intrinsik, iaitu hasil pengekstrakan ciri-ciri muka dari imej ekspresi muka, diikuti dengan lapisan depthwise convolution yang memainkan peranan sebagai operasi linear untuk menghasilkan ghost features daripada m peta ciri intrinsik yang dijanakan semasa operasi konvolusi biasa. Operasi penjanaan peta ciri oleh lapisan konvolusi boleh diwakili sebagai:

$$FM(output) = X * f + bias \quad (1)$$

| | |
|---|---|
| $\boldsymbol{FM(output)} \epsilon\, \mathbb{R}^{h' \times w' \times n}$: | Peta ciri yang dijanakan dengan jumlah saluran n |
| $*$ | Operasi konvolusi |
| $\boldsymbol{X} \epsilon\, \mathbb{R}^{c \times h \times w}$ | Imej ekspresi muka dengan dimensi ketinggian h, kelebaran w dan jumlah saluran c |
| $\boldsymbol{f} \epsilon\, \mathbb{R}^{c \times k \times k \times n}$ | *convolution filters* dengan saiz kernel kxk dan jumlah saluran |
| ***bias*** | Nilai berat sebelah |

Berdasarkan (1), jumlah FLOPs untuk operasi konvolusi biasa untuk menghasilkan peta ciri intrinsik adalah h'×w'×n×c×k×k dan ia adalah nilai yang besar disebabkan oleh nilai n dan c yang tinggi. Walau bagaimanapun, disebabkan oleh wujudnya peta ciri yang mengekstrakkan ciri-ciri muka dari imej ekspresi muka yang sama ataupun serupa, operasi penghasilan peta ciri intrinsik boleh dipermudahkan dan dikurangkan lagi melalui penggunaan operasi *depthwise convolution* yang ringan untuk menghasilkan *n* peta ciri *ghost* yang dimanifestasikan dengan*:*




*Rajah 4.3: Struktur keseluruhan and proses komputasi perhatian silang dimensi - Triplet Attention*
Sumber Rujukan: (Misra, D. et al. 2020)

$$y_{i,j} = \Phi_{i,j}(FM'(output)_i),\ \forall i=1,\ \ldots m;\ j=1\ldots s \quad (2)$$

| | |
|---|---|
| $\boldsymbol{y_{ij}}$ | Peta ciri *Ghost* yang ke-*j*, dihasilkan daripada proses konvolusi *depthwise* |
| $\boldsymbol{FM'(output)_i}$ | Peta intrinsik *i*, iaitu salah satu peta ciri daripada peta ciri *m* |
| $\boldsymbol{\Phi_{i,j}}$ | Operasi konvolusi *depthwise*, *j* pada *i* |

Dariapda peta ciri yang mempunyai persamaan yang tinggi dari segi ciri-ciri wajah ekspresi muka yang diekstrakkan oleh *m* dan ia dimanifestasikan dengan persamaan:

$$FM'(output) = X * f' + bias \quad (3)$$

| | |
|---|---|
| $\boldsymbol{FM'^{(output)} \epsilon\ \mathbb{R}^{h' \times w' \times m}, m < n}$: | Jumlah *m* peta intrinsik yang dijana daripada operasi konvolusi oleh lapisan konvolusi biasa |
| * | Operasi konvolusi |
| $\boldsymbol{f'\ \epsilon \mathbb{R}^{c \times k \times k \times m}}$ | *Filters* yang digunakan untuk menjana peta ciri *ghosts* |

Penjanaan peta ciri *ghost* yang mempunyai ciri-ciri ekspresi muka melalui operasi *depthwise convolution* merupakan proses yang lebih ringan kerana proses konvolusi yang diimplementasikan adalah berdasarkan satu filter kepada satu saluran sahaja, berbanding dengan proses konvolusi biasa yang melibatkan komputasi n jumlah saluran kepada n jumlah *filters.* Tambahan pula, *depthwise convolution* yang memanfaatkan kelewaan ciri-ciri muka dalam peta ciri ekspresi muka seperti yang dijelaskan merupakan sumbangan kepada pengurangan kos komputasi model FER dengan lebih lanjutnya.

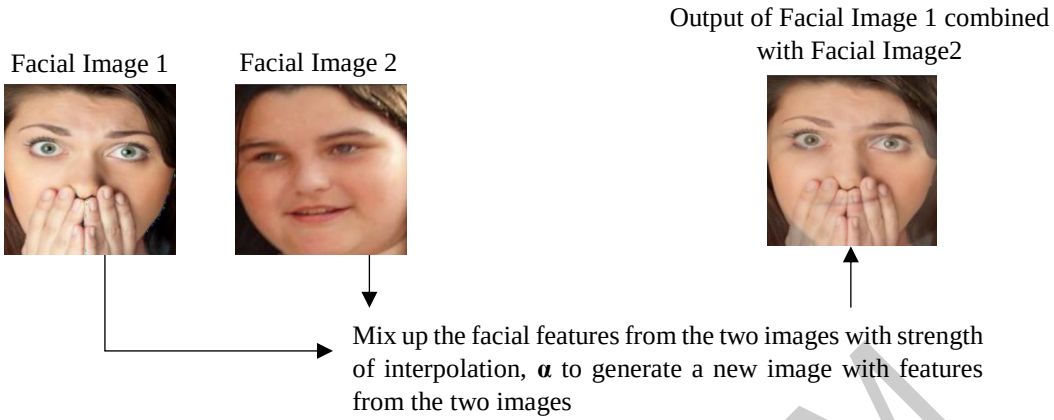
## TRIPLET ATTENTION

*Triplet Attention* (Rajah 4.3) diperkenalkan sebagai mekanisme perhatian GhostFace supaya ia memiliki penumpuan terpilih seperti penglihatan dan persepsi manusia untuk berfokus kepada bahagian wajah imej ekspresi muka yang penting dalam pengklasifikasian ekspresi muka yang

tepat dan mengabaikan maklumat- maklumat dalam imej yang tidak penting. Triplet Attention menggunakan interaksi silang dimensi (cross-dimensional interaction) untuk menangkapkan perwakilan ciri wajah ekspresi muka yang diskret di bawah kos komputasi overhead yang lebih rendah tanpa memerlukan GhostFace untuk mencapai threshold jumlah parameter model minimum untuk memanfaatkan faedah seperti yang dinyatakan. Dengan ini, GhostFace boleh mencapai keringanan dalam jumlah parameter dan pada masa yang sama berfokus kepada bahagian muka yang penting dalam pengklasifikasian ekspresi muka yang baik.

Triplet Attention terdiri daripada 3 cabang selari di mana 2 daripadanya masing-masing bertanggungjawab untuk menangkapkan interaksi silang dimensi antara dimensi saluran, C dengan dimensi ketinggian, H dan kelebaran, W. Cabang terakhirnya berfokus kepada menangkapkan perhatian ruangan (spatial attention) yang menentukan bahagian muka yang perlu difokuskan ("where") dalam saluran oleh model sebelum ia diagregatkan dengan 2 perhatian silang dimensi saluran (cross-dimensional channel attention) tersebut dan dipuratakan sebagai output perhatian terakhir yang dikirakan terhadap bahagian muka imej ekspresi muka yang memerlukan perhatian yang lebih signifikan demi mencapai ketepatan pengklasifikasian ekspresi muka yang tinggi.

