

Penggunaan Teknologi Citra Digital dan Metode *Convolutional Neural Network* untuk Mendeteksi Emosi Anak Berdasarkan Ekspresi Wajah

Putri Angraeni Badar ¹ and Putri Angraeni ²

¹ Affiliation 1; angraeniputri59@gmail.com

² Affiliation 2; putriangraeiny@gmail.com

Abstrak: Penelitian ini mengkaji penerapan teknologi citra digital dan metode *Convolutional Neural Network* (CNN) untuk mendeteksi emosi anak berdasarkan ekspresi wajah. Ekspresi wajah mencerminkan berbagai emosi tanpa kata-kata dan memiliki peran krusial dalam perkembangan anak. Metode konvensional seringkali kurang akurat, sehingga teknologi citra digital dan CNN menjadi pilihan efektif. Sumber data untuk penelitian ini berasal dari internet yang dikumpulkan oleh peneliti dan dataset *Facial Expression Recognition 2013* (FER2013). Hasil penelitian menunjukkan bahwa model CNN dengan 4 lapisan konvolusi dan 2 lapisan terhubung menghasilkan akurasi sebesar 82.37% dengan 50 epoch. Kualitas sistem dalam mendeteksi beberapa wajah sudah cukup akurat untuk mengenali kategori emosi pada anak.

Kata Kunci: Teknologi Citra Digital, *Convolutional Neural Network*, Emosi Anak, Ekspresi Wajah.

Abstract: This study examines the application of digital image technology and *Convolutional Neural Network* (CNN) methods to detect children's emotions based on facial expressions. Facial expressions reflect various emotions without words and play a crucial role in child development. Conventional methods are often less accurate, making digital image technology and CNN effective choices. The data source for this research comes from the internet, collected by the researchers, and the *Facial Expression Recognition 2013* (FER2013) dataset. The research results indicate that the CNN model with 4 convolutional layers and 2 connected layers achieves an accuracy of 82.37% with 50 epochs. The system's quality in detecting several faces is accurate enough to recognize emotion categories in children.

Keywords: Digital Image Technology, *Convolutional Neural Network*, Children's Emotions, Facial Expression.

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Last-name

Received: date

Revised: date

Accepted: date

Published: date



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Pendahuluan

Wajah merupakan aspek paling ekspresif dan komunikatif dari seorang manusia[1]. Ekspresi wajah memiliki kemampuan untuk menyampaikan beragam emosi tanpa memerlukan kata-kata [2]. Kemampuan untuk mengenali dan memberi nama pada emosi sendiri dan emosi orang lain berdasarkan petunjuk ekspresi wajah adalah kemampuan adaptif yang baik untuk bertahan hidup maupun berkembang di masyarakat. Kemampuan ini secara langsung terkait dengan cara individu berinteraksi dengan orang lain dan memahami perasaan dan emosi dalam setiap konteks. Keterampilan ini bahkan lebih penting pada masa kanak-kanak, ketika interaksi sosial pertama terjadi, sebelum kemampuan berbicara sepenuhnya berkembang [3]. Begitu melihat wajah seseorang, bisa diidentifikasi banyak informasi, termasuk kelompok usia, jenis kelamin, dan arah pandangan. Sebagian besar komunikasi non-verbal antara individu tercermin melalui ekspresi wajah. Penilaian otomatis dan penafsiran subjektif dari ekspresi wajah dapat dengan cepat memberikan pemahaman tentang emosi dan perilaku seseorang selama interaksi sosial.

Mendekode emosi dengan tepat dari ekspresi wajah tampaknya menjadi salah satu mekanisme kunci untuk memahami informasi sosial. Dalam penelitian ontogenik,

perkembangan signifikan dalam pengolahan emosi dari wajah telah dicatat pada tahun pertama kehidupan. Sebagai contoh, bayi yang baru lahir cenderung melihat lebih lama pada wajah yang tersenyum daripada wajah yang netral atau takut, sementara bayi yang berusia antara 5 dan 7 bulan menunjukkan kecenderungan untuk fokus perhatian pada wajah yang menakutkan [4]. Anak kecil seringkali belum mampu mengkomunikasikan perasaan mereka dengan kata-kata, sehingga pengamatan ekspresi wajah menjadi cara yang efektif untuk memahami keadaan emosional mereka. Studi sebelumnya menunjukkan bahwa ekspresi wajah dapat memberikan informasi penting tentang perasaan dan tanggapan emosional anak-anak [5]. Memahami emosi memiliki manfaat yang nyata bagi perkembangan sehat anak [6]. Kegagalan dalam mengidentifikasi ekspresi wajah yang mencerminkan emosi dapat menjadi faktor yang terkait erat dengan hambatan dalam perkembangan anak. Kegagalan ini juga merupakan tanda dari beberapa gangguan perkembangan, yang dapat mengakibatkan keterlambatan dalam mengembangkan keterampilan sosial dasar yang diperlukan untuk beradaptasi dengan kehidupan Masyarakat [7]. Ketidapahaman terhadap emosi pada anak-anak telah dikaitkan dengan dampak negatif, termasuk rendahnya kemampuan sosial, kinerja akademis yang kurang baik, dan masalah perilaku internal/eksternal. Oleh karena itu, diperlukan metode yang efektif untuk mendeteksi emosi anak berdasarkan ekspresi wajah mereka.

Pendekatan konvensional seperti pengamatan mata manusia memiliki keterbatasan dalam hal akurasi dan efisiensi [8]. Sehingga, penerapan teknologi citra digital dan metode *Convolutional Neural Network* (CNN) menjadi opsi menarik untuk mendeteksi emosi berdasarkan ekspresi wajah balita. Teknologi citra digital memungkinkan pemrosesan gambar secara digital, sementara CNN merupakan metode pembelajaran mesin yang dapat mengekstrak fitur kompleks dan mengenali pola pada gambar. CNN tergolong sebagai jenis *Deep Neural Network* karena memiliki struktur jaringan yang mendalam dan sering digunakan dalam pemrosesan data gambar [9].

