

# Perbandingan Kinerja VGG-16 dan ResNet untuk Pengenalan Ekspresi Wajah Mahasiswa Berbasis CNN pada Smart Learning Environment

*Comparison of VGG16 and ResNet Performance for CNN-Based Student Facial Expression Recognition in Smart Learning Environment*

Dian Ade Kurnia<sup>1</sup>, Fatihanursari Dikananda<sup>2</sup>, Saeful Anwar<sup>3</sup>, Dadang Sudrajat<sup>4</sup>, Abdul Ajiz<sup>5</sup>

<sup>1</sup>Program Studi Manajemen Informatika, STMIK IKMI Cirebon

<sup>2,3,4,5</sup>Program Studi Teknik Informatika, STMIK IKMI Cirebon

<sup>1</sup>dianade2014@gmail.com, <sup>2</sup>fatihah@ikmi.ac.id, <sup>3</sup>saefulanwar419@gmail.com, <sup>4</sup>dias\_sudrajat@yahoo.com,

<sup>5</sup>abdulaziz@gmail.com

### Abstract

*The advancement of artificial intelligence (AI) and computer vision technologies has created significant opportunities for facial expression recognition in diverse domains. In higher education, student engagement is an essential factor that remains difficult to measure objectively through conventional approaches. However, existing studies have rarely examined the performance of widely used convolutional neural network (CNN) architectures under real classroom conditions with varied lighting and pose challenges. This study contributes by comparing the performance of two prominent deep learning architectures, VGG-16 and ResNet, in classifying student facial expressions within a Smart Learning Environment. A quantitative experimental approach was applied through five stages: data collection from students in classrooms, preprocessing using cropping, resizing, and augmentation, CNN model development, training with an 80%-20% split for training and validation, and evaluation using accuracy, precision, recall, and F1-score. The experimental results indicate that VGG-16 outperforms ResNet in recognizing the "happy" expression with the highest F1-score of 85%, while ResNet shows relatively better performance in detecting the "bored" expression with an F1-score of 73.2%. Both models, however, demonstrate weaknesses in identifying the "dislike" expression. These findings suggest that VGG-16 is more suitable for real-time student engagement analysis in AI-based Smart Learning Environments.*

**Keywords:** VGG-16, ResNet, CNN, facial expression, Smart Learning Environment

### Abstrak

Perkembangan teknologi kecerdasan buatan (AI) dan visi komputer telah membuka peluang besar dalam penerapan pengenalan ekspresi wajah pada berbagai bidang. Dalam konteks pendidikan tinggi, keterlibatan mahasiswa selama proses belajar menjadi faktor penting yang masih sulit diukur secara objektif menggunakan metode konvensional. Namun pada kenyataannya, penelitian sebelumnya masih jarang menguji performa arsitektur CNN populer secara langsung di lingkungan pembelajaran nyata dengan kondisi pencahayaan dan pose yang beragam. Penelitian ini berkontribusi dengan membandingkan kinerja dua arsitektur deep learning, yaitu VGG-16 dan ResNet, dalam klasifikasi ekspresi wajah mahasiswa pada Smart Learning Environment. Penelitian dilakukan dengan pendekatan eksperimen kuantitatif melalui lima tahapan, yaitu pengumpulan data wajah mahasiswa di kelas, preprocessing berupa cropping, resizing, dan augmentasi, pengembangan model CNN, pelatihan menggunakan data split 80% training dan 20% validasi, serta evaluasi dengan metrik akurasi, presisi, recall, dan F1-score. Hasil eksperimen menunjukkan bahwa VGG-16 unggul dalam mengenali ekspresi suka dengan nilai F1-score tertinggi sebesar 85%, sedangkan ResNet relatif lebih baik pada ekspresi bosan dengan F1-score 73,2%. Sementara itu, keduanya sama-sama lemah dalam mengenali ekspresi tidak suka. Temuan ini mengimplikasikan bahwa VGG-16 lebih sesuai digunakan untuk mendukung analisis keterlibatan mahasiswa secara real-time dalam Smart Learning Environment berbasis AI.

Kata kunci: VGG-16, ResNet, CNN, ekspresi wajah, Smart Learning Environment.

### 1. Pendahuluan

Perkembangan teknologi kecerdasan buatan (AI) dan visi komputer telah membuka peluang luas dalam analisis ekspresi wajah, khususnya melalui pendekatan deep learning [1]. Pengenalan ekspresi wajah memiliki peran penting dalam mendukung interaksi manusia

komputer, termasuk pada Smart Learning Environment (SLE), yang membutuhkan adaptasi berbasis emosi mahasiswa [2]. Dengan memanfaatkan Convolutional Neural Network (CNN), sistem pembelajaran mampu menilai keterlibatan peserta didik melalui analisis real-time terhadap ekspresi wajah, sehingga dapat

meningkatkan kualitas interaksi dan personalisasi pembelajaran [3], [4].

Dalam konteks pendidikan tinggi, keterlibatan mahasiswa selama pembelajaran menjadi salah satu indikator penting keberhasilan proses belajar. Namun, metode konvensional dalam menilai keterlibatan mahasiswa seringkali subjektif dan sulit diukur secara akurat [5]. Teknologi pengenalan ekspresi wajah berbasis CNN menawarkan solusi dengan memberikan analisis objektif terhadap kondisi emosional mahasiswa, seperti ekspresi bosan, suka, dan tidak suka [6] [7]. Beberapa arsitektur CNN, seperti VGG-16 dan ResNet, telah terbukti unggul dalam ekstraksi fitur visual yang kompleks, sehingga relevan digunakan dalam mendeteksi ekspresi wajah pada SLE [8]. Perbandingan performa kedua arsitektur tersebut penting dilakukan untuk mengetahui model yang lebih tepat dalam mendukung implementasi sistem pembelajaran adaptif di lingkungan pendidikan tinggi [9].

Penelitian terbaru dalam pengenalan ekspresi wajah menunjukkan dominasi penggunaan arsitektur CNN dengan berbagai varian, termasuk VGG-16, ResNet, DenseNet, dan EfficientNet [10] [11]. VGG-16 dikenal memiliki struktur sederhana namun efektif dalam ekstraksi fitur hierarkis, sedangkan ResNet menggunakan mekanisme skip connection yang mampu mengatasi degradasi performa pada jaringan yang lebih dalam [12]. Studi oleh [13] menunjukkan bahwa ResNet memberikan akurasi tinggi dalam klasifikasi emosi pada citra statis, sementara penelitian [9] mengindikasikan bahwa VGG-16 tetap relevan dan dapat menghasilkan kinerja yang layak pada dataset wajah dengan resolusi terbatas. Dalam konteks pendidikan, [5] membuktikan bahwa ekspresi wajah dapat digunakan sebagai indikator keterlibatan mahasiswa. Penelitian lanjutan oleh [14] serta [15] mengintegrasikan FER berbasis CNN ke dalam SLE untuk mendukung personalisasi pembelajaran. Meskipun demikian, masih terdapat tantangan dalam mengatasi variasi pencahayaan, pose, dan ekspresi halus (subtle expressions) yang dapat memengaruhi performa model [16]. Oleh karena itu, studi perbandingan kinerja model pada lingkungan pembelajaran nyata menjadi relevan sebagai kontribusi praktis sekaligus akademis.

