

Sistem Deteksi Ekspresi Wajah Berbasis *Convolutional Neural Network* (CNN) Untuk Pengenalan Emosi Manusia

Facial Expression Detection System Based on Convolutional Neural Network (CNN) for Human Emotion Recognition

Fahrim Irhamna Rachman¹⁾, Rizki Yusliana Bakti²⁾, Titin Wahyuni³⁾, Arya Wibawa. Ar⁴⁾

^{1,2,3,4}Program Studi Informatika Universitas Muhammadiyah Makassar

^{1,2,3,4}Jl. Sultan Alauddin No 259, Kota Makassar, Sulawesi Selatan 90123

E-mail: fachrim141020@unismuh.ac.id¹⁾, Rizkiyusliana@unismuh.ac.id²⁾, Titinwahyuni@unismuh.ac.id³⁾, 105841109320@student.unismuh.ac.id⁴⁾

Abstrak – Pengenalan sistem deteksi ekspresi wajah manusia telah menjadi topik penelitian yang semakin berkembang, terutama dalam upaya mengembangkan aplikasi yang mampu memahami dan merespons emosi manusia secara otomatis. Pada penelitian ini bertujuan untuk mengembangkan dan mengevaluasi sistem deteksi ekspresi wajah manusia menggunakan metode *Convolutional Neural Network (CNN)*. Dataset yang digunakan terdiri dari gambar wajah dengan berbagai ekspresi yang diambil dari sumber yang beragam. Data ini diproses melalui beberapa tahap preprocessing, termasuk normalisasi, augmentasi, dan pembagian data menjadi set latih dan uji. Penelitian ini menggunakan beberapa arsitektur *CNN* untuk mengidentifikasi emosi seperti senang, sedih, marah, dan takut. Pengujian dilakukan menggunakan berbagai parameter, termasuk pembagian data latih dan uji, serta arsitektur *CNN* yang berbeda. Hasil penelitian menunjukkan bahwa model *CNN* mampu mencapai akurasi lebih dari 90% pada data pelatihan, dengan performa terbaik pada emosi "Senang" dengan *f1-score* 0.93. Namun, terdapat penurunan akurasi pada data validasi, dengan rata-rata akurasi keseluruhan sebesar 78%, menunjukkan adanya tantangan dalam generalisasi model. Selain itu, emosi "Sedih" memiliki *recall* terendah sebesar 0.49, mengindikasikan perlunya peningkatan model dalam klasifikasi emosi tertentu. Penelitian ini berkontribusi pada pengembangan sistem deteksi ekspresi wajah berbasis *CNN*, namun diperlukan eksplorasi arsitektur lebih kompleks, evaluasi dataset beragam, dan pengujian data *real-time* untuk peningkatan performa.

Kata Kunci: Ekspresi, Deteksi, dan *Convolutional Neural Network (CNN)*

Abstract – The development of human facial expression detection systems has become a growing research topic, particularly in efforts to create applications capable of automatically understanding and responding to human emotions. This research aims to develop and evaluate a human facial expression detection system using the *Convolutional Neural Network (CNN)* method. The dataset used consists of facial images with various expressions sourced from diverse origins. The data undergoes several preprocessing stages, including normalization, augmentation, and splitting into training and test sets. This study employs several *CNN* architectures to identify emotions such as happy, sad, angry, and scared. Testing is conducted using various parameters, including training and test data splits, as well as different *CNN* architectures. The results show that the *CNN* model can achieve over 90% accuracy on training data, with the best performance on the "Happy" emotion, achieving an *f1-score* of 0.93. However, there is a decrease in accuracy on validation data, with an overall average accuracy of 78%, indicating challenges in model generalization. Additionally, the "Sad" emotion has the lowest recall of 0.49, indicating the need for model improvement in classifying specific emotions. This study contributes to the development of *CNN*-based facial expression detection systems, but further exploration of more complex architectures, evaluation with diverse datasets, and real-time testing are needed to improve system performance.

Keywords: Expression, Detection, *Convolutional Neural Network (CNN)*

PENDAHULUAN

Wajah adalah fitur utama yang digunakan untuk mengenali seseorang. Ekspresi wajah manusia mencerminkan emosi dan perasaan mereka saat berinteraksi sosial. Ekspresi wajah dapat mengungkapkan emosi yang dirasakan seseorang (Widodo et al., 2022). Ekspresi wajah adalah bentuk komunikasi non-verbal yang sangat penting dalam

interaksi manusia. Wajah sering kali menjadi refleksi dari emosi, perasaan, dan pikiran seseorang. Melalui ekspresi wajah, seseorang bisa menunjukkan kebahagiaan, kesedihan, ketakutan, atau berbagai emosi lainnya tanpa harus mengucapkan kata-kata. Dengan memahami emosi yang dialami seseorang, kita dapat menilai kondisi psikologisnya (Achmad et al., 2019)).

