

Perbandingan Kinerja Model *Pre-Trained* CNN (VGG16, RESNET, dan INCEPTIONV3) untuk Aplikasi Pengenalan Wajah pada Sistem Absensi Karyawan

Muhammad Khatama Insani ¹, Dwi Budi Santoso ^{2*}

^{1,2*} Fakultas Teknologi Informasi dan Industri, Universitas Stikubank, Kota Semarang, Provinsi Jawa Tengah, Indonesia.

Email: muhammadkhatamainsani@mhs.unisbank.ac.id ¹, dbs@edu.unisbank.ac.id ^{2*}

Histori Artikel:

Dikirim 26 Mei 2024; *Diterima dalam bentuk revisi* 25 Juni 2024; *Diterima* 17 Juli 2024; *Diterbitkan* 20 September 2024. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STM IK Indonesia Banda Aceh.

Abstrak

Pengenalan wajah telah menjadi teknologi kunci dalam meningkatkan efisiensi dan keamanan sistem absensi karyawan modern. Studi ini membandingkan kinerja tiga model Convolutional Neural Network (CNN) pre-trained - VGG16, ResNet50, dan InceptionV3 - dalam konteks pengenalan wajah untuk sistem absensi karyawan. Penelitian mengevaluasi akurasi, konsistensi, dan generalisasi model pada dataset wajah karyawan, menggunakan metode pengukuran akurasi prediksi, confusion matrix, dan classification report. Hasil menunjukkan InceptionV3 memiliki performa terbaik secara keseluruhan, dengan konsistensi dan tingkat kepercayaan tinggi, mencapai akurasi hingga 99% pada data uji. ResNet50 menunjukkan kinerja konsisten dalam beberapa kasus namun membutuhkan fine-tuning lebih lanjut, sementara VGG16 menunjukkan performa terburuk. Temuan ini memiliki implikasi praktis signifikan bagi industri, merekomendasikan penggunaan InceptionV3 untuk implementasi sistem absensi berbasis pengenalan wajah yang andal, dengan pertimbangan penggunaan threshold kepercayaan untuk mengoptimalkan akurasi. Penelitian ini juga menyoroti pentingnya optimisasi lebih lanjut, termasuk fine-tuning hyperparameter dan strategi augmentasi data yang lebih canggih, untuk meningkatkan performa sistem dalam berbagai kondisi lingkungan kerja.

Kata Kunci: Pengenalan Wajah; Convolutional Neural Network (CNN); Sistem Absensi Karyawan; Perbandingan Model.

Abstract

Face recognition has become a key technology in improving the efficiency and security of modern employee attendance systems. This study compares the performance of three pre-trained Convolutional Neural Network (CNN) models - VGG16, ResNet50, and InceptionV3 - in the context of face recognition for employee attendance systems. The study evaluated the accuracy, consistency, and generalization of the models on a dataset of employee faces, using prediction accuracy, confusion matrix, and classification report measurement methods. Results showed InceptionV3 performed best overall, with high consistency and confidence, achieving up to 99% accuracy on the test data. ResNet50 showed consistent performance in some cases but required further fine-tuning, while VGG16 showed the worst performance. These findings have significant practical implications for the industry, recommending the use of InceptionV3 for the implementation of reliable face recognition-based attendance systems, with consideration of the use of confidence thresholds to optimize accuracy. This research also highlights the importance of further optimization, including hyperparameter fine-tuning and more sophisticated data augmentation strategies, to improve system performance under various work environment conditions.

Keyword: Facial Recognition; Convolutional Neural Network (CNN); Employee Attendance System; Model Comparison.

1. Pendahuluan

Pengenalan wajah telah menjadi teknologi yang semakin krusial dalam berbagai aplikasi, termasuk sistem absensi karyawan, dengan peningkatan signifikan dalam efisiensi operasional dan keamanan (Hartiwi *et al.*, 2020). Dalam era kecerdasan buatan (*artificial intelligence*) dan pembelajaran mesin (*machine learning*), model *Convolutional Neural Network* (CNN) yang telah dilatih sebelumnya (*pre-trained*) menawarkan solusi yang efektif untuk tugas pengenalan wajah yang kompleks (Saragih & To, 2022). Teknologi ini memanfaatkan model-model yang telah dilatih pada dataset berskala besar untuk mengekstraksi fitur penting dari gambar wajah, memungkinkan identifikasi individu dengan tingkat akurasi yang tinggi.

Kemajuan dalam teknologi pengenalan wajah, terutama melalui arsitektur *deep learning*, telah memperlihatkan peningkatan substansial dalam hal akurasi dan efisiensi. Gwyn *et al.* (2021) melakukan studi komparatif mengenai penggunaan arsitektur *deep learning* populer untuk pengenalan wajah dan menemukan bahwa model CNN memiliki keunggulan dalam menangani variasi pose dan kondisi pencahayaan yang berbeda. Penelitian lain oleh Kumar *et al.* (2023) menunjukkan efektivitas kombinasi CNN dengan jaringan *Siamese* dalam meningkatkan akurasi pengenalan wajah, terutama dalam kasus-kasus dengan keterbatasan dataset, yang seringkali menjadi tantangan signifikan dalam implementasi praktis.

