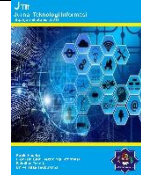


Terbit online pada laman: <http://jurnal.utu.ac.id/JTI>

## Jurnal Teknologi Informasi

ISSN (Online): 2829-8934



# Identifikasi Emosi Berdasarkan Citra Ekspresi Wajah Menggunakan Algoritma *Convolutional Neural Network* (CNN)

Melja Siska<sup>1</sup>, Rivansyah Suhendra<sup>2</sup>, Cut Mutia<sup>3</sup>, Gendhi Haris<sup>4</sup>, Nica Astrianda<sup>5</sup><sup>1,2,3,4</sup>Program Studi Teknologi Informasi, Universitas Teuku Umar, Aceh Barat, Indonesia

Jl. Alue Peunyareng, Ujong Tanoh Darat, Meurebo, Kabupaten Aceh Barat, Aceh, Indonesia

Email: [meljasiska07@gmail.com](mailto:meljasiska07@gmail.com), <sup>2</sup>[rivansyahsuhendra@utu.ac.id](mailto:rivansyahsuhendra@utu.ac.id), <sup>3</sup>[cutmutia@utu.ac.id](mailto:cutmutia@utu.ac.id), <sup>4</sup>[gendhi@utu.ac.id](mailto:gendhi@utu.ac.id), <sup>5</sup>[nicaastrianda@utu.ac.id](mailto:nicaastrianda@utu.ac.id)

## INFORMASI ARTIKEL

Sejarah Artikel:  
 Diterima: 6 Januari 2026  
 Revisi: 25 Mei 2026  
 Diterbitkan: 31 Mei 2026

Kata Kunci:  
 Identifikasi Emosi  
 Ekspresi Wajah  
*Convolutional Neural Network*  
*Visual Geometry Group 16*  
*Human Computer Interaction*

## ABSTRAK

Identifikasi emosi berdasarkan citra ekspresi wajah merupakan aspek penting dalam *Human-Computer Interaction* (HCI) karena mampu memberikan pemahaman terhadap respons emosional pengguna secara *non-verbal*. Penelitian ini bertujuan untuk menganalisis kinerja algoritma *Convolutional Neural Network* (CNN) dengan arsitektur VGG16 dalam mengklasifikasikan emosi wajah. Pertanyaan penelitian (*research question*) yang diajukan adalah: "Bagaimana kinerja arsitektur VGG16 dalam mengidentifikasi emosi dari citra ekspresi wajah?". Dataset yang digunakan terdiri dari 400 citra primer hasil pengambilan langsung dan 2.600 citra sekunder yang diperoleh dari *Kaggle*, dengan empat kelas emosi yaitu bahagia, sedih, marah, dan terkejut. Tahapan penelitian meliputi pra-pemrosesan citra berupa *cleaning* data, *resize* ke ukuran 224×224 piksel, *augmentasi*, dan *normalisasi*. Model dibangun melalui pendekatan *transfer learning* dengan membekukan *layer* blok 1-4 dan melatih ulang *layer* blok ke-5 VGG16, kemudian dilatih menggunakan *optimizer* Adam dan dievaluasi menggunakan metrik akurasi, *presisi*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa model CNN berbasis VGG16 mencapai akurasi pelatihan sebesar 91,03% dan akurasi validasi sebesar 88,14%. Model kemudian diimplementasikan ke dalam mini *website* sebagai media visualisasi hasil identifikasi emosi. Hasil penelitian ini menunjukkan bahwa CNN VGG16 efektif digunakan untuk identifikasi emosi wajah dan dapat dimanfaatkan sebagai evaluasi awal pengalaman pengguna dalam konteks HCI.

Copyright © 2026 Jurnal Teknologi Informasi UTU  
 All rights reserved

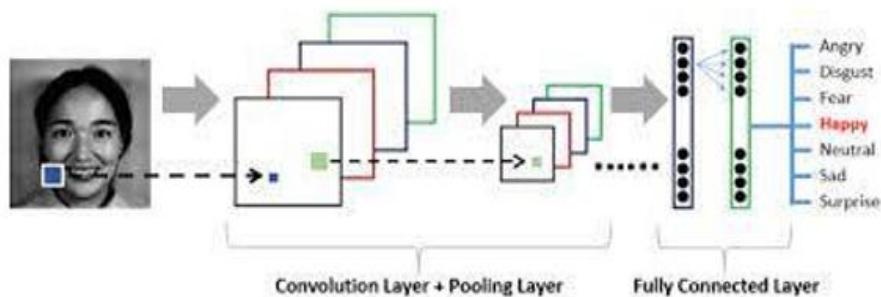
## 1. Pendahuluan

Perkembangan teknologi saat ini mendorong komputer untuk tidak hanya berfungsi sebagai alat pemrosesan data, tetapi juga mampu memahami kondisi dan respons pengguna. Salah satu aspek penting dalam bidang *Human-Computer Interaction* (HCI) adalah bagaimana sistem dapat merespons pengguna secara lebih alami melalui pengenalan ekspresi wajah. Ekspresi seperti bahagia, marah, sedih, dan terkejut merupakan bentuk komunikasi *non-verbal* yang umum muncul ketika seseorang berinteraksi dengan perangkat digital. HCI mempelajari bagaimana manusia berinteraksi dengan komputer, di mana antarmuka pengguna berperan sebagai media komunikasi yang memungkinkan manusia dan sistem saling bertukar informasi [1].

Wajah merupakan salah satu media utama manusia dalam menyampaikan informasi. Ekspresi wajah muncul sebagai respons terhadap perasaan dan situasi sosial yang dialami seseorang [2]. Secara psikologis, sejumlah emosi dasar bersifat universal dan dapat dikenali secara lintas budaya melalui

konfigurasi otot wajah, sebagaimana ditunjukkan pada studi klasik Ekman dan Friesen [12]. Dalam konteks HCI, ekspresi wajah menjadi indikator penting untuk memahami reaksi pengguna saat berinteraksi dengan sebuah sistem. Emosi senang dapat terlihat dari senyum atau ekspresi ceria ketika pengguna merasa nyaman dengan antarmuka, sedangkan emosi sedih atau tidak nyaman dapat tampak dari wajah yang murung atau alis yang berkerut ketika pengguna mengalami kebingungan. Pengenalan ekspresi wajah ini membantu sistem dalam memahami kebutuhan, tingkat kepuasan, serta kenyamanan pengguna secara lebih natural, sehingga mendukung peningkatan kualitas antarmuka dan pengalaman pengguna (user experience) [3].