Walaupun banyak penelitian membahas pengenalan ekspresi wajah berbasis CNN, sebagian besar studi berfokus pada skenario laboratorium dengan kondisi terkontrol [5]. Namun pada kenyataannya, implementasi di lingkungan pembelajaran nyata menghadapi tantangan pencahayaan tidak seragam, pose beragam, dan keterbatasan resolusi citra [16]. Masih jarang penelitian yang secara spesifik membandingkan kinerja arsitektur VGG-16 dan ResNet pada dataset ekspresi wajah mahasiswa di kelas

nyata untuk mendukung SLE [10]. Kekosongan inilah yang menjadi dasar urgensi penelitian ini.

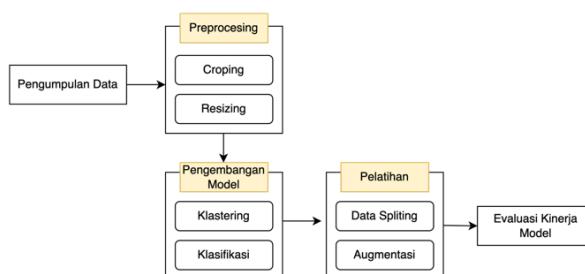
Penelitian ini bertujuan untuk mengevaluasi sekaligus membandingkan kinerja dua arsitektur deep learning yang banyak digunakan dalam pengenalan citra, yaitu VGG-16 dan ResNet, pada tugas klasifikasi ekspresi wajah mahasiswa di lingkungan pembelajaran cerdas (*Smart Learning Environment*). Secara khusus, penelitian diarahkan untuk membangun dataset ekspresi wajah mahasiswa yang dikumpulkan secara langsung di kelas dengan mempertimbangkan variasi pose, kondisi pencahayaan, serta ekspresi emosional yang beragam. Dataset tersebut kemudian diproses melalui tahapan cropping, resizing, dan data augmentation agar kualitas dan keragamannya meningkat serta siap digunakan untuk pelatihan model.

Selanjutnya, penelitian difokuskan pada pengembangan dan pelatihan model klasifikasi berbasis VGG-16 dan ResNet untuk mengenali tiga kategori ekspresi utama, yaitu bosan, suka, serta tidak suka. Kinerja kedua arsitektur tersebut dibandingkan menggunakan metrik evaluasi akurasi, presisi, recall, dan F1-score, sehingga dapat diketahui arsitektur mana yang lebih optimal dalam mendeteksi ekspresi wajah mahasiswa. Melalui tujuan ini, penelitian diharapkan mampu menghasilkan rekomendasi model yang efektif untuk mendukung sistem analisis ekspresi wajah secara real-time, sekaligus memberikan kontribusi praktis dalam pengembangan Smart Learning Environment berbasis kecerdasan buatan di perguruan tinggi.

Penelitian ini memberikan kontribusi dalam tiga aspek utama. Pertama, menyediakan dataset ekspresi wajah mahasiswa yang dikumpulkan pada kondisi pembelajaran nyata dengan variasi pencahayaan dan pose. Kedua, melakukan perbandingan kinerja antara VGG-16 dan ResNet dalam klasifikasi ekspresi wajah mahasiswa, menggunakan metrik akurasi, presisi, recall, dan F1-score. Ketiga, hasil penelitian ini diharapkan mendukung pengembangan Smart Learning Environment yang mampu menilai keterlibatan mahasiswa secara objektif dan real-time. Dengan demikian, studi ini tidak hanya berkontribusi pada pengembangan teori dalam bidang facial expression recognition, tetapi juga memiliki dampak praktis pada inovasi pendidikan berbasis AI.

## 2. Metode Penelitian

Metode penelitian menggunakan pendekatan eksperimen kuantitatif, terdiri dari 5 (lima) tahapan yang meliputi tahap pengumpulan data, tahap preprosesing data, tahap pengembangan model, pelatihan model dan tahap evaluasi model. Tahapan tersebut dapat disajikan pada Gambar 1.



Gambar 1. Tahapan Metode Penelitian



Gambar 2. Pengumpulan data diambil menggunakan kamera 15 menit pertama saat pembelajaran berlangsung

## 2.2. Preprocessing Data

Data wajah yang telah dikumpulkan diproses melalui tahapan cropping untuk memotong area citra sehingga hanya bagian wajah yang relevan yang dipertahankan, kemudian dilakukan resizing dengan ukuran 48 x 48 piksel guna menyeragamkan ukuran citra sesuai resolusi yang telah ditentukan, sehingga citra siap

## 2.1. Pengumpulan Data

Pengumpulan data berdasarkan dataset berupa ekspresi wajah mahasiswa dikumpulkan dengan menggunakan kamera yang dipasang secara strategis di ruang kelas. Pengambilan data dilakukan dalam berbagai kondisi pencahayaan dan sudut pandang sehingga dihasilkan dataset yang bervariasi dan representatif. Sampel data dari aktivitas tersebut dapat disajikan pada Gambar 2.



Gambar 3. Preprocessing Data dengan Cropping dan Resizing

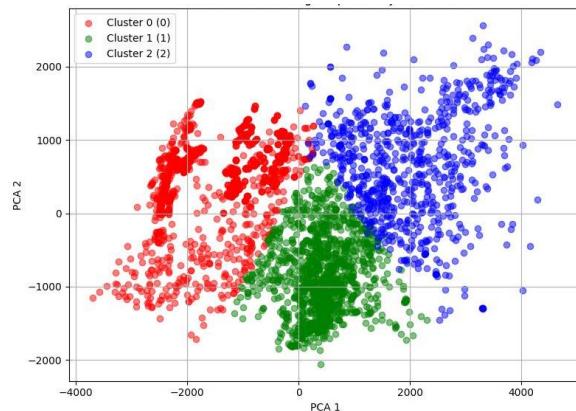
### 2.3. Pengembangan Model

Pada bagian ini, membahas proses penyusunan pendekatan komputasional yang digunakan untuk mengenali dan membedakan ekspresi wajah. Dalam tahap ini, model dikembangkan melalui dua pendekatan utama, yaitu clustering dan klasifikasi.

Clustering digunakan untuk mengelompokkan data wajah berdasarkan kemiripan fitur sehingga ekspresi dapat teridentifikasi dalam kelompok yang seragam tanpa label awal, sedangkan klasifikasi digunakan untuk memetakan data wajah ke dalam kategori ekspresi tertentu berdasarkan model yang telah dilatih. Melalui kombinasi kedua pendekatan ini, sistem mampu memberikan hasil analisis yang lebih akurat dan terstruktur dalam mendekripsi serta mengenali berbagai ekspresi wajah.

