

ARTICLE

Perbandingan Arsitektur ResNet50V2, InceptionV3, dan DenseNet121 Dalam Klasifikasi Pengenalan Ekspresi Wajah

Comparison of ResNet50V2, InceptionV3, and DenseNet121 Algorithms in Facial Expression Recognition Classification

Sebastianus Adi Santoso Mola, Benyamin Orison Darling Kana Wadu, Asnat Nofri Kenlopo,* dan Varra Chandrika Kumara Tungga

Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana, Kupang, Indonesia

*Penulis Korespondensi: asnatkenlopo@gmail.com

(Disubmit 09-12-24; Diterima 18-02-25; Dipublikasikan online pada 20-06-25)

Abstrak

Ekspresi wajah mampu menyampaikan perasaan seseorang, seperti kebahagiaan, kesedihan, atau kemarahan. Meski manusia secara alami mampu mengenali ekspresi wajah, pengklasifikasian ekspresi sering kali menjadi tantangan. Dengan kemajuan teknologi, analisis dan klasifikasi ekspresi wajah kini dapat dilakukan secara otomatis menggunakan pembelajaran mesin, terutama pada metode Convolutional Neural Network (CNN) seperti ResNet50V2, InceptionV3, dan DenseNet121. Penelitian ini bertujuan untuk membandingkan kemampuan dan efisiensi dari tiga model arsitektur CNN yaitu ResNetV50, InceptionV3, dan DenseNet121 dalam klasifikasi pengenalan ekspresi wajah. Penelitian ini menggunakan dataset ekspresi wajah berjumlah 14.248 gambar yang terbagi menjadi lima kelas: bahagia, marah, netral, sedih, dan terkejut. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk validasi. Hasil evaluasi menunjukkan bahwa ResNet50V2 memberikan performa terbaik dengan akurasi 79%, macro average F1-score 0,76, dan weighted average F1-score 0,75. Model ini unggul dalam menangani distribusi data tidak merata, terutama pada kelas dominan seperti Happy dan Neutral. DenseNet121 menempati posisi kedua dengan akurasi 75%, diikuti oleh InceptionV3 dengan akurasi terendah 65%. ResNet50V2 terbukti sebagai model yang paling efektif untuk klasifikasi ekspresi wajah

Kata kunci: ResNet50V2; IceptionV3; DenseNet121; Klasifikasi; Ekspresi Wajah

Abstract

Facial expressions can convey a person's feelings, such as happiness, sadness, or anger. Although humans are naturally able to recognize facial expressions, classifying expressions often becomes a challenge. With technological advancements, facial expression analysis and classification can now be performed automatically using Machine Learning, especially in Convolutional Neural Network (CNN) methods such as ResNet50V2, InceptionV3, and DenseNet121. This study aims to compare the capabilities and efficiency of three CNN architecture models, namely ResNetV50, InceptionV3, and DenseNet121, in facial expression recognition classification. This study uses a facial expression dataset consisting of 14,248 images divided into five classes: happy, angry, neutral, sad, and surprised. The data is divided into 80% for training and 20% for validation. The evaluation results show that ResNet50V2 provides the best performance with an accuracy of 79%, a macro average F1-score of 0.76, and a weighted average F1-score of 0.75. This model excels in handling uneven data distribution, especially in dominant classes such as Happy and Neutral. DenseNet121 ranks second with an accuracy of 75%, followed by InceptionV3 with the lowest accuracy of 65%. ResNet50V2 has proven to be the most effective model for facial expression classification.

KeyWords: ResNet50V2; IceptionV3; DenseNet121; Classification; Expression Recognition

1. Pendahuluan

Wajah manusia merupakan bagian tubuh yang unik, dan mencerminkan identitas individu sekaligus menjadi alat komunikasi nonverbal yang penting. Melalui ekspresi wajah, manusia dapat menyampaikan berbagai emosi, seperti kebahagiaan, kesedihan, atau kemarahan. Meskipun manusia secara alami mampu mengenali ekspresi wajah, proses pengklasifikasian ekspresi secara objektif sering menjadi tantangan, terutama ketika dilakukan secara manual. Namun, dengan kemajuan teknologi informasi saat ini, ekspresi wajah dapat dianalisis dan diklasifikasikan secara otomatis menggunakan teknologi seperti Machine Learning, khususnya metode Convolutional Neural Network (CNN) [1].

Berbagai model arsitektur CNN, seperti ResNet50V2, InceptionV3, dan DenseNet121, telah banyak digunakan untuk tugas klasifikasi gambar. Setiap model memiliki keunggulan tertentu dalam hal akurasi dan efisiensi. Terdapat penelitian yang menunjukkan bahwa DenseNet121 memiliki performa yang lebih baik dari model lain. Penelitian [2] menunjukkan bahwa model DenseNet121 mampu mengklasifikasikan gambar jenis hewan lebih baik dari model VGG16. Penelitian [3] menunjukkan bahwa DenseNet121 lebih baik dari EfficientNet, dengan akurasi validasi 0,9821 dan pengujian 0,9837. Demikian juga dengan penelitian [4] yang menunjukkan bahwa DensNet121 memiliki performa yang lebih baik dari ResNet34. Kemudian terdapat penelitian lain yang menunjukkan bahwa InceptionV3 lebih direkomendasikan seperti pada aplikasi pengenalan wajah pada sistem absensi karyawan [5] dan dalam klasifikasi jenis jerawat wajah [6]. Namun demikian, terdapat juga penelitian yang menunjukkan bahwa arsitektur Resnet50V2 memiliki performa yang lebih baik untuk mengklasifikasikan jenis rumah adat di NTT [7]. Penelitian [8] dan [9] juga menunjukkan bahwa ResNet50V2 memiliki performa yang lebih baik dibanding arsitektur lain dalam identifikasi penyakit pneumonia [8] dan klasifikasi jenis ras kucing [9].