Untuk melakukan identifikasi ekspresi wajah secara otomatis, diperlukan metode yang mampu mengekstraksi pola visual dari citra dengan akurat. Convolutional Neural Network (CNN) merupakan salah satu metode deep learning yang banyak digunakan dalam pengolahan citra karena kemampuannya dalam mengenali pola dan fitur penting pada gambar, seperti klasifikasi, deteksi objek, dan segmentasi citra [7], [8]. CNN memiliki tiga struktur utama, yaitu convolutional layer untuk mengekstraksi fitur, pooling layer untuk mereduksi dimensi, serta fully connected layer untuk menghasilkan prediksi akhir [4]. Proses kerja CNN memodelkan pola visual dalam bentuk jaringan tiga dimensi yang mencakup lebar, tinggi, dan kedalaman, di mana lebar dan tinggi merepresentasikan ukuran citra, sedangkan kedalaman menunjukkan jumlah filter yang digunakan dalam proses ekstraksi fitur [5]. Arsitektur umum Convolutional Neural Network dapat dilihat pada Gambar 1.



Gambar 1 Arsitektur Umum Convolutional Neural Network

Berdasarkan uraian tersebut, identifikasi emosi wajah berdasarkan citra ekspresi wajah penting dilakukan karena memungkinkan sistem memahami reaksi emosional pengguna secara *real-time*, sehingga kualitas interaksi manusia dan komputer dapat ditingkatkan. Dengan kemampuan ini, sebuah aplikasi dapat menilai tingkat kenyamanan, kebingungan, maupun kepuasan pengguna melalui ekspresi wajah yang ditampilkan. Pemahaman terhadap emosi pengguna tersebut menjadi dasar dalam pengembangan antarmuka yang lebih responsif, adaptif, dan berpusat pada pengguna.

Berdasarkan latar belakang tersebut, pertanyaan penelitian (*research question*) yang diajukan pada penelitian ini adalah: "Bagaimana kinerja arsitektur VGG16 dalam mengidentifikasi emosi dari citra ekspresi wajah?". Untuk menjawab pertanyaan tersebut, penelitian ini bertujuan untuk (1) membangun model klasifikasi emosi wajah berbasis CNN dengan arsitektur VGG16 melalui pendekatan *transfer learning*; (2) mengevaluasi kinerja model menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*; serta (3) mengimplementasikan model ke dalam mini *website* sebagai media visualisasi hasil identifikasi emosi. Dalam konteks penerapan HCI, ekspresi wajah berperan sebagai bentuk umpan balik *non-verbal* pengguna terhadap sistem dan dapat digunakan untuk mengevaluasi tingkat kepuasan pengguna saat berinteraksi dengan antarmuka aplikasi, layanan digital, maupun tampilan produk.

## 2. Tinjauan Pustaka

Penelitian mengenai identifikasi emosi berbasis ekspresi wajah (Facial Expression Recognition/ FER) telah berkembang seiring kemajuan metode machine learning dan deep learning. Beberapa pendekatan yang telah dilakukan dirangkum sebagai berikut.

Khaliqah dkk. [2] menerapkan algoritma K-Nearest Neighbor (K-NN) untuk mengklasifikasikan berbagai jenis ekspresi wajah manusia. Pendekatan berbasis machine learning klasik ini bergantung pada ekstraksi fitur secara manual sehingga kinerjanya terbatas ketika menghadapi variasi pencahayaan, pose, dan latar belakang. Bahri dkk. [3] mengembangkan pengenalan ekspresi wajah untuk identifikasi psikologis pengguna menggunakan Neural Network yang dipadukan dengan transformasi Ten Crops,

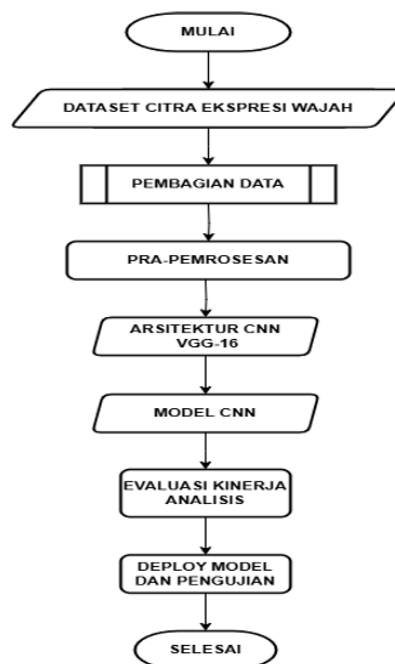
dan menunjukkan bahwa augmentasi pada tahap pra-pemrosesan berkontribusi terhadap peningkatan akurasi.

Guntoro dkk. [4] menggunakan Convolutional Neural Network untuk pengenalan ekspresi wajah dan menegaskan keunggulan CNN dalam mengekstraksi fitur citra secara otomatis tanpa perlu perancangan fitur manual. Pada skala yang lebih besar, Mollahosseini dkk. [11] memperlihatkan bahwa arsitektur CNN yang lebih dalam mampu meningkatkan kinerja FER pada beberapa dataset publik, sementara Li dan Deng [10] dalam studi survey-nya menyimpulkan bahwa tantangan utama FER berbasis deep learning meliputi keterbatasan jumlah data berlabel, variasi intra-kelas, serta ketidakseimbangan kelas. Dari sisi arsitektur, Simonyan dan Zisserman [6] memperkenalkan VGG16 yang membuktikan bahwa penggunaan filter konvolusi berukuran kecil ( $3 \times 3$ ) secara berlapis dapat meningkatkan kedalaman jaringan sekaligus akurasi klasifikasi citra. Keterbatasan jumlah data pada domain FER umumnya diatasi dengan pendekatan transfer learning [13], yaitu memanfaatkan bobot model yang telah dilatih pada dataset berskala besar seperti ImageNet [9] sebagai titik awal pelatihan.

Berbeda dengan penelitian-penelitian sebelumnya yang umumnya berhenti pada tahap evaluasi model, penelitian ini memberikan kontribusi berupa: (1) pemanfaatan transfer learning VGG16 dengan strategi pembekuan (freezing) layer blok 1–4 dan pelatihan ulang (fine-tuning) layer blok ke-5; (2) penggabungan data primer (pengambilan langsung) dan data sekunder (Kaggle) untuk empat kelas emosi; serta (3) implementasi model ke dalam mini website berbasis Streamlit sebagai media visualisasi hasil identifikasi emosi dalam konteks HCI. Rangkuman perbandingan dengan penelitian terdahulu disajikan pada Tabel 6.

