

KLASIFIKASI EKSPRESI WAJAH MENGGUNAKAN COVOLUTIONAL NEURAL NETWORK

Ahmad Taufiq Akbar^{*1}, Shoffan Saifullah², Hari Prapcoyo³

^{1,3}Universitas Pembangunan Nasional Veteran Yogyakarta, Kabupaten Sleman

²AGH Univerisity of Krakow, Poland

Email: ¹ahmadtaufiq.akbar@upnyk.ac.id, ²saifulla@agh.edu.pl, ³hari.prapcoyo@upnyk.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 23 Maret 2023, diterima untuk diterbitkan: 26 November 2024)

Abstrak

Pengenalan ekspresi wajah adalah tantangan penting dalam pengolahan citra dan interaksi manusia-komputer karena kompleksitas dan variasi yang ada. Penelitian ini mengusulkan arsitektur sederhana Convolutional Neural Network (CNN) untuk meningkatkan efisiensi klasifikasi emosi pada dataset kecil. Dataset yang digunakan adalah Jaffe, yang terdiri dari 213 citra berukuran 256x256 piksel dalam tujuh kategori ekspresi. Citra-citra tersebut di-resize menjadi 128x128 piksel untuk mempercepat pemrosesan. Data diproses menggunakan arsitektur CNN yang terdiri dari 3 lapisan konvolusi, 2 lapisan subsampling, dan 2 lapisan dense. Kami mengevaluasi model dengan 5-fold dan 10-fold cross-validation untuk estimasi kinerja yang robust, serta teknik hold-out (70:30, 80:20, 85:15, dan 90:10) untuk perbandingan hasil yang jelas. Hasil menunjukkan akurasi tertinggi sebesar 90.6% dengan *learning rate* 0.001 pada pembagian 85% data latih dan 15% data uji, melebihi model yang lebih kompleks. Meskipun tidak menggunakan transfer learning atau augmentasi data, model ini tetap unggul dibandingkan pendekatan tradisional seperti Local Binary Pattern (LBP) dan Histogram Oriented Gradient (HOG). Dengan demikian, arsitektur CNN yang sederhana ini terbukti efektif untuk pengenalan ekspresi wajah pada dataset kecil.

Kata kunci: CNN, ekspresi, wajah, deep learning, pengolahan citra

FACIAL EXPRESSION CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORKS

Abstract

Facial expression recognition is a significant challenge in image processing and human-computer interaction due to its inherent complexity and variability. This study proposes a simple Convolutional Neural Network (CNN) architecture to enhance the efficiency of emotion classification on small datasets. Jaffe's dataset consists of 213 images sized 256x256 pixels across seven expression categories. These images were resized to 128x128 pixels to accelerate processing. The data was processed using a CNN architecture comprising 3 convolutional layers, 2 subsampling layers, and 2 dense layers. We evaluated the model with 5-fold- and 10-fold cross-validation for robust performance estimation and hold-out techniques (70:30, 80:20, 85:15, and 90:10) for clear result comparison. The results indicated the highest accuracy of 90.6% with a learning rate of 0.001 using the 85% training and 15% testing data split, surpassing that of more complex models. Although the model does not employ transfer learning or data augmentation, it still outperforms traditional approaches such as Local Binary Pattern (LBP) and Histogram Oriented Gradient (HOG). Thus, this simple CNN architecture proves effective for facial expression recognition on small datasets.

Keywords: CNN, expression, facial, deep learning, image processing

1. PENDAHULUAN

Ekspresi wajah berperan penting dalam komunikasi interpersonal, membantu menafsirkan status emosi, aktivitas kognitif, kepribadian, dan psikopatologi seseorang. Ekspresi wajah juga berkontribusi besar terhadap komunikasi verbal (De La Torre and Cohn, 2011). Menurut (Swapna

Subudhiray et al., 2023), ekspresi wajah menyampaikan 55% informasi emosi, sementara suara menyumbang 38% dan bahasa hanya 7%. Kompleksitas dan keragaman ekspresi wajah membuatnya tetap menjadi topik penelitian yang menarik (Appasaheb Borgalli & Surve, 2022).

Penelitian terkini semakin berfokus pada interaksi manusia-komputer yang cerdas, sehingga

komputer dapat beroperasi dengan mempertimbangkan ekspresi wajah (Martinez & Valstar, 2016; McDuff et al., 2013). Berbagai implementasi sistem berbasis ekspresi wajah telah dikembangkan di berbagai bidang, termasuk pendidikan (Krithika et al., 2017; Tonguç & Ozaydin Ozkara, 2020), kesehatan (Altameem & Altameem, 2020; Leo et al., 2020), industri (Busso & Jain, 2012), dan hiburan (M. T. Akbar et al., 2019; Almeida et al., 2021). Namun, pengenalan ekspresi wajah tetap menjadi domain penelitian yang kompleks karena belum ditemukan fitur standar terbaik yang dapat mencirikan setiap jenis ekspresi, terutama dalam variasi kondisi lingkungan dan pencahayaan yang berbeda (Alreshidi & Ullah, 2020; Mehta et al., 2018; S Subudhiray et al., 2021). Selain itu, gambar dengan resolusi rendah dan sudut pemotretan wajah yang bervariasi juga menyulitkan pengenalan. Dataset kecil membatasi proses pelatihan dan dapat menyebabkan overfitting (Li et al., 2018).

Sistem pengenalan ekspresi wajah umumnya difokuskan pada tujuh kategori dasar ekspresi: netral, marah, jijik, takut, senang, sedih, dan terkejut, yang diterima secara luas oleh masyarakat (Li et al., 2018; Ng et al., 2015). Ketujuh ekspresi tersebut terdapat dalam dataset FER 2013 dan Jaffe. Dataset FER 2013 memiliki sampel yang cukup banyak dengan kondisi wajah dalam beragam lingkungan, posisi, dan pencahayaan. Sementara itu, Jaffe memuat ketujuh ekspresi dengan posisi wajah yang teratur di depan kamera (Li et al., 2018).

Dalam beberapa penelitian terbaru, metode deep learning telah menjadi populer dalam pengenalan ekspresi wajah. Dengan semakin banyaknya lapisan konvolusi, fitur yang diperoleh menjadi semakin kompleks (Jeong et al., 2021; S. Xie & Hu, 2017). Namun, kelemahan dari deep learning adalah kurangnya efektivitas pada dataset kecil, yang dapat mengakibatkan overfitting akibat kapasitas pelatihan yang terbatas (Appasaheb Borgalli & Surve, 2022; Li et al., 2018).

Beberapa penelitian sebelumnya telah menggunakan dataset Jaffe dengan berbagai teknik. Misalnya, beberapa peneliti melakukan cropping dengan OpenCV dan mengekstraksi fitur menggunakan arsitektur CNN gabungan dari model Tang (Tang, 2013) dan CaffeNet yang telah dilatih sebelumnya dengan ImageNet (Jia et al., 2014). Hasil ekstraksi fitur kemudian diklasifikasikan dengan L2-SVM, menghasilkan akurasi rata-rata 95,7% melalui 10-fold cross-validation (Li et al., 2018). Penelitian lain yang menggunakan metode ekstraksi fitur facial landmark pada dataset Jaffe, melaporkan akurasi tertinggi 76% menggunakan Extreme Learning Machine (ELM) dengan 5-fold cross-validation (Bachtiar & Wafi, 2021).

Penelitian oleh (Akhand et al., 2021) mengklasifikasikan dataset Jaffe dengan arsitektur CNN VGG16 yang dilatih dengan ImageNet. Pada

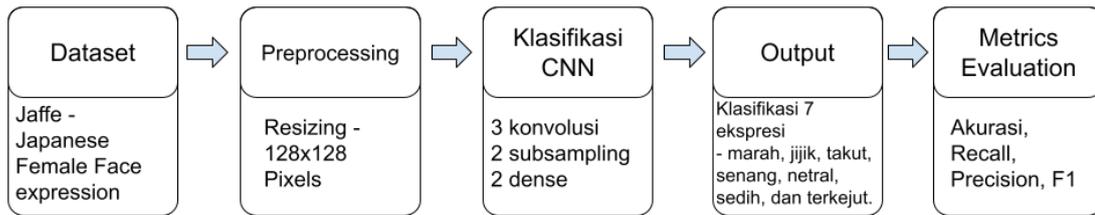
lapisan convolutional block 1 hingga 4, arsitektur VGG16 berfungsi sebagai feature extractor dalam kondisi frozen, sedangkan pada block 5 dalam kondisi fine-tuned untuk memperbarui bobot dengan dataset Jaffe. Penelitian tersebut melaporkan akurasi tertinggi 100% untuk 10% data testing dan 99,52% untuk 10-fold cross-validation.