Dengan kemajuan teknologi pengolahan citra dan kecerdasan buatan, menjadi mungkin untuk mengembangkan sistem yang lebih akurat dan efisien dalam mendeteksi ekspresi wajah. Salah satu metode yang telah terbukti efektif adalah penggunaan *Convolutional Neural Network (CNN)*, yang memiliki tingkat keakuratan tinggi (Muttaqiin et al., 2023). *CNN* ialah jenis arsitektur jaringan saraf tiruan yang sangat efektif untuk mengenali pola dalam data gambar, termasuk ekspresi wajah (Fasounaki et al., 2021).

CNN dirancang khusus untuk mengenali dan mengklasifikasikan gambar. *CNN* terdiri dari beberapa lapisan (layer) yang berfungsi mengekstrak informasi dari gambar dan menentukan klasifikasi gambar dalam bentuk skor klasifikasi (Nugroho et al., 2020). Sistem deteksi ekspresi wajah berbasis *CNN* memiliki berbagai aplikasi penting. Dalam interaksi manusia-komputer, teknologi ini memungkinkan pengembangan antarmuka yang lebih intuitif dan responsif, yang dapat menyesuaikan perilaku sistem berdasarkan emosi pengguna.

Penelitian ini dilatarbelakangi oleh tantangan dalam mendeteksi ekspresi wajah secara akurat, terutama ketika berhadapan dengan beragam kondisi pencahayaan, latar belakang, serta variasi individu dalam menampilkan emosi. *CNN* telah terbukti sebagai metode yang efektif dalam pengenalan gambar dan pola, sehingga diharapkan mampu mengatasi tantangan ini. Beberapa penelitian sebelumnya telah menggunakan berbagai arsitektur *CNN* untuk mendeteksi emosi dengan hasil yang bervariasi. Misalnya, penelitian oleh Musa et al. (2023) menggunakan arsitektur *Xception* dengan hasil akurasi sekitar 66%, sementara penelitian oleh Daffa Ulhaq et al. (2023) menggunakan metode *SSD MobileNet* untuk deteksi ekspresi wajah secara *real-time* dengan akurasi rata-rata 51.2%.

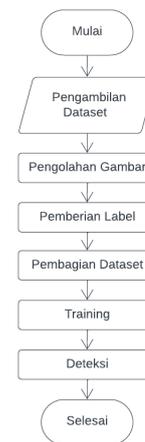
Namun, tantangan dalam generalisasi model masih menjadi isu yang signifikan. Dalam penelitian ini, dilakukan eksplorasi lebih lanjut terhadap arsitektur *CNN* menggunakan metode *MobileNet*, yang diharapkan dapat memberikan kinerja yang lebih baik dalam mengenali berbagai emosi manusia seperti senang, sedih, marah, dan takut.

Tujuan penelitian ini adalah untuk mengembangkan sistem yang mampu mendeteksi dan mengklasifikasikan ekspresi wajah manusia dengan menggunakan metode *CNN*. Selain itu, penelitian ini bertujuan untuk mengukur seberapa baik model *CNN* mampu mengenali emosi dari wajah manusia dalam berbagai kondisi. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan sistem interaksi manusia-komputer yang lebih intuitif serta membuka peluang untuk aplikasi yang lebih luas di berbagai bidang seperti kesehatan mental dan pendidikan.

METODOLOGI PENELITIAN

Convolutional Neural Network (CNN) adalah salah satu jenis deep learning yang mampu mengenali dan mengklasifikasikan gambar. *CNN* merupakan jenis neural network yang khusus menangani data dengan struktur topologi seperti grid, seperti gambar. *CNN* dapat diterapkan dalam berbagai tugas, seperti pengenalan wajah, analisis dokumen, klasifikasi gambar, klasifikasi video, dan lainnya. (Gupta et al., 2021). Metode *CNN* mengklasifikasikan gambar dengan memproses gambar yang diinput, lalu mengelompokkannya ke dalam kategori tertentu, misalnya pada citra ekspresi wajah, terdapat dahi, alis, mata, pipi, mulut dan lain-lain.

Adapun tahapan dalam perancangan sistem sebagai berikut:



Gambar 1 Perancangan Sistem

Pada gambar 1 proses dimulai dengan pengambilan dataset yang di ambil langsung pada mahasiswa Universitas Muhammadiyah Makassar dan beberapa orang di tembat umum. Pengambilan gambar dilakukan menggunakan smartphone, di mana subjek diminta mengekspresikan lima jenis emosi: netral, senang, sedih, marah, dan takut.