Komputasi Triplet Attention bermula dengan input tensor $\boldsymbol{X} \boldsymbol{\epsilon} \mathbb{R}^{\boldsymbol{h} \times \boldsymbol{w} \times \boldsymbol{c}}$ yang diputarkan ke 90◦ lawan arah jam sepanjang paksi H dan diteruskan dengan komputasi yang ringan, $Z - Pool\ (x) = [MaxPool_{od}(X), AvgPool_{od}(X)]$ untuk mengurangkan dimensi sifar tensor kepada dua demi menggambarkan perwakilan tensor sebenar dengan lebih diskret. Hasil daripada Z-Pooling, $\boldsymbol{X} \boldsymbol{\epsilon} \mathbb{R}^{\boldsymbol{2} \times \boldsymbol{h} \times \boldsymbol{c}}$, akan melalui lapisan konvolusi dengan saiz kernel KxK sebelum Batch Normalization dan pengiraan keberatan perhatian melalui pengaktifan Sigmoid, σ. Akhir sekali, keberatan perhatian yang dikira adalah digunakan oleh $\boldsymbol{X}$ dan ia diputarkan 90◦ arah jam sepanjang paksi H demi mengekalkan bentuk asal tensor. Proses yang sama ini berulang bagi saluran W. Bagi komputasi perhatian ruangan pula, C bagi $\boldsymbol{X} \boldsymbol{\epsilon} \mathbb{R}^{\boldsymbol{h} \times \boldsymbol{w} \times \boldsymbol{c}}$ adalah dikurangkan dengan Z-Pooling sebelum diproseskan oleh lapisan konvolusi, Batch Normalization dan pengiraan keberatan perhatian dengan pengaktifan Sigmoid. Akhirnya, ketiga-tiga output keberatan perhatian daripada cabangan adalah ditambah dan dipuratakan serta peta ciri yang dihasilkan pada lapisan di mana wujudnya mekanisme perhatian ini mengandungi ciri-ciri muka yang diskret dalam pengklasifikasian ekspresi muka yang tepat.




*Rajah 4.4: Mekanisma pembesaran data nyah-hafazan MixUp adalah divisualisasikan dengan menggunakan sampel imej ekspresi muka set data RAF-DB pembelajaran model. Ia menggabungkan ciri-ciri muka daripada imej muka kedua kepada imej muka pertama dengan kekuatan interpolasi, α sebelum digunakan dalam pembelajaran model FER. Hal ini untuk mengelakkan model FER daripada menghafazkan ciri-ciri ekspresi muka yang salah bagi sesuatu emosi ekspresi muka dengan memperkenalkan sifat ciri-ciri muka yang berlainan pada suatu imej muka yang ditetapkan.*

## 4.3 FUNGSI KERUGIAN

$$L = -\frac{1}{m}\sum_{i=1}^{m} y^i \cdot log(\hat{y}^i)$$

$y^i$ : Label bagi kelas ekspresi muka sebenar

$\hat{y}$ : Kebarangkalian softmax bagi kelas i

m : Jumlah kelas dalam klasifikasi ekspresi muka

Fungsi kerugian yang digunakan adalah kerugian cross entropy yang menghitung perbezaan antara dua taburan kebarangkalian ekspresi muka dalam format *one-hot encoding*, iaitu taburan kebarangkalian ekspresi muka sebenar (*ground truth*) dan taburan kebarangkalian ekspresi muka yang diramalkan oleh model. Hal ini kerana jumlah kelas pengklasifikasian melebihi dua kelas.

## 4.4 TEKNIK PEMBESARAN DATA NYAH-HAFAZAN

Teknik pembesaran data nyah-hafazan MixUp dan RandAug diperkenalkan sebelum bermulanya proses pengekstrakan ciri-ciri muka dan pengklasifikasian ekspresi muka oleh GhostFace demi mengelakkan ia daripada menghafazkan sifat ciri-ciri ekspresi muka yang salah bagi emosi ekspresi muka yang dilabelkan ataupun terlalu yakin dengan hasil pengklasifikasian ekspresi mukanya. Dengan ini, hasil pengklasifikasian ekspresi muka GhostFace terhadap imej ekspresi muka yang menyerupai senario kehidupan sebenar manusia dan yang bukan ahli kepada set data pembelajaran model adalah tepat dan konsisten mengikut emosi yang ditunjukkan.

(a) Imej ekspresi muka asal (sebelum teknik pembesaran data MixUp)

(b) Imej ekspresi muka selepas teknik pembesaran data MixUp

*Rajah 4.5: (a) Sampel imej-imej ekspresi muka dari set data pembelajaran model – RAF-DB yang diaplikasikan dengan pembesaran data MixUp untuk menghasilkan (b). Kotak sempadan berwarna merah menunjukkan menerapan salah satu ciri-ciri wajah imej ekspresi muka yang lain dalam set data pembelajaran kepada imej ekspresi muka asal.*

MixUp merupakan teknik pembesaran data agnostik yang dibina untuk mengatasi masalah penghafazan model yang berlaku disebabkan oleh peraturan pembelajaran bagi pengklasifikasian ekspresi muka yang merupakan pembelajaran diselia, iaitu Empirical Risk Minimization (ERM) (Zhang, Chiyuan. et al. 2017). Mekanisme MixUp adalah ditunjukkan dalam Rajah 3.4 dan dimanifestasikan dengan persamaan $\tilde{x} = \lambda x_i + (1-\lambda)x_j$ , $\tilde{y} = \lambda y_i + (1-\lambda)y_j$di mana $(x_i, y_i)$ dan $(x_j, y_j)$ merupakan vektor *feature-target* yang diperoleh daripada set data pembelajaran FER dan $\lambda \sim \text{Beta}(\alpha,\alpha)$ ataupun *Beta Distribution*, dimana $\alpha \in (0,\infty)$ dan α merupakan kekuatan interpolasi bagi kedua-dua vektor *feature-target* tersebut. Peningkatan dalam α melambangkan pengaruh $(x_j, y_j)$ terhadap $(x_i, y_i)$ juga meningkat. Lebih banyak contoh teknik MixUp terhadap imej ekspresi muka yang menyerupai senario kehidupan sebenar adalah ditunjukkan dalam Rajah 4.4.

RandAug merupakan teknik pembesaran data yang dibangunkan bukan sahaja untuk melemahkan kekuatan dan risiko penghafazan model FER terhadap sifat ciri-ciri muka yang tidak tepat bagi emosi ekspresi muka yang ditunjukkan akibat daripada masalah contoh musuh dan pelabelan salah dalam set data, tetapi juga menghapuskanataupun mengurangkan kerumitan proses pencarian pencarian teknik pembesaran data yang hanya disesuaikan kepada set data imej ekspresi muka yang spesifik sahaja, tanpa mengurangkan dinamik perubahan pada imej ekspresi muka. Hal ini direalisasikan dengan pemilihan bilangan penjelmaan ataupun teknik pembesaran data, *N* daripada jumlah penjelmaan yang wujud, *K* yang akan digunakan terhadap suatu imej ekspresi muka secara rawak dengan kebarangkalian $\frac{1}{K}$ dan magnitud herotan teknik pembesaran data, *M* tersebut terhadap imej ekspresi muka asal. Contoh operasi RandAug terhadap imej ekspresi muka yang menyerupai senario kehidupan sebenar adalah

*Jadual 4.1: Komponen matriks kekeliruan dan maknanya dalam konteks MIA.*

| Komponen Matriks Kekeliruan | Makna |
|---|---|
| Positif Sebenar (TP) | Sampel ahli kepada set data pembelajaran berjaya diramalkan sebagai ahli |
| Positif Palsu (FP) | Sampel bukan ahli kepada set data pembelajaran adalah diramalkan sebagai ahli |
| Negatif Sebenar (TN) | Sampel bukan ahli kepada set data pembelajaran adalah diramalkan sebagai bukan ahli |
| Negatif Palsu (FN) | Sampel ahli kepada set data pembelajaran adalah diramalkan sebagai bukan ahli |

ditunjukkan dalam Rajah 4.5.