Studi ini difokuskan untuk membandingkan kinerja tiga model CNN *pre-trained* yang banyak digunakan: *VGG16* (Simonyan & Zisserman, 2014), *ResNet50* (He *et al.*, 2015), dan *InceptionV3* (Szegedy *et al.*, 2015), dalam konteks sistem absensi karyawan. Ketiga model ini dipilih berdasarkan efektivitasnya dalam berbagai tugas pengenalan gambar dan perbedaan dalam arsitektur, yang memungkinkan analisis lebih rinci terhadap kelebihan dan kekurangan masing-masing (Ko *et al.*, 2020). *VGG16* dikenal dengan struktur arsitekturnya yang sederhana namun mendalam, yang memungkinkan ekstraksi fitur secara menyeluruh. Namun, model ini cenderung memiliki jumlah parameter yang besar, mengakibatkan waktu inferensi yang lebih lama dan kebutuhan komputasi yang tinggi.

Sementara itu, *ResNet50* memperkenalkan konsep koneksi residual yang membantu mengatasi masalah *vanishing gradient* pada jaringan yang sangat dalam, membuat pelatihan model lebih efisien dan stabil. Arsitektur ini memungkinkan *ResNet50* untuk mencapai performa yang konsisten, terutama dalam kondisi data yang kompleks. Di sisi lain, *InceptionV3* mengadopsi modul *inception*, yang mendukung pemrosesan multi-skala dan meningkatkan efisiensi penggunaan parameter, meskipun implementasinya lebih kompleks (Bimorogo, 2020). Kompleksitas ini sering kali menjadi tantangan dalam aplikasi nyata, terutama di lingkungan kerja yang beragam.

Penelitian ini mengevaluasi kemampuan setiap model berdasarkan akurasi, konsistensi, dan kemampuan generalisasi terhadap dataset wajah karyawan. Evaluasi dilakukan melalui pengukuran akurasi prediksi, tingkat kesalahan, dan performa generalisasi model pada data yang belum pernah dilihat sebelumnya. Teknik validasi silang juga diterapkan untuk memastikan keandalan (*robustness*) hasil evaluasi, mengacu pada praktik terbaik yang telah diidentifikasi oleh Ketut Adi Wirayasa *et al.* (2021) dalam studi komparatif mereka mengenai model CNN.

Berbeda dengan penelitian sebelumnya yang umumnya hanya berfokus pada satu arsitektur atau membandingkan performa pada dataset standar, studi ini memberikan kontribusi unik melalui pendekatan yang lebih menyeluruh. Pertama, penelitian ini melakukan analisis mendalam terhadap tiga arsitektur CNN populer dalam konteks khusus sistem absensi karyawan. Kedua, evaluasi kinerja model tidak hanya didasarkan pada akurasi, tetapi juga melibatkan analisis efisiensi komputasi dan kemampuan model untuk beradaptasi dengan variasi kondisi lingkungan kerja. Terakhir, teknik augmentasi data yang disesuaikan dengan karakteristik dataset wajah karyawan digunakan untuk meningkatkan daya tahan model terhadap berbagai kondisi pengambilan gambar, seperti pencahayaan yang kurang optimal atau perubahan ekspresi wajah.

Diharapkan, hasil dari penelitian ini dapat menjadi panduan berharga bagi pengembang dan praktisi di bidang visi komputer dan manajemen sumber daya manusia. Lebih lanjut, temuan studi ini juga dapat berfungsi sebagai landasan bagi pengembangan strategi optimasi yang lebih lanjut untuk sistem pengenalan wajah dalam aplikasi absensi karyawan. Dengan demikian, penelitian ini tidak hanya memperkaya literatur ilmiah tetapi juga memberikan implikasi praktis yang signifikan dalam meningkatkan efisiensi dan keamanan sistem absensi berbasis pengenalan wajah di berbagai industri.

2. Metode Penelitian

Metode penelitian ini dimulai dengan pengumpulan dan persiapan dataset yang dilakukan secara sistematis. Dataset yang digunakan terdiri dari rekaman absensi berupa foto wajah karyawan dari sebuah perusahaan di Semarang. Foto-foto ini mencakup lima individu, yaitu Azizul, Anwar, Eldra, Fajar, dan Tama. Langkah awal yang dilakukan adalah pembersihan data (*data cleaning*), di mana gambar wajah diseleksi dengan ketat untuk memastikan kualitas yang optimal. Gambar-gambar yang buram, terlalu kecil, atau yang mengandung objek selain wajah dihapus dari dataset untuk menjaga konsistensi dan akurasi dalam proses pelatihan model. Selanjutnya, setiap gambar yang lolos seleksi diberi label sesuai dengan identitas individu masing-masing, memastikan bahwa data siap digunakan untuk proses pelatihan (*training*) dan evaluasi model.