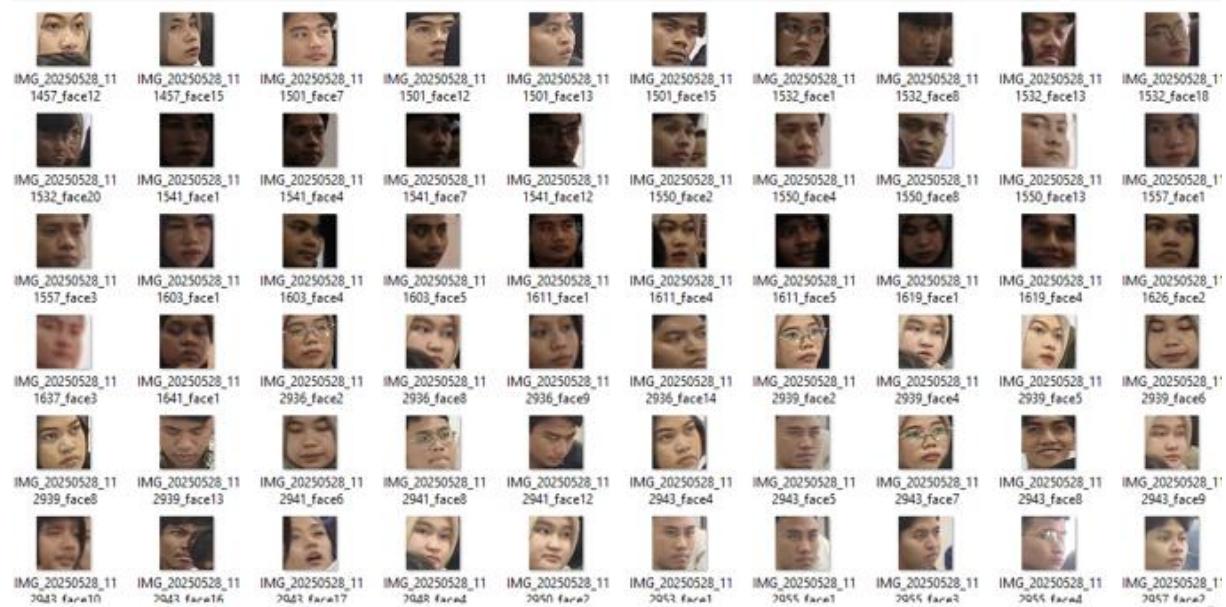
Gambar 4 menampilkan hasil visualisasi clustering ekspresi wajah menggunakan metode Principal Component Analysis (PCA) dalam ruang 2 dimensi (2D). Setiap titik merepresentasikan data ekspresi wajah yang telah direduksi dimensinya, kemudian dikelompokkan ke dalam tiga cluster berbeda. Warna merah, hijau, dan biru menunjukkan perbedaan cluster, yang masing-masing merepresentasikan pola distribusi ekspresi wajah tertentu berdasarkan kemiripan fitur. Visualisasi ini membantu memberikan gambaran yang

lebih jelas mengenai pemisahan antar kelompok ekspresi wajah pada dataset yang digunakan.



Gambar 4. Visualisasi Klastering Ekspresi Wajah

Klaster 0 merepresentasikan kumpulan ekspresi wajah yang cenderung memiliki pola tatapan kosong, minim senyuman, serta ekspresi datar, sehingga dapat diinterpretasikan sebagai kategori ekspresi bosan. Gambar 5 memvisualisasikan wajah dalam cluster ini menunjukkan konsistensi karakteristik yang seragam, di mana individu tampak kurang menunjukkan keterlibatan emosional maupun respons afektif yang kuat.



Gambar 5. Visualisasi Klaster 0

Berdasarkan hasil pengamatan terhadap kumpulan gambar wajah yang ditampilkan pada Gambar 5, mayoritas ekspresi yang muncul cenderung mengarah pada kategori bosan. Hal ini terlihat dari pola umum pada wajah-wajah tersebut yang memperlihatkan tatapan kosong, mata setengah terbuka, serta minimnya gerakan otot wajah yang menandakan adanya keterlibatan emosional. Bibir pada sebagian besar

individu berada dalam posisi netral atau sedikit menurun, tanpa indikasi senyum yang jelas (Babaei et al., 2020). Karakteristik ini merupakan indikator tipikal dari ekspresi bosan, di mana seseorang tidak menunjukkan ketertarikan maupun keterlibatan terhadap situasi yang sedang dihadapinya. Selain itu, ketidaaan ekspresi emosi positif maupun negatif yang kuat memperkuat dugaan bahwa dominasi ekspresi

dalam kumpulan data ini memang lebih banyak mengarah pada kebosanan (Aslan et al., 2019).

Keseragaman pola ekspresi wajah ini memperlihatkan bahwa kategori bosan lebih mendominasi dibandingkan kategori lain yaitu suka atau tidak suka. Meskipun terdapat beberapa individu yang menampilkan variasi ekspresi dengan sedikit senyum atau tanda ketidaknyamanan, jumlahnya relatif kecil dan tidak signifikan dalam keseluruhan dataset. Dengan demikian, dapat disimpulkan bahwa distribusi ekspresi wajah pada gambar-gambar tersebut menunjukkan kecenderungan kuat terhadap ekspresi bosan. Dominasi ini memberikan gambaran penting dalam analisis pengenalan ekspresi wajah, karena dapat menjadi dasar untuk menilai kondisi afektif seseorang dalam konteks tertentu, misalnya dalam pembelajaran atau interaksi sosial yang membutuhkan keterlibatan emosional.

Klaster 1 yang terdapat pada Gambar 6, memperlihatkan kumpulan ekspresi wajah yang didominasi oleh senyum lebar, mata yang lebih terbuka, serta raut wajah cerah, sehingga dapat diidentifikasi sebagai kategori ekspresi suka. Pola ekspresi dalam kelompok ini menegaskan adanya keterlibatan emosional positif yang konsisten, menggambarkan kondisi afektif yang penuh energi dan kegembiraan.



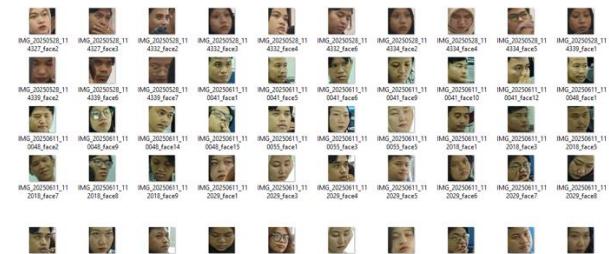
Gambar 6. Visualisasi Klaster 1

Berdasarkan hasil pengelompokan pada Cluster 1 pada Gambar 6, terlihat bahwa mayoritas wajah menunjukkan ekspresi yang jelas mengarah pada kategori suka. Hal ini tampak dari senyum yang mendominasi hampir seluruh individu, diiringi oleh sudut bibir yang terangkat serta raut wajah yang lebih terbuka dan bersahabat. Selain itu, ekspresi mata yang tampak lebih hidup dan keterlibatan otot wajah, khususnya di area pipi, memberikan indikasi kuat bahwa kelompok ini merepresentasikan kondisi emosional yang positif. Pola ekspresi yang konsisten ini menandakan adanya tingkat keseragaman afektif yang tinggi dalam cluster tersebut.

Oleh karena itu, visualisasi wajah dalam cluster ini memperlihatkan karakteristik afektif yang berlawanan dengan cluster lain seperti bosan atau tidak suka. Ekspresi suka dalam konteks ini dapat dihubungkan dengan keadaan psikologis yang mendukung

keterlibatan, motivasi, dan interaksi sosial yang baik. Dengan kata lain, individu dalam cluster ini cenderung menunjukkan respons emosional yang positif dan penuh energi, sehingga cluster ini dapat dijadikan representasi kuat dari ekspresi suka dalam analisis pengenalan ekspresi wajah.

Klaster 2 pada Gambar 7, memperlihatkan kumpulan ekspresi wajah yang didominasi oleh ciri-ciri ketidaknyamanan, penolakan, atau ketidaksenangan, seperti alis berkerut, tatapan tajam atau menyamping, serta bibir yang menurun atau mengerucut. Pola ekspresi ini mengindikasikan bahwa kelompok wajah dalam cluster ini secara konsisten merepresentasikan kategori ekspresi tidak suka.