Secara spesifik dalam aplikasi pengenalan ekspresi wajah, penelitian [10] membandingkan model DenseNet dan ResNet50 dengan dataset FER-2013. Dengan menggunakan metode transfer learning dan fine tuning, model DenseNet memiliki performa yang lebih baik dari ResNet50. Namun, penelitian [11] menemukan bahwa ResNet50 memiliki performa lebih baik dibandingkan dengan DenseNet12. Selanjutnya penelitian [12] membandingkan model VGG16 dengan model Inception dan menemukan bahwa Inception memiliki performa yang lebih baik dari VGG16.

Penelitian-penelitian tersebut menunjukkan bahwa ketiga model tersebut dianggap lebih baik dari model lainnya. Tujuan dari penelitian ini adalah untuk membandingkan performa dari tiga model arsitektur CNN yaitu ResNetV50, InceptionV3, dan DenseNet121 dalam klasifikasi pengenalan ekspresi wajah. Perbandingan ini dilakukan untuk menentukan model yang paling efektif dalam menyelesaikan klasifikasi ekspresi wajah secara otomatis.

2. Metode

2.1 Alur Penelitian



Gambar 1. Alur Penelitian

Gambar 1 menunjukkan alur penelitian. Secara umum, langkah-langkah penelitian ini adalah :

1. Pengambilan Dataset

Penelitian ini menggunakan gambar ekspresi wajah dengan 14.248 gambar, terdiri dari 5 kelas di dalamnya yaitu *angry*, *happy*, *neutral*, *sad*, dan *surprise*. Pengumpulan dataset ekspresi wajah terdiri dari 5 kelas yaitu marah, sedih, senang, terkejut dan netral. Dataset diambil dari website Kaggle [13]. Di dalam dataset tersebut terdapat 6 kelas tetapi yang digunakan hanya 5 kelas yaitu *Angry*, *Happy*,

Neutral, *Sad*, dan *Surprise*, karena kelas *Ahegao* merupakan ekspresi wajah yang menggambarkan pornografi.

2. *Preprocessing*

Tahapan *preprocessing* meliputi normalisasi piksel, aumentasi data, pengubahan ukuran gambar, dan pengisian area kosong akibat augmentasi. Langkah awal, normalisasi pixel ke rentang $[0,1]$ atau rescale dengan nilai sebesar $1/255$, yang bertujuan untuk menyelaraskan skala nilai pixel, untuk mempermudah proses optimasi model. Proses augmentasi dilakukan dengan rotasi acak pada gambar hingga ± 20 derajat untuk mensimulasikan variasi orientasi wajah yang umum terjadi dalam data nyata. Kemudian terdapat pergeseran gambar secara horizontal hingga 20% dari lebar gambar dan secara vertikal hingga 20% dari tinggi gambar dipilih untuk menangani variasi posisi wajah dalam bingkai gambar. Dan yang terakhir pembesaran atau pengecilan gambar hingga 20% dari ukuran aslinya yang digunakan untuk mensimulasikan variasi skala yang sering terjadi akibat jarak kamera dengan subjek. Tahapan *preprocessing* selanjutnya adalah mengubah ukuran gambar setiap dataset menjadi 224×224 , gambar dicerminkan secara horizontal dan ketika augmentasi seperti rotasi atau pergeseran menghasilkan area kosong di luar batas gambar, area tersebut diisi dengan warna piksel terdekat.

3. *Split data/ Pembagian Dataset*

Setelah dilakukan *preprocessing* dan augmentasi data maka data akan dibagi menjadi 2 yaitu data *training* dan data *validation*. Proses pembagian dataset 80% untuk *training* dan 20% untuk validasi. Dataset yang diambil berjumlah 14.248 dataset dengan 11.401 untuk *training* dan 2.847 untuk *validation* sedangkan untuk memprediksi gambar baru diambil gambar ekspresi wajah dari google untuk memprediksi gambar.

4. Klasifikasi

Dataset yang sudah bersih kemudian dilatih menggunakan 3 model arsitektur CNN yaitu *ResNet50V2*, *InceptionV3*, dan *DenseNet121* agar dapat mengklasifikasikan gambar dengan baik.

5. Hasil Klasifikasi

Hasil pelatihan dari ketiga model ini akan dievaluasi dan dibandingkan berdasarkan akurasi pelatihan serta akurasi validasi.