Pemilihan parameter terbaik berdasarkan teknik pembagian data sangat penting untuk memastikan model dapat memanfaatkan semua data yang tersedia dan meningkatkan generalisasi. Menurut beberapa studi, meskipun dataset kecil, model dapat berfungsi secara optimal dengan pemilihan parameter yang tepat dan teknik validasi yang sesuai (Vabalas et al., 2019; Xu & Goodacre, 2018; Zhang & Ling, 2018). Dengan menggunakan teknik cross-validation dan hold-out (Mattiev et al., 2022; Utami & Mustakim, 2021), penelitian ini bertujuan untuk mengoptimalkan pembelajaran model meskipun menggunakan dataset yang kecil.

Penelitian ini bertujuan untuk mengembangkan arsitektur CNN yang sederhana guna mengetahui seberapa optimal hasil klasifikasi pada dataset kecil seperti Jaffe, meskipun proses training tidak menggunakan augmentasi data. Augmentasi data dapat meningkatkan waktu komputasi dan kebutuhan sumber daya yang lebih besar. Selain itu, semakin kompleks arsitektur CNN, semakin besar kebutuhan data training atau augmentasi untuk mencapai performa klasifikasi yang optimal (Akhand et al., 2021). Penelitian ini akan melakukan pemilihan parameter terbaik untuk mengklasifikasikan dataset Jaffe dengan 5-fold dan 10-fold cross-validation serta pembagian data dengan komposisi 80% training dan 20% testing serta 90% dan 10%. Parameter yang akan dipilih mencakup nilai epoch dan *learning rate* pada arsitektur CNN. Hasil klasifikasi akan dievaluasi dan dibandingkan dengan penelitian sebelumnya, diharapkan penelitian ini dapat memberikan kontribusi pada pengembangan model selanjutnya.

2. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan utama, yang mencakup pengumpulan dataset, preprocessing data, pembangunan model arsitektur Convolutional Neural Network (CNN), pemilihan parameter yang sesuai, pembagian data untuk pelatihan dan pengujian, klasifikasi, serta evaluasi hasil klasifikasi. Gambar 1 menggambarkan alur penelitian secara keseluruhan, dimulai dari tahap pengumpulan dataset Jaffe, yang berisi citra ekspresi wajah, hingga tahapan akhir evaluasi menggunakan berbagai metrik seperti akurasi, precision, recall, dan F1-score. Setiap tahapan ini dirancang untuk memastikan bahwa model yang dihasilkan dapat mengenali ekspresi wajah dengan akurasi yang optimal.



Gambar 1. Alur Penelitian untuk Pengenalan Ekspresi Wajah Menggunakan CNN

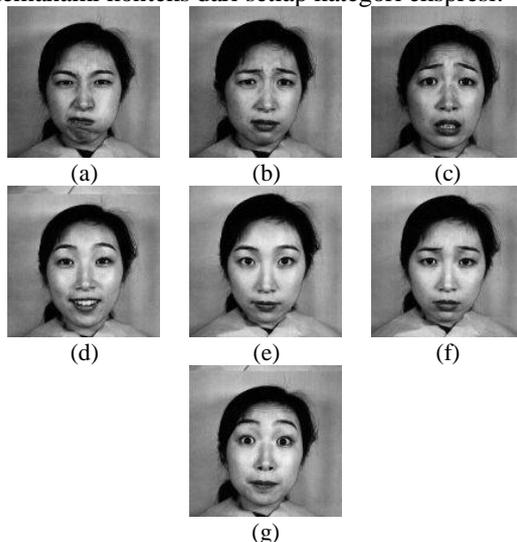
2.1. Dataset

Dataset yang digunakan dalam penelitian ini adalah Jaffe (Japanese Female Facial Expression), yang diperoleh dari situs <http://www.kasrl.org/Jaffe.html>. Dataset ini terdiri dari citra berukuran 256x256 piksel dalam tujuh kategori ekspresi wajah, yaitu marah (*anger*), jijik (*disgust*), takut (*fear*), senang (*happy*), netral (*neutral*), sedih (*sadness*), dan terkejut (*surprised*). Tabel 1 menyajikan jumlah sampel untuk setiap kategori ekspresi.

Tabel 1. Rancangan Analisis Komputasi (8pt, ditengah)

Kategori Ekspresi	Jumlah Citra
Marah (<i>Anger</i>)	30
Jijik (<i>Disgust</i>)	29
Takut (<i>Fear</i>)	32
Senang (<i>Happy</i>)	31
Netral (<i>Neutral</i>)	30
Sedih (<i>Sadness</i>)	31
Terkejut (<i>Surprised</i>)	30
Total	213

Setiap kategori dalam dataset mencerminkan variasi ekspresi yang penting untuk analisis emosi. Selain itu, Gambar 2 menunjukkan contoh citra untuk masing-masing kategori ekspresi, memberikan gambaran visual yang jelas tentang variasi dalam dataset. Citra ini menyoroti elemen-elemen penting, seperti posisi alis, bentuk mulut, dan sudut mata, yang membantu dalam identifikasi emosi. Dengan adanya contoh citra ini, pembaca dapat lebih memahami konteks dari setiap kategori ekspresi.



Gambar 2. Sampel citra dengan 7 ekspresi: (a) marah, (b) jijik, (c) takut, (d) senang, (e) netral, (f) sedih, dan (g) terkejut

2.2. Pra-pemrosesan

Pada tahap pra-pemrosesan, penelitian ini melakukan *resizing* citra untuk mengubah ukuran semua gambar menjadi 128x128 piksel. Ukuran ini dipilih berdasarkan studi sebelumnya yang menunjukkan bahwa penggunaan ukuran citra ini dapat menghasilkan akurasi yang cukup baik, yakni 87.5% pada rasio 85% data pelatihan dan 15% data pengujian (Akhand et al., 2021; Rere et al., 2019). Penelitian terdahulu mencatat bahwa model CNN tidak selalu optimal untuk citra dengan ukuran lebih besar. Semakin besar ukuran citra, semakin banyak informasi yang perlu dilatih, yang bisa menambah kompleksitas proses pelatihan (Akhand et al., 2021).

Dalam konteks ini, pemilihan parameter terbaik berdasarkan teknik pembagian data sangat penting untuk mengoptimalkan performa model (Almabdy & Elrefaei, 2019; Kahloot & Ekler, 2021). Penelitian menunjukkan bahwa penggunaan dataset yang lebih besar umumnya menghasilkan model yang lebih baik, tetapi teknik pembagian data yang tepat juga dapat membantu memaksimalkan efisiensi belajar dari dataset kecil (Althnian et al., 2021; Davila Delgado & Oyedele, 2021; Shaikhina & Khovanova, 2017). Oleh karena itu, ukuran citra 128x128 piksel dipertahankan untuk memastikan kesesuaian dengan arsitektur CNN yang diusulkan, sambil menghindari proses ekstraksi fitur tradisional seperti Local Binary Pattern (LBP) (Shan et al., 2009; Swapna Subudhiray et al., 2023), Facial Landmark (Bachtiar & Wafi, 2021), Histogram of Oriented Gradients (HOG) (Eng et al., 2019), dan Gabor (Swapna Subudhiray et al., 2023). Dengan tidak menggunakan teknik ekstraksi fitur ini, penelitian bertujuan untuk mendapatkan performa murni dari arsitektur CNN yang dikembangkan.