Beberapa penelitian telah dilakukan untuk pengenalan ekspresi wajah. Salah satunya adalah pengenalan ekspresi wajah dari *cross dataset* menggunakan CNN [10]. Hasil yang diperoleh dari studi ini mencakup tingkat akurasi data uji sebesar 91,29%, sensitivitas atau recall atau True Positive Rate (TPR) sebesar 91,29%, presisi atau Positive Predictive Value (PPV) sebesar 91,29%, serta akurasi keseluruhan sebesar 97,51%. Selain itu, penelitian *real time emotion recognition from facial expressions using CNN Architecture* [11]. Hasil akurasi sebesar 96.43% dan akurasi validasi sebesar 91.81% untuk klasifikasi 7 emosi berbeda melalui ekspresi wajah. Lalu, penelitian *CNN based emotion classification cognitive model for facial expression* [12]. Evaluasi model menunjukkan bahwa akurasi dari set data pengujian kondisi laboratorium dibandingkan dengan model yang diusulkan, dengan akurasi tertinggi untuk parameter bahagia mencapai 99%, diikuti oleh terkejut dengan 98%, netral dengan 96%, dan akurasi terendah untuk emosi takut sebesar 45%.

Berdasarkan hasil penelitian yang telah dilakukan, penerapan teknologi citra digital serta metode *Convolutional Neural Network* (CNN) terbukti sebagai pendekatan yang berhasil untuk mendeteksi emosi anak melalui ekspresi wajah mereka. CNN mampu mengekstrak fitur kompleks dan mengenali pola pada gambar dengan tingkat akurasi yang tinggi. Sejumlah penelitian telah memanfaatkan CNN untuk mengenali ekspresi wajah anak, dengan parameter emosi yang mencakup bahagia, terkejut, netral, dan takut. Dengan kecakapan CNN dalam mengenali serta mengklasifikasikan emosi anak berdasarkan ekspresi wajah, teknologi ini memiliki potensi besar dalam berbagai aplikasi, termasuk pengembangan metode pendeteksian emosi anak yang lebih akurat dan efisien.

2. Literature Review

Perkembangan emosi anak-anak sangat dipengaruhi oleh emosi yang mereka alami. Emosi ini memiliki dampak signifikan pada perilaku, pemikiran, dan interaksi sosial anak-anak. Selain itu, kondisi kesehatan mental anak dapat tercermin dalam ekspresi wajah mereka. Meskipun ekspresi wajah anak dapat memberikan petunjuk tentang pengalaman

emosional mereka, seringkali orang dewasa menghadapi kesulitan dalam memahaminya dengan tepat. Teknologi citra digital serta metode *Convolutional Neural Network* (CNN) bisa diterapkan untuk mengenali ekspresi wajah pada anak dan mengidentifikasi ragam emosi yang mereka tunjukkan [13].

Sejumlah studi telah memanfaatkan teknologi citra digital dan pendekatan CNN untuk mengidentifikasi emosi melalui ekspresi wajah. Penelitian yang dilakukan oleh Ajeng Restu Kusumastuti Et Al [14] yang berjudul “Klasifikasi Ketertarikan Anak PAUD Melalui Ekspresi Wajah Menggunakan Metode CNN” dalam mengembangkan model klasifikasi ketertarikan anak PAUD, keempat model menunjukkan peningkatan nilai akurasi dari epoch 25 hingga 100, dengan arsitektur Raj Mehrotra mencapai nilai tertinggi. Namun, saat dilakukan pengujian pada data, arsitektur yang digunakan dalam penelitian ini menunjukkan peningkatan yang lebih lambat, mencapai puncak akurasi 81,66%. Hasil tersebut menunjukkan peningkatan sebesar 3,33% dibandingkan dengan hasil yang diperoleh menggunakan arsitektur Raj Mehrotra, dan 1,66% lebih baik daripada arsitektur R. Cui.

Penelitian oleh Kuang Liu Et Al [15] dengan judul “Facial Expression Recognition with CNN Ensemble” Peneliti menggunakan dataset FER-2013 dalam penelitian ini. Arsitektur yang diterapkan terdiri dari 3 lapisan konvolusi, di mana lapisan pertama menggunakan filter 5x5 sebanyak 64 dengan fungsi aktivasi ReLU, diikuti dengan *max pooling* berukuran 2x2. Pada lapisan kedua, digunakan filter 5x5 sebanyak 128 dengan fungsi aktivasi yang sama dan *max pooling* berukuran 2x2. Selanjutnya, lapisan ketiga menerapkan filter 5x5 sebanyak 256 dengan fungsi aktivasi dan *max pooling* yang sama. Setelah melewati proses regularisasi, hasilnya dimasukkan ke dalam lapisan fully connected dengan 4096 kelas, yang kemudian diregularisasi lagi dengan *dropout* sebesar 0,5. Proses terakhir melibatkan pemrosesan ke dalam lapisan *fully connected* dengan 3 kelas. Arsitektur yang diuji dalam penelitian ini sesuai dengan subnet1.

Penelitian Ivan Azhari Et Al [16] berjudul “Implementasi algoritma Convolutional Neural Network dalam Deteksi Emosi Manusia Berdasarkan Ekspresi Wajah” di mana data yang digunakan untuk melatih model berasal dari sampel citra dalam dataset FER 2013. Hasil penelitian menunjukkan bahwa *Convolutional Neural Network* mampu mengenali ekspresi wajah dengan tingkat akurasi berkisar antara 67% hingga 72%, berdasarkan uji coba yang dilakukan.

Penelitian Mr. Rohan Appasaheb Borgalli Et Al [17] dengan judul “Deep learning for facial emotion recognition using custom CNN architecture” akurasi untuk dataset FER13 mendapatkan akurasi keseluruhan sebesar 86,78%, yang sedikit lebih rendah dibandingkan dengan akurasi CK+ sebesar 92,27% dan akurasi JAFFE sebesar 91,58%.