Gambar 7. Visualisasi Klaster 2

Berdasarkan hasil pengelompokan pada Cluster 2 pada Gambar 7, ekspresi wajah yang ditampilkan cenderung menunjukkan ciri-ciri ketidaknyamanan atau ketidaksenangan. Hal ini terlihat dari raut wajah yang lebih tegang, alis yang berkerut, serta tatapan mata yang tampak kurang ramah atau bahkan menyiratkan rasa penolakan. Selain itu, posisi bibir pada sebagian besar wajah tampak menurun atau mengerucut ke arah bawah, menandakan kondisi emosional yang negatif. Karakteristik tersebut sejalan dengan indikator umum ekspresi tidak suka, yaitu adanya tanda-tanda visual yang menunjukkan rasa tidak puas, terganggu, atau kurang nyaman dengan situasi tertentu.

Wajah-wajah dalam cluster ini menampilkan pola ekspresi yang berbeda secara jelas dari cluster sebelumnya, khususnya cluster suka. Ekspresi tidak suka yang konsisten ini memperlihatkan adanya keseragaman kondisi afektif di antara individu, sehingga cluster ini dapat dijadikan representasi yang kuat untuk kategori ekspresi tidak suka. Dalam konteks analisis pengenalan ekspresi wajah, temuan ini menegaskan bahwa fitur visual seperti kerutan dahi, tatapan menyamping, serta bentuk bibir yang menurun merupakan indikator penting yang mampu membedakan kategori tidak suka dari ekspresi emosional lainnya.

Dalam penelitian ini, label kelas untuk supervised learning diperoleh dari hasil clustering yang dilakukan pada tahap pra-pemrosesan data. Proses clustering menghasilkan tiga kelompok utama yang sesuai dengan kategori ekspresi emosional mahasiswa, yaitu bosan,

suka, dan tidak suka. Hasil dari clustering tersebut kemudian dijadikan sebagai target (ground truth) pada tahap klasifikasi dengan arsitektur CNN, VGG-16, dan ResNet, sehingga setiap model dapat dilatih dan dievaluasi secara konsisten terhadap kelas yang sama. Proses perbandingan ini melibatkan evaluasi menyeluruh terhadap performa masing-masing arsitektur dalam mendeteksi berbagai ekspresi

emosional mahasiswa yaitu bosan, suka dan tidak suka. Pemilihan arsitektur terbaik dilakukan dengan mempertimbangkan metrik evaluasi seperti akurasi, presisi, recall, serta kecepatan inferensi, sehingga model yang dihasilkan tidak hanya memiliki tingkat ketepatan tinggi, tetapi juga efisien untuk diimplementasikan dalam lingkungan pembelajaran yang dinamis. Kedua arsitektur tersebut disajikan pada Tabel 1 dan Tabel 2.

Tabel 1. Arsitektur VGG-16

Layer Name	Layer Type	Output Shape	Param #
input_1	InputLayer	(48, 48, 3)	0
block1_conv1	Conv2D	(48, 48, 64)	1,792
block1_conv2	Conv2D	(48, 48, 64)	36,928
block1_pool	MaxPooling2D	(24, 24, 64)	0
block2_conv1	Conv2D	(24, 24, 128)	73,856
block2_conv2	Conv2D	(24, 24, 128)	147,584
block2_pool	MaxPooling2D	(12, 12, 128)	0
block3_conv1	Conv2D	(12, 12, 256)	295,168
block3_conv2	Conv2D	(12, 12, 256)	590,080
block3_conv3	Conv2D	(12, 12, 256)	590,080
block3_pool	MaxPooling2D	(6, 6, 256)	0
block4_conv1	Conv2D	(6, 6, 512)	1,180,160
block4_conv2	Conv2D	(6, 6, 512)	2,359,808
block4_conv3	Conv2D	(6, 6, 512)	2,359,808
block4_pool	MaxPooling2D	(3, 3, 512)	0
block5_conv1	Conv2D	(3, 3, 512)	2,359,808
block5_conv2	Conv2D	(3, 3, 512)	2,359,808
block5_conv3	Conv2D	(3, 3, 512)	2,359,808
block5_pool	MaxPooling2D	(1, 1, 512)	0
flatten	Flatten	(512,)	0
dense	Dense (ReLU + L2)	(128,)	65,664
dropout	Dropout(0.5)	(128,)	0
dense_1	Dense (Softmax)	(3,)	387

Tabel 2. Arsitektur ResNet

Layer Name	Layer Type	Output Shape	Param #
Input_1	InputLayer	(48, 48, 3)	0
block1_conv1	Conv2D	(24, 24, 64)	9,472
block1_pool	MaxPooling2D	(12, 12, 64)	0
block2_conv1	Conv2D	(12, 12, 256)	16,896
block2_conv2	Conv2D	(12, 12, 256)	73,856
block2_conv3	Conv2D	(12, 12, 256)	147,584
block2_conv4	Conv2D	(6, 6, 512)	590,080
block2_conv5	Conv2D	(6, 6, 512)	1,180,160
block2_conv6	Conv2D	(6, 6, 512)	2,359,808
block2_conv7	Conv2D	(3, 3, 1024)	4,719,616
block2_conv8	Conv2D	(3, 3, 1024)	4,719,616
block2_conv9	Conv2D	(3, 3, 2048)	9,439,232
global	GlobalAveragePooling2D	(2048,)	0
dense	Dense	(128,)	262,272
dropout	Dropout	(128,)	0
dense	Dense (Output)	(n_classes,)	(128 × n + n)

Untuk mendukung proses pelatihan dan evaluasi model CNN berbasis arsitektur VGG-16 dan arsitektur ResNet dalam penelitian ini, dilakukan pengaturan parameter optimasi yang disesuaikan dengan karakteristik dataset dan tujuan penelitian. Pemilihan parameter meliputi aspek algoritma optimasi, fungsi kerugian, ukuran batch, jumlah epoch, hingga teknik regulasi seperti dropout dan early stopping. Pengaturan ini bertujuan untuk memperoleh performa model yang optimal sekaligus mencegah terjadinya overfitting. Rincian

parameter optimasi yang digunakan oleh kedua arsitektur tersebut ditunjukkan pada Tabel 3.

Tabel 3. Parameter Optimasi Arsitektur VGG-16 dan ResNet

Parameter	Value
Optimizer	Adam
Learning Rate	<b>0.00001</b>
Loss Function	categorical_crossentropy
Epoch (early stopping)	50
Batch size	32
Metric	Accuracy
Regularisasi	l2(0.001) pada Dense(128)

Dropout	0.5
Early Stopping	Patience = 7, restore_best_weights = True

## 2.4. Pelatihan

### a. Data Splitting

Pada tahapan ini, pembagian data (data splitting) dilakukan untuk memisahkan dataset menjadi beberapa bagian dengan tujuan memastikan bahwa model tidak hanya menghafal data, tetapi mampu melakukan generalisasi pada data baru. Dataset dibagi menjadi training set dan testing set. Training set digunakan untuk melatih model, sedangkan testing set digunakan untuk menilai kinerja model pada data yang belum pernah dilihat sebelumnya. Proses pembagian ini dilakukan secara proporsional, yaitu 80% untuk pelatihan, 20% untuk validasi, sehingga evaluasi terhadap model dapat dilakukan secara objektif [17].