Penelitian ini tidak menggunakan augmentasi data dan transfer learning (A. T. Akbar et al., 2024) pada model CNN agar dapat menguji performa murni dari arsitektur CNN dalam mengklasifikasikan dataset kecil seperti Jaffe. Beberapa penelitian menunjukkan bahwa augmentasi data dapat meningkatkan waktu komputasi dan kebutuhan sumber daya (Alomar et al., 2023)(Alomar et al., 2023; Appasaheb Borgalli & Surve, 2022). Selain itu, mengingat dataset yang kecil, augmentasi data bisa menjadi kontraproduktif karena dapat menyebabkan distorsi informasi yang tidak diinginkan, yang pada gilirannya dapat menurunkan akurasi model (Tatar et al., 2024).

Dengan demikian, meskipun augmentasi dapat bermanfaat dalam banyak kasus, untuk penelitian ini, kami memilih untuk fokus pada penggunaan data asli demi menjaga integritas fitur yang diidentifikasi oleh model. Selain itu, semakin kompleks arsitektur CNN, semakin besar kebutuhan data training atau augmentasi untuk mencapai performa klasifikasi yang optimal (Akhand et al., 2021).

2.3. Model Arsitektur CNN

Usulan model arsitektur CNN dalam penelitian ini dirancang dengan mempertimbangkan jumlah lapisan dan ukuran filter yang optimal untuk dataset kecil. Model ini terdiri dari tiga lapisan konvolusi dan dua lapisan dense, yang merupakan konfigurasi sederhana namun efektif. Dalam perbandingan, penelitian oleh (Akhand et al., 2021) menggunakan arsitektur VGG16 yang lebih kompleks dengan hasil akurasi yang tinggi, namun memerlukan lebih banyak data untuk pelatihan. Sementara itu, penelitian oleh (Rere et al., 2019) menunjukkan bahwa arsitektur yang lebih sederhana dapat memberikan hasil yang kompetitif pada dataset kecil.

Metode Convolutional Neural Network (CNN) merupakan pendekatan yang sangat penting dalam deep learning, khususnya untuk pengolahan citra (Rere et al., 2019; Saifullah et al., 2023). CNN menggabungkan lapisan-lapisan jaringan konvolusi untuk ekstraksi fitur citra dengan jaringan saraf untuk klasifikasi (Liew et al., 2016). Model arsitektur CNN yang diusulkan dalam penelitian ini, sebagaimana ditunjukkan pada Gambar 3, terdiri dari beberapa lapisan sebagai berikut:

1. **Lapisan Konvolusi 1:** Menggunakan 6 kernel filter berukuran 5x5 dengan padding 0 dan stride 1. Persamaan konvolusi untuk lapisan ini adalah (1).

$$x_j^1 = f\left(\sum_{i \in M_1} x_i^0 * k_{ij}^1 + b_j^1\right) \quad (1)$$

Di mana x_j^1 adalah output dari neuron j pada lapisan konvolusi pertama, dan f adalah fungsi aktivasi yang digunakan (tanh).

2. **Lapisan Subsampling 1:** Menggunakan pooling rata-rata (average pooling) dengan ukuran 2x2, padding 0, dan stride 2. Proses pooling dinyatakan dengan (2).

$$y_{m,n} = \frac{1}{4} \sum_{i=0}^1 \sum_{j=0}^1 x_{2m+1,2n+j} \quad (2)$$

3. **Lapisan Konvolusi 2:** Menggunakan 10 filter berukuran 5x5 dengan padding 0 dan stride 1. Rumus konvolusi untuk lapisan ini adalah (3).

$$x_j^2 = f\left(\sum_{i \in M_2} x_i^1 * k_{ij}^2 + b_j^2\right) \quad (3)$$



Gambar 3. Arsitektur Model CNN untuk Pengenalan Ekspresi Wajah Mendalam Penelitian ini

4. **Lapisan Subsampling 2:** Memiliki spesifikasi yang sama dengan lapisan subsampling pertama, dengan pooling rata-rata 2x2 dan stride 2.
5. **Lapisan Konvolusi 3:** Menggunakan 1 filter berukuran 5x5 dengan padding 0 dan stride 1. Rumusnya adalah (4).

$$x_j^3 = f\left(\sum_{i \in M_3} x_i^2 * k_{ij}^3 + b_j^3\right) \quad (4)$$

6. **Lapisan Dense 1 (Fully Connected):** Menghubungkan semua neuron dari lapisan sebelumnya ke 512 neuron. Persamaannya adalah (5).

$$z_j^1 = f\left(\sum_{i=1}^N x_i^3 \cdot w_{ij}^1 + b_j^1\right) \quad (5)$$

7. **Lapisan Dense 2 (Fully Connected):** Menghubungkan 512 neuron ke 128 neuron dengan rumus (6).

$$z_j^2 = f\left(\sum_{i=1}^{512} z_i^1 \cdot w_{ij}^2 + b_j^2\right) \quad (6)$$

8. **Lapisan Output:** Mengklasifikasikan output ke dalam 7 kelas ekspresi dengan fungsi aktivasi softmax (7).

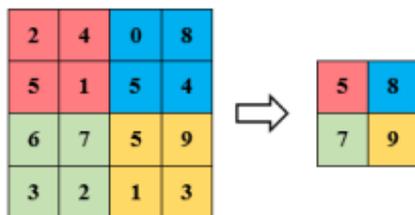
$$y_k = s(z_k^2) = \frac{e^{z_k^2}}{\sum_{j=1}^7 e^{z_j^2}} \quad (7)$$

Lapisan konvolusi memiliki peran penting dalam CNN sebagai pengekstraksi fitur (Suhirman et al., 2024). Proses konvolusi dilakukan antara nilai-nilai piksel citra dengan nilai-nilai dalam matriks filter (kernel), menghasilkan peta fitur (Rere et al., 2019). Semakin banyak jumlah kernel filter yang digunakan, semakin kompleks pula model arsitektur CNN (Khanday et al., 2021). Jumlah filter dalam model arsitektur CNN ini dibatasi untuk menjaga performa optimal dalam mengklasifikasikan dataset kecil seperti Jaffe tanpa menggunakan augmentasi data dan transfer learning.

Setiap keluaran fitur map dilengkapi dengan parameter bias b . Untuk beberapa keluaran fitur map, masukan fitur map akan dikonvolusi dengan kernel yang spesifik. Oleh karena itu, jumlah lapisan konvolusi dianggap sebagai hyperparameter (Liew et al., 2016; Wafaa Shihab & Abdul amir A., 2020).

Pooling Layer atau Lapisan Subsampling: Setelah proses aktivasi dalam lapisan konvolusi, dilakukan subsampling pada lapisan pooling. Proses pooling memperkecil ukuran matriks dari peta fitur lapisan sebelumnya dengan cara pooling maksimum atau pooling rata-rata (Zainudin et al., 2019). Hasil pooling ini berupa matriks berdimensi lebih kecil, dan proses ini penting untuk mendapatkan ukuran peta fitur yang memadai untuk diproses pada lapisan fully connected (Akhand et al., 2021; Rere et al., 2019).

Proses pooling dan konvolusi melibatkan ukuran stride. Stride adalah parameter yang menentukan pergeseran filter di dalam piksel citra. Nilai stride 1 menggeser filter satu pixel secara vertikal dan horizontal. Jika nilai stride kecil, fitur yang diperoleh semakin detail, tetapi memerlukan komputasi yang lebih besar. Nilai stride kecil dapat menghindari informasi piksel yang terbuang (Wafaa Shihab & Abdul amir A., 2020). Gambar 4 menunjukkan proses pooling berdasarkan nilai maksimum atau max pooling.



Gambar 4. *Subsampling atau pooling*

Pada Gambar 4, terlihat bahwa peta fitur berukuran 4x4, yang merupakan output dari lapisan konvolusi, terbagi dalam 4 grid karena menggunakan ukuran pooling 2x2. Setiap grid pooling 2x2 akan mengambil nilai piksel maksimum, kemudian bergeser sejauh 2 piksel untuk

mengambil nilai maksimum dalam grid berikutnya, menghasilkan peta fitur berukuran 2x2.

Fully Connected Layer: Lapisan ini sering disebut sebagai dense layer. Di dalamnya terdapat operasi pembelajaran fitur dengan metode backpropagation untuk klasifikasi (Liew et al., 2016). Penambahan lapisan ini meningkatkan kompleksitas pembelajaran dan membutuhkan sumber daya komputasi yang signifikan (Rere et al., 2019). Hasil dari proses pada dense layer terakhir kemudian diaktivasi dengan fungsi softmax, sehingga diperoleh probabilitas kelas tertinggi untuk data yang diuji.