2.2 *Convolutional Neural Network (CNN)*

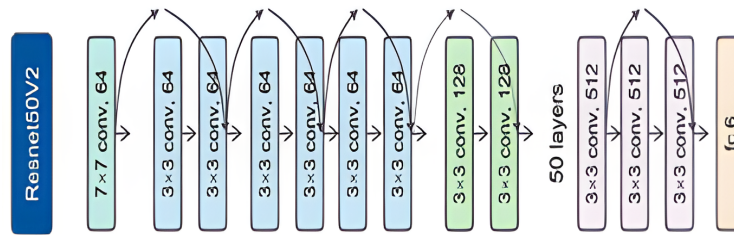
Convolutional Neural Network (CNN) adalah jenis jaringan saraf tiruan yang dikembangkan untuk mengolah data berstruktur jaringan dua dimensi, seperti gambar. CNN merupakan salah satu metode dalam pembelajaran mendalam (*deep learning*) yang dikembangkan berdasarkan *multilayer perceptron (MLP)*. CNN bekerja dengan cara mengekstraksi fitur-fitur penting dari gambar secara otomatis melalui lapisan-lapisan konvolusi yang menangkap pola visual, seperti tepi, tekstur, dan bentuk objek. Didalam CNN terdiri dari beberapa lapisan utama yaitu *convolutional layer*, *activation layer*, *polling layer*, *fully connected layer*, dan *dropout layer*.

2.3 *ResNet50V2*

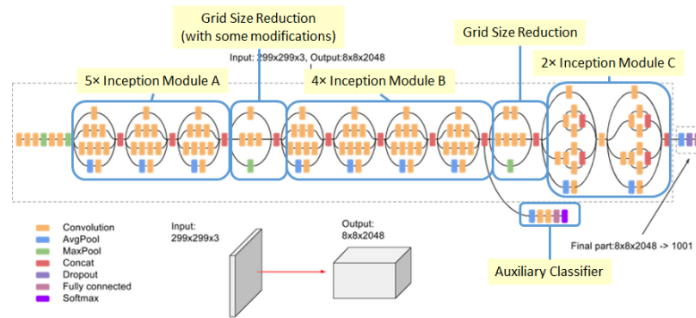
ResNet50V2 adalah arsitektur jaringan saraf konvolusional yang merupakan versi kedua dari model *ResNet50*, yang dirancang oleh tim Microsoft Research. *ResNet (Residual Network)* adalah jaringan yang menggunakan blok residual, yang memudahkan pelatihan jaringan saraf dalam yang sangat dalam. Setelah melewati seluruh blok residual, output dari *convolutional layers* diratakan menggunakan *Global Average Pooling*, yang kemudian diteruskan ke lapisan *fully connected (fc)* untuk menghasilkan prediksi akhir sesuai jumlah kelas. Dengan penerapan *pre-activation* dan *batch normalization* yang diperbaiki, arsitektur *ResNet50V2* memberikan kinerja yang lebih stabil dan efisien dibandingkan dengan versi sebelumnya, dapat dilihat pada Gambar 2.

2.4 *InceptionV3*

InceptionV3 adalah arsitektur jaringan saraf konvolusi mendalam yang dirancang untuk pengenalan gambar dengan efisiensi tinggi. *InceptionV3* dibangun berdasarkan versi arsitektur *Inception* sebelumnya dan menggabungkan berbagai penyempurnaan, sehingga lebih efisien dan efektif dalam komputasi dengan kinerja tinggi, dapat dilihat pada Gambar 3.



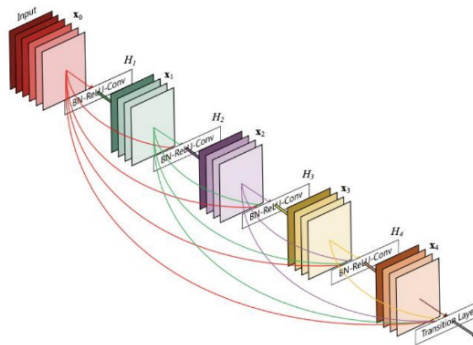
Gambar 2. Arsitektur ResNet50V2 [14]



Gambar 3. Arsitektur Inception V3 [14]

2.5 DenseNet121

DenseNet121 adalah salah satu varian dari arsitektur jaringan saraf konvolusi DenseNet (*Densely Connected Convolutional Networks*) yang dikenal karena konektivitas antar lapisannya. DenseNet121 menghubungkan setiap lapisan secara langsung ke semua lapisan setelahnya dalam blok tertentu. Pendekatan ini memungkinkan jaringan untuk menggunakan kembali fitur yang telah dipelajari sebelumnya, sehingga mengurangi jumlah parameter yang diperlukan dan mempercepat pelatihan, dapat dilihat pada Gambar 4.