3. Metode Penelitian

Dalam penelitian ini, diuraikan model implementasi metode *Convolutional Neural Network* (CNN) dalam mendeteksi emosi pada anak dengan memanfaatkan teknologi citra digital. Tahapan penelitian ditunjukkan pada gambar 1.



Gambar 1. Tahapan Penelitian

3.1 Dataset

Penelitian ini memanfaatkan data gambar emosi yang telah terlebih dahulu diberi label. Data tersebut mencakup label emosi yang akan digunakan untuk mengembangkan model dalam penelitian ini, yang melibatkan 7 kelas emosi seperti marah, sedih, senang, netral, jijik, takut, dan terkejut. Data pelatihan akan digunakan untuk melatih

model, sedangkan data uji akan digunakan untuk mengevaluasi akurasi model yang telah dibuat. Sumber data untuk penelitian ini berasal dari internet yang dikumpulkan oleh peneliti dan dataset Facial Expression Recognition 2013 (FER2013) dari situs web (<https://www.kaggle.com/datasets>). Dataset ini kemudian dipilah, dimana pada penelitian ini hanya menggunakan foto anak kecil. Dataset ini kemudian dibagi menjadi data pelatihan (80%) dan data uji (20%), dengan data pelatihan mencakup 4.229 gambar dan data uji mencakup 1.269 gambar dengan total data gambar 5.498. Ukuran gambar emosi ini adalah 48 x 48 piksel, dan citra tersebut berwarna greyscale atau dalam skala warna abu-abu dan putih. Dataset ini dapat dilihat pada gambar 2.



Gambar 2. Contoh kumpulan dataset

3.2 Daftar Kelas

Dataset yang digunakan terbagi menjadi 7 kelas, mencakup emosi marah, sedih, senang, netral, jijik, takut, dan terkejut.

Tabel 1. Jumlah gambar dalam setiap kategori pada dataset (data pelatihan)

Emosi	Jumlah Data
Marah	616
Sedih	597
Senang	1018
Netral	567
Jijik	305

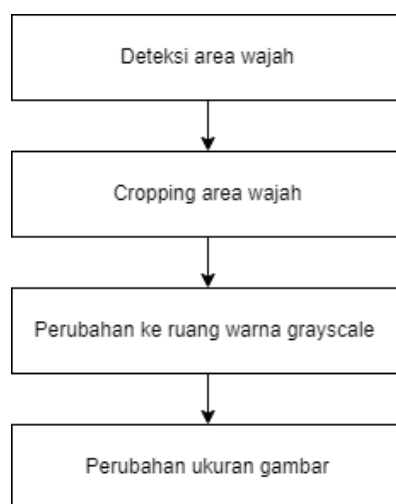
Takut	513
Terkejut	613

Tabel 2. Jumlah gambar dalam setiap kategori pada dataset (data Pengujian)

Emosi	Jumlah Data
Marah	185
Sedih	179
Senang	305
Netral	170
Jijik	92
Takut	154
Terkejut	184

3.3 Preprocessing

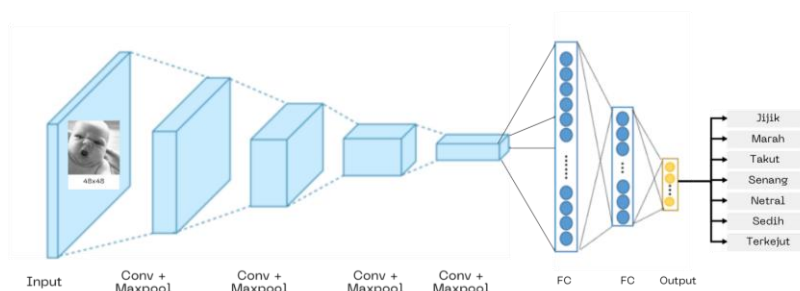
Proses pra-pemrosesan dilaksanakan untuk mempermudah tahap pelatihan, mengurangi waktu yang diperlukan pada setiap epoch. Pra-pemrosesan dimulai dengan mendeteksi wajah dalam gambar input menggunakan sistem, kemudian wajah yang terdeteksi dipotong dan mengalami transformasi ruang warna dari RGB ke skala keabuan, dengan normalisasi ukuran menjadi 48×48. Selanjutnya, gambar wajah yang telah diolah ini dijadikan input untuk *Convolutional Neural Network* (CNN). Pada akhirnya, hasilnya merupakan pengenalan ekspresi wajah, termasuk emosi seperti marah, senang, sedih, jijik, terkejut, atau netral. Struktur pendekatan pra-pemrosesan dapat dilihat pada Gambar 3.



Gambar 3. Tahapan Preprocessing

3.4 Membangun Arsitektur Sistem

Arsitektur yang akan dibangun terdiri dari beberapa lapisan seperti yang terlihat pada Gambar 4. Terdapat lapisan-lapisan konvolusional dan lapisan-lapisan terhubung penuh dalam arsitektur ini. Pada lapisan-lapisan konvolusional, terdapat fungsi aktivasi ReLU, *Max pooling*, *Batch normalization*, dan *Dropout*. Sedangkan pada lapisan-lapisan terhubung penuh, termasuk *Flatten*, *Dense*, *Batch normalization*, dan *Softmax*.



Gambar 4. Arsitektur CNN

3.4.1 Metode

1) Lapisan pertama

Mengandung *layer Conv2D* dengan 64 *filter* dan *kernel size* 3, diaktifkan dengan fungsi aktivasi *relu*. Selanjutnya, ditambahkan *layer MaxPool2D* sebagai lapisan tersembunyi.

2) Lapisan kedua

Mengandung *layer Conv2D* dengan 128 *filter* dan *kernel size* 3, diaktifkan dengan fungsi aktivasi *relu*. Selanjutnya, ditambahkan *layer MaxPool2D* sebagai lapisan tersembunyi.