### 3. Metodologi Penelitian

Penelitian ini dilakukan dengan beberapa tahapan utama yang disusun secara sistematis, dimulai dari pengumpulan data, pembagian data, hingga implementasi model dan pengujian. Keseluruhan rangkaian proses penelitian, termasuk tahapan pra-pemrosesan data, pembangunan model CNN, dan evaluasi kinerja, dapat dilihat pada Gambar 2.



Gambar 2 Alur Keseluruhan Penelitian

Gambar 2 menunjukkan alur keseluruhan proses penelitian identifikasi emosi berdasarkan citra ekspresi wajah, dimulai dari studi literatur, pengumpulan *dataset* citra ekspresi wajah, pembagian data, pra-pemrosesan, penerapan arsitektur CNN VGG16, pembangunan model CNN, evaluasi kinerja, dan terakhir implementasi model serta pengujian.

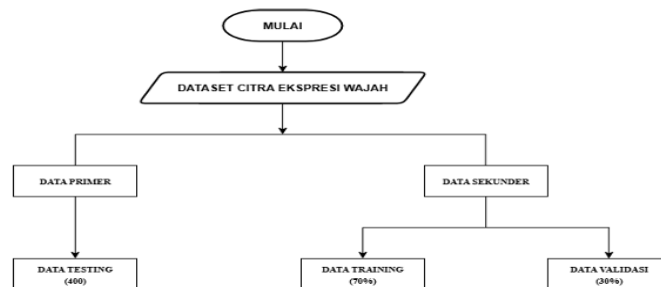
### 3.1. Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan melalui dua tahap, yaitu observasi dan studi literatur. Observasi dilakukan untuk memperoleh data citra ekspresi wajah secara langsung di lingkungan sekitar peneliti menggunakan kamera ponsel, serta melalui pengamatan terhadap gambar atau rekaman visual yang telah dikumpulkan. Selain itu, peneliti juga terlibat secara langsung dalam interaksi sosial guna menangkap ekspresi wajah yang sesuai dengan kebutuhan penelitian.

Studi literatur dilakukan dengan menelaah berbagai sumber ilmiah, seperti jurnal dan buku, yang berkaitan dengan jenis emosi, ekspresi wajah, serta teknik pengolahan citra dan identifikasi emosi. Studi ini bertujuan untuk memperkuat landasan teori dan mendukung proses analisis dalam pengembangan model identifikasi emosi.

### 3.2. Pembagian Data

Data pada penelitian ini dibagi menjadi data training, validasi, dan testing. Data training dan validasi menggunakan data sekunder sebanyak 2.600 citra dengan proporsi 70% (1.816 citra) untuk training dan 30% (784 citra) untuk validasi, sedangkan data testing menggunakan 400 citra data primer untuk pengujian kinerja model. Pembagian data dapat dilihat pada Gambar 3.



Gambar 3 Pembagian Data

### 3.3. Pra-Pemrosesan

Penelitian ini menerapkan beberapa tahapan pra-pemrosesan citra untuk mendukung proses identifikasi emosi berdasarkan ekspresi wajah menggunakan Convolutional Neural Network (CNN). Cleaning Data. Cleaning data dilakukan untuk menjaga kualitas citra dengan menghapus gambar yang buram, gelap, atau tidak mengandung wajah agar pelatihan model CNN lebih optimal. Resize. Resize citra dilakukan dengan mengubah ukuran gambar menjadi  $224 \times 224$  piksel agar sesuai dengan input CNN serta menjaga efisiensi komputasi dan kualitas fitur. Augmentasi Data. Tahap augmentasi dilakukan untuk menambah variasi data latih melalui teknik rotasi (rotation), flip horizontal, zoom, dan shift, sehingga mengurangi risiko overfitting dan meningkatkan kemampuan generalisasi model CNN. Normalisasi Data. Normalisasi citra dilakukan dengan mengubah nilai piksel ke rentang  $[0-1]$  melalui pembagian dengan 255 untuk menstabilkan dan mempercepat proses pelatihan model CNN.

### 3.4. Arsitektur VGG16 Dan Bangun Model CNN

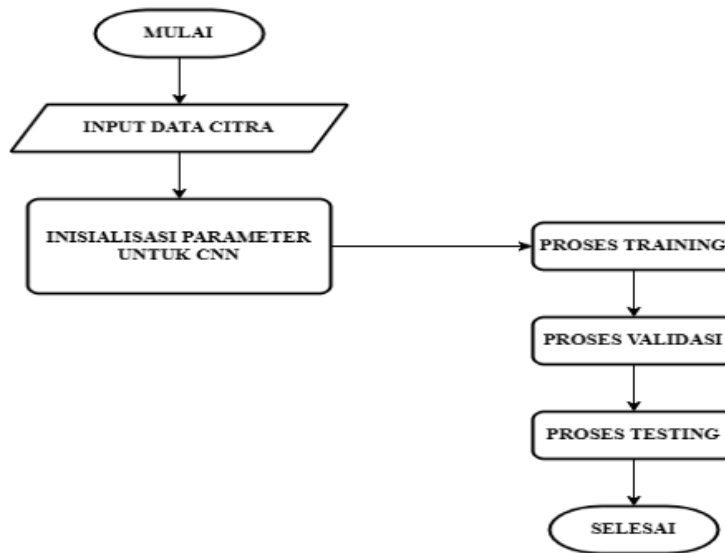
Arsitektur VGG16 yang dikembangkan oleh Visual Geometry Group Universitas Oxford [6] digunakan dalam penelitian ini karena kesederhanaan dan efektivitasnya dalam ekstraksi fitur citra.

Secara detail, VGG16 terdiri dari 16 layer berbobot, yakni 13 convolutional layer yang dikelompokkan ke dalam 5 blok konvolusi dan 3 fully connected layer. Seluruh convolutional layer menggunakan filter (kernel) berukuran  $3 \times 3$  dengan stride 1 dan padding "same", diikuti fungsi aktivasi ReLU. Setiap blok konvolusi diakhiri dengan operasi max pooling berukuran  $2 \times 2$  (stride 2) untuk mereduksi dimensi spasial secara bertahap. Susunan blok tersebut adalah: blok 1 dan 2 masing-masing berisi 2 layer konvolusi (64 dan 128 filter), sedangkan blok 3, 4, dan 5 masing-masing berisi 3 layer konvolusi (256, 512, dan 512 filter). Input jaringan berupa citra RGB berukuran  $224 \times 224 \times 3$ , sesuai dengan ukuran hasil tahap resize.