Tahap berikutnya melibatkan teknik augmentasi data (*data augmentation*) untuk memperkaya variasi dalam dataset. Teknik augmentasi yang diterapkan meliputi rotasi, flipping, zooming, cropping, dan penyesuaian kecerahan, yang bertujuan untuk mensimulasikan berbagai kondisi pengambilan gambar yang mungkin terjadi dalam situasi nyata. Dengan augmentasi ini, model menjadi lebih adaptif terhadap variasi input yang berbeda. Setelah proses augmentasi, dataset dibagi menjadi tiga bagian: data latih (*training set*) sebesar 70%, data validasi (*validation set*) sebesar 20%, dan data uji (*test set*) sebesar 10%. Pembagian ini dirancang untuk memastikan bahwa model tidak hanya belajar dari data yang ada, tetapi juga dapat memprediksi dengan baik pada data yang belum pernah dilihat sebelumnya.

Untuk pengembangan model, arsitektur *Convolutional Neural Network* (CNN) dibangun menggunakan framework *TensorFlow*. Model ini terdiri dari beberapa lapisan, termasuk lapisan konvolusi (*convolutional layers*) yang berfungsi untuk mengekstraksi fitur-fitur penting dari gambar wajah, lapisan pooling (*pooling layers*) untuk mengurangi dimensi spasial fitur dan meningkatkan efisiensi komputasi, serta lapisan perataan (*flatten layer*) yang mengubah fitur menjadi vektor satu dimensi. Selain itu, model ini dilengkapi dengan lapisan terhubung penuh (*fully connected layers*) yang bertugas melakukan klasifikasi berdasarkan fitur yang telah diekstraksi. Fungsi aktivasi yang digunakan adalah *Rectified Linear Unit* (ReLU) pada lapisan konvolusi dan terhubung penuh, yang meningkatkan kemampuan model dalam mengenali pola-pola kompleks. Fungsi kerugian (*loss function*) yang digunakan adalah *categorical cross-entropy*, yang bertujuan untuk mengukur tingkat kesalahan dalam proses klasifikasi, sementara algoritma optimasi *Adam* digunakan untuk memperbarui bobot model selama pelatihan.

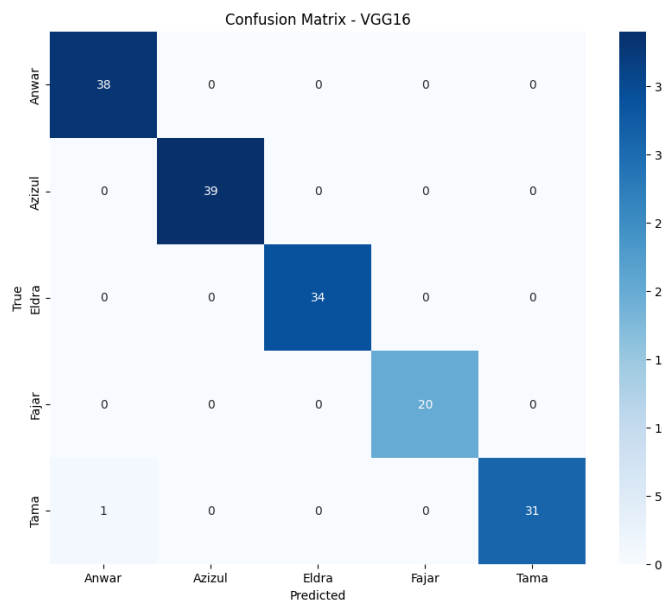
Proses pelatihan model dilakukan menggunakan data latih dengan pengaturan *epoch* sebanyak 5 dan ukuran *batch* sebesar 32. Validasi model dilakukan secara berkala menggunakan data validasi untuk memantau performa model dan mencegah *overfitting*, yaitu kondisi di mana model terlalu menyesuaikan dengan data latih dan gagal dalam generalisasi pada data baru. Setelah proses pelatihan selesai, model diuji dengan menggunakan data uji untuk mengevaluasi kemampuan generalisasi model terhadap data baru. Metrik evaluasi yang digunakan mencakup akurasi, presisi, *recall*, dan *F1-score*, yang memberikan gambaran komprehensif mengenai kinerja model dalam klasifikasi wajah. Evaluasi ini penting untuk memastikan bahwa model yang dikembangkan tidak hanya akurat, tetapi juga robust dalam berbagai kondisi operasional.

3. Hasil dan Pembahasan

3.1 Hasil

Evaluasi performa model menjadi bagian penting dalam menentukan efektivitas model *Convolutional Neural Network* (CNN) yang digunakan untuk sistem pengenalan wajah pada absensi karyawan. Dalam studi ini, kami mengevaluasi tiga model *pre-trained CNN* yang populer, yaitu

VGG16, *ResNet50*, dan *InceptionV3*, menggunakan *confusion matrix* dan *classification report* untuk menganalisis kemampuan klasifikasi masing-masing model.