3) Lapisan ketiga

Mengandung *layer Conv2D* dengan 512 *filter* dan *kernel size* 3, diaktifkan dengan fungsi aktivasi *relu*. Selanjutnya, ditambahkan *layer MaxPool2D* sebagai lapisan tersembunyi.

4) Lapisan keempat

Mengandung *layer Conv2D* dengan 512 *filter* dan *kernel size* 3, diaktifkan dengan fungsi aktivasi *relu*. Selanjutnya, ditambahkan *layer MaxPool2D* sebagai lapisan tersembunyi.

5) Lapisan kelima

Mengandung lapisan *flatten*.

6) Lapisan Keenam

Mengandung dua lapisan *Dense*, dimana lapisan pertama terdiri dari 256 unit dengan aktivasi *relu*, sementara lapisan kedua memiliki 512 unit dengan aktivasi *relu*.

3.4.2 Fungsi CNN

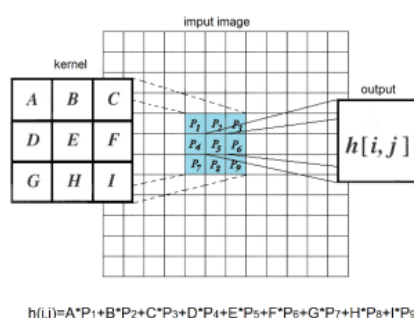
Metode CNN yang diterapkan menggunakan fungsi *categorical_crossentropy*, yang umumnya digunakan dalam konteks pemilihan item ke dalam satu kategori dari beberapa kategori yang ada. Untuk optimalisasi, algoritma Adam yang telah disediakan oleh *Tensorflow* digunakan.

```
opt = Adam(learning_rate=0.0001)
model.compile(optimizer=opt, loss='categorical_crossentropy', metrics=['accuracy'])
```

Gambar 5. Kode yang mengimplementasikan konsep cross entropy.

3.4.3 Lapisan Arsitektur 2D Convolution

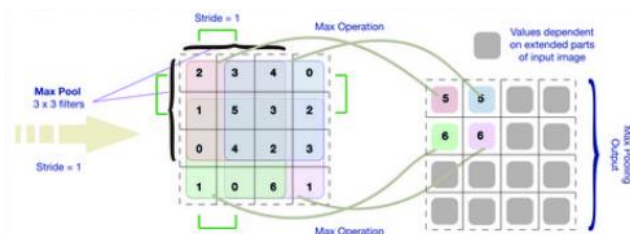
Lapisan konvolusional adalah bagian dari jaringan saraf yang terdiri dari serangkaian operasi matematika, seperti konvolusi, yang berfungsi untuk mengekstraksi fitur lokal seperti tepi, sudut, dan tekstur dari gambar input. Lapisan ini memegang peran kunci dalam *Convolutional Neural Network* (CNN) dan merupakan elemen inti yang membangun struktur CNN. Proses konvolusi melibatkan perkalian sekumpulan bobot dengan input, dan hasil dari konvolusi yang diterapkan pada satu input tetap berupa piksel tunggal. Namun, jika konvolusi diterapkan pada seluruh gambar 3 dimensi, hasilnya akan menjadi gambar 2 dimensi. Selain itu, lapisan konvolusional dapat diikuti oleh lapisan konvolusional tambahan atau lapisan penyatuan, sehingga kompleksitas CNN meningkat dengan setiap lapisannya untuk mengenali bagian yang lebih besar dari gambar [18].



Gambar 6. Proses konvolusi untuk kernel konvolusi berukuran 3×3

3.4.4 Lapisan Arsitektur Max Pooling

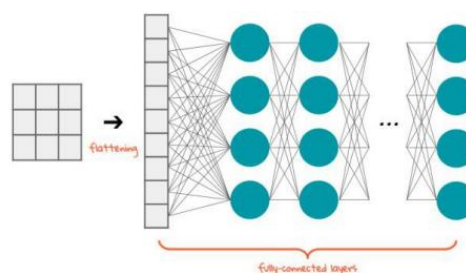
Lapisan *max pooling* merupakan komponen dalam *Convolutional Neural Network* (CNN) yang berfungsi mengurangi dimensi spasial dari peta fitur hasil lapisan konvolusi. Tujuannya adalah untuk mempertahankan fitur-fitur yang paling penting. Operasi *max pooling* dilakukan dengan menghitung nilai maksimum dari setiap kelompok data pada peta fitur, sehingga fitur yang paling menonjol di setiap wilayah dapat dipertahankan. Fungsi ini membantu mengendalikan kompleksitas perhitungan CNN, serta meningkatkan kinerja jaringan secara keseluruhan [19]. Dalam penelitian ini, max pooling yang digunakan adalah 3×3. Dapat dilihat pada gambar berikut.



Gambar 7. Max Pooling ukuran 3×3

3.4.5 Lapisan Arsitektur Flatten

Lapisan flatten dalam arsitektur *Convolutional Neural Network* (CNN) berfungsi untuk mengubah hasil keluaran dari lapisan konvolusi dan lapisan pooling menjadi vektor satu dimensi. Vektor tersebut kemudian dapat diproses oleh lapisan klasifikasi. Dalam proses ini, output dari lapisan konvolusi dan lapisan pooling disusun menjadi satu vektor panjang dengan menggabungkan semua fitur yang dihasilkan. Umumnya, lapisan flatten ditempatkan setelah lapisan konvolusi dan lapisan pooling, namun sebelum lapisan klasifikasi [20].



