

Pengenalan Ekspresi Wajah Menggunakan Convolutional Neural Network

AL Sigit Guntoro¹, Edy Julianto², Djoko Budiyanto³

Program Studi Informatika, Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta
 Jl. Babarsari No.43, Kec. Depok, Kabupaten Sleman 55281, Daerah Istimewah Yogyakarta, Indonesia
 Email: ¹alsigitguntoro@gmail.com, ²eddy.julianto@uajy.ac.id, ³djoko.budiyanto@uajy.ac.id

Abstrak. Wajah adalah bagian tubuh manusia yang berfungsi sebagai pusat ekspresi, pengenalan dan juga komunikasi. Dalam bersosialisasi juga wajah merupakan alat utama yang digunakan untuk pendukung secara langsung. Ekspresi wajah sendiri timbul dari emosi yang dialami seseorang yang mengakibatkan pergerakan otot pada wajah. Unikunya perbedaan emosi dapat menimbulkan perbedaan pergerakan otot pada wajah. Bentuk tersebut lalu dikalsifikasikan menjadi tujuh macam ekspresi universal menurut sebuah penelitian. Klasifikasi tersebut menjadi dasar pengolahan dataset yang nantinya akan dilatih untuk mengenali pola ekspresi wajah pada manusia. Pada penelitian ini dataset dilatih menggunakan model Convolutional Neural Network. Metode ini adalah bagian dari machine learning supervised yang memiliki 3 hidden layer. Machine learning sendiri adalah bagian dari artificial intelligence dengan ciri khasnya mempelajari sendiri dari data. Kasus pengenalan ekspresi pada wajah manusia menggunakan data berupa citra. Alur dari pengembangan model machine learning ini adalah data pre-prosesing, pembuatan model, pelatihan dan evaluasi, dan terakhir pengujian, model.

Kata kunci: Pengenalan Ekspresi Wajah Manusia, Convolutional Neural Network, Machine learning, Deep Learning, Dataset

1. Pendahuluan

Wajah adalah bagian tubuh manusia yang berfungsi sebagai pusat ekspresi, pengenalan dan juga komunikasi. Dalam bersosialisasi wajah merupakan alat pendukung secara langsung. Seseorang dapat menggunakan ekspresi wajah dari akibat emosi yang dialami seseorang. Emosi merupakan perasaan yang dapat mendorong individu untuk bertindak atau merupakan respon dari suatu stimulus[1]. Emosi dapat diketahui berdasarkan ekspresi mikro seseorang dan ekspresi mikro yang dilakukan manusia merupakan hal yang bersifat universal [2]. Dengan mengetahui emosi seseorang pada waktu dan kondisi tertentu, kita dapat menilai kondisi psikologi orang tersebut.

Emosi yang dihasilkan tersebut dapat mempengaruhi gerak otot pada wajah yang diidentifikasi sebagai ekspresi wajah. Gerak otot pada wajah dapat berbeda-beda tergantung pada emosi yang sedang dirasakan seseorang. Secara universal Ekman mengklasifikasikan ekspresi tersebut ke dalam tujuh klasifikasi diantaranya yaitu senang, sedih, jijik, marah, terkejut, dan biasa saja [2]. Tujuh klasifikasi ekspresi wajah tersebut selanjutnya dapat dipergunakan sebagai landasan untuk melakukan pengenalan ekspresi wajah menggunakan metode CNN (*Convolutional Neural Network*) dengan kumpulan citra yang sudah dilabelkan menurut tujuh klasifikasi wajah. Penelitian ini bertujuan membuat sebuah model yang dapat memprediksi ekspresi wajah seseorang secara *real-time*.

CNN sendiri adalah bagian dari *machine learning* yang menggunakan tiga *hidden layer* yang biasa disebut juga *deep learning*. Melalui *machine learning* diharapkan, tidak hanya manusia saja yang dapat membedakan ekspresi wajah seseorang. Mesin juga akan dapat membedakan ekspresi wajah melalui set data citra yang akan dilatih menggunakan metode CNN. Untuk mengubah gambar ke bentuk yang dapat dikenali oleh mesin pembelajaran dibutuhkan filter khusus yang terdapat pada CNN. Filter tersebut dapat mengidentifikasi tepian, warna, bagian parsial dari objek berupa citra. Oleh karena itu metode CNN digunakan pada penelitian ini karena dinilai cocok dengan kasus yang ada.