Proses dimulai dengan input citra ekspresi wajah yang telah melalui tahap normalisasi dan augmentasi, kemudian diproses menggunakan serangkaian convolutional layer dan max pooling untuk mengekstrak fitur penting seperti bentuk wajah, mata, dan bibir. Selanjutnya, fitur yang diperoleh diproses pada lapisan fully connected untuk menghasilkan probabilitas tiap kelas emosi.

Pada penelitian ini diterapkan pendekatan transfer learning dengan memanfaatkan bobot VGG16 hasil pra-pelatihan pada dataset ImageNet [9], [13]. Lapisan klasifikasi bawaan (top layer) VGG16 tidak digunakan (`include_top=False`) dan digantikan dengan lapisan klasifikasi baru. Dari 13 layer konvolusi, 10 layer konvolusi pada blok 1 hingga blok 4 dibekukan (`frozen`) sehingga bobotnya tidak diperbarui selama pelatihan, sedangkan 3 layer konvolusi pada blok ke-5 dilatih ulang (`fine-tuning`) agar model dapat menyesuaikan fitur tingkat tinggi dengan karakteristik citra ekspresi wajah. Di atas base model tersebut ditambahkan lapisan klasifikasi baru berupa: Flatten → Dense 128 neuron (aktivasi ReLU) → Dropout 0,5 → Dense 4 neuron (aktivasi softmax) untuk mengklasifikasikan empat kelas emosi.

Model dilatih menggunakan data training untuk mempelajari pola ekspresi wajah melalui proses konvolusi, aktivasi, dan pooling sehingga mampu mengenali karakteristik visual emosi secara akurat. Alur proses pembangunan model CNN dapat dilihat pada Gambar 4.



Gambar 4 Alur Proses Bangun Model CNN

Gambar 4 menunjukkan alur kerja pembangunan model CNN yang dimulai dari *input* data citra dan inialisasi parameter, dilanjutkan dengan proses pelatihan untuk mengekstraksi fitur, serta evaluasi melalui validasi dan pengujian. Hasil akhir berupa keluaran identifikasi emosi yang disimpan atau ditampilkan, menunjukkan proses yang sistematis dalam membangun dan mengevaluasi kinerja model CNN.

### 3.5. Evaluasi Kinerja Analisis

Tahap evaluasi dilakukan untuk menilai kinerja model CNN dalam mengidentifikasi emosi berdasarkan citra ekspresi wajah. Evaluasi dilakukan menggunakan data validasi dan data uji guna memastikan kemampuan generalisasi model. Kinerja model diukur menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*, serta didukung oleh analisis *confusion matrix* untuk mengevaluasi kesalahan klasifikasi pada tiap kelas emosi. Akurasi prediksi dari suatu pengidentifikasi dapat ditampilkan dalam bentuk tabel kontingensi atau *confusion matrix* seperti pada Tabel 1.

Tabel 1 *Confusion Matrix*

	Prediksi (+)	Prediksi (-)

Actual (+)	TP (True Positif)	FN (False Negatif)
Actual (-)	FP (False Positif)	TN (True Negatif)

**1. Akurasi**

Akurasi digunakan untuk mengukur tingkat ketepatan model dalam mengklasifikasikan emosi secara keseluruhan.

$$\text{Akurasi} = \frac{\text{jumlah identifikasi benar}}{\text{jumlah total data}} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

**2. Presisi**

Presisi menunjukkan proporsi prediksi positif yang benar terhadap seluruh prediksi positif yang dihasilkan model.

$$\text{Presisi} = \frac{TP}{TP+FP} \tag{2}$$

**3. Recall**

Recall mengukur kemampuan model dalam mengenali seluruh data positif dari suatu kelas emosi.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

**4. F1-Score**

F1-Score merupakan rata-rata harmonik antara presisi dan recall untuk menilai keseimbangan kinerja model.

$$\text{F1-Score} = F1 \frac{2PR}{P+R} \tag{4}$$

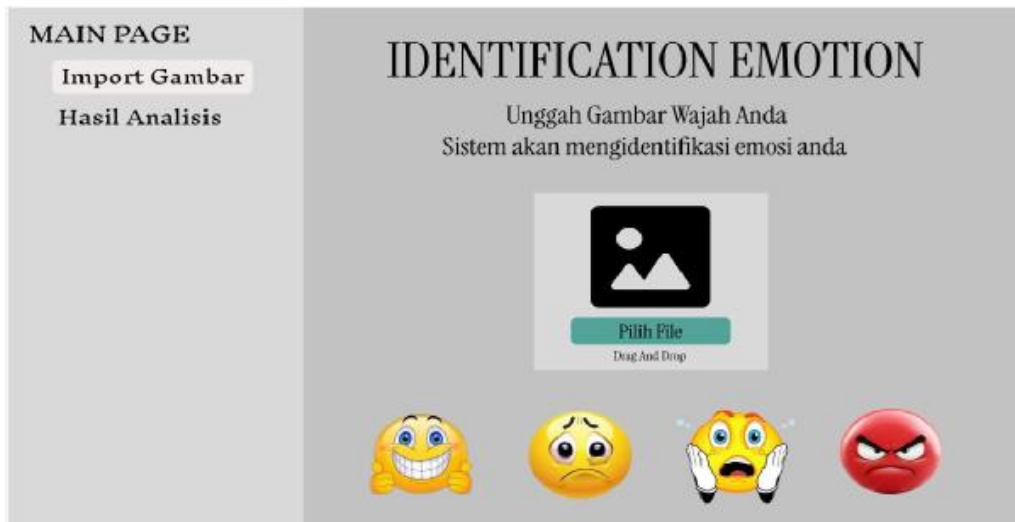
**3.6. Implementasi Model dan Pengujian**

Implementasi model dilakukan dengan menyematkan model CNN terlatih ke dalam sebuah mini *website* sederhana untuk memvisualisasikan hasil identifikasi emosi; mini *website* ini berfungsi sebagai media uji dan tidak ditujukan sebagai sistem aplikasi produksi. Model diuji menggunakan 400 citra data *testing* yang terpisah dari data *training* untuk menilai kemampuan generalisasi model terhadap data baru, dengan hasil identifikasi emosi ditampilkan melalui antarmuka sederhana. Ilustrasi rancangan implementasi model dapat dilihat pada Gambar 5.

