

Klasifikasi Ekspresi Wajah Menggunakan Metode CNN: Studi Kasus Dataset Kaggle

Wayan Restama Yasa^{a1}, AAIN Eka Karyawati ^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹wayanrestama20@gmail.com
²eka.karyawati@unud.ac.id

Abstract

This research aims to implement a Convolutional Neural Network (CNN) in facial expression classification using the Kaggle dataset which consists of five types of facial expressions, namely anger, disgust, fear, happiness and sadness. This method is considered important in supporting various applications such as emotion detection, facial recognition, and better human-machine communication. In this research, data preprocessing and augmentation were carried out using ImageDataGenerator to increase data diversity and prevent overfitting. Next, a CNN architecture is built which consists of convolution layers, pooling layers, and Dense layers. The model was trained using the Adam optimizer with a categorical crossentropy loss function for 50 epochs. The results show that the model achieves approximately 51% accuracy on the validation set. However, further analysis showed variations in model performance among facial expression classes, with some classes performing better than others.

Keywords: Facial Expression Classification, Convolutional Neural Network, Kaggle Dataset, Data Augmentation, Image Processing.

1. Pendahuluan

Wajah memiliki peran yang sentral dan penting dalam menyampaikan pesan emosional, mengenali individu, dan berkomunikasi. Dalam interaksi sosial, ekspresi wajah menjadi sarana komunikasi yang sangat penting, yang mencerminkan emosi yang dirasakan oleh seseorang. Ekspresi wajah tersebut bukan hanya sekedar pantulan dari keadaan emosi seseorang, tetapi juga dapat memberikan petunjuk terhadap perasaan emosional seseorang pada saat tertentu, sehingga memungkinkan kita untuk memahami kondisi psikologis mereka [1]. Ekspresi wajah dapat menjadi penanda yang khas dari berbagai jenis emosi yang dialami oleh seseorang. Melalui gerakan otot yang halus, wajah yang dapat menggambarkan kegembiraan, kesedihan, kejjjikan, kemarahan, ketakutan dan lainnya [2]. Pengenalan ekspresi wajah, sebagai salah satu aplikasi dari *Convolutional Neural Network* (CNN), menjadi fokus utama dalam upaya memahami dan menginterpretasikan bahasa visual manusia dengan memanfaatkan teknologi deep learning [3]. Penelitian ini bertujuan untuk mengembangkan model prediktif yang mampu secara real-time mengidentifikasi ekspresi wajah manusia. Melalui pendekatan ini, diharapkan mesin juga dapat menjadi semakin cakap dalam memahami dimensi ekspresif manusia, berkontribusi dalam berbagai bidang seperti deteksi emosi, analisis sentiment, dan antarmuka manusia-mesin yang lebih responsive. *Convolutional Neural Network* (CNN) digunakan karena kemampuannya didalam mengidentifikasi fitur – fitur penting secara otomatis dari data citra, seperti tepi, warna, dan bagian-bagian parsial dari objek. Dengan demikian, penelitian ini memusatkan perhatian pada pengembangan model CNN untuk mengenali ekspresi wajah manusia, sebagai langkah dalam memahami dimensi ekspresif manusia dengan lebih mendalam serta meningkatkan kemampuan teknologi dalam memahami bahasa visual yang kompleks.

2. Metode Penelitian

2.1 Desain Penelitian



Gambar 1. Diagram Flow Penelitian

Diagram flow penelitian diatas mengilustrasikan langkah langkah dari pengumpulan dataset Kaggle hingga evaluasi model CNN. Langkah – langkah tersebut mencakup pengumpulan dataset, *preprocessing*, augmentasi data, Pembangunan arsitektur CNN, pelatihan model, dan evaluasi model CNNnya.

2.2 Dataset

Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle yang memuat gambar-gambar ekspresi wajah manusia. Dataset ini mencakup lima jenis ekspresi emosi yang berbeda, yaitu marah, sedih, senang, jijik, dan takut. Tabel 1 menampilkan contoh data gambar dari masing-masing ekspresi, disertai dengan labelnya.