Gambar 8. Proses Flattening dan Densifikasi

3.4.6 Lapisan Arsitektur Dense

Lapisan *dense* adalah jenis lapisan yang bertindak sebagai lapisan sepenuhnya terhubung, di mana setiap *neuron* dalam lapisan tersebut memiliki koneksi dengan setiap *neuron* dalam lapisan *dropout* yang sesuai. Lapisan *dense* menciptakan hubungan antara setiap *neuron* dalam lapisan sebelumnya dengan setiap *neuron* dalam lapisan berikutnya. Secara umum, lapisan *dense* ditempatkan setelah lapisan konvolusi, lapisan *pooling*, dan lapisan *flatten* [21].

3.5 Pseudocode

Pseudocode menampilkan proses algoritma CNN dapat dilihat pada gambar 9.

```

Inisialisasi:
- Membaca dataset gambar ekspresi wajah anak
- Membagi dataset menjadi data pelatihan dan data pengujian
- Menentukan jumlah kelas emosi (senang, sedih, terkejut, jijik, marah, takut, netral)
- Menentukan arsitektur Convolutional Neural Network (CNN)

Pelatihan CNN:
1. Membangun model CNN:
  - Input layer dengan dimensi gambar ekspresi wajah
  - Convolutional layers untuk mengekstrak fitur
  - Pooling layers untuk mereduksi resolusi spatial
  - Fully connected layers untuk klasifikasi

2. Menentukan fungsi kerugian (loss function) dan algoritma optimasi (optimizer)

3. Melakukan pelatihan model menggunakan data pelatihan:
  - Iterasi sejumlah epoch
  - Memasukkan gambar ke dalam model
  - Menghitung nilai loss
  - Melakukan backpropagation dan pembaharuan bobot

Pengujian:
1. Menggunakan model yang telah dilatih untuk memprediksi emosi pada data pengujian:
  - Memasukkan gambar pengujian ke dalam model
  - Menerima distribusi probabilitas untuk setiap kelas emosi
  - Memilih kelas dengan probabilitas tertinggi sebagai prediksi

2. Evaluasi performa model:
  - Membandingkan prediksi dengan label sebenarnya pada data pengujian
  - Menghitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score

Penggunaan pada Data Baru:
1. Menerima gambar wajah anak sebagai input

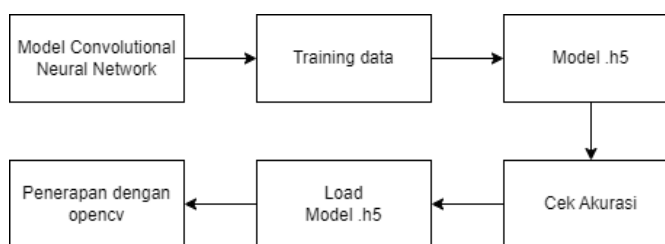
2. Menggunakan model untuk memprediksi emosi:
  - Memasukkan gambar ke dalam model
  - Menerima distribusi probabilitas untuk setiap kelas emosi
  - Memilih kelas dengan probabilitas tertinggi sebagai prediksi

3. Menampilkan hasil prediksi kepada pengguna
  
```

Gambar 9. Pseudocode Algoritma CNN

3.6 Scenarion Testing

Setelah mengembangkan model CNN dan melakukan pelatihan data, langkah selanjutnya adalah mengimplementasikan model CNN. Tahapan uji skenario tergambar dalam Gambar 10.



Gambar 10. Tahapan Scenario Testing

Proses ini melibatkan pelatihan data dengan menggunakan model CNN yang telah dibuat, yang terdiri dari 4 lapisan konvolusi dan 2 lapisan sepenuhnya terhubung (dense) dengan aktivasi ReLu. Setelah data uji dilatih, hasilnya disimpan dalam file berekstensi .h5. File ini mengandung informasi hasil pembelajaran mesin yang telah dilakukan. Selanjutnya, langkah selanjutnya adalah memeriksa akurasi dari data uji, dan untuk menerapkannya dengan menggunakan *library* opencv, diperlukan file dengan ekstensi .h5 yang berisi hasil pembelajaran mesin.

4. Hasil dan Pembahasan

4.1 Model CNN

Model CNN yang telah disusun mencakup 4 lapisan konvolusi dan 2 lapisan *fully connected* terhubung dengan fungsi aktivasi *ReLU*. Untuk keterangan model yang lebih jelasnya dapat ditemukan dalam Tabel 3.

Tabel 3. Lapisan Model

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 48, 48, 64)	640
batch_normalization_4 (BatchNormalization)	(None, 48, 48, 64)	256
activation_4 (Activation)	(None, 48, 48, 64)	0
max_pooling2d_4 (MaxPooling2D)	(None, 16, 16, 64)	0
dropout_3 (Dropout)	(None, 16, 16, 64)	0
conv2d_5 (Conv2D)	(None, 16, 16, 128)	73856
batch_normalization_5 (BatchNormalization)	(None, 16, 16, 128)	512
activation_5 (Activation)	(None, 16, 16, 128)	0
max_pooling2d_5 (MaxPooling2D)	(None, 5, 5, 128)	0
...		
Total params: 3692295 (14.08 MB)		
Trainable params: 3688327 (14.07 MB)		
Non-trainable params: 3968 (15.50 KB)		

4.2 Dataset

Masing-masing dataset terbagi menjadi beberapa kelas diantaranya marah, sedih, senang, netral, jijik, takut, dan terkejut.