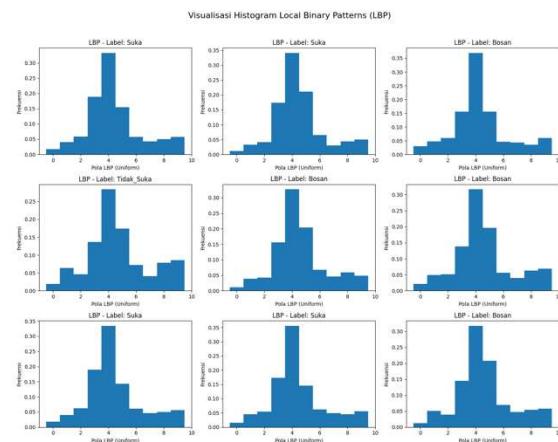
### b. Data Augmentasi

Dalam penelitian ini, proses preprocessing tidak hanya dilakukan melalui normalisasi dan penyesuaian dimensi gambar, tetapi juga dengan menerapkan berbagai teknik data augmentation. Penerapan augmentation bertujuan untuk memperkaya variasi data latih tanpa menambah jumlah data secara manual, sehingga model dapat lebih robust dan mampu mengenali pola ekspresi wajah dalam kondisi berbeda [18]. Teknik augmentasi yang digunakan meliputi rotasi, pergeseran, pencerminan, serta normalisasi piksel. Rincian parameter teknik augmentasi yang diterapkan dapat dilihat pada Tabel 4.

Tabel 1. Parameter Teknik Augmentasi

Teknik Augmentasi	Value	Deskripsi
Rescale	1. / 255	normalisasi piksel
Rotation_Range	15	Rotasi acak maksimal 15 derajat
Width_shift_range	0.1	Geser horizontal
Height_shift_range	0.1	Geser vertikal
Zoom_range	0.1	Zoom acak
Shear_range	0.1	Distorsi miring
Horizontal_flip	True	Pencerminan horizontal

Sebagai hasil dari penerapan teknik data augmentation pada tahapan preprocessing, dilakukan analisis tekstur menggunakan metode Local Binary Patterns (LBP) untuk melihat distribusi pola fitur yang terbentuk pada setiap kelas ekspresi wajah [19]. Visualisasi histogram LBP ini bertujuan untuk menggambarkan frekuensi kemunculan pola LBP (uniform) pada masing-masing label ekspresi, seperti Suka, Tidak Suka, dan Bosan. Dengan adanya variasi distribusi pola ini, dapat diamati bahwa proses augmentasi berhasil memperkaya representasi data latih, sehingga model diharapkan mampu mengenali ekspresi wajah dengan lebih baik. Hasil visualisasi histogram LBP ditunjukkan pada Gambar 9.



Gambar 8. Histogram Hasil Augmentasi

## 2.5. Evaluasi Kinerja Model

Setelah pelatihan selesai, kinerja model dievaluasi menggunakan metrik performa seperti Accuracy, Precision, Recall, dan F1-Score, yang masing-masing digunakan untuk menilai tingkat ketepatan dan kemampuan model dalam mendeteksi kelas target [20]. Adapun rumus metrik performa yang digunakan tersebut adalah:

Akurasi mengukur proporsi prediksi yang benar dibandingkan dengan seluruh prediksi. Semakin tinggi nilainya, semakin akurat model secara keseluruhan, sebagaimana disajikan pada rumus (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

TP (True Positive): Kasus positif yang diprediksi benar; TN (True Negative): Kasus negatif yang diprediksi benar; FP (False Positive): Kasus negatif yang diprediksi salah sebagai positif; FN (False Negative): Kasus positif yang diprediksi salah sebagai negatif.

Precision mengukur ketepatan model dalam mengidentifikasi kelas positif. Nilai yang tinggi menunjukkan bahwa kesalahan prediksi positif (FP) sangat sedikit, sebagaimana disajikan pada rumus (2).

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

TP (True Positive): Kasus positif yang diprediksi benar; FP (False Positive): Kasus negatif yang diprediksi salah sebagai positif.

Recall mengukur kemampuan model menemukan semua kasus positif yang sebenarnya. Nilai tinggi menunjukkan model jarang melewatkannya sampel positif, sebagaimana disajikan pada rumus (3).

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

TP (True Positive): Kasus positif yang diprediksi benar; FN (False Negative): Kasus positif yang diprediksi salah sebagai negatif.

F1-Score adalah rata-rata harmonis antara precision dan recall. Digunakan ketika diperlukan keseimbangan antara keduanya, khususnya pada dataset yang tidak seimbang (imbalanced dataset), sebagaimana disajikan pada rumus (4).

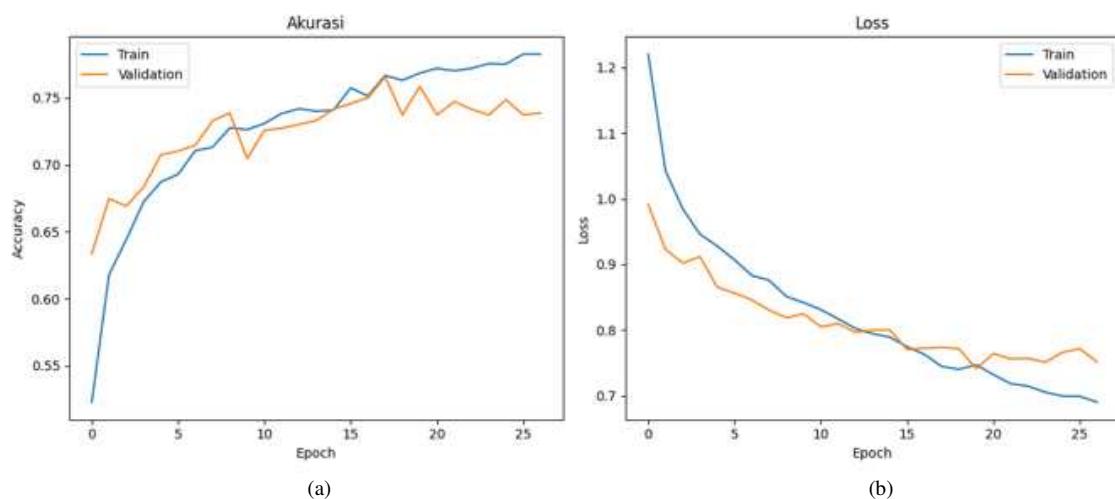
$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

### 3. Hasil dan Pembahasan

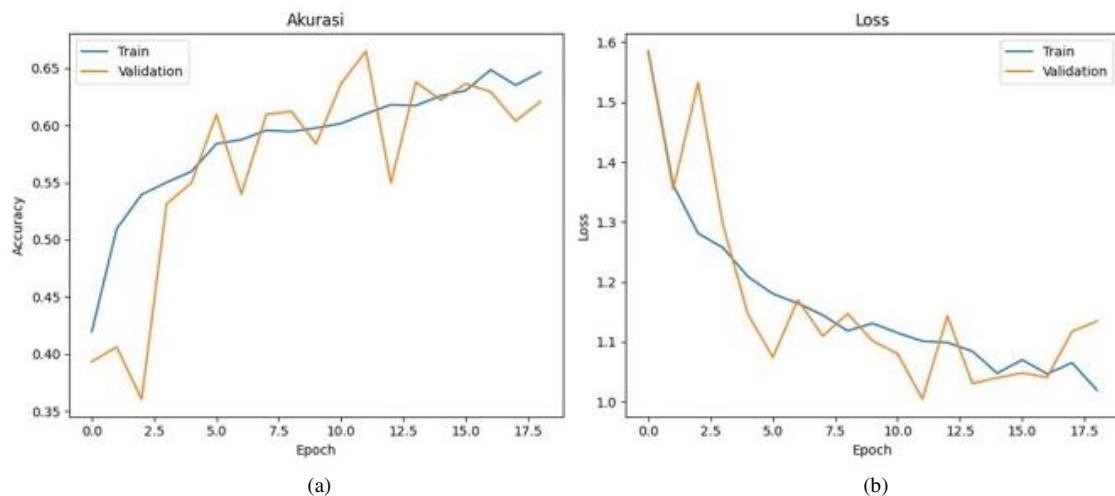
#### 3.1. Hasil Pemodelan

Untuk mengevaluasi kinerja model CNN berbasis arsitektur VGG-16 dan arsitektur ResNet dalam