Gambar 4. Arsitektur DenseNet121 [15]

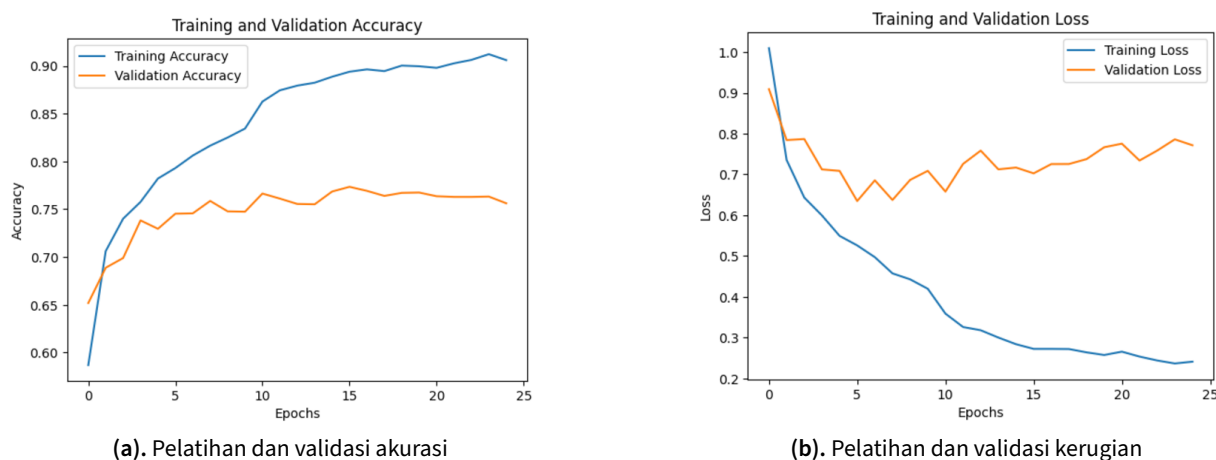
3. Hasil

Dalam setiap model dilakukan preprosecing yang sama. Model dasar ini kemudian ditambahkan lapisan-lapisan kustom, yaitu GlobalAveragePooling2D untuk mereduksi dimensi, lapisan dense dengan 1024 neuron menggunakan fungsi aktivasi ReLU, dan lapisan *dropout* dengan rasio 0.5 untuk mengurangi *overfitting*. Setiap model dikompilasi menggunakan *optimizer* Adam dengan *learning rate* yang cukup kecil yaitu sebesar 0.0001 untuk memungkinkan pelatihan yang stabil dan menghindari langkah pembaruan parameter yang terlalu besar, Fungsi loss *categorical crossentropy* yang cocok untuk klasifikasi multi-kelas, dan metrik akurasi untuk memonitor kinerja selama pelatihan. Selain itu, dua *callbacks* diterapkan untuk meningkatkan efisiensi pelatihan. *EarlyStopping* menghentikan pelatihan jika tidak ada peningkatan pada *validation loss* setelah sejumlah *epoch*, dan memuat kembali bobot terbaik sebelum pelatihan dihentikan. Setiap model dilakukan iterasi hingga mencapai 25 *epoch*. Pemilihan jumlah epoch didasarkan pada hasil

eksperimen awal dengan ResNet50V2, yang menunjukkan bahwa pada *epoch* ke-25, model sudah memberikan hasil pelatihan dan validasi yang baik. Oleh karena itu, jumlah *epoch* yang sama diterapkan pada model lain untuk memastikan perbandingan kinerja antar model dilakukan secara adil dan konsisten.

3.1 Akurasi dan Loss

Akurasi dan loss dari ketiga model dapat dilihat pada Gambar 5, Gambar 5 dan Gambar 5. Pada Gambar 5, grafik bagian (a) menunjukkan bahwa akurasi pelatihan meningkat secara konsisten seiring bertambahnya jumlah *epoch*. Pada akhir pelatihan, akurasi mencapai angka lebih dari 0.90, yang menunjukkan model belajar dengan baik terhadap data pelatihan. Begitupun dengan Akurasi validasi yang juga meningkat pada awal pelatihan, namun cenderung mendatar setelah sekitar *epoch* ke-10, stabil di sekitar angka 0.75. Sedangkan pada gambar bagian (b) menunjukkan bahwa loss pelatihan terus menurun secara signifikan hingga mencapai angka mendekati 0.2 di akhir pelatihan. Ini menunjukkan bahwa model berhasil meminimalkan error pada data pelatihan. Begitu juga dengan loss validasi menurun di awal pelatihan hingga sekitar *epoch* ke-5, namun mulai stabil atau bahkan sedikit meningkat setelahnya.



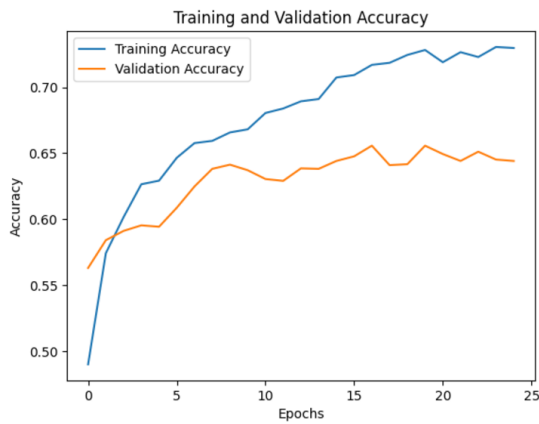
Gambar 5. Grafik pelatihan, validasi akurasi dan kerugian ResNet50V2