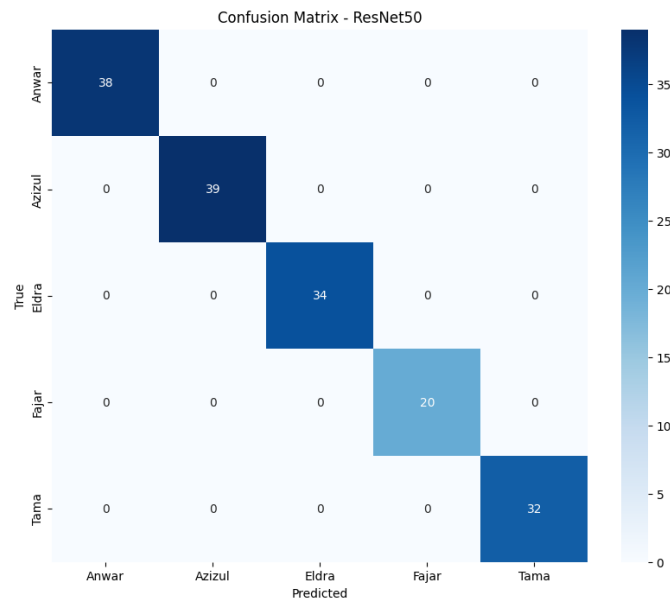
Gambar 1. *Confusion Matrix* untuk Model VGG16

	precision	recall	f1-score	support
Anwar	0.97	1.00	0.99	38
Azizul	1.00	1.00	1.00	39
Eldra	1.00	1.00	1.00	34
Fajar	1.00	1.00	1.00	20
Tama	1.00	0.97	0.98	32
accuracy			0.99	163
macro avg	0.99	0.99	0.99	163
weighted avg	0.99	0.99	0.99	163

Gambar 2. *Classification Report* untuk Model VGG16

Pada model *VGG16*, *confusion matrix* menunjukkan bahwa model ini mampu mengklasifikasikan kelima kelas dengan cukup baik, meskipun terdapat beberapa kelemahan. Berdasarkan *classification report*, model ini mencapai akurasi keseluruhan sebesar 99%, yang menunjukkan performa yang sangat tinggi. *Precision*, *recall*, dan *F1-score* untuk semua kelas berada pada rentang 0,97 hingga 1,00, mengindikasikan performa yang konsisten dalam sebagian besar klasifikasi. Namun, terdapat sedikit penurunan pada *recall* untuk kelas Tama (0,97), yang mengindikasikan adanya beberapa *false negatives*, sehingga kinerja model tidak sempurna pada kelas tertentu.

Untuk model *ResNet50*, hasil evaluasi menunjukkan performa yang hampir sempurna. *Confusion matrix* dari model ini mengindikasikan tidak adanya kesalahan dalam klasifikasi, sementara *classification report* menunjukkan akurasi 100% untuk semua kelas. *Precision*, *recall*, dan *F1-score* mencapai nilai maksimum 1,00, menandakan bahwa model ini tidak hanya akurat, tetapi juga sangat konsisten dalam klasifikasinya tanpa adanya kesalahan *false positives* atau *false negatives*. Hal ini menempatkan *ResNet50* sebagai model dengan performa terbaik di antara ketiga model yang dievaluasi.



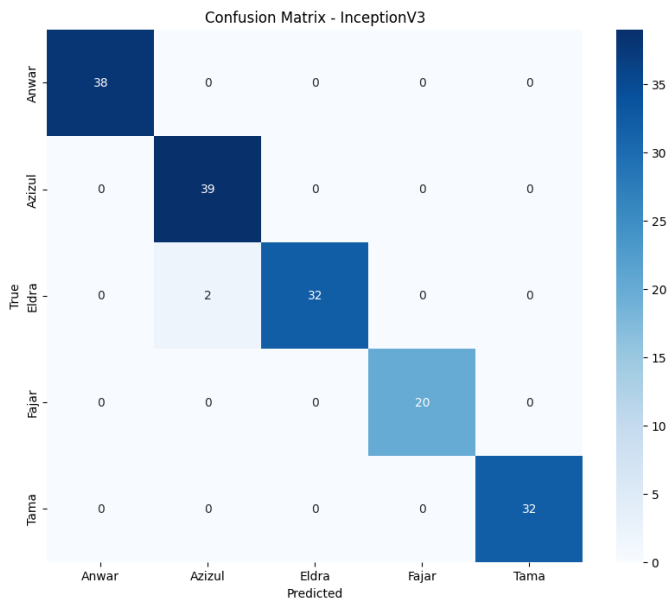
Gambar 3. *Confusion Matrix* untuk Model ResNet50

Classification Report - ResNet50

	precision	recall	f1-score	support
Anwar	1.00	1.00	1.00	38
Azizul	1.00	1.00	1.00	39
Eldra	1.00	1.00	1.00	34
Fajar	1.00	1.00	1.00	20
Tama	1.00	1.00	1.00	32
accuracy			1.00	163
macro avg	1.00	1.00	1.00	163
weighted avg	1.00	1.00	1.00	163

Gambar 4. *Classification Report* untuk Model ResNet50

Model *InceptionV3* juga menunjukkan performa yang sangat baik, dengan akurasi keseluruhan mencapai 99%. *Confusion matrix* menunjukkan bahwa model ini berhasil dalam sebagian besar klasifikasi, namun terdapat sedikit penurunan pada *precision* untuk kelas Azizul (0,95) dan *recall* untuk kelas Eldra (0,94). Hal ini mengindikasikan adanya beberapa *false positives* dan *false negatives* dalam klasifikasi, meskipun secara keseluruhan model ini tetap dapat diandalkan. Dalam membandingkan ketiga model berdasarkan *confusion matrix* dan *classification report*, *ResNet50* menonjol dengan akurasi sempurna sebesar 100%, menjadikannya model dengan performa terbaik. Meskipun *VGG16* dan *InceptionV3* juga menunjukkan akurasi tinggi sebesar 99%, perbedaan performa antar model sangat kecil, sehingga ketiganya tetap layak diimplementasikan dalam sistem absensi berbasis pengenalan wajah.