Fungsi Aktivasi: Semua lapisan sebelum layer terakhir menggunakan aktivasi tanh seperti pada persamaan (8).

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

Semua lapisan sebelum layer terakhir menggunakan aktivasi tanh seperti pada persamaan 8. Jenis aktivasi ini dipilih karena merupakan aktivasi dalam arsitektur Lenet 5 (Z. Xie et al., 2019). yang merupakan arsitektur CNN dengan komputasi lebih ringan dibandingkan arsitektur lainnya. Sedangkan aktivasi softmax seperti dalam persamaan 9, terdapat pada lapisan yang terdapat pada *dense layer* terakhir.

$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (9)$$

Dimana, x merupakan input dari fungsi aktivasi, sedangkan s adalah output probabilitas berdasarkan kelas i dan n totalnya.

2.1. Evaluasi Model

Setelah proses pelatihan model CNN selesai, langkah selanjutnya adalah melakukan evaluasi untuk menilai kinerja model dalam mengklasifikasikan ekspresi wajah berdasarkan dataset Jaffe. Evaluasi dilakukan dengan menggunakan metrik yang umum dalam pengolahan citra dan pembelajaran mesin, seperti akurasi, precision, recall, dan F1-score (Saeed et al., 2022). Metrik-metrik ini memberikan gambaran menyeluruh tentang performa model dalam mengenali berbagai ekspresi wajah.

1. **Akurasi** persentase prediksi yang benar dari total prediksi yang dilakukan oleh model. Dihitung dengan rumus (10).

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Prediksi}} \times 100\% \quad (10)$$

2. **Precision** menunjukkan seberapa banyak prediksi positif yang benar dari seluruh prediksi positif yang dibuat. Rumusnya adalah (11).

$$\text{Precision} = \frac{TP}{TP+FP} \tag{11}$$

Dimana *TP* adalah True Positives dan *FP* adalah False Positives.

- Recall** mengukur kemampuan model untuk menemukan semua instance positif. Rumusnya adalah (12)

$$\text{Recall} = \frac{TP}{TP+FN} \tag{12}$$

Di mana *FN* adalah False Negatives.

- F1-Score** rata-rata harmonis antara precision dan recall, memberikan gambaran tentang keseimbangan antara keduanya. Rumusnya adalah (13)

$$F1 = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 2 \tag{13}$$

3. HASIL DAN PEMBAHASAN

Pada bagian ini, kami menganalisis hasil pengujian model arsitektur Convolutional Neural Network (CNN) yang diusulkan (Gambar 5) untuk pengenalan ekspresi wajah menggunakan dataset Jaffe. Berbagai kombinasi hyperparameter, seperti *learning rate* dan jumlah epoch, dievaluasi untuk menilai akurasi model. Hasil akan dibandingkan dengan penelitian sebelumnya, serta dievaluasi menggunakan metrik umum dalam pengolahan citra, termasuk akurasi, recall, precision, dan F1-score. Bab ini bertujuan untuk memberikan wawasan tentang kinerja model dalam klasifikasi emosi berdasarkan ekspresi wajah.

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 124, 124, 6)	456
average_pooling2d (AveragePooling2D)	(None, 62, 62, 6)	0
conv2d_1 (Conv2D)	(None, 58, 58, 10)	1510
average_pooling2d_1 (AveragePooling2D)	(None, 29, 29, 10)	0
conv2d_2 (Conv2D)	(None, 25, 25, 1)	251
flatten (Flatten)	(None, 625)	0
dense (Dense)	(None, 512)	320512
dense_1 (Dense)	(None, 128)	65664
dense_2 (Dense)	(None, 7)	903

```

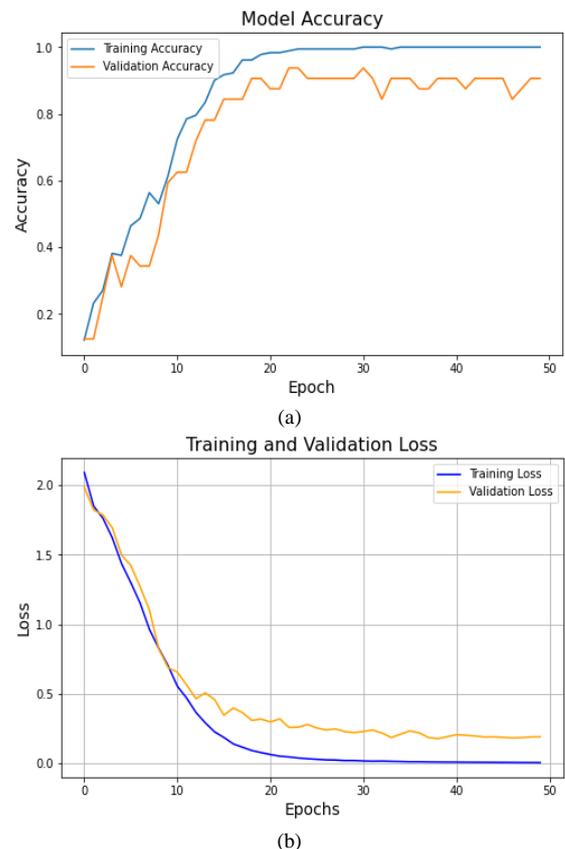
Total params: 389,296
Trainable params: 389,296
Non-trainable params: 0
    
```

Gambar 5. Model Arsitektur yang diusulkan

3.1. Hasil Pengujian

Hasil penelitian pada model yang diusulkan menunjukkan variasi akurasi yang tergantung pada kombinasi parameter yang digunakan, sebagaimana

ditunjukkan dalam Tabel 2. Pengujian pada skenario 9, yang membagi data dengan proporsi 85% untuk pelatihan dan 15% untuk pengujian, mencapai nilai akurasi tertinggi sebesar 90,625%, seperti pada grafik pada Gambar 6. Gambar 6 menunjukkan grafik akurasi dan loss model selama pelatihan. Subplot (a) menggambarkan peningkatan akurasi untuk data pelatihan (biru) dan validasi (oranye) seiring bertambahnya epoch, menandakan kemampuan model dalam belajar. Subplot (b) menunjukkan penurunan loss untuk kedua data, mengindikasikan efektivitas model dalam mengurangi kesalahan prediksi. Keduanya mencerminkan kinerja model yang baik dalam pengenalan ekspresi wajah. Hal ini menunjukkan bahwa model dapat memanfaatkan sebagian besar data untuk pelatihan, menghasilkan kinerja yang optimal pada data yang tidak terlihat sebelumnya.



Gambar 6. Grafik Akurasi dan Loss Model

Skenario 1 (50 epoch) menunjukkan akurasi 86.385%, sedangkan skenario 2 (250 epoch) memiliki akurasi 87.3239%. Selisih antara kedua skenario ini adalah 0.94%, yang mengindikasikan bahwa peningkatan jumlah epoch tidak selalu berbanding lurus dengan peningkatan akurasi. Sebaliknya, skenario 3, yang menggunakan 10-fold cross-validation dengan 65 epoch, mencapai akurasi 89.2019%. Ini menunjukkan bahwa metode validasi silang dapat memberikan estimasi kinerja yang lebih stabil dan lebih representatif dibandingkan dengan pembagian data yang sederhana.