Pada Gambar 6, grafik bagian (a) menunjukkan akurasi untuk data pelatihan dan validasi yang terus meningkat seiring dengan jumlah *epoch*. Pada akurasi pelatihan, akurasi terus meningkat secara konsisten, ini menandakan bahwa model belajar lebih baik terhadap data pelatihan. Begitu juga dengan akurasi validasi meningkat pada awal pelatihan, tetapi mulai stabil atau bahkan sedikit menurun setelah beberapa *epoch*. Sedangkan pada grafik bagian (b) menunjukkan nilai *loss* pada data pelatihan dan validasi yang terus berubah selama pelatihan. *Loss* pelatihan terus menurun secara konsisten, yang merupakan indikator bahwa model berhasil meminimalkan *error* terhadap data pelatihan. *Loss* validasi menurun pada awal pelatihan tetapi kemudian mulai meningkat atau stabil setelah beberapa *epoch*, menunjukkan bahwa model mulai kehilangan kemampuan untuk generalisasi pada data validasi.

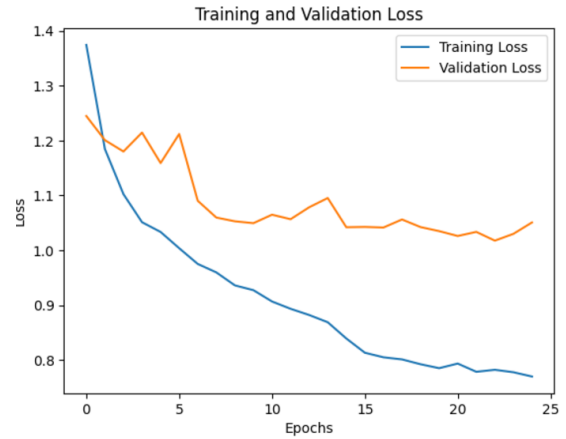
Pada Gambar 7, grafik bagian (a) menunjukkan bahwa akurasi pelatihan meningkat secara konsisten dan stabil hingga mendekati angka 0.80. Ini menunjukkan bahwa model berhasil belajar dari data pelatihan. Begitupun pada akurasi validasi yang meningkat di awal pelatihan, tetapi cenderung mendatar setelah beberapa *epoch* terakhir, dan stabil di sekitar angka 0.70. Sedangkan pada gambar (b) *loss* pelatihan terus menurun secara konsisten, menunjukkan bahwa model semakin baik meminimalkan error pada data pelatihan. *Loss* validasi menurun pada awal pelatihan, namun mulai stabil dan sedikit meningkat pada *epoch* terakhir. Ini adalah tanda bahwa model mulai kehilangan kemampuan generalisasi terhadap data validasi.

3.2 3.2 Classification Report/ Laporan Klasifikasi

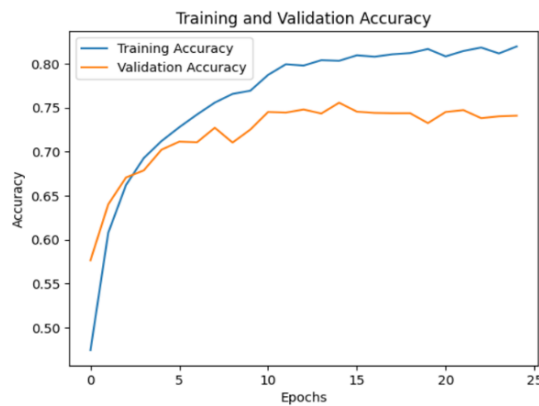
Gambar 8 merupakan laporan klasifikasi (*classification report*) dari tiga model CNN yaitu ResNet50V2, InceptionV3, dan DenseNet121, berdasarkan metrik *precision*, *recall*, *f1-score*, dan akurasi pada masing-masing kelas emosi (*Angry*, *Happy*, *Neutral*, *Sad*, dan *Surprise*). Pada *InceptionV3*, model ini memiliki kinerja baik untuk kelas *Happy* (*precision* 0.77, *recall* 0.81, *f1-score* 0.79), namun performa rendah pada kelas *Sur-*