## 4.5 EKSPERIMEN

### METRIK PENILAIAN

Kos komputasi yang berkaitan dengan keringanan model FER dalam pembelajaran dan pengklasifikasian (pengekstrakan dan pengklasifikasian) ekspresi muka yang menyerupai senario kehidupan sebenar, keteguhannya terhadap contoh musuh dan pelabelan salah serta hubungan kekuatan penghafazan model FER dengan kejayaan MIA untuk mengenal pasti jikalau suatu imej ekspresi muka merupakan ahli kepada set data pembelajaran model merupakan aspek-aspek dalam skop kajian projek ini yang boleh diukur secara kuantitatifnya dengan menggunakan metrik penanda aras. Pengukuran keringanan model FER adalah diwakili oleh dua metrik, iaitu *floating point operation per second* (*FLOPs*) dan jumlah parameter ataupun saiz model FER (Zhao, Zengqun. et al. 2021) (Laha Ale. et al. 2019). Keteguhan model FER terhadap contoh musuh dan pelabelan salah boleh diukur melalui nilai ralat ujian terbaik dan yang terakhir bagi epok pembelajaran model (Zhang, Hongyi. et al.2018). Kejayaan MIA adalah diukur berdasarkan nilai Kadar Positif Benar (TP) dan Kadar Positif Palsu (FP) dari MIA yang diperolehi daripada matriks kekeliruan (Murakonda, S. K. et al. 2020) yang merupakan komponen utama dalam matriks kekeliruan MIA yang membawa makna tentang keberkesanannya untuk mengenal pasti imej ekspresi muka adalah ahli dan bukan ahli kepada set data pembelajaran GhostFace dengan betul ataupun salah. Makna kepada komponen-komponen matriks kekeliruan adalah ditunjukkan dalam Jadual 4.1.

Seterusnya, ketepatan GhostFace untuk mengklasifikasikan imej ekspresi muka ke kelas yang betul adalah diukur bagi setiap tetapan eksperimen dalam fasa pembelajaran model yang berlainan seperti pembelajaran model dengan teknik pembesaran data nyah-hafazan ataupun sebaliknya. Lebih-lebih lagi, keteguhan GhostFace terhadap contoh musuh dan

masalah pelabelan salah dalam set data FER adalah dijustifikasikan melalui pendekatan kualitatif, iaitu visualisasi yang melibatkan imej ekspresi muka di mana nilai ground truth yang diberikan adalah salah bagi emosi ekspresi muka tersebut tetapi berjaya diramalkan dengan ekspresi muka yang betul oleh GhostFace. Imej ekspresi muka yang tidak diklasifikasikan kepada kelas emosi yang betul sebelum penggunaan teknik pembesaran data nyah-hafazan dan hasil pengklasifikasiannya terhadap imej ekspresi muka yang sama selepas penggunaan teknik ini tersebut adalah divisualisasikan demi menunjukkan kesannya terhadap keupayaan pengklasifikasian ekspresi muka oleh GhostFace. Akhirnya sekali, perhatian GhostFace terhadap ciri-ciri muka untuk mengklasifikasikan ekspresi muka adalah divisualisasi dan ia juga merupakan penilaian kualitatif terhadap keupayaan GhostFace untuk berfokus kepada ciri-ciri muka yang munasabah digunakan dalam pengklasifikasian ekspresi muka.

**SET DATA**

Eksperimen bagi projek ini menggunakan set data Real-world Affective Faces (RAF-DB) (Li, Shan, et. al, 2018) yang mengandungi sejumlah 15,339 imej ekspresi muka RGB statik, berdimensi berlainan yang dikumpulkan daripada Internet dan ekspresi muka yang ditunjukkan mengandungi tujuh ekspresi muka universal dan ia menyerupai senario kehidupan sebenar manusia. Hal ini kerana subjek dalam imej-imej ekspresi muka mempunyai kepelbagaian dari aspek umur, jantina dan etnik, posisi kepala, keadaan pencahayaan, oklusi seperti cermin mata, rambut muka atau oklusi diri, serta operasi pasca pemprosesan seperti pelbagai penapis dan kesan khas. Daripada 15,339 imej ekspresi muka RAF-DB asas, 12,271 daripadanya adalah set data latihan dan 3,068 adalah set data ujian.

**KETETAPAN EKSPERIMEN**

Pembelajaran GhostFace menggunakan pengoptimuman Stochastic Gradient Descent (SGD), kadar pembelajaran 0.1, latihan kumpulan 16 dan 100 epok, bersama dengan Early Stopping dan Reduce Learning Rate on Plateau demi mengurangkan masalah *overfitting* dan memastikan model mencapai global minimum dengan lebih pantas. Seterusnya, beberapa teknik pembesaran data asas seperti *Horizontal Flip, Normalization, Random Rotation, Height and Width Shift*, digunakan sebagai kaedah asas untuk mengatasi masalah *overfitting* dengan memperkenalkan variasi dalam imej-imej ekspresi muka. Seterusnya, teknik pembesaran data RandAug menggunakan 0 hingga 3 daripada transformasinya dengan magnitud rawak minimum 2 dan maksimum 4. Teknik pembesaran data MixUp hanya dikenakan pada 25% imej dalam set data pembelajaran dengan kelegapan 0.2. Penyiasatan hubungan antara

*Jadual 5.1: Perbandingan jumlah parameter, FLOPs, ralat ujian dan ketepatan pengklasifikasian antara GhostFace dan model daripada kajian sebelumnya.*

| Kaedah | Jumlah Parameter (M) | FLOPS (M) | Ralat Ujian | Ketepatan (%) |
|---|---|---|---|---|
| EfficientFace (Zhao, Zengqun, 2021) | 1.28 | 154.18 | - | 88.36 |
| GhostFace | 1.28 | **134.66** | 0.674 | 77.97 |
| GhostFace + RandAug | | | 0.669 | 77.21 |
| GhostFace + MixUp | | | 0.655 | 77.60 |
| GhostFace + RandAug + MixUp | | | 0.679 | 76.96 |

kekuatan penghafazan GhostFace dan tahap pemeliharaan privasi imej ekspresi muka yang digunakan dalam pembelajaran model adalah diukur dengan memperkenalkan MIA - *Population Attack* terhadap GhostFace yang hanya melibatkan sejumlah 2000 imej ekspresi muka daripada set data pembelajaran dan 2000 imej ekspresi muka daripada set data ujian. Ketetapan teknik pembesaran data asas dan nyah-hafazan adalah sama dalam penyiasatan ini dan 100% imej ekspresi muka akan dikenakan dengan teknik pembesaran data yang dispesifikasikan. Keseluruhan proses kajian ini dijalankan di persekitaran maya dalam Paperspace gradient dengan GPU NVIDIA RTX5000.

## 5 HASIL KAJIAN

### KERINGANAN MODEL

Keringanan dan ketepatan yang dicapai oleh GhostFace adalah dibandingkan dengan kajian yang menyerupainya seperti yang ditunjukkan dalam Jadual 5.1.