Tabel 2. Rancangan Analisis Komputasi

Skenario	Metode Pengujian	Kombinasi Hyperparameter		Parameter Hasil Pengujian			
		<i>Learning rate</i>	Epoch	Akurasi (%)	Recall(%)	Precision (%)	F1 (%)
1	5-fold cross validation	0.001	50	86.385	86.4	86.5	86.2
2	5-fold cross validation	0.001	250	87.3239	87.3	87.4	87.2
3	10-fold cross validation	0.001	65	89.2019	89.2	89.4	89.1
4	10-fold cross validation	0.001	50	87.7934	87.8	87.9	87.7
5	Split 70% : 30%	0.001	70	82.8125	82.8	84.9	82.7
6	Split 70% : 30%	0.0001	50	64.0625	64.1	67.0	63.2
7	Split 70% : 30%	0.001	50	82.8125	82.8	85.0	82.6
8	Split 80% : 20%	0.001	50	83.7209	83.7	87.7	84
9	Split 85% : 15%	0.001	50	90.625	90.6	92.4	90.6
10	Split 85% : 15%	0.001	90	90.625	90.6	92.4	90.6
11	Split 90% : 10%	0.001	65	85.7143	85.7	88.5	85.7
12	Split 90% : 10%	0.001	50	76.1905	76.2	84.5	76.6
13	Split 90% : 10%	0.001	90	80.9524	81.0	86.1	80.7

Pada skenario 5, di mana pelatihan dilakukan dengan 70% data dan pengujian 30%, didapatkan akurasi sebesar 82.8125% pada epoch 70. Namun, skenario 6 dengan *learning rate* 0.0001 dan epoch 50 untuk pembagian yang sama hanya menghasilkan akurasi 64.0625%. Penurunan akurasi ini menunjukkan bahwa pengaturan *learning rate* yang terlalu kecil dapat mengakibatkan model tidak belajar dengan baik dari data yang tersedia, sehingga menghasilkan performa yang buruk.

Dari hasil di atas, nilai *learning rate* yang kecil tidak selalu meningkatkan pembelajaran dalam mengurangi overfitting. Penelitian menunjukkan bahwa semakin kecil *learning rate*, semakin detail proses pembelajaran yang dilakukan, namun hal ini juga dapat menyebabkan penurunan kinerja pada fungsi loss (Zainorzuli et al., 2022). Skenario 9, yang menunjukkan akurasi 90.625%, mirip dengan pengujian yang dilakukan oleh Rere, Usna, dan Soegijanto (2019), di mana akurasi penelitian ini lebih tinggi dibandingkan dengan akurasi penelitian sebelumnya yang mencapai 87%.

Dengan data uji yang lebih sedikit, seperti pada skenario 11 hingga 13 (90% data latih dan 10% data uji), akurasi tercatat sebesar 85.7% pada epoch 65. Namun, pada epoch 50, nilai akurasi turun menjadi 76.19%. Ketika epoch ditingkatkan menjadi 90, akurasi kembali meningkat menjadi 80.95%, tetapi tetap di bawah akurasi skenario 7. Hal ini menunjukkan bahwa hiperparameter epoch dapat berfluktuasi dan tidak selalu menghasilkan akurasi yang lebih baik pada nilai epoch yang lebih tinggi. Hasil pengujian dalam penelitian ini lebih lanjut dapat dilihat pada Gambar 5.

Grafik pada Gambar 7 menunjukkan kolom paling kiri yang mewakili skenario 1 hingga 13, di mana skenario tanpa mencantumkan epoch dan *learning rate* menggunakan epoch 50 dan *learning rate* 0.001 (default). Setiap skenario memiliki hasil nilai akurasi sesuai dengan warna grafik masing-masing, dengan akurasi tertinggi pada skenario 9 dan 10.

3.2. Analisis Parameter Model dan *Learning rate*

Dalam penelitian ini, analisis parameter model berfokus pada jumlah filter dalam lapisan konvolusi dan pengaturan *learning rate*, yang keduanya merupakan elemen kunci dalam mempengaruhi kinerja model CNN yang diusulkan.

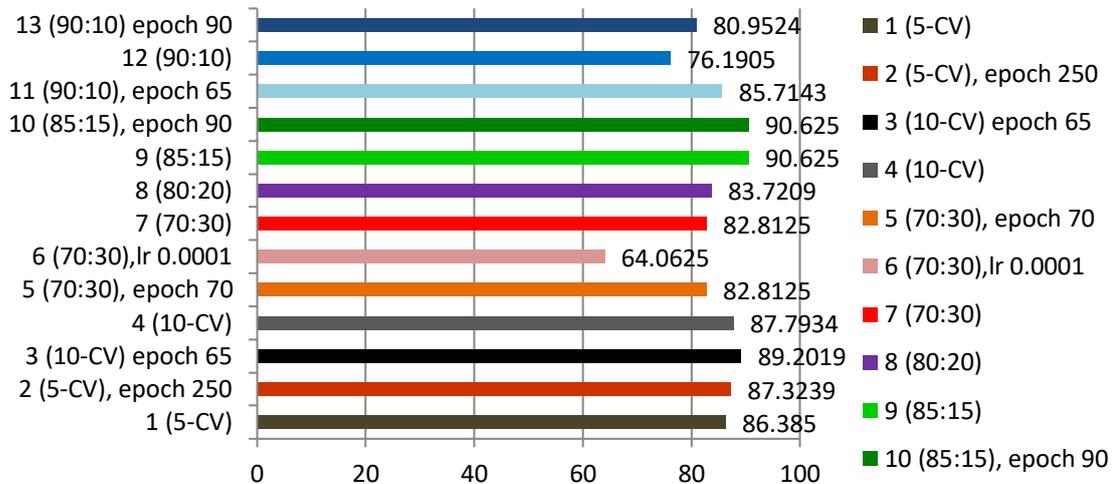
Jumlah Filter

Parameter jumlah filter pada lapisan konvolusi terakhir ditetapkan sebanyak satu. Hal ini bertujuan untuk memperkecil total parameter model, sehingga mengurangi kompleksitas komputasi dan risiko overfitting. Penelitian sebelumnya (Rere et al., 2019) menunjukkan bahwa penggunaan jumlah filter konvolusi yang lebih banyak tidak memberikan peningkatan akurasi yang signifikan dibandingkan dengan model yang diusulkan dalam penelitian ini. Sebaliknya, meningkatnya jumlah filter justru berpotensi memperbesar total parameter dan kebutuhan komputasi, yang dapat menjadi beban pada sumber daya yang tersedia (Liew et al., 2016).

Penentuan jumlah filter yang optimal sangat bergantung pada karakteristik spesifik dari tugas dan dataset yang diproses. Penambahan filter dapat membantu model dalam mempelajari representasi fitur yang lebih kompleks, yang penting untuk mengidentifikasi pola dalam data. Namun, jika terlalu banyak filter digunakan, ada risiko bahwa model dapat menghafal data pelatihan tanpa dapat menggeneralisasi dengan baik pada data baru, sebuah kondisi yang dikenal sebagai overfitting (Wafaa Shihab & Abdul amir A., 2020; Z. Xie et al., 2019).

Oleh karena itu, penting untuk melakukan eksplorasi beberapa konfigurasi jumlah filter melalui validasi silang (cross-validation). Ini akan memungkinkan evaluasi kinerja model dengan variasi jumlah filter, memastikan bahwa model tidak terlalu kompleks untuk dataset yang relatif kecil. Dalam konteks dataset besar, seperti ImageNet, model dapat memanfaatkan kompleksitas arsitektur CNN yang lebih tinggi, seperti VGG16 (Akhand et al., 2021), tanpa mengkhawatirkan overfitting.

Nilai Akurasi (%)



Gambar 7. Grafik Pengujian dari 13 Skenario

Learning rate

Pengaturan *learning rate* juga menjadi faktor krusial dalam proses pelatihan model. Dalam penelitian ini, *learning rate* terbaik yang digunakan adalah 0.001. Pada skenario 6, ketika *learning rate* diturunkan menjadi 0.0001, akurasi model justru menurun. Hal ini menunjukkan bahwa semakin kecil *learning rate* tidak selalu menjamin optimalisasi pembelajaran, dan dalam banyak kasus dapat mengakibatkan kinerja yang kurang baik pada fungsi loss.

Dengan demikian, untuk penelitian ini, pemilihan kombinasi parameter model—termasuk jumlah filter dan *learning rate*—merupakan langkah kritis dalam mencapai akurasi yang diinginkan, serta menjaga keseimbangan antara kompleksitas model dan kemampuan generalisasi terhadap data baru.