(a). Pelatihan dan validasi akurasi



(b). Pelatihan dan validasi kerugian

Gambar 6. Grafik pelatihan, validasi akurasi dan kerugian InceptionV3

(a). Pelatihan dan validasi akurasi



(b). Pelatihan dan validasi kerugian

Gambar 7. Grafik pelatihan, validasi akurasi dan kerugian DenseNet121

prise ($f1$ -score 0.33) dan *Angry* ($f1$ -score 0.45). Akurasi keseluruhan model adalah 65%, dengan rata-rata tertimbang (*weighted average*) $f1$ -score sebesar 0.64. Pada DenseNet121 menunjukkan peningkatan signifikan, terutama pada kelas *Happy* (precision 0.90, recall 0.91, $f1$ -score 0.91) dan *Neutral* ($f1$ -score 0.71). Namun, kelas *Surprise* masih menjadi kelemahan, meskipun ada sedikit perbaikan dibandingkan InceptionV3 ($f1$ -score 0.62). Akurasi totalnya adalah 75%, dengan *weighted average* $f1$ -score mencapai 0.75. Sedangkan pada model ResNet50V2, model ini menjadi model dengan performa terbaik di antara ketiganya. Model ini mencapai hasil unggul pada kelas *Happy* (precision 0.92, recall 0.91, $f1$ -score 0.91) dan *Surprise* ($f1$ -score 0.76). Rata-rata tertimbang $f1$ -score model ini adalah 0.76, dan akurasi keseluruhan mencapai 79%. Model ini juga menunjukkan keseimbangan yang lebih baik di semua kelas, meskipun kelas *Angry* masih memiliki kelemahan ($f1$ -score 0.68). Secara keseluruhan, ResNet50V2 menunjukkan performa paling handal, diikuti oleh DenseNet121, sedangkan InceptionV3 memiliki performa paling rendah untuk klasifikasi ekspresi wajah.

3.3 Confusion Matrix

Pada Gambar 9, Confusion matrix dari ketiga model CNN (ResNet50V2, InceptionV3, dan DenseNet121) menunjukkan performa masing-masing dalam mengklasifikasikan lima kelas emosi (*Angry*, *Happy*, *Neutral*, *Sad*, dan *Surprise*). InceptionV3 memiliki performa yang baik dalam mengenali kelas *Happy* dan *Neutral* tetapi lemah pada kelas *Surprise*. DenseNet121 menunjukkan akurasi tinggi untuk kelas *Happy* dan *Neutral*, namun masih mengalami kesulitan dalam mengenali *Surprise*. ResNet50V2 adalah model dengan performa terbaik, unggul dalam mengenali kelas *Surprise* dan menunjukkan hasil yang seimbang pada kelas lainnya. Secara keseluruhan, semua model memiliki kecenderungan salah prediksi pada kelas *Neutral* yang sering diklasifikasikan sebagai *Sad*.

Classification Report:				
	precision	recall	f1-score	support
Angry	0.80	0.59	0.68	262
Happy	0.90	0.92	0.91	748
Neutral	0.66	0.77	0.71	805
Sad	0.71	0.70	0.70	786
Surprise	0.86	0.62	0.72	246
accuracy			0.76	2847
macro avg	0.79	0.72	0.75	2847
weighted avg	0.77	0.76	0.76	2847

(a). Precision Recall ResNet50V2

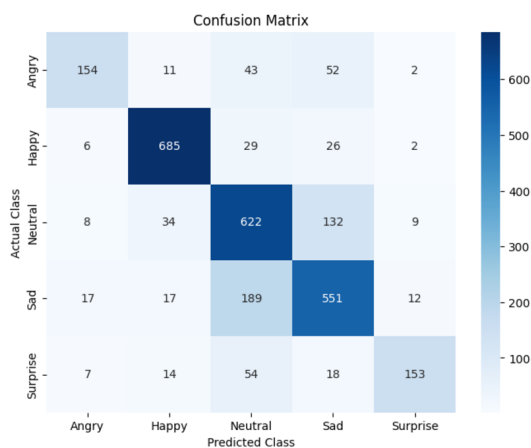
Classification Report:				
	precision	recall	f1-score	support
Angry	0.63	0.35	0.45	262
Happy	0.77	0.81	0.79	748
Neutral	0.58	0.69	0.63	805
Sad	0.61	0.67	0.64	786
Surprise	0.77	0.33	0.46	246
accuracy			0.65	2847
macro avg	0.67	0.57	0.60	2847
weighted avg	0.66	0.65	0.64	2847

(b). Precision Recall InceptionV3

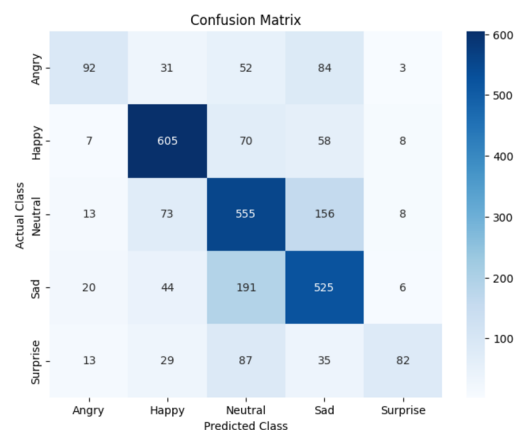
Classification Report:				
	precision	recall	f1-score	support
Angry	0.72	0.54	0.62	262
Happy	0.90	0.91	0.91	748
Neutral	0.64	0.81	0.71	805
Sad	0.73	0.69	0.71	786
Surprise	0.92	0.46	0.62	246
accuracy			0.75	2847
macro avg	0.78	0.68	0.71	2847
weighted avg	0.76	0.75	0.75	2847