1) Dataset untuk pelatihan

Gambar dengan kategori marah sebanyak 616 citra. Ekspresi untuk anak yang sedang marah diekspresikan dengan alis mengkerut dan mata yang menatap tajam. Dataset marah dapat dilihat sebagai berikut:



Gambar 11. Dataset latih untuk kelas marah

Gambar dengan kategori sedih sebanyak 597 citra. Ekspresi untuk anak yang sedang sedih diekspresikan dengan mata yang tidak fokus menatap objek serta sudut bibir yang sedikit ditarik ke bawah. Dataset sedih dapat dilihat sebagai berikut:



Gambar 12. Dataset latih untuk kelas sedih

Gambar dengan kategori senang sebanyak 1018 citra. Ekspresi untuk anak yang sedang senang diekspresikan dengan wajah yang tersenyum. Dataset senang dapat dilihat sebagai berikut:



Gambar 13. Dataset latih untuk kelas senang

Gambar dengan kategori netral sebanyak 567 citra. Ekspresi untuk anak yang sedang netral diekspresikan dengan wajah biasa saja. Dataset netral dapat dilihat sebagai berikut:



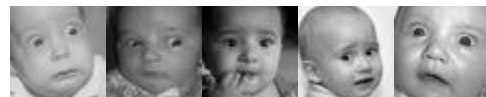
Gambar 14. Dataset latih untuk kelas netral

Gambar dengan kategori jijik sebanyak 305 citra. Ekspresi untuk anak yang sedang jijik diekspresikan dengan matap menatap tajam ke bawah ke objek yang membuat jijik. Dataset jijik dapat dilihat sebagai berikut:



Gambar 15. Dataset latih untuk kelas jijik

Gambar dengan kategori takut sebanyak 512 citra. Ekspresi untuk anak yang sedang takut diekspresikan dengan wajah terlihat ragu. Dataset takut dapat dilihat sebagai berikut:



Gambar 16. Dataset latih untuk kelas takut

Gambar dengan kategori terkejut sebanyak 613 citra. Ekspresi untuk anak yang sedang terkejut diekspresikan dengan mulut menganga. Dataset terkejut dapat dilihat sebagai berikut:



Gambar 17. Dataset latih untuk kelas terkejut

2) Dataset untuk pengujian

Gambar dengan kategori marah sebanyak 185 citra. Dataset ini dapat dilihat sebagai berikut:



Gambar 18. Dataset uji untuk kelas marah

Gambar dengan kategori sedih sebanyak 179 citra. Dataset ini dapat dilihat sebagai berikut:



Gambar 19. Dataset uji untuk kelas sedih

Gambar dengan kategori senang sebanyak 305 citra. Dataset ini dapat dilihat sebagai berikut:



Gambar 20. Dataset uji untuk kelas senang

Gambar dengan kategori netral sebanyak 170 citra. Dataset ini dapat dilihat sebagai berikut:



Gambar 21. Dataset uji untuk kelas netral

Gambar dengan kategori jijik sebanyak 92 citra. Dataset ini dapat dilihat sebagai berikut:



Gambar 22. Dataset uji untuk kelas jijik

Gambar dengan kategori takut sebanyak 154 citra. Untuk contoh dataset takut sebagai berikut:



Gambar 23. Dataset uji untuk kelas takut

Gambar dengan kategori terkejut sebanyak 184 citra. Untuk contoh dataset terkejut sebagai berikut:

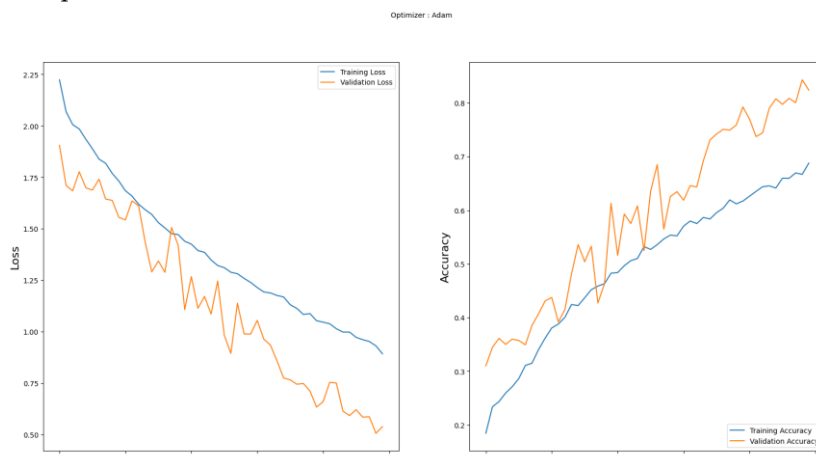


Gambar 24. Dataset uji untuk kelas terkejut

4.3 Data Latih dan Data Uji

Dalam penelitian ini, data dibagi menjadi dua bagian, yaitu data latih dan data uji. Jumlah data latih mencakup 4.229 data, sementara data uji 1.269 data dari total

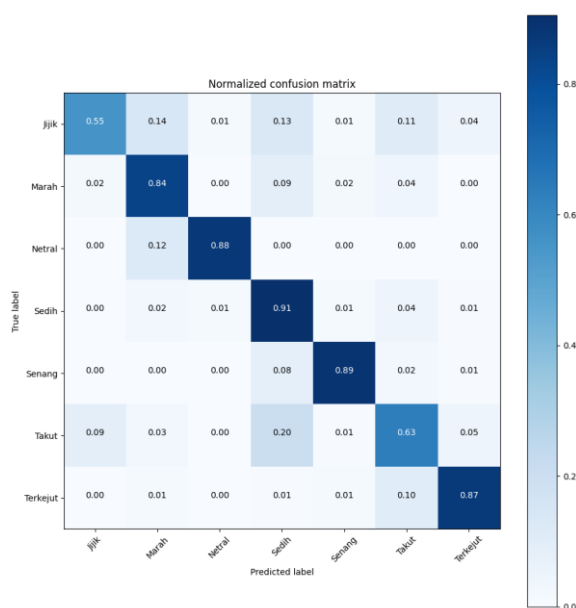
5.498 data. Proses pelatihan model CNN menggunakan data uji dilakukan sebanyak 50 *epoch*, dengan setiap *epoch* melibatkan 225 langkah. Model terbaik yang dihasilkan mencapai akurasi validasi sekitar 82%, suatu pencapaian yang baik mengingat terdapat 7 kemungkinan nilai untuk kelas target. Pada setiap *epoch*, perpustakaan Keras melakukan pengecekan apakah kinerja model meningkat dibandingkan dengan *epoch* sebelumnya. Jika ya, bobot model optimal baru disimpan dalam file dengan ekstensi .h5. Hal ini memungkinkan penggunaan bobot model tanpa perlu melatih ulang kapan pun diperlukan. Visualisasi hasil akurasi dapat ditemukan dalam Gambar 25.