**4. Hasil dan Pembahasan**

**4.1. Hasil Pengumpulan Data**

Data primer berupa citra ekspresi wajah dengan 4 kelas emosi (bahagia, sedih, marah, dan terkejut) yang diambil langsung menggunakan kamera ponsel. Data *testing* terdiri dari 400 citra. Selain itu, digunakan data sekunder dari *Kaggle* sebagai pelengkap karena diperlukan jumlah data yang besar untuk memperoleh akurasi yang lebih baik. Adapun jumlah citra dapat dilihat pada Tabel 2.




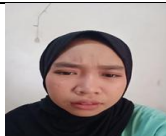






Gambar 5 Ilustrasi Rancangan *Deploy Model Mini Website*

Tabel 2 Jumlah Citra

No	Kelas Ekspresi Wajah	Data Sekunder	Data Primer
1	Emosi Bahagia	650	400
2	Emosi Marah	650	
3	Emosi Sedih	650	
4	Emosi Terkejut	650	
Total		3000	

Tabel 3 Citra Dataset

Kelas	Data Sekunder	Data Primer
Bahagia		
Marah		
Sedih		
Terkejut		

## 4.2. Hasil *Pra*-Pemrosesan

### 1. *Cleanning* Data

Data *cleaning* adalah proses membersihkan data citra wajah agar siap digunakan dalam penelitian. Tahapan ini meliputi seleksi data, penghapusan gambar yang rusak atau berkualitas rendah, penyesuaian pencahayaan dan warna, pemotongan gambar agar fokus pada wajah, serta penghapusan artefak yang mengganggu. Data yang telah dibersihkan kemudian disimpan secara rapi dan siap digunakan untuk pelatihan CNN.



(a)



(b)

Gambar 6 (a) Citra yang tidak digunakan, (b) Citra yang digunakan

## 2. *Resize*

Citra di-*resize* menggunakan metode *bilinear interpolation* agar ukuran seragam dan tetap halus. Proses *resize* dilakukan dengan *library OpenCV* pada *Python*, sehingga dihasilkan citra wajah berukuran  $224 \times 224$  piksel dalam format RGB yang siap digunakan untuk pelatihan CNN.



(a)



(b)

Gambar 7 (a) Sebelum *Resize* (960x1280), (b) Sesudah *Resize* (224x224)

## 3. Augmentasi Data

Teknik augmentasi yang digunakan meliputi *rotation*, *horizontal flip*, *zooming*, dan *shift*. Proses ini bertujuan untuk meningkatkan keragaman citra sehingga model tidak mudah mengalami *overfitting*.

## 4. Normalisasi Data

Normalisasi citra dilakukan dengan mengubah nilai piksel ke dalam rentang  $[0-1]$ , yaitu dengan membagi nilai intensitas setiap piksel menggunakan angka 255. Tahap ini berfungsi untuk menyeragamkan skala data sehingga dapat mempercepat proses pelatihan model CNN.

### 4.3. Hasil Penerapan Model CNN VGG16 *Layer* Tambahan

Proses identifikasi emosi dilakukan menggunakan CNN dengan arsitektur VGG16 sebagai ekstraktor fitur citra wajah  $224 \times 224$  piksel. Model digunakan tanpa *top layer*, dengan sebagian besar *layer* awal dibekukan dan beberapa *layer* akhir dilatih. Selanjutnya ditambahkan *layer* klasifikasi berupa *Flatten*, *Dense 128 neuron* dengan *ReLU*, *Dropout 0.5*, serta *Dense* akhir 4 *neuron* dengan *softmax* untuk mengklasifikasikan empat emosi.

Model: "sequential"

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 7, 7, 512)	14,714,688
Flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 128)	3,211,392
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 4)	516

Total params: 17,926,596 (68.38 MB)  
 Trainable params: 10,291,332 (39.26 MB)  
 Non-trainable params: 7,635,264 (29.13 MB)

Gambar 8 Model Summary VGG16 Dengan Layer Tambahan

Arsitektur pada gambar 9 dirancang dengan penyesuaian *hyperparameter* melalui proses *trial and error* untuk memperoleh konfigurasi VGG16 yang optimal dalam identifikasi emosi. Hasil *tuning* tersebut menghasilkan kombinasi *hyperparameter* yang digunakan pada pelatihan model, sebagaimana ditampilkan pada Tabel 4.

Tabel 4 *Hyperparameter* model CNN VGG16 Layer Tambahan

<i>Hyperparameter</i>	
<i>Split Data</i>	70% Train, 30% Val
<i>Layers</i>	5
<i>Epoch</i>	10
<i>Optimizer</i>	Adam
<i>Learning Rate</i>	0.0001

Konfigurasi *hyperparameter* pada pelatihan model VGG16 meliputi pembagian data 70% untuk *training* dan 30% untuk validasi, dengan *batch size* 64. Arsitektur terdiri dari VGG16 sebagai *ekstraktor* fitur, *Flatten*, *Dense* beraktivasi *ReLU*, *Dropout*, dan *Dense* akhir beraktivasi *softmax* untuk empat kelas emosi. Pelatihan dilakukan selama 10 epoch menggunakan *optimizer Adam* dengan *learning rate* 0.0001 agar pembelajaran berlangsung stabil dan optimal.

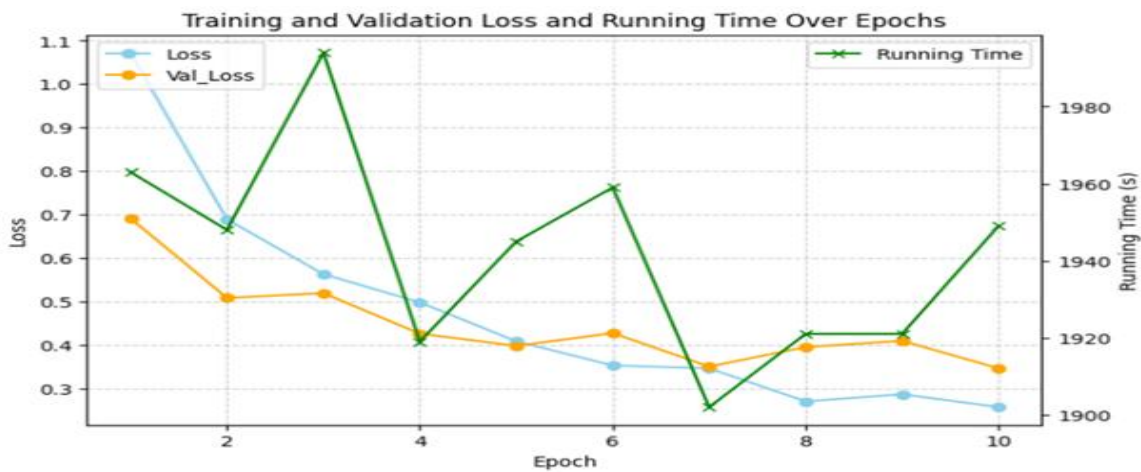
#### 4.4. Hasil Proses Training Model CNN VGG16 Layer Tambahan

Pelatihan model CNN berbasis VGG16 dilakukan selama 10 epoch menggunakan data latih dan validasi. Akurasi validasi tertinggi dicapai pada *epoch* ke-10 sebesar 88.14% dengan *val\_loss* 0.3472, sehingga model pada *epoch* ini disimpan sebagai model terbaik. Akurasi pelatihan tertinggi juga terjadi pada *epoch* ke-10, yaitu 91.03% dengan *loss* 0.2708.