Tabel 1. Dataset Ekspresi Wajah

No	Data	Label
1		Ekspresi marah
2		Ekspresi jijik
3		Ekspresi takut

No	Data	Label
4		Ekspresi senang
5		Ekspresi sedih

Kemudian setelah seluruh dataset terkumpul dengan jumlah 1174 gambar dari seluruh ekspresi emosi. Dataset tersebut dibagi menjadi lima jenis ekspresi emosi: marah, jijik, takut, senang, dan sedih. Setiap jenis ekspresi memiliki jumlah gambar yang berbeda. Untuk keperluan pelatihan model, dataset dibagi menjadi dua bagian, yaitu data training dan data validation. Proporsi pembagian dataset untuk training dan validation adalah 70% dan 30% secara berurutan. Rincian pembagian datasetnya dapat terlihat pada Tabel 2.

Tabel 2. Spesifikasi Dataset

Ekspresi	Jumlah	Training	Validation
Marah	165	110	55
Jijik	360	240	120
Takut	324	216	108
Senang	157	105	52
Sedih	168	112	56

2.3 Preprocessing Data

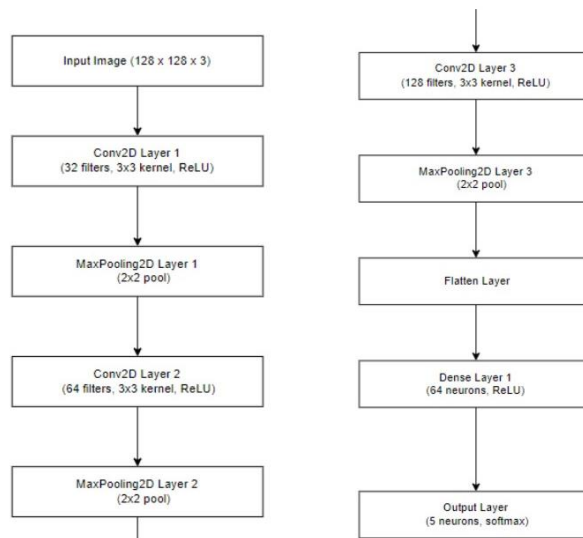
Sebelum dilakukan pelatihan model, dilakukan tahap preprocessing data terlebih dahulu. Tahap ini meliputi dua langkah utama, yaitu pengubahan ukuran gambar dan normalisasi intensitas piksel. Pertama, semua gambar dalam dataset diubah ukurannya menjadi 128x128 piksel. Hal ini dilakukan untuk memastikan bahwa semua gambar memiliki ukuran yang konsisten, memudahkan proses pelatihan model. Selanjutnya, intensitas piksel dalam gambar dinormalisasi ke rentang [0,1]. Dengan melakukan normalisasi ini, nilai piksel dalam gambar diubah sedemikian rupa sehingga nilai piksel tertinggi menjadi 1 dan nilai piksel terendah menjadi 0

2.4 Augmentasi Data

Augmentasi data dilakukan menggunakan ImageDataGenerator, sebuah alat bawaan TensorFlow yang memungkinkan untuk melakukan augmentasi data secara otomatis pada gambar-gambar dalam dataset. Teknik augmentasi yang diterapkan mencakup rotasi, pergeseran, pemotongan, dan pembalikan horizontal. Augmentasi data bertujuan untuk meningkatkan keragaman data yang tersedia untuk pelatihan model, sehingga model dapat belajar dari berbagai variasi gambar dan menjadi lebih robust terhadap variasi yang mungkin ada dalam data yang sebenarnya.

2.5 Arsitektur CNN

Arsitektur Convolutional Neural Network (CNN) telah dibangun untuk mengekstrak fitur dari gambar-gambar ekspresi wajah. Arsitektur ini terdiri dari lapisan-lapisan konvolusi yang bertugas untuk mengekstrak fitur-fitur penting dari gambar, diikuti oleh lapisan pooling untuk mereduksi dimensi fitur, dan dilanjutkan dengan lapisan-lapisan dense sebagai pengklasifikasi. Lapisan-lapisan konvolusi menggunakan filter untuk mendeteksi pola-pola visual, sedangkan lapisan pooling digunakan untuk mereduksi dimensi peta fitur. Lapisan-lapisan dense bertugas sebagai pengklasifikasi untuk menghasilkan prediksi akhir terkait dengan kategori ekspresi wajah.