3.3. Perbandingan dengan Penelitian Sebelumnya

Dalam konteks pengembangan model pengenalan ekspresi wajah, hasil penelitian ini dibandingkan dengan hasil dari berbagai penelitian sebelumnya untuk memberikan wawasan yang lebih baik mengenai efektivitas dan kinerja model yang diusulkan (Tabel 3). Sebagaimana ditunjukkan pada Tabel 2, model yang diusulkan mencapai nilai akurasi maksimum sebesar 90.625% pada skenario 9, di mana data pelatihan dan pengujian dibagi dalam proporsi 85% dan 15%. Meskipun pencapaian ini signifikan, akurasi ini masih lebih rendah dibandingkan dengan beberapa penelitian lain yang menerapkan pendekatan lebih kompleks. Contohnya, penelitian oleh (Akhand et al., 2021) menggunakan model *pre-trained* yang menghasilkan akurasi hingga 100%. Penggunaan model yang telah dilatih sebelumnya menunjukkan bahwa memanfaatkan bobot dari dataset yang lebih besar dan kompleks

dapat meningkatkan kinerja model pada dataset yang lebih kecil.

Namun, salah satu kelebihan utama penelitian ini adalah model sederhana yang diusulkan mampu menghasilkan kinerja kompetitif meskipun tidak menggunakan teknik augmentasi data atau transfer learning. Model ini berhasil mencapai akurasi 90.625%, yang lebih baik daripada beberapa model lain yang menggunakan jumlah filter konvolusi yang lebih banyak tetapi tanpa augmentasi data. Hal ini menunjukkan bahwa, dalam kondisi tertentu, model CNN sederhana mampu bersaing dengan model yang lebih kompleks dalam pengenalan ekspresi wajah, terutama ketika data terbatas, seperti pada dataset Jaffe.

Perbandingan dengan penelitian lain menunjukkan bahwa meskipun model sederhana yang diusulkan dalam penelitian ini menunjukkan akurasi yang kompetitif, terdapat kekuatan yang jelas pada penelitian lain yang menggunakan teknik augmentasi data dan transfer learning. Misalnya, penelitian oleh (Appasaheb Borgalli & Surve, 2022) berhasil mencapai akurasi 91.58% dengan menggunakan teknik cropping citra menggunakan Haar Cascade dan augmentasi data, yang berkontribusi pada kemampuan model untuk belajar dari variasi yang lebih luas dalam dataset. Ini menunjukkan bahwa meskipun arsitektur CNN menawarkan kemampuan untuk belajar fitur secara otomatis, pendekatan tradisional seperti Gabor dan K-NN, yang digunakan dalam penelitian oleh (Swapna Subudhiray et al., 2023), juga dapat memberikan hasil yang kompetitif, mencapai akurasi 94.8% pada dataset yang sama.

Tabel 3. Perbandingan dengan Penelitian sebelumnya

Referensi,tahun	Data training dan data testing	Akurasi (%)	Metode yang digunakan
(Shan et al., 2009)	213: 10-Fold CV	81	LBP dan SVM
(Rere et al., 2019)	213: 85%:15%	87.5	CNN
	213: 70%:30%	78.1	CNN
(Akhand et al., 2021)	213:90%:10%	100	Pretrained Deep CNN
	213: 10-Fold CV	99.52	Pretrained Deep CNN
(Li et al., 2018)	213: 10-Fold CV	95.7	Transfer learning on model CNN MNF network (Hybrid Caffe-imagenet) and L2-SVM
(Swapna Subudhiray et al., 2023)	213: 70%:30%	94.8	Gabor +K-NN
(Eng et al., 2019)	213: 70%:30%	76.19	HOG dan SVM
(Appasaheb Borgalli & Surve, 2022)	213: 10-Fold CV	91.58	CNN dengan haar cascade dan augmentasi data
(Bachtiar & Wafi, 2021)	213: 5-Fold CV	76 %	Facial landmark, dengan ELM
Penelitian ini 2024	5-fold cross validation	87.3239	CNN 8 layer
Penelitian ini 2024	10-fold cross validation	89.2	CNN 8 layer
Penelitian ini 2024	Split 70% : 30%	82.8125	CNN 8 layer
Penelitian ini 2024	Split 80% : 20%	83.7209	CNN 8 layer
Penelitian ini 2024	Split 85% : 15%	90.625	CNN 8 layer
Penelitian ini 2024	Split 90% : 10%	85.7143	CNN 8 layer

Selain itu, penelitian ini memberikan nilai lebih dalam hal kepraktisan dan efisiensi komputasi. Meskipun tidak menggunakan augmentasi data atau transfer learning, penelitian ini tetap berhasil mencapai hasil yang lebih baik dibandingkan model yang menggunakan metode ekstraksi fitur tradisional seperti LBP dan HOG, yang terbukti memiliki keterbatasan dalam mempelajari fitur kompleks dari citra wajah. Sebagai contoh, penelitian oleh (Eng et al., 2019) menggunakan HOG dan SVM, menghasilkan akurasi yang lebih rendah, yaitu 76.19%, dibandingkan dengan model yang diusulkan.

Pengujian oleh (Akhand et al., 2021; Appasaheb Borgalli & Surve, 2022; Li et al., 2018) menunjukkan bahwa semakin kompleks arsitektur CNN, semakin besar kebutuhan akan augmentasi data dan transfer learning untuk mengurangi overfitting. Hal ini terbukti dapat menghasilkan akurasi yang lebih tinggi dibandingkan dengan penelitian yang diusulkan. Namun, keunggulan utama dari penelitian ini adalah pada kemampuannya untuk mencapai performa yang cukup tinggi tanpa mengandalkan kompleksitas arsitektur yang berlebihan atau metode pelatihan lanjutan, sehingga meminimalkan kebutuhan komputasi dan sumber daya.

Dengan membandingkan hasil yang diperoleh, terlihat adanya potensi untuk meningkatkan kinerja model yang diusulkan. Penelitian mendatang dapat mempertimbangkan integrasi metode pre-trained atau augmentasi data untuk mengatasi kelemahan yang ada. Misalnya, penelitian oleh (Akhand et al., 2021) menunjukkan bahwa penggunaan arsitektur yang lebih kompleks dan pre-trained dapat menghasilkan akurasi yang jauh lebih tinggi, bahkan ketika diterapkan pada dataset yang lebih kecil.

Perbandingan ini juga menyoroti pentingnya pendekatan yang cermat dalam pemilihan metode yang sesuai dengan kebutuhan spesifik dataset. Sementara penelitian ini menggunakan model yang lebih sederhana, hasilnya menunjukkan bahwa

model tersebut mampu memberikan hasil yang kompetitif. Penelitian ini menunjukkan bahwa model sederhana juga dapat memberikan hasil yang signifikan dengan efisiensi komputasi yang lebih tinggi, terutama dalam konteks dataset kecil seperti Jaffe.

3.3. Limitasi Riset

Meskipun penelitian ini memberikan kontribusi signifikan dalam pengenalan ekspresi wajah menggunakan arsitektur CNN yang sederhana, terdapat beberapa limitasi yang perlu diakui. Pertama, ukuran dataset yang digunakan dalam penelitian ini terbatas, yaitu hanya 213 citra. Meskipun dataset Jaffe mencakup berbagai ekspresi wajah, jumlah yang relatif kecil dapat membatasi kemampuan model untuk belajar dan menggeneralisasi dengan baik pada data baru. Penggunaan dataset yang lebih besar dapat meningkatkan robustnes model dan mengurangi risiko overfitting, terutama ketika menerapkan arsitektur yang lebih kompleks.

Kedua, penelitian ini tidak menggunakan teknik augmentasi data. Augmentasi data dapat membantu memperluas variasi dalam dataset, sehingga model dapat belajar dari berbagai kondisi pencahayaan, posisi wajah, dan ekspresi yang mungkin tidak tersedia dalam data asli. Tanpa augmentasi, model mungkin tidak mampu menangkap fitur yang lebih kompleks yang muncul dalam situasi dunia nyata.

Ketiga, meskipun model ini mencapai akurasi yang memuaskan, penggunaan arsitektur yang lebih sederhana juga membatasi kemampuannya dalam belajar representasi fitur yang sangat kompleks. Penelitian sebelumnya menunjukkan bahwa model CNN yang lebih dalam dan kompleks, seperti VGG16 dan ResNet, mampu memberikan hasil yang lebih baik dalam tugas-tugas pengenalan citra. Namun, penelitian ini bertujuan untuk mengeksplorasi kemampuan model yang lebih ringan dan sederhana untuk aplikasi yang lebih

efisien, yang mungkin tidak selalu memberikan hasil yang optimal dibandingkan model yang lebih rumit.