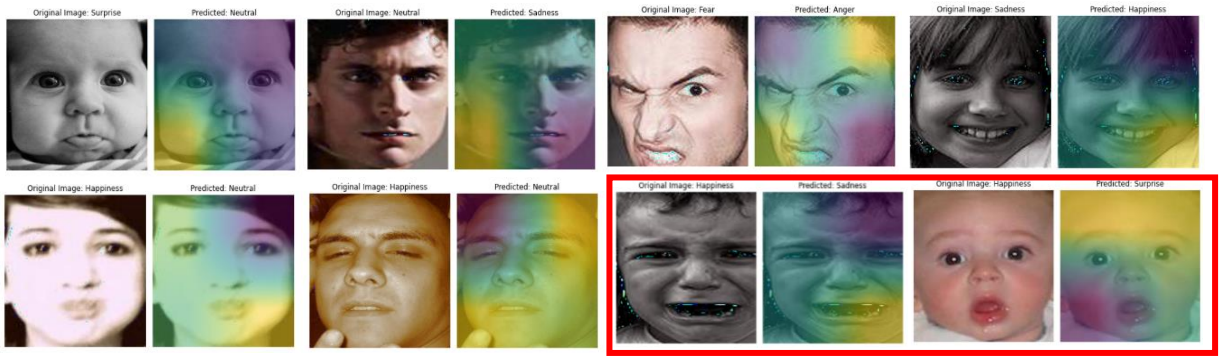
Berdasarkan Jadual 5.1, EfficientFace mencapai ketepatan 88.36% dalam pengklasifikasian imej ekspresi muka bagi set data RAF-DB dan ia adalah lebih tinggi daripada ketepatan yang dicapai oleh GhostFace. EfficientFace merupakan model yang dibangunkan dalam kajian pembangunan model pengecaman ekspresi muka yang ringan dan teguh. Ia menggunakan ShuffleNet-V2 sebagai tulang belakang model FER dan memperkenalkan 2 komponen auksilari perhatian, iaitu blok pengekstrakan ciri tempatan (Local-Feature Extractor) dan blok modulator saluran-ruangan (Channel-Spatial Modulator) yang masing-masing digunakan untuk mengekstrakkan wajah tempatan dan global imej ekspresi muka dengan baik walaupun terdapat kewujudan kepelbagaian yang dinamik dan realistik seperti dalam senario kehidupan seharian manusia pada imej ekspresi muka. Selain itu, kaedah pembelajaran pengagihan label (Label Distributing Learning, LDL) yang mengambil seni bina ResNet50 dan EfficientFace adalah dipra-latihkan dengan set data Microsoft Celeb (MS-CELEB-1M) yang mengandungi 10 juta imej ekspresi muka yang dikumpulkan daripada *Internet* (Guo et al. 2016).

Akhirnya, LDL adalah disertakan bersama EfficientFace pada fasa pembelajaran model terhadap set data RAF-DB. Proses ini menyebabkan EfficientFace mencapai ketepatan pengklasifikasikan ekspresi muka *state-of-the-art* dengan tanpa memperkenalkan jumlah parameter dan FLOPS secara melampau. Namun begitu, EfficientFace melibatkan teknik pra-latihan model dengan set data imej muka yang besar, MSCeleb-1M dan penambahan komponen auksilari secara berlebihan. Hal ini meningkatkan kerumitan seni bina model FERnya dan ia kebergantungan kepada teknik pra-latihan untuk mencapai ketepatan pengklasifikasian ekspresi muka yang tinggi. Seterusnya, set data imej muka yang besar ini merupakan sumber yang mahal untuk diperolehi disebabkan oleh masalah pelanggaran terma privasi asas yang menyebabkan akses kepada set data imej adalah ditamatkan (Harvey, A., n.d). Walaupun ketepatan GhostFace adalah tidak setinggi seperti EfficientFace, GhostFace berjaya mencapai keringanan model dari segi FLOPs yang lebih baik berbanding dengan EfficientFace, hanya dengan menggunakan konsep depthwise convolution dan triplet attention. GhostFace mencapai 12.66% FLOPS yang lebih rendah berbanding dengan EfficientFace di bawah jumlah parameter yang sama. Selain itu, GhostFace tidak bergantung kepada pra-latihan model ataupun memperkenalkan komponen tambahan secara berlebihan yang meningkatkan kerumitan struktur model FER. Ini menjustifikasikan GhostFace merupakan calon yang lebih baik dalam pengklasifikasian ekspresi muka pada masa nyata di mana hasil pengklasifikasian model adalah tepat dan sepadan dengan perubahan ekspresi muka manusia yang cepat dan halus dalam kehidupan sebenar.

**KETEGUHAN MODEL**

Dua teknik pembesaran data nyah-fahazan RandAug, MixUp dan gabungan antaranya telah dikaji sebagai pendekatan untuk meningkatkan keteguhan model terhadap kewujudan contoh musuh dan masalah pelabelan salah di mana imej ekspresi muka adalah disalah labelkan kepada kelas emosi ekspresi yang tidak sepadan dengan ekspresi yang ditunjukkan oleh muka subjek dalam set data FER. Berdasarkan Jadual 5.1, GhostFace tanpa teknik pembesaran data RandAug dan MixUp telah mencapai ketepatan pengklasifikasian ekspresi muka tertinggi dan ketepatan pengklasifikasian ekspresi muka menurun sekiranya kerumitan yang diperkenalkan oleh teknik pembesaran data terhadap imej ekspresi muka meningkat. Akan tetapi, ralat ujian yang dicapai oleh GhostFace sekiranya pembelajarannya disertakan dengan teknik pembesaran data nyah-hafazan adalah lebih rendah berbanding dengan pembelajarannya tanpa teknik ini.

Berdasarkan Jadual 5.1, pembelajaran GhostFace dengan teknik pembesaran data MixUp mencapai ralat ujian yang paling rendah, diikuti dengan RandAug. Ini melambangkan



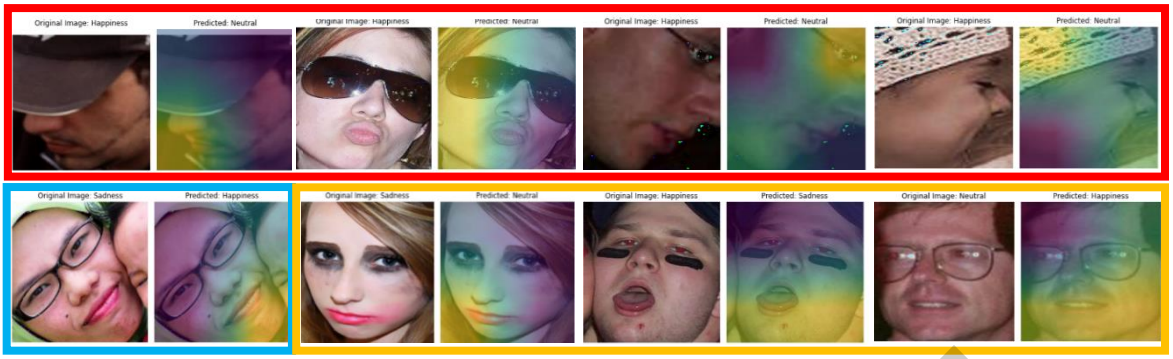

*Rajah 5.1: Contoh imej ekspresi muka dalam set data RAF-DB di mana label ekspresi muka asal adalah tidak sepadan dengan emosi ekspresi muka yang ditunjukkan oleh subjek pada imej (kiri). Akan tetapi, dengan penambahan data nyah-hafazan, GhostFace berjaya mengklasifikasikan imej ekspresi dengan emosi ekspresi muka yang lebih sesuai dengan ekspresi muka yang ditunjukkan (kanan).*