Gambar 2 Ekpresi senang



Gambar 3 Ekpresi sedih



Gambar 4 Ekspresi marah



Gambar 5 Ekspresi takut

Adapun jumlah dan pembagian gambar yang di kumpulkan sebagai berikut:

Tabel 1 Jumlah dataset

Tempat pengambilan data	Objek Data Yang di ambil	Jumlah Data Gambar
Universitas Muhammadiyah Makassar	Ekspresi Senang	200
	Ekspresi Sedih	200
	Ekspresi Marah	200
	Ekspresi Takut	200
	Jumlah Data Gambar	800

Dataset yang diperoleh kemudian melalui tahap preprocessing, yang mencakup normalisasi gambar, pengubahan ukuran gambar menjadi 320x320 piksel, dan augmentasi data jika diperlukan. Setelah preprocessing, dataset diberi label menggunakan platform Roboflow, di mana setiap gambar dilengkapi dengan kotak pembatas (bounding box) untuk objek yang relevan. Pembagian dataset dalam penelitian ini dilakukan dengan metode pembagian data latih (80%) dan data uji (20%). Tujuan utama dari pembagian ini adalah untuk memastikan bahwa model yang dibangun dapat belajar dari data yang memadai serta diuji secara adil dengan data yang belum pernah dilihat sebelumnya.

Model *CNN* yang digunakan pada penelitian ini adalah *MobileNet*, yang dipilih karena efisiensinya dalam mendeteksi dan mengklasifikasikan gambar dengan beban komputasi yang lebih rendah. Proses pelatihan model dilakukan dengan membagi dataset menjadi tiga bagian: data latih, data validasi, dan data uji. Data latih digunakan untuk melatih model, sementara data validasi digunakan untuk menyesuaikan parameter model guna menghindari *overfitting*.

HASIL DAN PEMBAHASAN

Adapun hasil akurasi dengan menggunakan metode *CNN* model *MobileNet*.

```

5/5 [=====] - 0s 69ms/step
Confusion Matrix:
[[27  1  1  8]
 [ 9 17  1  8]
 [ 2  1 46  1]
 [ 2  1  1 34]]
Classification Report:

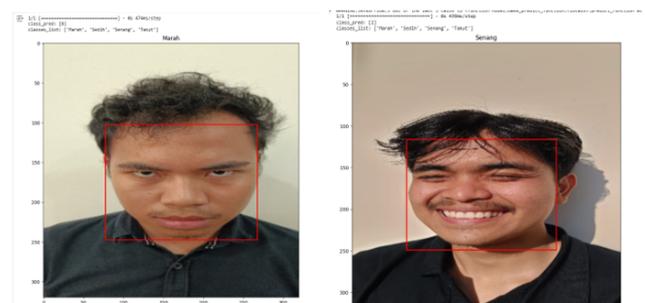
```

	precision	recall	f1-score	support
Marah	0.68	0.73	0.70	37
Sedih	0.85	0.49	0.62	35
Senang	0.94	0.92	0.93	50
Takut	0.67	0.89	0.76	38
accuracy			0.78	160
macro avg	0.78	0.76	0.75	160
weighted avg	0.79	0.78	0.77	160

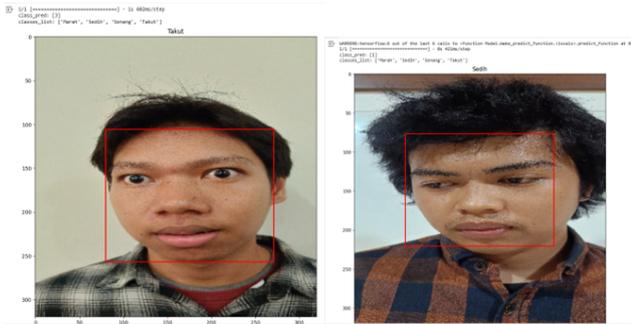
Gambar 6 Hasil training

Confusion matrix dengan epoch 150 menggambarkan distribusi prediksi model untuk setiap kelas, dengan diagonal utama menunjukkan jumlah prediksi yang benar untuk masing-masing kelas. *Precision*, *recall*, dan *f1-score* untuk setiap kelas memberikan detail lebih lanjut mengenai kinerja model. Untuk kelas Marah, *precision* adalah 0.68, *recall* 0.73, dan *f1-score* 0.70. Pada kelas Sedih, *precision* mencapai 0.85, *recall* 0.49, dan *f1-score* 0.62. Untuk kelas Senang, *precision* adalah 0.94, *recall* 0.92, dan *f1-score* 0.93. Sedangkan untuk kelas Takut, *precision* adalah 0.67, *recall* 0.89, dan *f1-score* 0.76. Model memiliki performa terbaik dalam mengklasifikasikan emosi "Senang" dengan *f1-score* 0.93, sedangkan kinerja terburuk dalam mengklasifikasikan emosi "Sedih" dengan *recall* hanya 0.49, yang menunjukkan model sering salah dalam mengidentifikasi emosi ini. Rata-rata akurasi keseluruhan model adalah 0.78, dengan *precision* rata-rata (macro avg) 0.78, *recall* rata-rata 0.76, dan *f1-score* rata-rata 0.75.