Keempat, parameter yang digunakan dalam model ini, seperti jumlah filter dan *learning rate*, ditetapkan berdasarkan percobaan awal. Meskipun analisis hyperparameter telah dilakukan, terdapat kemungkinan bahwa kombinasi parameter lain yang belum dieksplorasi dapat menghasilkan performa yang lebih baik. Oleh karena itu, eksplorasi lebih lanjut mengenai variasi parameter dapat menjadi langkah penting dalam penelitian di masa depan.

Akhirnya, meskipun penelitian ini memberikan wawasan penting dalam klasifikasi ekspresi wajah, hasilnya mungkin tidak sepenuhnya representatif untuk aplikasi di dunia nyata. Variasi yang lebih besar dalam data, seperti berbagai kelompok usia, etnis, dan gender, perlu dipertimbangkan untuk memastikan model dapat diimplementasikan secara efektif dalam konteks yang lebih luas.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa ukuran citra, jumlah filter, *learning rate*, dan epoch memiliki pengaruh signifikan terhadap performa klasifikasi model CNN yang diusulkan. Ukuran citra yang lebih besar tidak selalu memberikan hasil terbaik, karena memerlukan komputasi yang lebih kompleks dan pembelajaran dari data pelatihan yang lebih besar. Dalam penelitian ini, ukuran citra 128x128 piksel diidentifikasi sebagai ukuran optimal, sejalan dengan temuan dari sejumlah penelitian sebelumnya. Analisis perbandingan dengan penelitian terdahulu menunjukkan bahwa penerapan transfer learning secara signifikan dapat meningkatkan akurasi klasifikasi, terutama pada dataset kecil seperti Jaffe, di mana model-model CNN yang menggunakan teknik ini menunjukkan hasil akurasi yang lebih tinggi dibandingkan dengan model yang diusulkan.

Meskipun penelitian ini tidak menggunakan teknik augmentasi data dan transfer learning, model yang diusulkan tetap berhasil mencapai akurasi 90.6%, yang lebih tinggi dibandingkan dengan model-model sebelumnya yang menggunakan arsitektur CNN dengan jumlah filter konvolusi yang lebih besar. Selain itu, akurasi model ini juga lebih baik dibandingkan dengan berbagai model yang menggunakan teknik ekstraksi fitur tradisional dan metode pembelajaran mesin. Untuk penelitian mendatang, penerapan transfer learning sebagai fitur extractor pada model CNN, yang kemudian diikuti oleh klasifikasi menggunakan metode pembelajaran mesin, sangat direkomendasikan. Penggunaan dataset yang berbeda dari Jaffe juga akan memberikan wawasan lebih dalam mengenai performa model yang diusulkan dan kemampuannya untuk generalisasi pada variasi ekspresi wajah dalam konteks yang lebih luas.

DAFTAR PUSTAKA

AKBAR, A. T., SAIFULLAH, S., PRAPCOYO, H.,

- HUSAINI, R., & AKBAR, B. M. 2024. EfficientNet B0-Based RLDA for Beef and Pork Image Classification BT. *Proceedings of the 2023 1st International Conference on Advanced Informatics and Intelligent Information Systems (ICAI3S 2023)*, 136–145. https://doi.org/10.2991/978-94-6463-366-5_13
- AKBAR, M. T., ILMI, M. N., RUMAYAR, I. V., MONIAGA, J., CHEN, T.-K., & CHOWANDA, A. 2019. Enhancing Game Experience with Facial Expression Recognition as Dynamic Balancing. *Procedia Computer Science*, 157, 388–395. <https://doi.org/10.1016/j.procs.2019.08.230>
- AKHAND, M. A. H., ROY, S., SIDDIQUE, N., KAMAL, M. A. S., & SHIMAMURA, T. 2021. Facial Emotion Recognition Using Transfer Learning in the Deep CNN. *Electronics*, 10(9), 1036. <https://doi.org/10.3390/electronics10091036>
- ALMABDY, S., & ELREFAEI, L. 2019. Deep Convolutional Neural Network-Based Approaches for Face Recognition. *Applied Sciences*, 9(20), 4397. <https://doi.org/10.3390/app9204397>
- ALMEIDA, J., VILAÇA, L., TEIXEIRA, I. N., & VIANA, P. 2021. Emotion Identification in Movies through Facial Expression Recognition. *Applied Sciences*, 11(15), 6827. <https://doi.org/10.3390/app11156827>
- ALOMAR, K., AYSEL, H. I., & CAI, X. 2023. Data Augmentation in Classification and Segmentation: A Survey and New Strategies. *Journal of Imaging*, 9(2), 46. <https://doi.org/10.3390/jimaging9020046>
- ALRESHIDI, A., & ULLAH, M. 2020. Facial Emotion Recognition Using Hybrid Features. *Informatics*, 7(1), 6. <https://doi.org/10.3390/informatics7010006>
- ALTAMEEM, T., & ALTAMEEM, A. 2020. Facial expression recognition using human machine interaction and multi-modal visualization analysis for healthcare applications. *Image and Vision Computing*, 103, 104044. <https://doi.org/10.1016/j.imavis.2020.104044>
- ALTHNIAN, A., ALSAEED, D., AL-BAITY, H., SAMHA, A., DRIS, A. BIN, ALZAKARI, N., ABOU ELWAFI, A., & KURDI, H. 2021. Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain. *Applied Sciences*, 11(2), 796. <https://doi.org/10.3390/app11020796>
- APPASAHEB BORGALLI, M. R., & SURVE, D. S. 2022. Deep learning for facial emotion recognition using custom CNN architecture. *Journal of Physics: Conference Series*, 2236(1), 012004. <https://doi.org/10.1088/1742->