Gambar 2. Arsitektur CNN

Gambar 2 menunjukkan arsitektur CNN yang digunakan dalam penelitian ini. Setiap lapisan dalam arsitektur ini dirancang untuk secara bertahap mengurangi dimensi data sambil mempertahankan informasi penting, yang akhirnya diolah oleh lapisan dense untuk menghasilkan prediksi akhir. Lapisan konvolusi dan pooling secara berurutan mengolah gambar input, mengekstrak fitur penting, dan mengurangi dimensi data sebelum data tersebut masuk ke lapisan dense untuk klasifikasi akhir.

2.6 Pelatihan dan Evaluasi Model

Setelah arsitektur CNN dibangun, selanjutnya akan dilakukan tahap tuning hyperparameter pada dataset training. Pada tahap ini, beberapa hyperparameter yang diatur dan diuji meliputi batch size dan jumlah neuron pada lapisan dense.

Tabel 3. Hasil Tuning Hyperparameter

No	Batch Size	Neuron Dense Layer	Epochs	Validation Accuracy
1	32	64	10	0.31
2	64	64	10	0.35
3	32	64	25	0.38
4	64	64	25	0.30
5	32	64	40	0.45
6	64	64	40	0.31
7	32	64	50	0.51
8	64	64	50	0.40

Dari hasil tuning hyperparameter yang dilakukan, model terbaik yang dipilih memiliki berdasarkan nilai akurasi yang tinggi yaitu pada tuning ke delapan dengan nilai akurasi 0.51%

3. Hasil dan Diskusi

3.1 Evaluasi Kinerja Model

Model yang dikembangkan mencapai akurasi sekitar 51% pada set validasi. Namun, analisis lebih lanjut menunjukkan variasi kinerja model di antara kelas-kelas ekspresi wajah, seperti yang terlihat pada Tabel 1

Table 4. Clasification Report

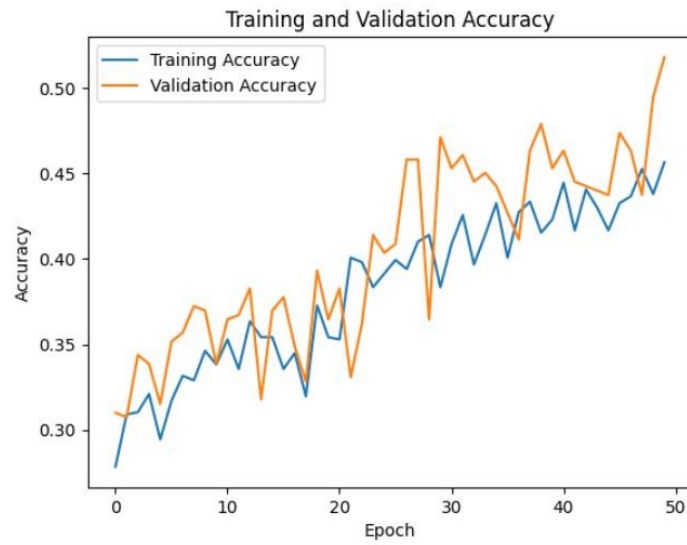
	precision	recall	f1-score	support
angry	0.07	0.02	0.03	55
disgusted	0.21	0.17	0.19	120
fearful	0.26	0.45	0.33	108
happy	0.11	0.15	0.13	52
sad	0.09	0.02	0.03	56
accuracy			0.21	391
macro avg	0.15	0.17	0.14	391
weighted avg	0.17	0.21	0.17	391

Variasi kinerja antar kelas menunjukkan bahwa beberapa kelas memiliki performa yang lebih baik daripada yang lain, sementara kinerja model cenderung rendah pada kelas dengan representasi data yang lebih sedikit atau fitur yang sulit untuk diekstrak. Tantangan dalam pengenalan ekspresi wajah muncul karena model kesulitan untuk mempelajari pola-pola yang ada dalam kelas-kelas tersebut. Hal ini mungkin memerlukan strategi tambahan untuk meningkatkan kinerja model pada kelas-kelas tersebut.