Adapun hasil dari deteksi ekspresi wajah sebagai berikut:



Gambar 7 Hasil deteksi Marah dan Senang



Gambar 8 Hasil deteksi takut dan sedih

KESIMPULAN

Metode *CNN* telah diterapkan secara efektif pada sistem deteksi ekspresi wajah, dengan akurasi klasifikasi pada data pelatihan yang meningkat dan mencapai lebih dari 90%. Namun, model menunjukkan tanda-tanda *overfitting*, terlihat dari perbedaan antara akurasi yang tinggi pada data pelatihan dan akurasi yang lebih bervariasi pada data validasi, yang menunjukkan tantangan dalam generalisasi. Confusion matrix menunjukkan distribusi prediksi model untuk setiap kelas, dengan kinerja terbaik pada emosi "Senang" (*f1-score* 0.93) dan kinerja terburuk pada emosi "Sedih" (*recall* 0.49). Secara keseluruhan, model memiliki akurasi rata-rata 0.78, *precision* rata-rata 0.78, *recall* rata-rata 0.76, dan *f1-score* rata-rata 0.75. Penelitian selanjutnya disarankan untuk mengeksplorasi arsitektur *CNN* seperti *ResNet*, *Inception*, atau *EfficientNet* guna menemukan model deteksi ekspresi wajah yang lebih optimal. Evaluasi sebaiknya dilakukan menggunakan dataset yang lebih beragam, mencakup variasi ekspresi, etnis, usia, dan kondisi pencahayaan untuk meningkatkan kemampuan generalisasi model. Selain itu, penting untuk menguji model pada data video *real-time* untuk menilai kinerja dalam kondisi dinamis dan efisiensi operasional dalam aplikasi dunia nyata.

DAFTAR PUSTAKA

- Achmad, Y., Wihandika, R. C., & Dewi, C. (2019). Klasifikasi emosi berdasarkan ciri wajah menggunakan convolutional neural network. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(11), 10595–10604.
- Daffa Ulhaq, M. R., Zaidan, M. A., & Firdaus, D. (2023). Pengenalan Ekspresi Wajah Secara *Real-time* Menggunakan Metode *SSD MobileNet* Berbasis Android. *Journal of Technology and Informatics (JoTI)*, 5(1), 48–52. <https://doi.org/10.37802/joti.v5i1.387>

- Fasounaki, M., Yüce, E. B., Öncül, S., & Ince, G. (2021). *CNN*-based Text-independent Automatic Speaker Identification Using Short Utterances. *Proceedings - 6th International Conference on Computer Science and Engineering, UBMK 2021*, 01, 413–418. <https://doi.org/10.1109/UBMK52708.2021.9559031>
- Gupta, Akanksha, Narwaria, & Pratap, D. R. (2021). Review on Deep Learning Handwritten Digit Recognition using Convolutional Neural Network. *International Journal of Recent Technology and Engineering (IJRTE)*, 9(5), 245–247. <https://doi.org/10.35940/ijrte.e5287.019521>
- Musa, P., Anam, W. K., Musa, S. B., Aryunani, W., Senjaya, R., & Sularsih, P. (2023). Pembelajaran Mendalam Pengklasifikasi Ekspresi Wajah Manusia dengan Model Arsitektur Xception pada Metode Convolutional Neural Network. *Rekayasa*, 16(1), 65–73. <https://doi.org/10.21107/rekayasa.v16i1.16974>
- Muttaqiin, A. K., Yuana, H., & Chulkamdi, M. T. (2023). Implementasi Algoritma *Convolutional Neural Network* Untuk Pengenalan Ekspresi Wajah. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, 8(2), 772–792. <https://tunasbangsa.ac.id/ejurnal/index.php/jurasi>
- Nugroho, P. A., Fenriana, I., & Ariyanto, R. (2020). Implementasi Deep Learning Menggunakan *Convolutional Neural Network (CNN)* Pada Ekspresi Manusia. *Algor*, 2(1), 12–21.
- Widodo, S., Setiawan, D., Ridwan, T., & Ambari, R. (2022). Perancangan Deteksi Emosi Manusia berdasarkan Ekspresi Wajah Menggunakan Algoritma VGG16. *Syntax : Jurnal Informatika*, 11(01), 01–12. <https://doi.org/10.35706/syji.v11i01.6594>