mengenali ekspresi wajah, dilakukan proses pelatihan dan validasi dengan memantau nilai akurasi dan loss pada setiap epoch. Grafik akurasi menggambarkan perkembangan kemampuan model dalam melakukan prediksi yang benar pada data latih maupun data validasi, sedangkan grafik loss menunjukkan penurunan kesalahan prediksi yang dihasilkan oleh model. Dengan visualisasi ini, dapat diamati dinamika proses pembelajaran model serta indikasi adanya overfitting atau underfitting yang mungkin terjadi. Hasil visualisasi akurasi dan loss selama pelatihan ditunjukkan pada Gambar 10 dan Gambar 11.



Gambar 9. Loss Akurasi (a) dan Loss Validasi (b) Arsitektur VGG-16



Gambar 10. Loss Akurasi (a) dan Loss Validasi (b) Arsitektur ResNet

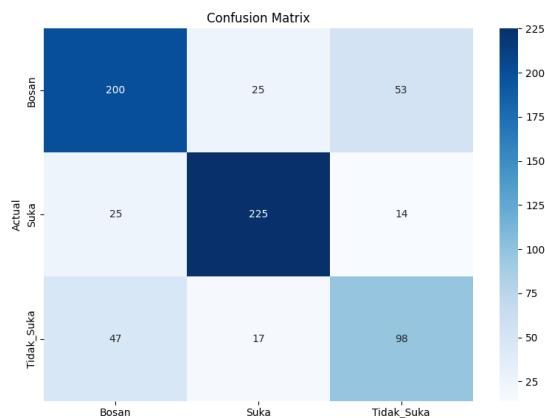
#### 3.2. Hasil Evaluasi Kinerja Model

Sebagai bagian dari evaluasi performa model, dilakukan analisis menggunakan confusion matrix untuk melihat tingkat akurasi prediksi pada masing-masing kelas ekspresi wajah. Confusion matrix ini memberikan gambaran detail mengenai jumlah prediksi yang benar maupun salah pada setiap label, sehingga dapat diketahui sejauh mana model mampu

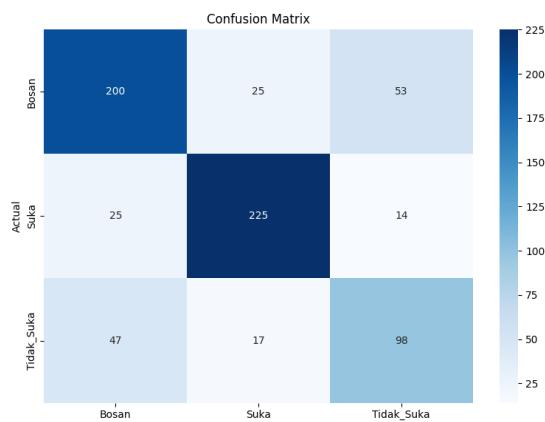
melakukan klasifikasi dengan tepat. Pada Gambar 12 Dan Gambar 13 berikut menampilkan hasil confusion matrix dari pemodelan menggunakan arsitektur VGG-16 dan ResNet.

Untuk menilai kinerja arsitektur VGG-16 dalam melakukan klasifikasi ekspresi wajah, digunakan metrik evaluasi berupa Precision, Recall, dan F1-Score pada masing-masing kelas. Precision digunakan untuk

mengukur ketepatan prediksi model terhadap kelas tertentu, Recall menunjukkan sejauh mana model mampu mengenali seluruh data yang benar pada kelas tersebut, sedangkan F1-Score merupakan rata-rata harmonis antara Precision dan Recall. Hasil evaluasi ini memberikan gambaran mengenai kekuatan dan kelemahan model dalam membedakan tiap ekspresi wajah. Rincian nilai evaluasi pada setiap kategori ekspresi dapat dilihat pada Tabel 5.



Gambar 11. Confusion Matrix Arsitektur VGG-16



Gambar 12. Confusion Matrix Arsitektur ResNet

Tabel 2. Matrik Evaluasi Arsitektur VGG-16

Ekspresi	Nilai		
	Precision	Recall	F1-Score
Bosan	74.00%	72.00%	73.00%
Suka	84.00%	85.00%	85.00%
Tidak Suka	59.00%	60.00%	60.00%
		Akurasi	74.00%

Sementara itu, untuk arsitektur ResNet, proses evaluasi juga dilakukan menggunakan metrik Precision, Recall, dan F1-Score pada masing-masing kelas ekspresi wajah. Sama halnya dengan VGG-16, hasil ini memberikan gambaran tentang kemampuan ResNet dalam mengenali pola ekspresi, sekaligus memperlihatkan perbedaan kinerja antar kelas. Rincian nilai evaluasi performa arsitektur ResNet pada setiap kategori ekspresi wajah dapat dilihat pada Tabel 6.

Tabel 3. Matrik Evaluasi Arsitektur ResNet

Ekspresi	Nilai		
	Precision	Recall	F1-Score
Bosan	70.90%	75.70%	73.20%
Suka	64.30%	73.80%	34.36%
Tidak Suka	58.60%	41.90%	48.90%
		Akurasi	66.36%

### 3.3. Hasil Perbandingan Evaluasi Kinerja Model

Untuk memperoleh gambaran yang lebih jelas mengenai kinerja masing-masing arsitektur, dilakukan tabulasi hasil evaluasi menggunakan metrik Precision, Recall, dan F1-Score pada setiap kelas ekspresi wajah. Tabel 7 ini memberikan informasi detail terkait keunggulan dan kelemahan model dalam mengenali tiap kategori ekspresi, serta memudahkan analisis performa antara arsitektur VGG-16 dan ResNet.

Tabel 4. Perbandingan Evaluasi Kinerja Model

Ekspresi	Arsitektur	Precision	Recall	F1-Score
Bosan	VGG-16	74.00%	72.00%	73.00%
	ResNet	70.90%	75.70%	73.20%
Suka	VGG-16	84.00%	85.00%	85.00%
	ResNet	64.30%	73.80%	34.36%
Tidak Suka	VGG-16	59.00%	60.00%	60.00%
	ResNet	58.60%	41.90%	48.90%

## 4. Pembahasan

### 4.1. Analisis Performa Model VGG-16

Grafik akurasi menunjukkan bahwa model VGG-16 yang terdapat pada Gambar 10 berhasil mengalami proses pembelajaran yang stabil sepanjang epoch. Pada tahap awal, akurasi training meningkat cukup tajam dari sekitar 0.55 hingga mencapai lebih dari 0.70 pada epoch kelima, kemudian terus mengalami peningkatan meskipun dengan laju yang lebih lambat hingga mendekati 0.82 pada epoch ke-25. Sementara itu, akurasi validasi juga mengalami tren kenaikan yang konsisten hingga mencapai kisaran 0.76, walaupun terlihat adanya fluktuasi di beberapa titik. Kedekatan kurva akurasi training dan validasi ini menandakan bahwa model mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat, serta tidak menunjukkan adanya gejala overfitting yang signifikan.