Gambar 5. *Confusion Matrix* untuk Model InceptionV3

Classification Report - InceptionV3

	precision	recall	f1-score	support
Anwar	1.00	1.00	1.00	38
Azizul	0.95	1.00	0.97	39
Eldra	1.00	0.94	0.97	34
Fajar	1.00	1.00	1.00	20
Tama	1.00	1.00	1.00	32
accuracy			0.99	163
macro avg	0.99	0.99	0.99	163
weighted avg	0.99	0.99	0.99	163

Gambar 6. *Classification Report* untuk Model InceptionV3



Gambar 7. Sampel foto kelas Tama dan hasil prediksi model

Evaluasi prediksi lebih lanjut dilakukan dengan menguji ketiga model *pre-trained CNN* pada tiga sampel foto baru yang belum pernah dilihat oleh model sebelumnya. Pada sampel pertama (foto baru kelas Tama), *InceptionV3* menunjukkan performa yang sangat baik dengan mengidentifikasi Tama dengan tingkat kepercayaan 98,74%. Sebaliknya, *VGG16* dan *ResNet50* memberikan prediksi yang salah dengan tingkat kepercayaan yang lebih rendah, masing-masing sebesar 32,62% dan 59,02%.



Gambar 8. Sampel foto kelas Anwar dan hasil prediksi mode

Pada sampel kedua (foto baru kelas Anwar), *InceptionV3* kembali berhasil mengidentifikasi Anwar meskipun dengan tingkat kepercayaan yang moderat sebesar 52,99%. Kedua model lainnya, *VGG16* dan *ResNet50*, memberikan prediksi yang salah dengan tingkat kepercayaan masing-masing sebesar 31,50% dan 59,52%. Hasil ini mengindikasikan bahwa meskipun *InceptionV3* memiliki tingkat kepercayaan yang tidak terlalu tinggi, model ini tetap lebih andal dibandingkan model lainnya. Untuk sampel ketiga (foto baru kelas Azizul), baik *ResNet50* maupun *InceptionV3* berhasil mengidentifikasi Azizul dengan tingkat kepercayaan masing-masing 60,12% dan 68,64%, dengan *InceptionV3* menunjukkan sedikit keunggulan. Sementara itu, *VGG16* kembali gagal dengan memberikan prediksi yang salah dan tingkat kepercayaan yang rendah (31,09%).



Gambar 9. Sampel foto kelas Azizul dan hasil prediksi model

Analisis mendalam dari hasil evaluasi ini menunjukkan bahwa *InceptionV3* memiliki performa terbaik dalam hal akurasi tertinggi dan konsistensi yang baik dalam pengujian pada data baru. Meskipun *ResNet50* menunjukkan performa moderat dengan akurasi yang baik untuk kelas Azizul, model ini kurang konsisten dalam mengidentifikasi kelas lainnya. *VGG16*, di sisi lain, menunjukkan performa yang kurang memuaskan dengan prediksi yang sering salah dan tingkat kepercayaan yang rendah. Beberapa faktor yang mempengaruhi performa model ini telah diidentifikasi. Dari sisi arsitektur model, *InceptionV3* menggunakan modul *inception* yang kompleks, memungkinkan ekstraksi fitur yang lebih efektif dan berkontribusi pada performa unggulnya. *ResNet50* dengan koneksi residualnya membantu mengatasi masalah *vanishing gradient*, tetapi mungkin kurang optimal untuk dataset yang digunakan dalam penelitian ini. Sementara itu, *VGG16*, dengan arsitektur yang lebih sederhana, mungkin kurang mampu menangkap kompleksitas fitur wajah dalam dataset yang digunakan.