(c). Precision Recall DenseNet121

Gambar 8. Hasil Klasifikasi ResNet50V2, InceptionV3, dan DenseNet121



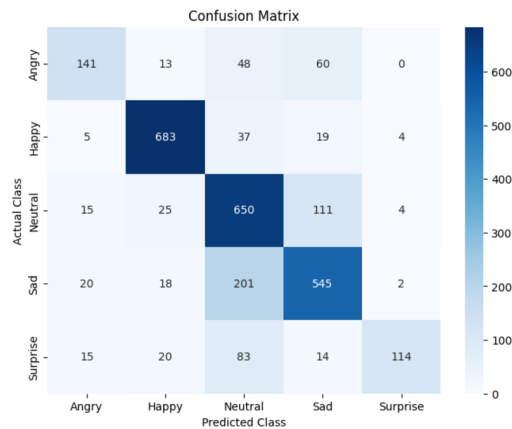
(a). Confusion Matrix ResNet50V2



(b). Confusion Matrix InceptionV3

3.4 Hasil Klasifikasi

Hasil analisis menunjukkan bahwa ResNet50V2 memiliki performa yang lebih baik dibandingkan kedua model lainnya, diikuti oleh DenseNet121, sedangkan InceptionV3 menunjukkan performa paling rendah untuk klasifikasi ekspresi wajah. DenseNet121, meskipun memiliki konektivitas yang baik antar lapisan, masih kalah optimal dibandingkan ResNet50V2, sedangkan InceptionV3 cenderung mengalami kesulitan



(c). Confusion Matrix DenseNet121

Gambar 9. Hasil Klasifikasi ResNet50V2, InceptionV3, dan DenseNet121

dalam menggeneralisasi fitur untuk dataset ini karena arsitekturnya yang lebih terfragmentasi.

4. Pembahasan

Evaluasi terhadap tiga model CNN, yaitu InceptionV3, DenseNet121, dan ResNet50V2, menunjukkan bahwa ResNet50V2 memiliki hasil yang baik dalam klasifikasi emosi. ResNet50V2 memberikan performa terbaik dengan akurasi total 79%, dengan *gap* kecil antara akurasi pelatihan dan validasi serta tren loss yang konsisten menurun. Keunggulan ini berasal dari arsitektur residual yang memungkinkan aliran gradien lebih lancar dan fitur *skip connection* yang membantu model mempelajari pola kompleks tanpa *overfitting*. DenseNet121 menunjukkan performa yang cukup baik tetapi kurang optimal dalam mengenali emosi seperti *Surprise*. Hal ini dapat disebabkan oleh koneksi padat antar lapisan yang kadang mengganggu fokus pada fitur relevan. Meskipun efisien dalam penggunaan parameter, model ini membutuhkan *fine-tuning* lebih lanjut atau *augmentasi* data untuk meningkatkan kinerja. Sedangkan InceptionV3 mengalami masalah *overfitting*, terlihat dari akurasi validasi yang stagnan meski akurasi pelatihan meningkat. Kompleksitas arsitektur *multi-kernel* pada model ini cenderung membuatnya terlalu sensitif terhadap data pelatihan, terutama pada dataset yang tidak seimbang atau memiliki variasi tinggi. Regularisasi tambahan atau data augmentation dapat membantu mengatasi masalah ini.

Analisis *confusion* matrix dan laporan klasifikasi memperkuat temuan ini, di mana ResNet50V2 menunjukkan performa yang lebih unggul dalam mengenali semua kelas dengan nilai *precision*, *recall*, dan *f1-score* yang konsisten dibandingkan model lainnya. Sebaliknya, DenseNet121 dan InceptionV3 memiliki performa yang lebih rendah, khususnya pada kelas dengan variasi pola yang lebih kompleks seperti *Surprise*. Walaupun dalam penelitian sebelumnya InceptionV3 dan DenseNet121 memiliki performa yang baik untuk menangani kasus-kasus tertentu, namun pada kasus ini Inception dan DenseNet kurang cocok dan memperoleh nilai yang lebih rendah.

Faktor utama yang memengaruhi perbedaan performa di antara model-model ini adalah arsitektur, kemampuan generalisasi, dan karakteristik dataset. ResNet50V2 lebih stabil dalam menangkap pola kompleks dibandingkan DenseNet121 dan InceptionV3 yang lebih terpengaruh oleh variasi data. Untuk meningkatkan performa di masa depan, disarankan melakukan *balancing* dataset, *fine-tuning hyperparameter*, *ensemble learning*, dan *augmentasi* data. Secara keseluruhan, ResNet50V2 menjadi pilihan terbaik karena kemampuannya menghasilkan prediksi yang akurat dan seimbang, meskipun optimalisasi lebih lanjut tetap diperlukan.