Gambar 25. Visualisasi training dan validation loss

4.4 Confusion Matrix

Matriks konfusi adalah tabel yang meresmikan kinerja model klasifikasi dengan menampilkan jumlah *instance* yang diklasifikasikan dengan benar dan yang diklasifikasikan dengan salah untuk setiap kelas. Matriks konfusi dapat ditemukan dalam Gambar 26.



Gambar 26. Konfusi Matriks

Matriks ini memiliki 7 baris dan 7 kolom yang merepresentasikan 7 kelas emosi wajah. Baris mencerminkan kelas sebenarnya dalam data pengujian, sedangkan kolom mencerminkan kelas yang diprediksi oleh model. Angka-angka dalam

matriks mencerminkan jumlah data uji yang sebenarnya termasuk dalam kelas tertentu dan diklasifikasikan oleh model.

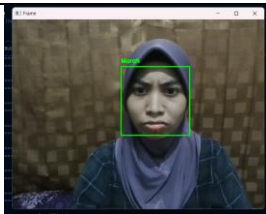

Pada gambar 26, terdapat 55 citra "jijik" yang terprediksi "jijik", sementara 84 citra "marah" terprediksi "marah", 91 citra "sedih" terprediksi "sedih", 89 citra "senang" terprediksi senang, 63 citra "takut" terprediksi "takut" dan 87 citra "terkejut" terprediksi "terkejut." Dari hasil pengujian arsitektur CNN, didapatkan akurasi data validasi sebesar 82.37%. Angka-angka yang terdapat dalam matriks dapat digunakan untuk mengkomputasi berbagai metrik performa, seperti *precision*, *recall*, *F1-score*, dan *support*. Informasi lebih lanjut dapat ditemukan dalam Tabel 4.

Tabel 3. Precision, recall, F1-score, dan support

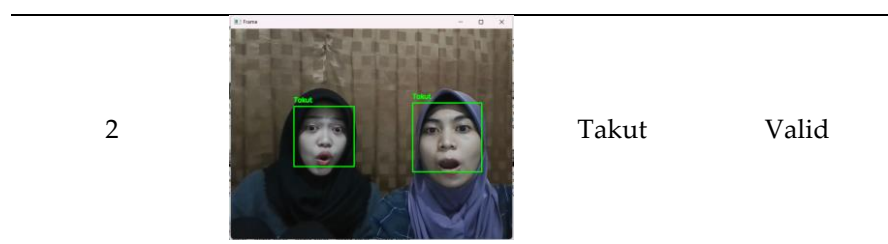
Class	Precision	Recall	F1-Score	Support
Jijik	0.75	0.55	0.64	92
Marah	0.78	0.84	0.81	185
Netral	0.98	0.88	0.93	170
Sedih	0.66	0.91	0.76	179
Senang	0.97	0.89	0.93	305
Takut	0.66	0.63	0.64	154
Terkejut	0.92	0.87	0.89	184
Akurasi			0.82	1269
Macro avg	0.81	0.79	0.80	1269
Weighted avg	0.84	0.82	0.83	1269

4.5 Implementasi

Penggunaan model CNN yang sudah disusun, dilatih, dan kemudian hasil pelatihannya disimpan dalam file model.h5 dapat dijalankan menggunakan modul opencv. Hasil uji coba dari implementasi pada modul CNN dapat dilihat pada tabel berikut.

Jumlah Wajah	Hasil Deteksi	Label	Keterangan
1		Marah	Valid
1		Netral	Valid

1		Senang	Valid
1		Takut	Valid
1		Sedih	Valid
1		Terkejut	Valid
1		Jijik	Tidak Valid
2		Senang	Valid
2		Marah	Valid



Dari hasil pengujian kinerja sistem pada tabel 4 dalam mendeteksi beberapa wajah, dapat disimpulkan bahwa kualitas deteksi untuk beberapa wajah sudah cukup akurat. Namun hasil pengujian pada label "jijik" diprediksi oleh sistem sebagai label sedih, oleh karena itu hasil pengujian tersebut tidak valid. Selain itu, dalam pengujian dengan beberapa wajah, sistem mampu mendeteksi semua wajah yang tertangkap kamera dengan cukup baik.

5. Kesimpulan dan Saran

A. Kesimpulan

Berdasarkan temuan penelitian, penggunaan metode *Convolutional Neural Network* (CNN) dalam mengklasifikasikan emosi anak telah berhasil diaplikasikan untuk mengenali kategori emosi pada anak. Berdasarkan hasil pada epoch ke-50, diperoleh nilai loss pada data pelatihan sebesar 0.8927 dengan akurasi sebesar 68.79%. Kemudian, pada data validasi, nilai loss mencapai 0.5388 dengan akurasi sebesar 82.37%. Model memiliki kinerja yang baik dalam mempelajari pola dari data pelatihan, namun pada label "jijik," sistem belum dapat mendeteksi dengan akurat karena kurangnya dataset. Pengujian model CNN menghasilkan akurasi yang baik pada beberapa kelas emosi. Emosi "Jijik" dengan akurasi sebesar 64%, sementara emosi "Takut" memiliki akurasi sebesar 64%, emosi "Marah" 81%, emosi "Senang" 93%, dan "Terkejut" 89%. Secara umum, model klasifikasi ini sudah cukup baik dalam memprediksi emosi dari data uji. Namun, ada beberapa kelas emosi yang perlu ditingkatkan performanya, seperti "jijik" dan "takutt."