Faktor lain yang berpengaruh adalah *hyperparameter*, termasuk *learning rate*, ukuran *batch*, dan jumlah *epoch*, yang mungkin memerlukan penyesuaian lebih lanjut, terutama untuk *VGG16* dan *ResNet50*. Optimisasi *hyperparameter* dapat meningkatkan performa, terutama dalam hal generalisasi pada data baru. Selain itu, teknik augmentasi data yang digunakan mungkin lebih menguntungkan beberapa model dibandingkan yang lain. Penerapan augmentasi yang lebih beragam, seperti perubahan pencahayaan, rotasi, atau oklusi parsial, dapat meningkatkan *robustness* model terhadap berbagai kondisi. Perbandingan dengan metode lain juga menunjukkan bahwa pendekatan CNN yang dikembangkan dalam studi ini menawarkan peningkatan signifikan dibandingkan dengan metode pengenalan wajah tradisional seperti *Eigenfaces* atau *Fisherfaces*. Selain itu, model CNN lain seperti *MobileNet* atau *EfficientNet* dapat menjadi alternatif dalam mengevaluasi trade-off antara akurasi dan efisiensi komputasi. Penggunaan metode *ensemble*, seperti *voting* atau *weighted averaging*, juga dapat menghasilkan performa yang lebih baik dibandingkan penggunaan model individual. Kesimpulannya, evaluasi mendalam menunjukkan bahwa meskipun ketiga model CNN memiliki kelebihan masing-masing, *InceptionV3* dan *ResNet50* menonjol dalam aplikasi pengenalan wajah untuk absensi karyawan. Namun, optimisasi dan adaptasi lebih lanjut tetap diperlukan untuk mencapai performa yang maksimal dalam berbagai kondisi lingkungan kerja.

3.2 Pembahasan

Pembahasan ini mengkaji performa tiga model *pre-trained Convolutional Neural Network (CNN)*, yaitu *VGG16*, *ResNet50*, dan *InceptionV3*, dalam konteks sistem absensi berbasis pengenalan wajah. Berdasarkan evaluasi menggunakan *confusion matrix* dan *classification report*, ditemukan bahwa setiap model memiliki kelebihan dan kekurangan yang dapat berdampak pada efektivitas implementasi dalam aplikasi nyata. Model *ResNet50* menunjukkan kinerja paling unggul dengan akurasi sempurna sebesar 100% tanpa kesalahan klasifikasi (*false positives* atau *false negatives*). Keunggulan ini didukung oleh arsitektur residual yang memungkinkan *ResNet50* mengatasi masalah *vanishing gradient* pada jaringan yang sangat dalam, sebagaimana dijelaskan oleh He *et al.* (2015). Pendekatan ini memungkinkan model untuk melakukan pembelajaran lebih efisien dan mengurangi kesalahan klasifikasi yang sering terjadi pada jaringan konvensional. Keberhasilan *ResNet50* dalam menghasilkan prediksi yang sangat konsisten sejalan dengan temuan Gwyn *et al.* (2021a), yang menyatakan bahwa model dengan arsitektur residual menunjukkan performa yang superior dalam pengenalan wajah.

Sementara itu, *InceptionV3* mencapai akurasi keseluruhan sebesar 99%, dengan beberapa kesalahan pada *precision* dan *recall* untuk kelas tertentu. Kompleksitas arsitektur *InceptionV3*, yang mencakup modul *inception* untuk ekstraksi fitur multi-skala, memberikan fleksibilitas dan efisiensi dalam menangkap variasi fitur wajah. Namun, seperti yang diidentifikasi oleh Szegedy *et al.* (2015), kompleksitas ini dapat meningkatkan kesalahan prediksi dalam situasi tertentu. Meskipun demikian, performa *InceptionV3* tetap lebih unggul dibandingkan *VGG16* dalam menangani kondisi variatif, sesuai dengan hasil yang dilaporkan oleh Gwyn *et al.* (2021b). Model *VGG16*, meskipun memiliki arsitektur yang lebih sederhana dan jumlah parameter yang besar, menunjukkan akurasi tinggi sebesar 99%, namun kurang konsisten dalam klasifikasi. Model ini sering kali gagal menangkap kompleksitas

fitur yang diperlukan, terutama dalam kondisi pencahayaan atau pose yang bervariasi. Simonyan & Zisserman (2014) mencatat bahwa meskipun *VGG16* mampu mengekstraksi fitur mendalam, ketergantungan pada jumlah parameter yang besar membuat model ini kurang efisien dan lebih rentan terhadap overfitting, khususnya dalam tugas pengenalan wajah yang lebih menantang.

Pengujian prediksi pada foto baru menunjukkan bahwa *InceptionV3* mampu mempertahankan performa terbaiknya dengan tingkat kepercayaan tinggi dalam mengidentifikasi individu, meskipun pada beberapa kasus tingkat kepercayaannya hanya moderat. Sebaliknya, *ResNet50* dan *VGG16* sering kali memberikan prediksi yang salah, terutama pada data baru yang tidak termasuk dalam data pelatihan. Hasil ini menunjukkan kemampuan generalisasi *InceptionV3* yang lebih baik, sejalan dengan penelitian Ko *et al.* (2020), yang menekankan pentingnya arsitektur multi-skala dalam menangani variasi data dalam aplikasi pengenalan wajah. Hal ini juga sesuai dengan temuan Bimorogo (2020), yang menunjukkan bahwa arsitektur *Inception* unggul dalam menangani variasi pada data karena kemampuannya untuk melakukan ekstraksi fitur secara simultan pada skala yang berbeda. Sebaliknya, *ResNet50*, meskipun menunjukkan performa yang sangat baik dalam skenario data pelatihan, kurang adaptif terhadap variasi baru yang tidak terduga, sebagaimana juga diungkapkan oleh Ketut Adi Wirayasa *et al.* (2021).