*Nota: Semua label asal dan label peramalan GhostFace adalah di atas setiap imej yang dipaparkan*


*Rajah 5.2: Contoh imej ekspresi muka dengan label ekspresi muka yang sepadan dengan ekspresi muka yang ditunjukkan oleh subjek muka tetapi diklasifikasikan ke kelas ekspresi muka yang salah oleh GhostFace yang dilatih tanpa teknik pembesaran data nyah-hafazan. Walau bagaimanapun, penjelmaan imej ekspresi muka dengan teknik pembesaran data nyah-hafazan membolehkan GhostFace untuk mengklasifikasikan imej ekspresi muka kepada kelas ekspresi muka yang betul.*

*Nota: Semua label asal dan label peramalan GhostFace adalah di atas setiap imej yang dipaparkan*

keteguhan model terhadap masalah pelabelan salah dan contoh musuh meningkat kerana ralat ujian yang lebih rendah melambangkan varians yang rendah, generalisasi model yang lebih baik ketika mengklasifikasikan ekspresi muka di luar set data pembelajarannya serta risiko penghafazan GhostFace terhadap sifat ciri-ciri muka yang salah bagi sesuatu kelas ekspresi muka adalah lebih rendah. Dengan ini, hasil pengklasifikasian ekspresi muka di luar set data pembelajaran adalah munasabah dengan ekspresi muka yang ditunjukkan oleh muka subjek dan ini boleh dijustifikasikan lagi dengan Rajah 5.1 dan Rajah 5.2. Merujuk kepada imej ekspresi muka dalam kotak sempadan berwarna merah dalam Rajah 5.1, imej ekspresi muka budak lelaki dan bayi adalah dilabelkan sebagai kegembiraan padahal emosi ekspresi muka




*Rajah 5.3: Imej-imej ekspresi muka dari set data RAF-DB dimana ekspresi muka yang ditunjukkan oleh subjek imej adalah sukar dikenal pasti bukan sahaja bagi GhostFace, tetapi juga bagi manusia. Hal ini disebabkan oleh kewujudan oklusi (heavily-occluded) dan variasi posisi kepala yang drastik dalam imej-imej ekspresi muka seperti yang ditunjukkan dalam kotak sempadan berwarna merah rajah ini, kewujudan dua jenis ekspresi muka yang wujud dalam satu imej ekspresi muka seperti imej dalam kotak sempadan berwarna biru dan masalah penyamaran melalui solekan dan ekspresi muka yang kabur telah menutupkan ekspresi muka sebenar subjek seperti imej-imej ekspresi muka dalam kotak sempadan berwarna kuning.*

*Nota: Semua label asal dan label peramalan GhostFace adalah di atas setiap imej yang dipaparkan*

yang ditunjukkan adalah bukan seperti yang dilabelkan. Ekspresi muka sebenar bagi budak lelaki adalah kesedihan manakala bayi tersebut adalah kejutan, seperti yang diramalkan oleh GhostFace selepas teknik pembesaran data RandAug dan MixUp. Ini menjelaskan ketepatan pengklasifikasian ekspresi muka yang lebih rendah sekiranya pembelajaran GhostFace melibatkan teknik pembesaran data nyah-hafazan kerana ekspresi muka yang diramalkan adalah tidak sama dengan label asal yang diberikan oleh ahli anotasi. Akan tetapi, ekspresi muka yang diramalkan adalah tepat dan sepadan dengan situasi kehidupan sebenar walaupun ia tidak dicerminkan secara kuantitatifnya. Selain itu, teknik pembesaran data juga meningkatkan kebolehan dan keteguhan GhostFace untuk mengklasifikasikan ekspresi muka yang susah diklasifikasikan (*Rajah 5.2).* Rajah 5.2 menggambarkan pengklasifikasian GhostFace terhadap imej ekspresi muka tanpa teknik pembesaran data RandAug dan MixUp gagal mengklasifikasikan imej ekspresi muka dengan tepat. Akan tetapi, ia berjaya mengklasifikasikan imej ekspresi muka kepada kelas ekspresi muka yang betul selepas penggunaan teknik pembesaran data RandAug dan MixUp kerana ia memanipulasikan dan meningkatkan kepelbagaian dalam perwakilan imej ekspresi muka dalam kelas ekspresi muka yang sama. Dengan ini, keteguhan dan ketepatan model FER untuk mengklasifikasikan imej ekspresi muka yang mempunyai perwakilan yang berbeza bagi kelas ekspresi muka yang sama akan meningkat.

Walau bagaimanapun, penggunaan teknik pembesaran data nyah-hafazan secara berlebihan ataupun meningkatkan tahap pengaruhnya atau interpolasi terhadap imej ekspresi muka sebaliknya mengurangkan ketepatan pengklasifikasikan ekspresi muka seperti yang ditunjukkan dalam Jadual 5.1 dimana pembelajaran GhostFace dengan gabunggan teknik

*Jadual 5.2: Hasil nilai-nilai bagi komponen matrik kekeliruan daripada MIA berdasarkan teknik pembesaran data yang digunakan (asas, RandAug, MixUp) dan jumlah lapisan gaussian noise (regularizer) dengan nilai sisihan piawai gaussian noise 0.1.*

| Ketetapan Eksperimen | | TP (%) | FP (%) | TN (%) | FN (%) |
|---|---|---|---|---|---|
| Lapisan *Gaussian Noise* | Teknik Pembesaran Data | | | | |
| 1 | Tiada | 8.3 | 5.5 | 45 | 42 |
| 1 | Asas | 5.9 | 4.8 | 45 | 44 |
| 1 | RandAug | 5 | 4.3 | 46 | 45 |
| **1** | **MixUp** | **5.5** | **5.3** | **45** | **44** |
| 2 | Tiada | 6 | 5.5 | 44 | 44 |
| 2 | Asas | 5.4 | 4.5 | 46 | 45 |
| 2 | RandAug | 5.8 | 5.5 | 44 | 44 |
| **2** | **MixUp** | **4.9** | **8.7** | **45** | **41** |

pembesaran data nyah MixUp dan RandAug mempunyai ralat ujian yang hampir sama dengan pembelajarannya tanpa teknik ini dan ketepatan yang lebih rendah. Selain itu, terdapat imej-imej ekspresi muka dalam set data RAF-DB yang terjejas secara teruknya oleh kewujudan oklusi, variasi posisi kepala yang melampau, kewujudan dua jenis emosi ekspresi muka yang berbeza pada imej ekspresi muka yang sama, penyamaran ekspresi muka melalui solekan dan lain-lain seperti yang ditunjukkan dalam Rajah 5.3. Oleh itu, ia menyebabkan majoriti ciri-ciri muka yang penting bagi pengklasifikasian ekspresi muka ataupun sifat sebenar ciri-ciri muka adalah tertutup yang akhirnya meningkatkan kesukaran bukan sahaja bagi GhostFace, tetapi juga kepada manusia untuk mengenal pasti emosi ekspresi muka yang ditunjukkan oleh subjek imej ekspresi tersebut dengan tepat.