Secara umum, akurasi meningkat signifikan pada *epoch* awal, stabil pada *epoch* menengah, dan melambat pada *epoch* akhir hingga mencapai performa optimal tanpa indikasi *overfitting* yang berarti. Hal ini menunjukkan bahwa model mampu belajar dengan baik dan memiliki kemampuan generalisasi yang baik. Rincian hasil pelatihan dan akurasi per *epoch* ditampilkan pada tabel 5.

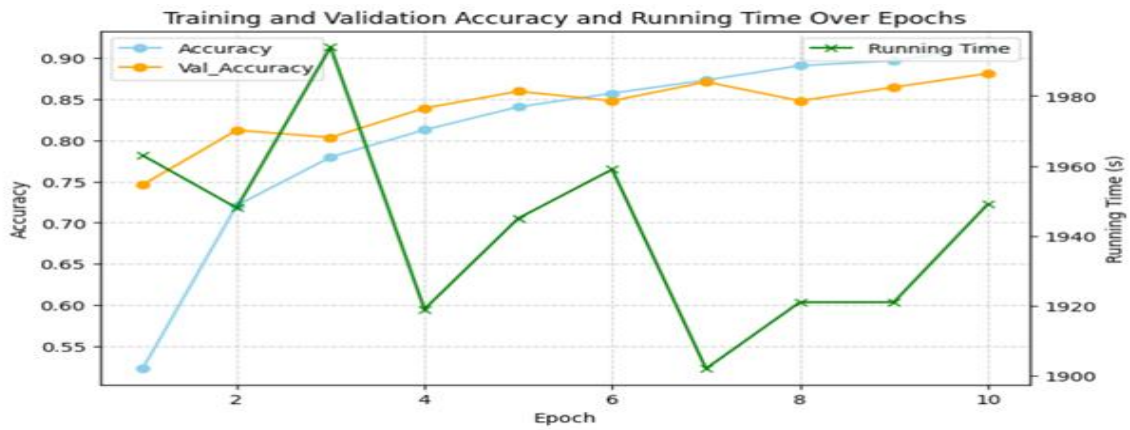
Tabel 5 *Training Model CNN VGG16 Layer Tambahan*

NO	Epoch	Running Time	Training		Validasi	
			Accuracy	Loss	Val-Accuracy	Val-Loss
1	Epoch 1	1963s	0.4143	1.2785	0.7462	0.6920
2	Epoch 2	1948s	0.7045	0.7140	0.8125	0.5088
3	Epoch 3	1994s	0.7787	0.5691	0.8036	0.5200
4	Epoch 4	1919s	0.7993	0.5359	0.8393	0.4268
5	Epoch 5	1945s	0.8388	0.4217	0.8597	0.3985
6	Epoch 6	1959s	0.8508	0.3721	0.8482	0.4280
7	Epoch 7	1902s	0.8773	0.3409	0.8712	0.3511
8	Epoch 8	1921s	0.9130	0.2250	0.8482	0.3959
9	Epoch 9	1921s	0.9055	0.2657	0.8648	0.4101
10	Epoch 10	1949s	0.9103	0.2708	0.8814	0.3472
			Total training akurasi = 91.03%		Total Validasi Akurasi = 88.14%	



Gambar 9 Hasil *Training And Val-Loss And Running Time Over Epochs*

Gambar 9 memperlihatkan hubungan *training loss*, *validation loss*, dan *running time* selama pelatihan model VGG16 selama 10 *epoch*. *Training loss* menurun secara konsisten, sedangkan *validation loss* turun tajam di awal, sedikit berfluktuasi, lalu kembali menurun di akhir, menunjukkan kemampuan generalisasi yang baik. *Running time per epoch* relatif stabil pada kisaran 1900–1990 detik, dengan waktu tercepat pada *epoch* ke-7. Secara keseluruhan, hasil ini menunjukkan proses pelatihan yang efektif, stabil, dan tanpa indikasi *overfitting*.

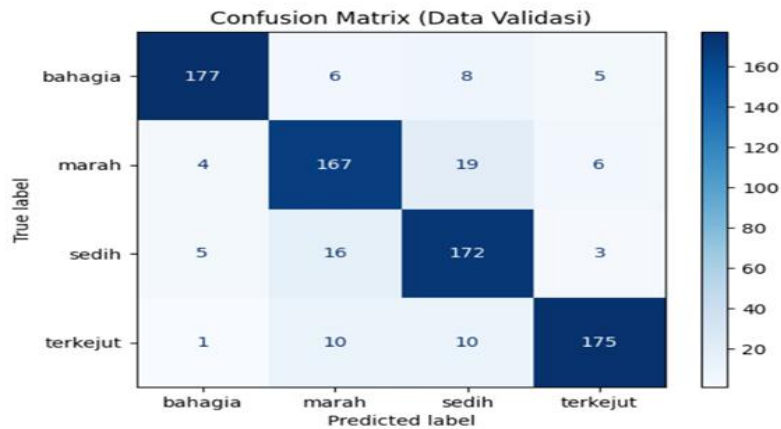


Gambar 10 Hasil Training And Validation Accuracy And Running Time Over Epoch

Gambar 10 menunjukkan peningkatan *training accuracy* dari sekitar 52% menjadi 91% dan *validation accuracy* dari sekitar 75% menjadi 88% selama 10 *epoch*, menandakan proses pembelajaran dan generalisasi model yang baik. *Running time per epoch* relatif stabil pada kisaran 1900–1990 detik, dengan waktu tercepat pada *epoch* ke-7. Secara keseluruhan, grafik ini menunjukkan pelatihan yang stabil, efektif, dan tanpa indikasi *overfitting*.