Beberapa faktor kunci yang mempengaruhi performa model ini meliputi arsitektur model, *hyperparameter*, dan teknik augmentasi data yang digunakan. Arsitektur *InceptionV3* yang kompleks memungkinkan ekstraksi fitur yang lebih efektif, sedangkan arsitektur residual *ResNet50* membantu stabilitas pelatihan pada jaringan yang lebih dalam (He *et al.*, 2015). Sebaliknya, *VGG16* dengan arsitektur yang lebih sederhana kurang mampu menangkap variasi kompleks dalam data wajah, sebagaimana dicatat oleh Simonyan & Zisserman (2014). Pemilihan *hyperparameter* seperti *learning rate*, ukuran *batch*, dan jumlah *epoch* juga sangat berpengaruh terhadap performa model. Optimisasi *hyperparameter* dapat meningkatkan kemampuan model dalam generalisasi pada data baru, seperti yang diidentifikasi oleh Sulistya *et al.* (2023), yang menunjukkan bahwa penyesuaian parameter yang tepat dapat meningkatkan akurasi secara signifikan. Teknik augmentasi data yang digunakan dalam penelitian ini, seperti rotasi, flipping, dan penyesuaian kecerahan, memungkinkan model untuk lebih beradaptasi dengan kondisi variatif yang mungkin terjadi dalam lingkungan nyata. Augmentasi yang lebih kompleks, seperti oklusi atau variasi pencahayaan ekstrem, dapat lebih lanjut meningkatkan *robustness* model terhadap berbagai kondisi, sebagaimana diungkapkan oleh Kumar *et al.* (2023a), yang menunjukkan bahwa augmentasi data yang lebih kaya dapat mengurangi kesalahan klasifikasi dalam skenario dunia nyata.

Pendekatan *CNN* dalam penelitian ini menawarkan peningkatan performa signifikan dibandingkan dengan metode pengenalan wajah tradisional seperti *Eigenfaces* dan *Fisherfaces*, terutama dalam hal akurasi dan ketahanan terhadap perubahan pose dan pencahayaan (Kamencay *et al.*, 2017). Model *InceptionV3* dapat lebih lanjut dieksplorasi dengan membandingkannya dengan model *CNN* lain seperti *MobileNet* atau *EfficientNet*, yang mungkin menawarkan trade-off yang berbeda antara akurasi dan efisiensi komputasi (Bimorogo, 2020). Metode *ensemble*, yang menggabungkan kekuatan beberapa model melalui *voting* atau *weighted averaging*, juga dapat menjadi strategi efektif untuk lebih meningkatkan akurasi dan konsistensi, seperti yang dibahas oleh Yaman *et al.* (2021). Pendekatan ini berpotensi mengurangi kelemahan model individual, menghasilkan klasifikasi yang lebih stabil dan akurat.

Secara keseluruhan, penelitian ini menunjukkan bahwa meskipun *ResNet50* menunjukkan performa terbaik dalam hal akurasi sempurna, *InceptionV3* menawarkan keunggulan dalam hal kemampuan generalisasi dan fleksibilitas menghadapi variasi data. Sementara itu, *VGG16* menunjukkan keterbatasan dalam menangkap fitur yang lebih kompleks. Optimisasi lebih lanjut melalui penyesuaian *hyperparameter*, augmentasi data yang lebih komprehensif, atau penggunaan metode *ensemble* dapat meningkatkan performa model, menjadikannya lebih andal dalam aplikasi sistem absensi berbasis pengenalan wajah. Hasil ini memberikan panduan bagi pengembangan dan implementasi teknologi pengenalan wajah yang lebih efisien dan akurat di berbagai industri, sejalan dengan temuan dalam literatur terkait.

4. Kesimpulan

Kesimpulan dari penelitian ini menunjukkan adanya variabilitas performa yang signifikan di antara ketiga model yang dievaluasi. *InceptionV3* tampil sebagai model dengan performa terbaik secara keseluruhan, sedangkan *VGG16* konsisten memberikan prediksi yang salah. Dari segi konsistensi, *InceptionV3* menampilkan akurasi tertinggi dan konsistensi dalam prediksi, sementara *ResNet50* menunjukkan konsistensi hanya pada kelas tertentu, seperti Azizul, tetapi sering mengalami kesalahan pada kelas lainnya. *VGG16* memperlihatkan performa yang paling buruk dengan konsistensi dalam memberikan prediksi yang salah, terutama memprediksi Eldra untuk semua sampel. Dari sisi tingkat kepercayaan, *InceptionV3* menunjukkan kepercayaan yang lebih tinggi dibandingkan model lainnya, terutama dalam kasus Tama dengan tingkat kepercayaan sebesar 98,74%. Secara keseluruhan, urutan keakuratan relatif di antara model adalah $InceptionV3 > ResNet50 > VGG16$.