5. Simpulan

Berdasarkan hasil evaluasi, model ResNet50V2 menunjukkan performa terbaik dengan akurasi tertinggi sebesar 79%, *macro average F1-score* sebesar 0.76, dan *weighted average F1-score* sebesar 0.75. Model ini ung-

gul dalam mengimbangi kelas-kelas dengan jumlah data yang tidak merata, terutama untuk kelas dengan jumlah data dominan seperti Happy dan Neutral. Kemudian untuk model DenseNet121 berada di posisi kedua dengan akurasi 75%, diikuti oleh InceptionV3 dengan akurasi terendah, yaitu 65%. Walaupun InceptionV3 menghasilkan hasil yang baik untuk kelas tertentu (Happy), performa keseluruhannya lebih rendah karena rendahnya recall pada beberapa kelas minoritas seperti Angry dan Surprise. Dari hasil pelatihan dan evaluasi model, terlihat bahwa fine-tuning parameter seperti learning rate dan jumlah *epoch* sangat berpengaruh terhadap hasil akhir. Proses pelatihan yang stabil dan strategi penurunan learning rate memungkinkan model mencapai akurasi yang optimal. Dari hasil penelitian diperoleh bahwa model ResNet50V2 dapat direkomendasikan untuk klasifikasi ekspresi wajah pada dataset ini, dengan potensi untuk dikembangkan lebih lanjut.

Implikasi praktis dari penelitian ini mencakup potensi penerapan model ResNet50V2 dalam sistem pengenalan ekspresi wajah berbasis kamera untuk aplikasi seperti pengawasan, interaksi manusia-mesin, atau analisis perilaku pelanggan. Penelitian masa depan dapat berfokus pada optimasi *hyperparameter* tambahan, penggunaan dataset yang lebih besar dan beragam, serta menguji keandalan model terhadap data yang mengandung noise, seperti gambar buram, kondisi pencahayaan ekstrem, atau sebagian wajah yang terhalang, untuk menilai performa dalam kondisi dunia nyata.

Pustaka

- [1] P. A. Nugroho, I. Fenriana, and R. Arijanto, "Implementasi deep learning menggunakan convolutional neural network (cnn) pada ekspresi manusia," *Algor*, vol. 2, no. 1, pp. 12–21, 2020.
- [2] W. W. Kusuma, R. R. Isnanto, A. Fauzi, and P. Korespondensi, "Densenet121 menggunakan kerangka kerja tensorflow untuk deteksi jenis hewan," *Jurnal Teknik Komputer*, vol. 1, no. 4, pp. 141–147, 2023.
- [3] M. A. Mutasodirin and F. M. Falakh, "Efficient weather classification using densenet and efficientnet," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 9, no. 2, pp. 173–179, 2024.
- [4] A. K. Putra, H. Bunyamin, and K. M. J. drg Surya Sumantri No, "Pengenalan simbol matematika dengan metode convolutional neural network (cnn)," *Jurnal Strategi*, vol. 2, no. November, p. 426, 2020.
- [5] M. K. Insani and D. B. Santoso, "Perbandingan kinerja model pre-trained cnn (vgg16, resnet, dan inceptionv3) untuk aplikasi pengenalan wajah pada sistem absensi karyawan," *Jurnal Indonesia: Manajemen Informatika dan Komunikasi (JIMIK)*, vol. 5, no. 3, pp. 2612–2622, 2024.
- [6] A. A. SHELEMO, "No title," *Nucl. Phys.*, vol. 13, no. 1, pp. 104–116, 2023.
- [7] Y. F. Riti, Y. Wahyuningsih, J. Roosandriantini, and P. W. Siswanto, "Perbandingan algoritma convolutional neural network dan capsule network dalam klasifikasi jenis rumah adat," *Teknika*, vol. 12, no. 3, pp. 243–251, 2023.
- [8] M. A. Izzulhaq and Alamsyah, "Penerapan algoritma convolutional neural network arsitektur resnet50v2 untuk mengidentifikasi penyakit pneumonia," *Indonesian Journal of Mathematics and Natural Sciences*, vol. 47, no. 1, pp. 12–22, 2024.
- [9] U. Kulsum and A. Cherid, "Penerapan convolutional neural network pada klasifikasi tanaman menggunakan resnet50," *Simkom*, vol. 8, no. 2, pp. 221–228, 2023.
- [10] G. E. P. Purba, S. H. Wijoyo, and N. Y. Setiawan, "Pengaruh transfer learning resnet dan densenet terhadap performa klasifikasi ekspresi wajah menggunakan dataset fer-2013," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 7, 2024, accessed: Feb. 03, 2025. [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/13931>
- [11] J. PARDEDE and S. S. KLEB, "Face race classification using resnet-152 and densenet-121," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 12, no. 3, p. 798, 2024.

- [12] K. R. R. Wardani, H. Suryalim, V. J. L. Engel, and H. Christian, "Analisis pemilihan optimizer dalam arsitektur convolution neural network vgg16 dan inception untuk sistem pengenalan wajah," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 9, no. 2, pp. 186–194, 2023.
- [13] S. Kapadnis, "Facial emotion recognition image dataset," 2024, accessed: Oct. 03, 2024. [Online]. Available: <https://www.kaggle.com/datasets/sujaykapadnis/emotion-recognition-dataset>
- [14] T. S. Winanto, C. Rozikin, and A. Jamaludin, "Analisa performa arsitektur transfer learning untuk mengidentifikasi penyakit daun pada tanaman pangan," *Journal of Applied Informatics and Computing*, vol. 7, no. 1, pp. 68–81, 2023.
- [15] U. Khultsum and G. Taufik, "Komparasi kinerja densenet 121 dan mobilenet untuk klasifikasi citra penyakit daun kentang," *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, p. 558, 2023.