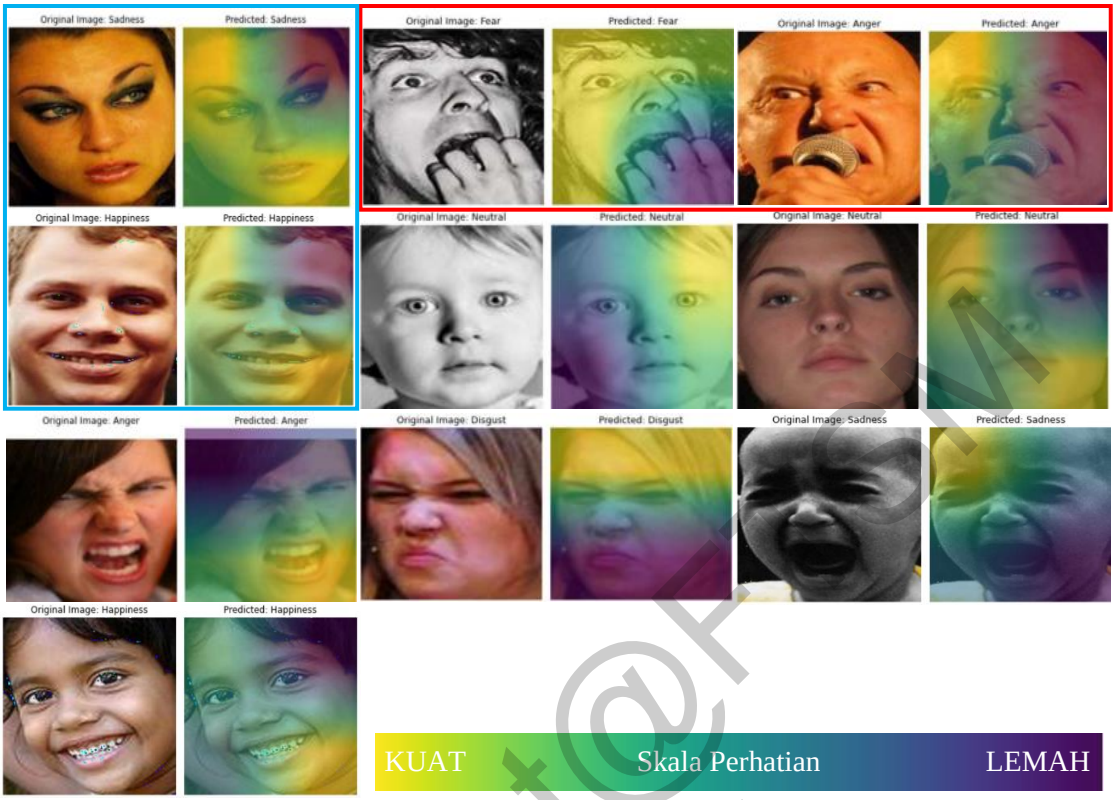
## PEMELIHARAAN PRIVASI

Berdasarkan (Song, Liwei, 2020) dan (Zhang, Chiyuan. et al. 2017), model rangkaian saraf mendalam adalah cenderung untuk menghafaz data sensitif sebaik sahaja jumlah parameter model melebihi jumlah data dalam set data, menjadikannya terdedah kepada MIA (Carlini, N., 2021). Dengan ini, penyiasatan hubungan antara kekuatan penghafazan GhostFace dan tahap pemeliharaan privasi terhadap imej ekspresi muka yang digunakan dalam pembelajaran model adalah dilaksanakan. Kekuatan pemeliharaan privasi terhadap imej ekspresi muka mempunyai kebergantungan yang lebih signifikan kepada TP dan FP yang dicapai daripada MIA yang diperkenalkan secara sengaja terhadap GhostFace. Tahap penghafazan GhostFace pula dipengaruhi oleh kewujudan dan ketidakwujudan teknik pembesaran data dalam pembelajarannya tentang ekspresi muka. Berdasarkan Jadual 5.2, pembelajaran ekspresi muka GhostFace yang melibatkan teknik pembesaran data asas dan nyah-hafazan telah mencapai pemeliharaan privasi imej ekspresi muka yang lebih baik terhadap MIA berbanding dengan pembelajarannya tanpa teknik pembesaran data disebabkan oleh nilai TP yang tinggi.

Walaupun kekuatan *regularizer* (lapisan gaussian noise) meningkat, pembelajaran ekspresi muka tanpa teknik pembesaran data mempunyai kerentaan yang lebih tinggi terhadap MIA.

Antara tiga jenis teknik penambahan data, MixUp, sebagai teknik penambahan nyah-hafazan terhadap data sensitif (imej ekspresi muka) menunjukkan perlindungan privasi terhadap data sensitif yang lebih baik disebabkan oleh nilai TP yang secara relatifnya rendah di bawah nilai FP yang tinggi. Ini menunjukkan bahawa ahli-ahli yang diramalkan oleh MIA sebagai ahli kepada imej ekspresi muka set data pembelajaran model adalah salah (Carlini, N., 2021) dan jumlah ahli kepada set data pembelajaran yang berjaya dikenal pastikan adalah relatifnya rendah bagi 1 lapisan gaussian noise dan paling rendah bagi 2 lapisan gaussian noise (Jadual 5.2). Ini menunjukkan kekuatan serangan daripada MIA adalah lebih lemah dan keteguhan GhostFace terhadap MIA serta pemeliharaan privasi imej ekspresi muka dalam set data pembelajaran menjadi lebih baik. Hal ini berlaku disebabkan oleh gabungan rawak beberapa ciri-ciri muka bagi sesuatu imej ekspresi muka kepada imej ekspresi muka yang lain yang menyebabkan MIA gagal dalam mengenal pasti pemilik atau identiti sebenar suatu imej tersebut. Bagi purata TN dan FN, ia adalah tinggi kerana kebanyakan sampel adalah bukan ahli dalam aplikasi dunia sebenar. Seterusnya, penambahan lapisan gaussian noise yang bertugas sebagai regularizer kepada GhostFace juga mempengaruhi kejayaan MIA. Daripada Jadual 5.2, dapat disimpulkan bahawa peningkatan dalam regularization GhostFace, secara keseluruhannya akan meningkatkan pemeliharaan privasi imej ekspresi muka seperti yang ditunjukkan dalam pengurangan TP dan peningkatan dalam FP bagi pembelajaran model dengan ataupun tanpa teknik pembesaran data. Hal ini demikian kerana filters yang mengandungi maklumat yang diekstrakkan daripada imej ekspresi muka dan yang penting untuk mengenal pasti identiti individu dan emosi imej ekspresi muka tersebut adalah diputarbelitkan ataupun rosak daripada imej ekspresi muka asal. Kerosakan yang diperkenalkan adalah aktif pada fasa pembelajaran model sahaja dan ia menyebabkan maklumat ciri-ciri muka yang dipelajari oleh GhostFace adalah tidak sama dengan perwakilan asal imej ekspresi muka tersebut dan ia akan dianggap sebagai imej yang berbeza oleh MIA. Oleh itu, serangan GhostFace daripada MIA akan menjadi lebih lemah di mana lebih banyak imej ekspresi muka yang diramalkan merupakan ahli kepada set data pembelajaran sebenarnya bukannya ahli. Oleh itu, dengan menggunakan teknik pembesaran data nyah-hafazan dan meningkatkan regularization semasa pembelajaran GhostFace terhadap ekspresi muka, ia mampu mengurangkan kerentaan model FER terhadap MIA dan meningkatkan pemerliharaan privasi imej ekspresi muka yang digunakan dalam pembelajaran model.