#### 4.5. Hasil Pengujian Model

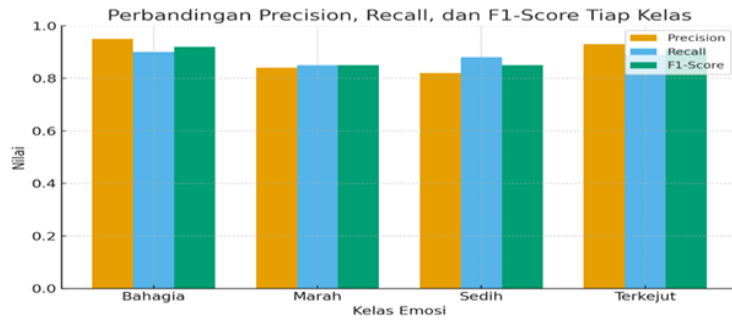
Evaluasi performa model dilakukan menggunakan *confusion matrix* dengan 784 data validasi. Pengujian ini menghasilkan metrik akurasi, presisi, *recall*, dan *F1-score* untuk menilai kemampuan model VGG16 dalam mengklasifikasikan emosi. Visualisasi hasil evaluasi ditampilkan pada Gambar 11.



Laporan Klasifikasi (Data Validasi):

	precision	recall	f1-score	support
bahagia	0.95	0.90	0.92	196
marah	0.84	0.85	0.85	196
sedih	0.82	0.88	0.85	196
terkejut	0.93	0.89	0.91	196
accuracy			0.88	784
macro avg	0.88	0.88	0.88	784
weighted avg	0.88	0.88	0.88	784

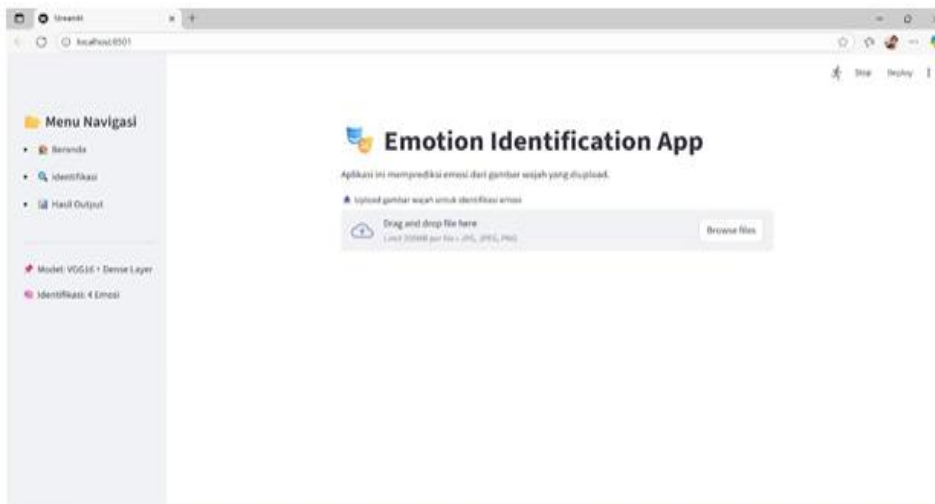
Gambar 11 Hasil Evaluasi Pengujian Model



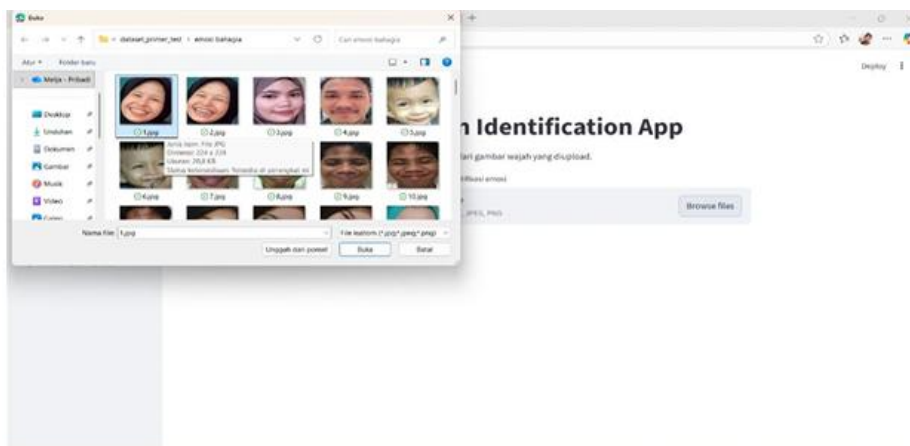
Gambar 12 Perbandingan *Precision*, *Recal*, Dan *F1-Score* Tiap Kelas

#### 4.6. Hasil Implementasi Model

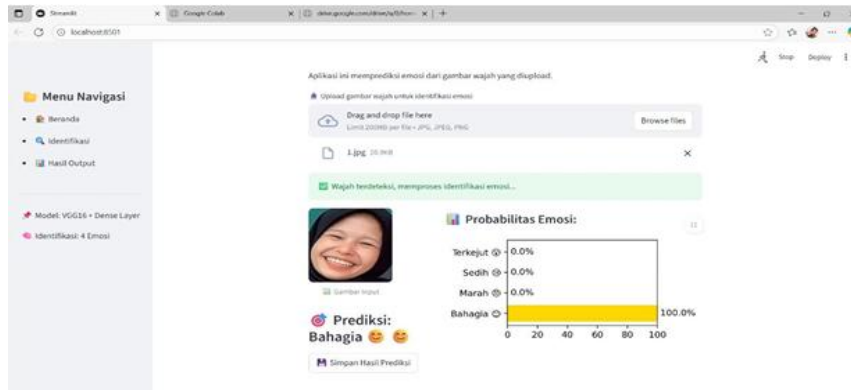
Pada tahap ini ditampilkan hasil penerapan model melalui antarmuka uji yang dibuat menggunakan library Streamlit pada Python. Antarmuka ini menyediakan menu identifikasi emosi untuk menampilkan hasil prediksi emosi dari citra wajah, dengan tampilan beranda aplikasi emotion identification ditunjukkan pada Gambar 13.



Gambar 13 Tampilan Beranda *Deploy Model Emotion Identification App*



Gambar 14 Tampilan *Import Citra Bahagia Deploy Model Emotion Identification app*



Gambar 15 Hasil Prediksi Bahagia *Deploy Model Emotion Identification App*

Gambar 15 menunjukkan hasil prediksi *deploy* model yang mengklasifikasikan citra wajah ke dalam empat emosi. Citra uji teridentifikasi sebagai emosi bahagia dengan probabilitas 100%, sedangkan kelas marah, sedih, dan terkejut bernilai 0%, sehingga sistem menyimpulkan emosi bahagia.