Pada grafik loss yang terdapat pada Gambar 10, terlihat penurunan yang cukup baik pada data training maupun validasi. Nilai loss training menurun dari sekitar 1.2 di awal menjadi sekitar 0.68 pada akhir epoch, sedangkan loss validasi juga menunjukkan tren serupa dengan nilai akhir di kisaran 0.72. Perbedaan nilai loss antara data training dan validasi relatif kecil, sehingga dapat disimpulkan bahwa model berhasil mempelajari pola data dengan efektif tanpa kehilangan kemampuan generalisasi. Pola ini menunjukkan bahwa penerapan parameter optimasi, regulasi, serta teknik augmentasi yang digunakan dalam penelitian ini mampu

mendukung kinerja VGG-16 sehingga menghasilkan model yang cukup andal dalam tugas pengenalan ekspresi wajah mahasiswa pada Smart Learning Environment

#### 4.2. Analisis Performa Model ResNet

Grafik akurasi pada model ResNet yang disajikan pada Gambar 11 menunjukkan adanya peningkatan performa model secara bertahap selama proses pelatihan. Pada awal epoch, akurasi training berada di kisaran 0.43 dan terus meningkat hingga mencapai sekitar 0.65 pada epoch ke-18. Akurasi validasi juga mengalami tren serupa, meskipun terlihat fluktuatif pada beberapa titik, namun tetap stabil di kisaran 0.60 - 0.64. Kedekatan antara kurva akurasi training dan validasi menandakan bahwa model ResNet memiliki kemampuan generalisasi yang cukup baik terhadap data validasi, serta tidak menunjukkan perbedaan signifikan yang mengindikasikan adanya overfitting.

Pada grafik loss model ResNet Gambar 11, terlihat penurunan yang konsisten baik pada data training maupun validasi. Nilai loss training awalnya cukup tinggi, yaitu sekitar 1.6, kemudian menurun hingga mendekati 1.0 pada epoch ke-18. Sementara itu, loss validasi juga memperlihatkan pola penurunan yang sejalan dengan training, meskipun terdapat fluktuasi kecil di beberapa epoch. Kondisi ini menunjukkan bahwa model ResNet mampu mempelajari pola data dengan efektif, di mana penurunan loss yang stabil disertai peningkatan akurasi membuktikan bahwa parameter optimasi yang digunakan sudah sesuai dalam mendukung kinerja model pada tugas pengenalan ekspresi wajah.

#### 4.3. Analisis Evaluasi Model VGG-16

Berdasarkan hasil confusion matrix pada Gambar 12, model VGG-16 menunjukkan performa yang relatif baik dalam mengenali ekspresi suka, di mana sebagian besar data berhasil diklasifikasikan dengan benar. Hal ini juga tercermin pada nilai evaluasi yang terdapat pada Tabel 5, di mana kelas Suka memperoleh nilai Precision 84%, Recall 85%, dan F1-Score 85%, yang merupakan nilai tertinggi dibandingkan kelas lainnya. Sementara itu, ekspresi Bosan juga mampu dikenali dengan cukup baik dengan nilai Precision 74%, Recall 72%, dan F1-Score 73%, meskipun masih terdapat kesalahan prediksi ke kelas lain. Kondisi ini menunjukkan bahwa VGG-16 cukup efektif untuk membedakan ekspresi positif (Suka) maupun netral (Bosan).

Namun, pada kelas tidak suka, performa VGG-16 masih cukup rendah, dengan nilai Precision 59%, Recall 60%, dan F1-Score 60%. Confusion matrix memperlihatkan bahwa sebagian data dari kelas ini salah diklasifikasikan sebagai kelas Bosan atau Suka. Hal ini mengindikasikan bahwa VGG-16 mengalami kesulitan dalam mengenali ekspresi negatif yang

mungkin memiliki pola visual lebih beragam atau cenderung mirip dengan ekspresi lainnya. Meski demikian, secara keseluruhan, VGG-16 tetap menunjukkan kemampuan generalisasi yang baik karena tidak terdapat gap besar antara performa training dan validasi, serta mampu menjaga keseimbangan dalam mengenali sebagian besar ekspresi.

#### 4.4. Analisis Evaluasi Model ResNet

Pada hasil evaluasi arsitektur ResNet, confusion matrix pada Gambar 13, menunjukkan bahwa kelas Bosan masih menjadi kategori dengan performa relatif lebih baik dibandingkan dua kelas lainnya. Hal ini tercermin pada Tabel 6, dari nilai Precision 70.90%, Recall 75.70%, dan F1-Score 73.20%, yang mendekati hasil yang dicapai VGG-16. Artinya, ResNet mampu mengenali pola visual ekspresi Bosan dengan cukup konsisten meskipun terdapat beberapa kesalahan prediksi ke kelas lain. Namun, kinerja model mulai menurun pada kelas Suka yang hanya memperoleh Precision 64.30%, Recall 73.80%, dan F1-Score 41.90%. Rendahnya nilai F1-Score pada kelas ini mengindikasikan adanya ketidakseimbangan antara Precision dan Recall, di mana model lebih banyak mendeteksi data kelas Suka namun sering salah prediksi sehingga menurunkan ketepatan klasifikasinya.

Kinerja ResNet yang paling lemah terlihat pada kelas Tidak Suka dengan nilai Precision 58.60%, Recall 41.90%, dan F1-Score 48.90%. Confusion matrix memperlihatkan bahwa banyak data dari kelas ini salah diklasifikasikan ke kelas Bosan maupun Suka. Hal ini menandakan bahwa ResNet mengalami kesulitan dalam membedakan ekspresi negatif dari ekspresi lainnya, lebih parah dibandingkan dengan performa VGG-16. Dengan demikian, meskipun ResNet dikenal sebagai arsitektur yang lebih dalam dan kompleks, pada penelitian ini hasilnya belum mampu melampaui kinerja VGG-16, khususnya pada kelas ekspresi Suka dan Tidak Suka. Hal ini bisa menjadi indikasi bahwa ResNet membutuhkan fine-tuning lebih lanjut, penyesuaian parameter, atau augmentasi tambahan agar dapat menghasilkan performa yang lebih optimal.

#### 4.5. Analisis Perbandingan Model VGG-16 dan ResNet

Hasil evaluasi memperlihatkan bahwa VGG-16 lebih unggul dibandingkan ResNet dalam tugas pengenalan ekspresi wajah mahasiswa. VGG-16 menunjukkan performa yang stabil di semua kelas, terutama pada kelas Suka, sementara ResNet hanya setara pada kelas Bosan namun jauh tertinggal pada kelas Suka dan Tidak Suka.

Temuan ini mengindikasikan bahwa VGG-16 lebih sesuai digunakan pada penelitian ini, sedangkan ResNet memerlukan fine-tuning parameter atau strategi augmentasi tambahan untuk meningkatkan kemampuannya dalam mengenali ekspresi wajah.

## 5. Kesimpulan

Penelitian ini berhasil membandingkan kinerja arsitektur VGG-16 dan ResNet dalam tugas klasifikasi ekspresi wajah mahasiswa pada Smart Learning Environment. Berdasarkan hasil evaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score, diperoleh fakta bahwa arsitektur VGG-16 menunjukkan performa yang lebih stabil dan unggul, khususnya dalam mengenali ekspresi suka dengan nilai F1-score sebesar 85%, sedangkan ResNet relatif lebih baik dalam mendekripsi ekspresi bosan namun masih lemah pada kategori tidak suka. Hasil ini menegaskan bahwa VGG-16 lebih sesuai digunakan pada skenario pembelajaran nyata yang memerlukan deteksi emosional secara real-time dan konsisten.