3.2 Diskusi

Hasil evaluasi model CNN memberikan wawasan dalam pengembangan teknologi pengenalan ekspresi wajah. Namun, tantangan yang dihadapi, terutama dalam mengenali ekspresi wajah, menunjukkan bahwa masih ada ruang untuk peningkatan kinerja model secara keseluruhan. Variasi kinerja antar kelas menunjukkan bahwa beberapa kelas memiliki performa yang lebih baik daripada yang lain, sementara kinerja model cenderung rendah pada kelas dengan representasi data yang lebih sedikit atau fitur yang sulit untuk diekstrak. Tantangan dalam pengenalan ekspresi wajah muncul karena model kesulitan untuk mempelajari pola-pola yang ada dalam kelas-kelas tersebut. Hal ini mungkin memerlukan strategi tambahan untuk meningkatkan kinerja model pada kelas-kelas tersebut.

Grafik berikut menggambarkan tren akurasi dan loss selama pelatihan dan validasi model:



Gambar 3. Grafik Training and Validation Accuracy

Gambar 3 menunjukkan grafik akurasi pelatihan dan validasi pada epoch 50. Dari grafik ini dapat dilihat bahwa akurasi pelatihan dan validasi meningkat pada setiap epoch, namun pada pertengahan terlihat bahwa akurasi validation mengalami penurunan namun bisa stabil kembali.



Gambar 4. Grafik Training and Validation Loss

Gambar 4 menunjukkan grafik loss pelatihan dan validasi selama 50 epoch. Loss pada data pelatihan terus menurun seiring bertambahnya epoch menunjukkan bahwa model semakin baik dalam memprediksi data pelatihan. Namun, berbeda dengan nilai pada validasi yang sempat mengalami kenaikan yang signifikan dan kembali turun.

4. Kesimpulan

Penelitian ini berhasil menerapkan Convolutional Neural Network dalam klasifikasi ekspresi wajah menggunakan dataset Kaggle. Model yang dikembangkan mampu mengklasifikasikan ekspresi wajah dengan akurasi sekitar 51% pada set validasi. Meskipun demikian, masih diperlukan penelitian lebih lanjut untuk meningkatkan kinerja model, terutama dalam mengenali ekspresi wajah yang lebih kompleks atau jarang terjadi. Ini menunjukkan bahwa model masih memiliki kesulitan dalam mengidentifikasi pola-pola yang rumit atau jarang muncul dalam dataset.

Daftar Pustaka

- [1] E. J. D. B. AL Sigit Guntoro, "Pengenalan Ekspresi Wajah Menggunakan Convolutional Neural Network," *Jurnal Informatika Atma Jogja*, vol. 3, pp. 155-160, 2022.
- [2] D. P. Derry Alamsyah, "Implementasi Convolutional Neural Networks (Cnn) Untuk Klasifikasi Ekspresi Citra Wajah Pada Fer-2013 Dataset," *Jurnal Teknologi Informasi*, vol. 4, no. 2615-2738, 2020.
- [3] I. F. R. A. Pulung Adi Nugroho, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Pada Ekspresi Manusia," *JURNAL ALGOR*, vol. 2, 2020.
- [4] D. I. M. D. L. Soleh Ependi, "Klasifikasi Pendeteksi Wajah Berhijab Menggunakan Metode CNN," vol. 6, no. 2614-3097, pp. 3157-3164, 2022.
- [5] I. G. A. W. Ni Putu Suci Paramita, "Penerapan Metode Kompresi Wavelet dalam Pengolahan Data Gambar untuk Mengurangi Ukuran File," *JNATIA*, vol. 1, no. 2986-3929, 2023.
- [6] I. N. P. K. Q. F. I Putu Agus Aryawan, "Analisis Perbandingan Algoritma Cnn Dan Svm Pada Klasifikasi Ekspresi Wajah," *Jurnal Teknologi Informasi dan Komputer*, vol. 9, 2023.

Halaman ini sengaja dibiarkan kosong