#### 4.7. Perbandingan dengan Penelitian Lain

Untuk menempatkan hasil penelitian ini dalam konteks yang lebih luas, dilakukan perbandingan pendekatan dengan beberapa penelitian terdahulu pada bidang pengenalan ekspresi wajah, sebagaimana disajikan pada Tabel 6.

Tabel 6. Perbandingan Pendekatan dengan Penelitian Terdahulu

Peneliti (Tahun)	Metode	Karakteristik Pendekatan
Khaliqah dkk. (2024) [2]	<i>K-Nearest Neighbor (K-NN)</i>	<i>Machine learning</i> klasik; bergantung pada ekstraksi fitur manual.
Bahri dkk. (2022) [3]	<i>Neural Network + Ten Crops</i>	Fokus identifikasi psikologis; memanfaatkan augmentasi <i>Ten Crops</i> .
Guntoro dkk. (2022) [4]	<i>Convolutional Neural Network</i>	Ekstraksi fitur otomatis melalui CNN tanpa perancangan fitur manual.
Mollahosseini dkk. (2016) [11]	<i>Deep CNN (Inception)</i>	Jaringan lebih dalam untuk FER pada beberapa <i>dataset</i> publik.
Penelitian ini	CNN VGG16 ( <i>transfer learning</i> )	4 kelas emosi; blok 1–4 dibekukan, blok 5 di- <i>fine-tune</i> ; diimplementasikan ke mini <i>website</i> berbasis <i>Streamlit</i> .

Berdasarkan Tabel 6, pendekatan CNN berbasis VGG16 dengan *transfer learning* pada penelitian ini menunjukkan kinerja yang kompetitif untuk klasifikasi empat kelas emosi, dengan akurasi validasi mencapai 88,14%. Dibandingkan dengan metode *machine learning* klasik seperti K-NN [2] yang mengandalkan ekstraksi fitur manual, pendekatan berbasis CNN mampu mengekstraksi fitur secara otomatis dan hierarkis sehingga lebih tahan terhadap variasi citra. Penggunaan *transfer learning* dari bobot *ImageNet* [9], [13] juga membantu model mencapai akurasi tinggi meskipun jumlah data relatif terbatas — sebuah tantangan yang umum ditemui pada domain FER [10]. Selain itu, penelitian ini melangkah lebih jauh dari sekadar evaluasi model dengan mengimplementasikan hasil ke dalam mini *website* sebagai media visualisasi dalam konteks HCI. Perlu dicatat bahwa perbandingan langsung antar penelitian bersifat indikatif karena adanya perbedaan *dataset*, jumlah kelas, dan skema evaluasi.

## 5. Kesimpulan

Hasil penelitian menunjukkan bahwa *Convolutional Neural Network* (CNN) dengan arsitektur VGG16 mampu mengidentifikasi emosi berdasarkan citra ekspresi wajah secara efektif. Penerapan tahapan *pra*-pemrosesan, yaitu *cleaning* data, *resize* 224×224 piksel, augmentasi, dan normalisasi, mendukung ekstraksi fitur wajah yang relevan untuk klasifikasi empat emosi dasar (bahagia, sedih, marah, dan terkejut). Model mencapai akurasi pelatihan sebesar 91,03%, akurasi validasi 88,14%, dan nilai *F1-score* rata-rata 0,88, serta menunjukkan performa yang konsisten pada pengujian menggunakan 400 citra *primer*. Implementasi model melalui mini *website* berbasis *Streamlit* membuktikan bahwa sistem dapat memvisualisasikan hasil identifikasi emosi secara interaktif dan mendukung penerapan dalam konteks *Human Computer Interaction* (HCI)

## Daftar Pustaka

- [1] J. Preece, H. Sharp, dan Y. Rogers, *Interaction Design: Beyond Human-Computer Interaction*, 5th ed. Hoboken, NJ, USA: John Wiley & Sons, 2019.
- [2] M. Khaliqah, L. Sarifah, dan S. Khotijah, "Implementasi Algoritma K-Nearest Neighbor (K-NN) dalam Mengklasifikasikan Berbagai Jenis Ekspresi Wajah Manusia," *Zeta – Math J.*, vol. 9, no. 1, pp. 10–20, 2024, doi: 10.31102/zeta.2024.9.1.10-20.
- [3] S. Bahri, R. Samsinar, dan P. S. Denta, "Pengenalan Ekspresi Wajah untuk Identifikasi Psikologis Pengguna dengan Neural Network dan Transformasi Ten Crops," *Resistor (Elektronika Kendali Telekomunikasi Tenaga Listrik Komputer)*, vol. 5, no. 1, pp. 15–20, 2022, doi: 10.24853/resistor.5.1.15-20.
- [4] A. L. Sigit Guntoro, E. Julianto, dan D. Budiayanto, "Pengenalan Ekspresi Wajah Menggunakan Convolutional Neural Network," *J. Inform. Atma Jogja*, vol. 3, no. 2, pp. 155–160, 2022, doi: 10.24002/jiaj.v3i2.6790.
- [5] Y. LeCun, L. Bottou, Y. Bengio, dan P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.
- [6] K. Simonyan dan A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–14.
- [7] Y. LeCun, Y. Bengio, dan G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [8] I. Goodfellow, Y. Bengio, dan A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [9] A. Krizhevsky, I. Sutskever, dan G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 25, 2012, pp. 1097–1105.
- [10] S. Li dan W. Deng, "Deep Facial Expression Recognition: A Survey," *IEEE Trans. Affective Comput.*, vol. 13, no. 3, pp. 1195–1215, 2022, doi: 10.1109/TAFFC.2020.2981446.
- [11] A. Mollahosseini, D. Chan, dan M. H. Mahoor, "Going Deeper in Facial Expression Recognition using Deep Neural Networks," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, 2016, pp. 1–10, doi: 10.1109/WACV.2016.7477450.
- [12] P. Ekman dan W. V. Friesen, "Constants across cultures in the face and emotion," *J. Pers. Soc. Psychol.*, vol. 17, no. 2, pp. 124–129, 1971, doi: 10.1037/h0030377.
- [13] S. J. Pan dan Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.
- [14] D. P. Kingma dan J. Ba, "Adam: A Method for Stochastic Optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–15.
- [15] I. J. Goodfellow *et al.*, "Challenges in representation learning: A report on three machine learning contests," *Neural Networks*, vol. 64, pp. 59–63, 2015, doi: 10.1016/j.neunet.2014.09.005.