Kemungkinan aplikasi dari penelitian ini adalah integrasi sistem analisis ekspresi wajah ke dalam platform pembelajaran berbasis AI, sehingga dosen atau sistem dapat menilai tingkat keterlibatan mahasiswa secara objektif dan memberikan intervensi adaptif sesuai kondisi emosional. Implikasi dari temuan ini menunjukkan bahwa model CNN dapat menjadi salah satu komponen penting dalam mendukung inovasi pendidikan digital yang lebih personal. Meskipun demikian, terdapat spekulasi bahwa kinerja ResNet dapat ditingkatkan melalui fine-tuning, penggunaan transfer learning dengan dataset yang lebih besar, atau integrasi mekanisme attention untuk mengatasi kesulitan dalam mengenali ekspresi negatif.

Penelitian ini memiliki beberapa keterbatasan, antara lain ukuran dataset yang relatif kecil sehingga generalisasi model masih terbatas, variasi ekspresi emosi yang diuji hanya mencakup tiga kategori (suka, bosan, dan tidak suka) sehingga belum sepenuhnya merepresentasikan keragaman emosi mahasiswa, serta kondisi pengambilan data yang relatif terkendali dan belum sepenuhnya mencerminkan dinamika lingkungan pembelajaran nyata.

Sebagai saran untuk penelitian selanjutnya, disarankan untuk memperluas jumlah kelas emosi yang dikenali, menggunakan dataset yang lebih besar dan bervariasi, serta menguji performa model dalam kondisi nyata dengan tingkat dinamika lebih kompleks, seperti kelas daring atau pembelajaran kolaboratif. Selain itu, integrasi dengan sensor multimodal seperti detak jantung atau suara juga dapat menjadi arah pengembangan agar sistem mampu memberikan analisis afektif yang lebih komprehensif dan akurat.

## Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Kementerian Pendidikan Tinggi, Sains, dan Teknologi Republik Indonesia, melalui Direktorat Jenderal Penelitian dan Pengembangan, atas dana yang diberikan dalam rangka Program Dana Penelitian

Dosen Pemula (PDP) 2025. Ucapan terima kasih khusus juga disampaikan kepada STMIK IKMI Cirebon atas dukungan, bantuan, dan fasilitas yang diberikan sepanjang proses penelitian.

## Daftar Rujukan

- [1] M. Shafiq and Z. Gu, "Deep Residual Learning for Image Recognition: A Survey," *Appl. Sci.*, vol. 12, no. 18, pp. 1–43, 2022, doi: 10.3390/app12188972.
- [2] S. Aslan *et al.*, "Investigating the Impact of a Real-time, Multimodal Student Engagement Analytics Technology in Authentic Classrooms," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–12, 2019, doi: 10.1145/3290605.3300534.
- [3] E. C. Olivetti, M. G. Violante, E. Vezzetti, F. Marcolin, and B. Eynar, "Engagement evaluation in a virtual learning environment via facial expression recognition and self-reports: A preliminary approach," *Appl. Sci.*, vol. 10, no. 1, 2020, doi: 10.3390/app10010314.
- [4] W. Alarefah, S. K. Jarraya, and N. Abuzinadah, "Transformer-Based Student Engagement Recognition Using Few-Shot Learning," *Computers*, vol. 14, no. 3, 2025, doi: 10.3390/computers14030109.
- [5] S. Wang, A. S. Shibghatullah, K. H. Keoy, and J. Iqbal, "YOLOv5 Based Student Engagement and Emotional States Detection in E-Classes," *J. Robot. Netw. Artif. Life*, vol. 10, no. 4, pp. 357–361, 2024, doi: 10.57417/jrnal.10.4\_357.
- [6] D. Arumugam and D. S., "Emotion Classification Using Facial Expression," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 7, 2011, doi: 10.14569/ijacsa.2011.020714.
- [7] T. Kopalidis, V. Solachidis, N. Vretos, and P. Daras, "Advances in Facial Expression Recognition: A Survey of Methods, Benchmarks, Models, and Datasets," *Inf.*, vol. 15, no. 3, 2024, doi: 10.3390/info15030135.
- [8] Z. Ullah, L. Qi, A. Hasan, and M. Asim, "Improved Deep CNN-based Two Stream Super Resolution and Hybrid Deep Model-based Facial Emotion Recognition," *Eng. Appl. Artif. Intell.*, vol. 116, no. September, 2022, doi: 10.1016/j.engappai.2022.105486.
- [9] L. L. X. Wei and N. S. Sani, "Enhanced Facial Expression Recognition Based on ResNet50 with a Convolutional Block Attention Module," *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 1, pp. 695–711, 2025, doi: 10.14569/IJACSA.2025.0160167.
- [10] M. Sajjad *et al.*, "A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines," *Alexandria Eng. J.*, vol. 68, pp. 817–840, 2023, doi: 10.1016/j.aej.2023.01.017.
- [11] S. Oh and D. K. Kim, "Comparative Analysis of Emotion Classification Based on Facial Expression and Physiological Signals Using Deep Learning," *Appl. Sci.*, vol. 12, no. 3, 2022, doi: 10.3390/app12031286.
- [12] I. Haider, H. J. Yang, G. S. Lee, and S. H. Kim, "Robust Human Face Emotion Classification Using Triplet-Loss-Based Deep CNN Features and SVM," *Sensors*, vol. 23, no. 10, 2023, doi: 10.3390/s23104770.
- [13] J. Shao and Y. Qian, "Three convolutional neural network models for facial expression recognition in the wild," *Neurocomputing*, vol. 355, pp. 82–92, 2019, doi: 10.1016/j.neucom.2019.05.005.
- [14] H. V. Manalu and A. P. Rifai, "Detection of human emotions through facial expressions using hybrid convolutional neural network-recurrent neural network algorithm," *Intell. Syst. with Appl.*, vol. 21, no. December 2023, p. 200339, 2024, doi: 10.1016/j.iswa.2024.200339.
- [15] N. Feng, Y. Chen, and C. Ran, "Automatic Student Engagement Recognition in Online Classrooms Based on Deep Image Segmentation and Attention Mechanisms," vol. 42, no. 4, pp. 2169–2180, 2025.
- [16] H. A. Shehu, W. N. Browne, and H. Eisenbarth, "Emotion

- categorization from facial expressions: A review of datasets, methods, and research directions,” *Neurocomputing*, vol. 624, no. February 2024, 2025, doi: 10.1016/j.neucom.2025.129367. [19]
- [17] M. M. Santoni, T. Basaruddin, and K. Junus, “Convolutional Neural Network Model based Students’ Engagement Detection in Imbalanced DAiSEE Dataset,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 3, pp. 617–626, 2023, doi: 10.14569/IJACSA.2023.0140371. [20]
- [18] S. Singh and S. V. A. V. Prasad, “Techniques and challenges of face recognition: A critical review,” *Procedia Comput. Sci.*, vol. 143, pp. 536–543, 2018, doi: 10.1016/j.procs.2018.10.427.
- L. Li, X. Mu, S. Li, and H. Peng, “A Review of Face Recognition Technology,” *IEEE Access*, vol. 8, pp. 139110–139120, 2020, doi: 10.1109/ACCESS.2020.3011028.
- D. Sunaryono, J. Siswantoro, and R. Anggoro, “An android based course attendance system using face recognition,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 3, pp. 304–312, Mar. 2021, doi: 10.1016/j.jksuci.2019.01.006.