Implikasi dari temuan ini mencakup beberapa aspek penting. Pertama, *InceptionV3* memiliki kemampuan generalisasi yang lebih baik pada data baru dibandingkan model lainnya, menjadikannya model yang lebih adaptif terhadap kondisi variatif. Kedua, *VGG16* memerlukan penyesuaian yang signifikan atau bahkan perlu dipertimbangkan untuk diganti, mengingat performanya yang kurang memuaskan. Sementara itu, *ResNet50* mungkin masih memerlukan *fine-tuning* tambahan untuk memperbaiki konsistensi dalam prediksi. Dalam aplikasi praktis, *InceptionV3* muncul sebagai pilihan terbaik untuk implementasi sistem absensi berbasis pengenalan wajah, namun perlu dipertimbangkan penggunaan *threshold* kepercayaan untuk menghindari *false positives*. Untuk meningkatkan performa model, beberapa langkah dapat dilakukan, termasuk melakukan *fine-tuning* lebih lanjut pada *InceptionV3* dengan data baru, mempertimbangkan penggunaan metode *ensemble* yang menggabungkan kekuatan *InceptionV3* dan *ResNet50*, serta mengevaluasi kembali arsitektur dan pelatihan *VGG16*.

Hasil penelitian ini menekankan pentingnya evaluasi yang menyeluruh pada data baru dan penyesuaian berkelanjutan dalam pengembangan sistem pengenalan wajah untuk aplikasi absensi karyawan. Meskipun *InceptionV3* menunjukkan potensi yang baik, variabilitas dalam hasil prediksi menunjukkan bahwa masih ada ruang untuk peningkatan, terutama dalam hal konsistensi dan kemampuan generalisasi pada berbagai kondisi dan subjek yang berbeda. Analisis mendalam juga menunjukkan bahwa *InceptionV3* unggul dalam hal akurasi dan kemampuan generalisasi untuk tugas pengenalan wajah dalam konteks sistem absensi, namun konsistensi performa *ResNet50* dan *VGG16* masih perlu ditingkatkan. Optimisasi lebih lanjut, termasuk *fine-tuning hyperparameter* dan strategi augmentasi data yang lebih canggih, dapat meningkatkan performa keseluruhan sistem. Selain itu, perbandingan dengan metode lain menegaskan potensi pendekatan berbasis *deep learning* dalam aplikasi pengenalan wajah yang lebih praktis dan andal di masa depan.

5. Daftar Pustaka

- Bimorogo, S. D., & Kusuma, G. P. (2020). A comparative study of pretrained convolutional neural network model to identify plant diseases on android mobile device. *International Journal*, 9(3).
- Gwyn, T., Roy, K., & Atay, M. (2021). Face recognition using popular deep net architectures: A brief comparative study. *Future Internet*, 13(7), 164. DOI: <https://doi.org/10.3390/fi13070164>.
- Hartiwi, Y., Rasywir, E., Pratama, Y., & Jusia, P. A. (2020). Sistem Manajemen Absensi dengan Fitur Pengenalan Wajah dan GPS Menggunakan YOLO pada Platform Android. *Jurnal Media Informatika Budidarma*, 4(4), 1235-1242. DOI: <http://dx.doi.org/10.30865/mib.v4i4.2522>.

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Kamencay, P., Benco, M., Mizdos, T., & Radil, R. (2017). A new method for face recognition using convolutional neural network. *Advances in Electrical and Electronic Engineering*, 15(4), 663-672.
- Kang, B. N., Kim, Y., & Kim, D. (2017). Deep convolutional neural network using triplets of faces, deep ensemble, and score-level fusion for face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 109-116).
- Ko, H., Chung, H., Kang, W. S., Kim, K. W., Shin, Y., Kang, S. J., ... & Lee, J. (2020). COVID-19 pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation. *Journal of medical Internet research*, 22(6), e19569.
- Kumar, C. R., Saranya, N., Priyadharshini, M., & Gilchrist, D. (2023). Face recognition using CNN and siamese network. *Measurement: Sensors*, 27, 100800.
- Lin, S., Chen, L., Chen, W., & Chang, B. (2024, January). Face Recognition via Thermal Imaging: A Comparative Study of Traditional and CNN-Based Approaches. In *Proceedings of the 2024 8th International Conference on Machine Learning and Soft Computing* (pp. 132-138). DOI: <https://doi.org/10.1145/3647750.3647771>.
- Saragih, R. E., & To, Q. H. (2022). A survey of face recognition based on convolutional neural network. *Indonesian Journal of Information Systems*, 4(2).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. DOI: <https://doi.org/10.48550/arXiv.1409.1556>.
- Sulistya, Y. I., Bangun, E. T. B., & Tyas, D. A. (2023). CNN Ensemble Learning Method for Transfer learning: A Review. *ILKOM Jurnal Ilmiah*, 15(1), 45-63.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).
- Wirayasa, I. K. A., Santoso, H., & Indrajit, E. (2021). Comparison of Convolutional Neural Networks Model Using Different Optimizers for Image Classification. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 60(2), 116-126.
- Yaman, M. A., Rattay, F., & Subasi, A. (2021). Comparison of bagging and boosting ensemble machine learning methods for face recognition. *Procedia Computer Science*, 194, 202-209.