- 6596/2236/1/012004
- BACHTIAR, F. A., & WAFI, M. 2021. Komparasi Metode Klasifikasi untuk Deteksi Ekspresi Wajah Dengan Fitur Facial Landmark. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(5), 949–956. <https://doi.org/10.25126/jtiik.2021834434>
- BUSSO, C., & JAIN, J. 2012. *Advances in multimodal tracking of driver distraction* (pp. 253–270 BT-Digital Signal Processing for In-Veh). Springer.
- DAVILA DELGADO, J. M., & OYEDELE, L. 2021. Deep learning with small datasets: using autoencoders to address limited datasets in construction management. *Applied Soft Computing*, 112, 107836. <https://doi.org/10.1016/j.asoc.2021.107836>
- DE LA TORRE, F., & COHN, J. F. 2011. Facial Expression Analysis. *Visual Analysis of Humans*, 9, 377–409. https://doi.org/10.1007/978-0-85729-997-0_19
- ENG, S. K., ALI, H., CHEAH, A. Y., & CHONG, Y. F. 2019. Facial expression recognition in JAFFE and KDEF Datasets using histogram of oriented gradients and support vector machine. *IOP Conference Series: Materials Science and Engineering*, 705(1). <https://doi.org/10.1088/1757-899X/705/1/012031>
- JEONG, J., CHO, J. H., & LEE, J. G. 2021. Filter combination learning for CNN model compression. *ICT Express*, 7(1), 5–9. <https://doi.org/10.1016/j.icte.2021.01.001>
- JIA, Y., SHELHAMER, E., DONAHUE, J., KARAYEV, S., LONG, J., GIRSHICK, R., GUADARRAMA, S., & DARRELL, T. 2014. Caffe: Convolutional Architecture for Fast Feature Embedding. *Proceedings of the 22nd ACM International Conference on Multimedia*, 675–678. <https://doi.org/10.1145/2647868.2654889>
- KAHLOOT, K. M., & EKLER, P. 2021. Algorithmic Splitting: A Method for Dataset Preparation. *IEEE Access*, 9, 125229–125237. <https://doi.org/10.1109/ACCESS.2021.3110745>
- KHANDAY, O. M., DADVANDIPOUR, S., & LONE, M. A. 2021. Effect of filter sizes on image classification in CNN: A case study on CFIR10 and fashion-MNIST datasets. *IAES International Journal of Artificial Intelligence*, 10(4), 872–878. <https://doi.org/10.11591/ijai.v10.i4.pp872-878>
- KRITHIKA, L. B., VENKATESH, K., RATHORE, S., & HARISH KUMAR, M. 2017. Facial recognition in education system. *IOP Conference Series: Materials Science and Engineering*, 263(4), 042021. <https://doi.org/10.1088/1757-899X/263/4/042021>
- LEO, M., CARCAGNÌ, P., MAZZEO, P. L., SPAGNOLO, P., CAZZATO, D., & DISTANTE, C. 2020. Analysis of Facial Information for Healthcare Applications: A Survey on Computer Vision-Based Approaches. *Information*, 11(3), 128. <https://doi.org/10.3390/info11030128>
- LI, C., MA, N., & DENG, Y. 2018. Multi-Network Fusion Based on CNN for Facial Expression Recognition. *Proceedings of the 2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE 2018)*, 80(Csece), 166–169. <https://doi.org/10.2991/csece-18.2018.35>
- LI EW, S. S., KHALIL-HANI, M., AHMAD RADZI, S., & BAKHTERI, R. 2016. Gender classification: A convolutional neural network approach. *Turkish Journal of Electrical Engineering and Computer Sciences*, 24(3), 1248–1264. <https://doi.org/10.3906/elk-1311-58>
- MARTINEZ, B., & VALSTAR, M. F. 2016. Advances, Challenges, and Opportunities in Automatic Facial Expression Recognition. In *Advances in Face Detection and Facial Image Analysis* (pp. 63–100). Springer International Publishing. https://doi.org/10.1007/978-3-319-25958-1_4
- MATTIEV, J., SAJOVIC, J., DREVENŠEK, G., & ROGELJ, P. 2022. Assessment of Model Accuracy in Eyes Open and Closed EEG Data: Effect of Data Pre-Processing and Validation Methods. *Bioengineering*, 10(1), 42. <https://doi.org/10.3390/bioengineering10010042>
- MCDUFF, D., EL KALIOUBY, R., DEMIRDJIAN, D., & PICARD, R. 2013. Predicting online media effectiveness based on smile responses gathered over the Internet. *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 1–7. <https://doi.org/10.1109/FG.2013.6553750>
- MEHTA, D., SIDDIQUI, M., & JAVAID, A. 2018. Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality. *Sensors*, 18(2), 416. <https://doi.org/10.3390/s18020416>
- NG, H.-W., NGUYEN, V. D., VONIKAKIS, V., & WINKLER, S. 2015. Deep Learning for Emotion Recognition on Small Datasets using Transfer Learning. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 443–449. <https://doi.org/10.1145/2818346.2830593>
- RERE, L. M. R., USNA, S., & SOEGIJANTO, D.

2019. Studi Pengenalan Ekspresi Wajah Berbasis Convolutional Neural Network. *Seminar Nasional Teknologi Informasi Dan Komunikasi STI&K (SeNTIK)*, 3, 71–78.
- SAEED, S., SHAH, A. A., EHSAN, M. K., AMIRZADA, M. R., MAHMOOD, A., & MEZGEBO, T. 2022. Automated Facial Expression Recognition Framework Using Deep Learning. *Journal of Healthcare Engineering*, 2022, 1–11. <https://doi.org/10.1155/2022/5707930>
- SAIFULLAH, S., DREZEWSKI, R., YUDHANA, A., PRANOLO, A., KASWIJANTI, W., SURYOTOMO, A. P., PUTRA, S. A., KHALIDUZZAMAN, A., PRABUWONO, A. S., & JAPKOWICZ, N. 2023. Nondestructive chicken egg fertility detection using CNN-transfer learning algorithms. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, 9(3), 854–871. <https://doi.org/10.26555/jiteki.v9i3.26722>
- SHAIKHINA, T., & KHOVANOVA, N. A. 2017. Handling limited datasets with neural networks in medical applications: A small-data approach. *Artificial Intelligence in Medicine*, 75, 51–63. <https://doi.org/10.1016/j.artmed.2016.12.003>
- SHAN, C., GONG, S., & MCOWAN, P. W. 2009. Facial expression recognition based on Local Binary Patterns: A comprehensive study. *Image Vis. Comput.*, 27, 803–816.
- SUBUDHIRAY, S, PALO, H. K., DAS, N., & CHANDRA, M. 2021. Comparison of Facial Emotion Recognition Using Effective Features. *Ann. Rom. Soc. Cell Biol.*, 25(6), 5241–5252.
- SUBUDHIRAY, SWAPNA, PALO, H. K., & DAS, N. 2023. K-nearest neighbor based facial emotion recognition using effective features. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 12(1), 57. <https://doi.org/10.11591/ijai.v12.i1.pp57-65>
- SUHIRMAN, S., SAIFULLAH, S., HIDAYAT, A. T., KUSUMA, M. A., & DREZEWSKI, R. 2024. Real-time mask-wearing detection in video streams using deep convolutional neural networks for face recognition. *International Journal of Electrical and Computer Engineering (IJECE)*, 14(1), 1005. <https://doi.org/10.11591/ijece.v14i1.pp1005-1014>
- TANG, Y. 2013. Deep Learning using Support Vector Machines. In *Eprint Arxiv*.
- TATAR, A., HAGHIGHI, M., & ZEINIJAHRAMI, A. 2024. Experiments on image data augmentation techniques for geological rock type classification with convolutional neural networks. *Journal of Rock Mechanics and Geotechnical Engineering*. <https://doi.org/10.1016/j.jrmge.2024.02.015>
- TONGUÇ, G., & OZAYDIN OZKARA, B. 2020. Automatic recognition of student emotions from facial expressions during a lecture. *Computers & Education*, 148(November 2019), 103797. <https://doi.org/10.1016/j.compedu.2019.103797>
- UTAMI, N., & MUSTAKIM. 2021. Data Distribution Modelling in Supervised Learning Algorithm is for The classification of Prospective Recipient Candidate. *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 593–597. <https://doi.org/10.1109/ISRITI54043.2021.9702783>
- VABALAS, A., GOWEN, E., POLIAKOFF, E., & CASSON, A. J. 2019. Machine learning algorithm validation with a limited sample size. *PLOS ONE*, 14(11), e0224365. <https://doi.org/10.1371/journal.pone.0224365>
- WAFAA SHIHAB, A., & ABDUL AMIR A., K. 2020. *The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN - Proceeding of the 2020 IEEE International Conference on Computer Science and Software Engineering: 16-18 April 2020, University of Duhok, Duhok, Kurdistan Region, Iraq.* 88–93. <https://doi.org/10.1109/CSASE48920.2020.9142089>
- XIE, S., & HU, H. 2017. Facial expression recognition with FRR-CNN. *Electronics Letters*, 53(4), 235–237. <https://doi.org/10.1049/el.2016.4328>
- XIE, Z., LI, J., & SHI, H. 2019. A Face Recognition Method Based on CNN. *Journal of Physics: Conference Series*, 1395(1). <https://doi.org/10.1088/1742-6596/1395/1/012006>
- XU, Y., & GOODACRE, R. 2018. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *Journal of Analysis and Testing*, 2(3), 249–262. <https://doi.org/10.1007/s41664-018-0068-2>
- ZAINORZULI, S. M., CHE ABDULLAH, S. A., ZAINOL ABIDIN, H., & AHMAT RUSLAN, F. 2022. Comparison Study on Convolution Neural Network (CNN) Techniques for Image Classification. *Journal of Electrical & Electronic Systems Research*, 20(APR2022), 11–17. <https://doi.org/10.24191/jeesr.v20i1.002>

- ZAINUDIN, Z., SHAMSUDDIN, S. M., & HASAN, S. 2019. Deep learning for image processing in WEKA environment. *International Journal of Advances in Soft Computing and Its Applications*, 11(1), 1–21.
- ZHANG, Y., & LING, C. 2018. A strategy to apply machine learning to small datasets in materials science. *Npj Computational Materials*, 4(1), 25. <https://doi.org/10.1038/s41524-018-0081-z>

Halaman ini sengaja dikosongkan