## VISUALISASI PERHATIAN MODEL




*Rajah 5.4: Visualisasi wajah muka yang diperhatikan oleh model FER semasa pengklasifikasian imej ekspresi muka RAF-DB dengan menggunakan GRAD-CAM. Peta warna jet yang digunakan dalam GRAD-CAM adalah viridis dan ia mewakili tahap ataupun skala perhatian yang dikenakan oleh model FER terhadap imej ekspresi muka. Bahagian muka dalam imej ekspresi muka yang mempunyai warna yang dekat dengan kategori skala perhatian kuat melambangkan ia merupakan ciri-ciri muka yang difokuskan oleh GhostFace semasa mengklasifikasikan kelas ekspresi muka bagi imej ekspresi muka tersebut. Akan tetapi, warna yang dekat dengan kategori skala perhatian lemah melambangkan ciri-ciri muka tersebut mempunyai sumbangan yang kecil dalam pengklasifikasian ekspresi muka oleh model.*

*Nota: Semua label asal dan label peramalan GhostFace adalah di atas setiap imej yang dipaparkan*

Kaedah GRAD-CAM (Selvaraju, R. et al. 2017) telah digunakan dalam visualisasi bahagian wajah yang difokuskan oleh GhostFace semasa pengklasifikasikan ekspresi muka dan ia direalisasikan dengan menggunakan peta pengaktifan (activation map) yang menggambarkan bahagian imej yang relevan terhadap kelas ekspresi tertentu. Visualisasi ini telah menggunakan lapisan konvolusi terakhir dalam GhostFace yang dibangunkan. Berdasarkan Rajah 5.4, GhostFace berjaya untuk berfokus kepada ciri-ciri muka penting yang lain (*non-occluded region*) bagi sesetengah imej ekspresi muka untuk mengklasifikasikan ekspresi muka walaupun wujudnya masalah oklusi daripada kewujudan objek ataupun *self-occlusion* yang menyorokkan ciri-ciri muka yang mungkin penting dalam pengklasifikasian ekspresi muka. Ini ditunjukkan oleh imej- imej dalam kotak sempadan berwarna merah dalam Rajah 5.4 dimana mulut subjek bagi imej ekspresi ketakutan (fear) dan kemarahan (anger) masing-masing adalah tertutup

dengan tangan sendiri dan mikrofon. Walau bagaimanapun, GhostFace berjaya berfokus kepada bahagian mata mereka untuk mengklasifikasikan imej ekspresi muka kepada kelas ekspresi muka yang betul, iaitu ketakutan dan kemarahan. Selain itu, terdapat imej ekspresi muka di mana GhostFace telah menggunakan 2 bahagian muka yang penting demi mengklasifikasikan ekspresi muka dengan betul, iaitu mata dan mulut seperti yang ditunjukkan oleh imej ekspresi muka dalam kotak sempadan berwarna biru di Rajah 5.4. Akhir sekali, terdapat situasi di mana GhostFace hanya berfokus kepada satu bahagian muka untuk mengklasifikasikan ekspresi muka dengan betul seperti imej ekspresi muka kesedihan bayi paling kanan ke bawah Rajah 5.4.

# 6 KESIMPULAN

Projek ini berjaya mencapai objektif kajian pertamanya, iaitu membangunkan model FER baru – GhostFace yang menggunakan kos komputasi yang lebih rendah ataupun mencapai peningkatan keringanan dari segi FLOPs sebanyak 12.66% dalam pembelajaran dan peramalan ekspresi muka yang menyerupai senario kehidupan sebenar manusia dari set data RAF-DB. Ini direalisasikan dengan menggunakan komputasi konvolusi yang lebih murah daripada konvolusi konvensional, iaitu dengan depthwise convolution yang juga memanfaatkan kewujudan kelewahan ciri-ciri muka yang diekstrakkan dalam peta ciri dalam pembelajarannya tentang ekspresi muka. GhostFace juga mengambil kira interaksi silang dimensi antara dimensi dalam imej yang dibekalkan oleh triplet attention supaya GhostFace mampu berfokus kepada ciri-ciri muka yang spesifik untuk mencapai ketepatan pengklasifikasian ekspresi muka yang baik. Kekuatan GhostFace adalah ia mempunyai seni bina yang ringkas dan hanya menggunakan FLOPs yang lebih rendah untuk mengklasifikasikan ekspresi muka dengan ketepatan yang baik tanpa bergantung kepada teknik pra-latihan model dengan set data imej muka yang besar dan mahal untuk diperolehi. Disebabkan ia mencapai keringanan yang lebih rendah, ini menjadikannya sesuai digunakan untuk mengklasifikasikan ekspresi muka manusia pada masa nyata kerana hasil pengklasifikasian ekspresi mukanya adalah lebih cepat dan sepadan dengan perubahan ekspresi muka manusia yang cepat dan halus seperti dalam kehidupan sebenar.

Seterusnya, projek ini juga berjaya mencapai objektif kajian keduannya, iaitu untuk meningkatkan keteguhan model FER cadangan – GhostFace terhadap contoh musuh dan masalah pelabelan salah yang wujud dalam set data FER dengan menggunakan teknik pembesaran data nyah-hafazan RandAug dan MixUp. Ini dibuktikan dengan nilai ralat ujian dicapai yang lebih rendah (Jadual 5.1) dan kejayaan GhostFace untuk mengklasifikasikan imej

ekspresi muka yang disalah labelkan dengan kelas emosi ekspresi muka yang tidak sepadan dengan ekspresi muka yang ditunjukkan oleh muka subjek pada imej ekspresi muka (Rajah 5.1). Selain itu, GhostFace bersama dengan teknik pembesaran data nyah-hafazan mampu mengklasifikasikan imej ekspresi muka kepada kelas ekspresi mukanya yang betul dimana ia gagal mengklasifikasinya dengan betul sebelum ataupun tanpa teknik pembesaran data nyahhafazan (Rajah 5.2). Walaupun berlakunya peningkatan dalam keteguhan model untuk mengklasifikasikan imej ekspresi muka yang disalah labelkan kepada kelas ekspresi muka yang betul dan hal ini tidak dicerminkan secara kuantitatifnya, tetapi keupayaannya untuk mengklasifikasikan imej ekspresi muka yang menyerupai senario kehidupan sebenar manusia dan yang bukannya ahli kepada set data pembelajaran model adalah lebih tepat.

Projek ini juga berjaya menyiasat hubungan antara kekuatan penghafazan model FER cadangan - GhostFace dan pemeliharaan privasi imej ekspresi muka yang digunakan dalam pembelajaran model terhadap MIA. Berdasarkan Jadual 5.2, disimpulkan bahawa pembelajaran GhostFace dengan teknik pembesaran data asas ataupun nyah-hafazan mempunyai pemeliharaan privasi imej ekspresi muka dalam set data pembelajaran yang lebih baik berbanding pembelajarannya tanpa teknik pembesaran data. Selepas itu, antara 3 teknik pembesaran data yang berbeza – asas, RandAug, dan MixUp, MixUp yang bertujuan untuk mengurangkan penghafazan GhostFace terhadap ekspresi muka merupakan kaedah yang lebih baik dalam memelihara privasi imej ekspresi muka yang digunakan dalam pembelajaran model daripada MIA. Peningkatan dalam kekuatan regularization juga membantu model untuk mengurangkan kerentaannya terhadap MIA. Oleh itu, dapat disimpulkan peningkatan keupayaan model untuk mengklasifikasikan ekspresi muka dengan tepat di luar set data pembelajarannya dan pengurangan dalam kerentaannya terhadap serangan adversarial seperti MIA boleh direalisasi dengan teknik pembesaran data nyah-hafazan yang mengurangkan kekuatan model dalam menghafazkan imej ekspresi muka dan risiko penghafazannya terhadap sifat ciri-ciri muka yang tidak sepadan dengan ekspresi muka yang ditunjukkan oleh muka subjek dalam imej ekspresi muka.