B. Saran

- 1) Jumlah data uji yang digunakan masih cukup sedikit. Dengan meningkatkan jumlah data uji, model klasifikasi dapat belajar dengan lebih baik untuk memprediksi emosi yang lebih beragam secara akurat.
- 2) Model klasifikasi CNN yang digunakan saat ini adalah model CNN sederhana. Dengan menggunakan model klasifikasi yang lebih kompleks, seperti CNN dengan lebih banyak layer atau model lain seperti RNN, performa model klasifikasi dapat ditingkatkan.

Referensi

- [1] I. Lasri, A. R. Solh, and M. El Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network," *2019 3rd Int. Conf. Intell. Comput. Data Sci. ICDS 2019*, no. October 2019, pp. 1–6, 2019, doi: 10.1109/ICDS47004.2019.8942386.
- [2] F. Azizi, "Deteksi Emosi Menggunakan Citra Ekspresi Wajah Secara Otomatis," *Informatics Eng.*, pp. 1–44, 2021, [Online]. Available: <https://dspace.uui.ac.id/handle/123456789/35834>
- [3] J. G. Negrão *et al.*, "The Child Emotion Facial Expression Set: A Database for Emotion Recognition in Children," *Front. Psychol.*, vol. 12, no. April, pp. 1–9, 2021, doi: 10.3389/fpsyg.2021.666245.
- [4] L. Bayet and C. A. Nelson, *Handbook of Emotional Development*. 2019. doi: 10.1007/978-3-030-17332-6.

- [5] S. R. M. R. E. A. and N. D., “基因的改变NIH Public Access,” *Cancer Prev Res*, vol. 6, no. 5, pp. 477–482, 201AD, doi: 10.1002/imhj.20175.SIGNS.
- [6] “Aspek Fisik Dalam Perkembangan Otak Si Kecil - Nutriclub.” [Online]. Available: <https://www.nutriclub.co.id/article-bayi/stimulasi/tumbuh-kembang-anak/aspek-fisik-dalam-perkembangan-otak-si-kecil>
- [7] H. M. Milojevich, K. A. Lindquist, and M. A. Sheridan, “Adversity and Emotional Functioning,” *Affect. Sci.*, vol. 2, no. 3, pp. 324–344, 2021, doi: 10.1007/s42761-021-00054-w.
- [8] Y. Sazaki *et al.*, “Pelatihan Pemakaian Aplikasi Permainan Untuk Menilai Karakter Anak Usia Dini Di TK Bintang Kecil Palembang Sumatera Selatan,” pp. 95–106, 2023.
- [9] S. R. Suartika E. P, I Wayan, Wijaya Arya Yudhi, “Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) Pada Caltech 101,” *J. Tek. ITS*, vol. 5, no. 1, p. 76, 2016, [Online]. Available: <http://repository.its.ac.id/48842/>
- [10] B. C. L. Adiatma, E. Utami, and A. D. Hartanto, “Pengenalan Ekspresi Wajah Menggunakan Deep Convolutional Neural Network,” *Explore*, vol. 11, no. 2, p. 75, 2021, doi: 10.35200/explore.v11i2.478.
- [11] M. A. Ozdemir, B. Elagoz, A. Alaybeyoglu, R. Sadighzadeh, and A. Akan, “Real time emotion recognition from facial expressions using CNN architecture,” *TIPTEKNO 2019 - Tip Teknol. Kongresi*, pp. 1–4, 2019, doi: 10.1109/TIPTEKNO.2019.8895215.
- [12] T. Patgar and Triveni, “Convolution Neural Network Based Emotion Classification Cognitive Model for Facial Expression,” *Turkish J. Comput. Math. Educ.*, vol. Vol.12 No., no. 13, pp. 6718–6739, 2021.
- [13] U. Laraib, A. Shaukat, R. A. Khan, Z. Mustansar, M. U. Akram, and U. Asgher, “Recognition of Children’s Facial Expressions Using Deep Learned Features,” *Electron.*, vol. 12, no. 11, pp. 1–16, 2023, doi: 10.3390/electronics12112416.
- [14] A. R. Kusumastuti, Y. Kristian, and E. Setyati, “Klasifikasi Ketertarikan Anak PAUD Melalui Ekspresi Wajah Menggunakan Metode CNN,” *J. Teknol. Inf. dan Terap.*, vol. 7, no. 2, pp. 92–96, 2020, doi: 10.25047/jtit.v7i2.176.
- [15] K. Liu, M. Zhang, and Z. Pan, “Facial Expression Recognition with CNN Ensemble,” *Proc. - 2016 Int. Conf. Cyberworlds, CW 2016*, pp. 163–166, 2016, doi: 10.1109/CW.2016.34.
- [16] I. Azhari, “Implementasi Algoritma Convolutional Neural Network Dalam Deteksi Emosi Manusia Berdasarkan Ekspresi Wajah,” *eProsiding Tek. Inform.*, vol. 1, no. 1, pp. 112–118, 2021, [Online]. Available: <http://eprosiding.ars.ac.id/index.php/pti/article/view/198%0Ahttps://eprosiding.ars.ac.id/index.php/pti/article/download/198/113>
- [17] M. R. Appasaheb Borgalli and D. S. Surve, “Deep learning for facial emotion recognition using custom CNN architecture,” *J. Phys. Conf. Ser.*, vol. 2236, no. 1, 2022, doi: 10.1088/1742-6596/2236/1/012004.
- [18] “Convolutional Neural Networks (CNNs) - Definisi dan Penjelasan.”
- [19] SuperDataScience Team, “Convolutional Neural Networks (CNN): Step 4 - Full Connection - Blogs - SuperDataScience | Machine Learning | AI | Data Science Career | Analytics | Success,” *SuperDataScience*. 2018. [Online]. Available: <https://www.superdatascience.com/blogs/convolutional-neural-networks-cnn-step-4-full-connection>
- [20] “What is a max pooling layer in CNN_.”
- [21] K. Azmi, S. Defit, and S. Sumijan, “Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat,” *J. Unitek*, vol. 16, no. 1, pp. 28–40, 2023, doi: 10.52072/unitek.v16i1.504.