## 7 RUJUKAN

Carlini, N., Chien, S., Nasr, M., Song, S., Terzis, A., & Tramer, F. 2021. Membership inference attacks from first principles. arXiv

Cubuk, E. D., Zoph, B., Shlens, J., & Le, Q. V. 2020. Randaugment: Practical automated data augmentation with a reduced search space. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. pp. 702-703

Dino, H.I., & Abdulrazzaq, M.B. 2019. Facial Expression Classification Based on SVM, KNN and MLP Classifiers. *2019 International Conference on Advanced Science and Engineering (ICOASE)*, pp. 70-75.

Ekman, P. 2009. Darwin's Contributions to Our Understanding of Emotional Expressions. *Philosophical Transactions of the Royal Society of London.* Series B, Biological Sciences, 364(1535): 3449-3451.

Ferro-P'erez, R., & Mitre-Hernández, H.A. 2020. ResMoNet: A Residual Mobile-based Network for Facial Emotion Recognition in Resource-Limited Systems. *ArXiv, abs/2005.07649.*

Haber, N., Voss, C., & Wall, D. 2020. Upgraded Google Glass Helps Autistic Kids "See" Emotions. *IEEE Spectrum* https://spectrum.ieee.org/upgraded-google-glass-helps-autistic-kids-see-emotions/particle-1 [17 Oktober 2021]

Han, Kai., Wang, Yunhe., Tian, Qi., Guo, Jianyuan., Xu, Chunjing., & Xu, Chang. 2020. Ghostnet: More features from cheap operations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 1580-1589.

Harvey, A. (n.d.). *MS-celeb-1m*. Exposing.ai. https://exposing.ai/msceleb/ [10 Julai 2022]

Hui, Ding., Peng, Zhou., & Rama Chellappa. 2020. Occlusion-Adaptive Deep Network for Robust Facial Expression Recognition. *CoRR*. pp. 1-9.

Katsis, C. D., Rigas, G., Goletsis, Y., & Fotiadis, D. I.. 2015. Emotion Recognition in Car Industry. In A. Konar, & A. Chakraborty. *Emotion Recognition: A Pattern Analysis Approach*, pp. 515-544. John Wiley & Sons, Inc.

Kret, M. E. 2015. Emotional Expressions Beyond Facial Muscle Actions. A Call for Studying Autonomic Signals and Their Impact on Social Perception. *Frontiers in Psychology* 6: 711.

Laha Ale, Xiaojie Fang, Dajiang Chen, Ye Wang, Ning Zhang. 2019. Lightweight Deep Learning Model For Facial Expression Recognition. *13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE).*pp. 707-712

Leino, K., & Fredrikson, M. 2020. Stolen Memories: Leveraging Model Memorization for Calibrated {White-Box} Membership Inference. *29th USENIX security symposium (USENIX Security 20).* pp. 1605-1622.

Li, HanTing. Sui, MingZhe. Feng, Zhao. Zha, ZhengJun. & Feng, Wu. 2021. MVT: Mask vision transformer for facial expression recognition in the wild. arVix. pp. 1-11.

Li, Shan. Deng, Weihong and Du, JunPing. 2017. Reliable Crowdsourcing and Deep Locality-Preserving Learning for Expression Recognition in the Wild 2017. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR*) *2017*. pp. 2584-2593

Li, Yong & Zeng, JiaBei & Shan, ShiGuang & Chen XiLin. 2019. Occlusion Aware Facial Expression Recognition Using CNN With Attention Mechanism. IEEE Transactions on Image Processing 28(5): 2439-2450.

Ma, H., Celik, T., & Li, H.-C. 2021. Lightweight attention convolutional neural network through network slimming for robust facial expression recognition. *Signal, Image and Video Processing*, 15(7), 1507–1515.

Mahawaga Arachchige, P., Bertók, P., Khalil, I., Liu, D., & Çamtepe, S.A. 2020. Privacy Preserving Face Recognition Utilizing Differential Privacy. *Computer & Security (97):* 101951.

Martinez, B., & Valstar, M. F. 2016. Advances, Challenges, and Opportunities in Automatic Facial Expression Recognition. In C. M. Kawulok M., *Advances in Face Detection and Facial Image Analysis*. pp. 63-100. Springer, Cham.

Martinez-Martin, N. 2019. What Are Important Ethical Implications of Using Facial Recognition Technology. AMA Journal of Ethics, 21(2): 180-187. doi:10.1001/amajethics.2019.180.

Misra, D., Nalamada, T., Arasanipalai, A. U., & Hou, Q. 2021. Rotate to attend: Convolutional triplet attention module. *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 3139-3148

Murakonda, S. K., & Shokri, R. 2020. Ml privacy meter: Aiding regulatory compliance by quantifying the privacy risks of machine learning. arXiv

Nor, M. A., Tasrib, M. A., Francis, B., Hesham, N. I., & Othman, M. B. 2021. A Study on the Laws Governing Facial Recognition Technology and Data Privacy in Malaysia. Malaysian Journal of Social Sciences and Humanities (MJSSH), 6(10), 480-487. doi:https://doi.org/10.47405/mjssh.v6i10.1086

Noor, J., Daud, M., Rashid, R., Mir, H., Nazir, S., & Velastin, S. A. 2020. Facial Expression Recognition using Hand-Crafted Features and Supervised Feature Encoding. 2nd International Conference on Electrical, Communication and Computer Engineering (ICECCE): 1-5.

Pantic, M., & Patras, I. 2005. Detecting Facial Actions and their Temporal Segments in Nearly Frontal-View Face Image Sequences. *2005 IEEE International Conference on Systems, Man and Cybernetics*. pp. 3358-3363.

Rabin, M. R. I., Hussain, A., Hellendoorn, V. J., & Alipour, M. A. 2021. Memorization and Generalization in Neural Code Intelligence Models. arXiv.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. 2017. Grad-cam: Visual explanations from deep models via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision.* pp. 618-626.

Sheaffer, B. L., A.Gold, J., & Averett, P. 2009. Facial Expression Recognition Deficits and Faulty Learning: Implications for Theoretical Models and Clinical Applications. *International Journal of Behavioral Consultation and Therapy* 5(3-4): 31-55.

Singh, A., Fan, S., & Kankanhalli, M. 2021. Human Attributes Prediction under Privacy-preserving Conditions. *MM '21 Proceedings of the 29th ACM International Conference on Multimedia*: 4698-4706. NY, USA. Association for Computing Machinery.

Song, Liwei, and Prateek Mittal. 2021. Systematic evaluation of privacy risks of machine learning models. *30th USENIX Security Symposium (USENIX Security 21)*. pp. 2615-2632

Tan, MingXing, & Le, Quoc V. 2021. Efficientnetv2: Smaller models and faster training. *International Conference on Machine Learning*, pp. 10096-10106.

Vemou., K., & Horvath, A. 2021. TechDispatch #1/2021 - Facial Emotion Recognition. (T. Zerdick, Editor) Retrieved from European Data Protection Supervisor: https://edps.europa.eu/data-protection/our-work/publications/techdispatch/techdispatch-12021-facial-emotion-recognition_en

Zhang, Chiyuan., Bengio. S., Hardt, M., Recht, B., & Vinyals, O. 2021. Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM* 64(3), 107-115.

Zhang, Hongyi, Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. 2017. mixup: Beyond empirical risk minimization. arXiv.

Zhao, ZengQun & Liu, QingShan & Zhou, Feng. 2021. Robust Lightweight Facial Expression Recognition Network with Label Distribution Training. *Proceedings of the AAAI Conference on Artificial Intelligence* 35(4): 3510-3519.