Berdasarkan permasalahan di atas, penelitian ini berfokus dalam pengembangan model CNN untuk pengenalan ekspresi wajah pada manusia dan pembangunan sistem berbasis *website video real-time* dengan ekspresi wajah senang, sedih, takut, jijik, marah, terkejut, dan biasa saja.

2. Tinjauan Pustaka

Machine learning merupakan salah satu cabang dari kecerdasan buatan yang dikembangkan untuk dapat belajar dengan sendiri tanpa ada intrupsi dari pengguna[3]. Kemampuan yang terdapat pada *machine learning* berasal dari analisa data yang dimiliki untuk dapat melakukan tugas tertentu. Pada penelitian ini khususnya adalah ekstraksi wajah untuk mengenali pola yang terdapat didalamnya. Saat ini pengembangan *machine learning* sudah sangat banyak dilakukan dan menggunakan berbagai macam platform AI. Pada bab ini dibahas tentang penelitian-penelitian yang sudah dilakukan sebelumnya oleh peneliti terdahulu. Penelitian tersebut dijadikan sebagai bahan referensi bagi penulis untuk menunjang penelitian yang dilakukan dengan topik tugas akhir. Penelitian ini diharapkan dapat membantu penulis dalam melakukan penelitian dan membangun system pengenalan ekspresi wajah.

Pengembangan *machine learning* untuk mengklasifikasikan ekspresi wajah dengan mesin dengan judul “Implementasi Deep Learning Menggunakan Convolutional Neural Network(CNN) Pada Ekspresi Manusia” pernah dilakukan oleh Pulung Adi Nugroho, Indah Fenriana, dan Rudy Ariyanto, M.Kom untuk mengenali ekspresi wajah pada manusia pada penelitian tersebut menggunakan metode CNN dengan jumlah dataset sebanyak 7066, dan memiliki tingkat akurasi sebesar 65%[4]

Pengembangan *machine learning* untuk pengenalan wajah dengan judul “Analisis Implementasi Preprocessing dengan Otsu-Gaussian pada Pengenalan Wajah” pernah dilakukan oleh Annahl Riadi dan Ruhmi Sulaehani. Penelitian tersebut terfokus pada ekspresi wajah untuk deteksi kepuasan pelanggan pada minimarket yang berada di kecamatan Marisa Kabupaten Pohuwanto yang tingkat pelayanannya dianggap kurang maksimal.[5]

Pengembangan *machine learning* dengan kombinasi mean denoising dan CNN juga pernah dilakukan oleh Putu Agus Eka Darma Udayana dan kadek Dwi Gandika Supartha yang bertujuan untuk mengidentifikasi titik focus pada wajah berdasarkan ciri penampakan bagian wajah yang cenderung menonjol.[6]

Dari penelitian-penelitian yang telah saya sebutkan sangat bermanfaat untuk mendukung membangun model *machine learning* untuk pengenalan ekspresi wajah menggunakan convolutional neural network.

3. Metode Penelitian

3.1. Studi Literatur

Tahapan pertama, penulis mencari referensi beberapa penelitian yang pernah dilakukan oleh peneliti terdahulu sehubungan dengan penelitian yang dilakukan dalam penelitian ini, yaitu artikel, buku, jurnal tesis, dan annual review. Tahapan ini bertujuan untuk memberikan pengetahuan sebagai dasar teori kepada penulis mengenai penelitian yang dilakukan.

3.2. Analisis Masalah

Tahapan kedua, penulis melakukan Analisa literatur yang sudah didapatkan dan observasi pada lingkungan sekitar. Berdasarkan observasi yang dilakukan untuk mengenali ekspresi wajah dibutuhkan dataset berupa citra wajah manusia dengan berbagai macam ekspresi dalam 7 klasifikasi dan juga model untuk melatih data set tersebut. Namun, dibutuhkan dataset yang sudah diolah dengan tepat, regulasi dan juga tuning parameter yang tepat agar terhindar dari underfitting dan juga overfitting.

3.3. Pengumpulan Data

Pada tahapan ketiga, mengumpulkan data sebanyak 22000 citra yang diambil dari internet. Data yang sudah dikumpulkan akan diolah terlebih dahulu sebelum nantinya dilatih dengan model CNN yang sudah dibuat.

3.4. Perancangan Algoritma

Pada tahap keempat, penulis melakukan Analisa terhadap model yang digunakan dalam membangun sistem prediksi citra ekspresi wajah. Analisa ini dilakukan dengan menggunakan Langkah kerja model CNN. Untuk membuat model dibutuhkan beberapa komponen jaringan dari CNN.

3.5. Perancangan Model

Berdasarkan pengetahuan yang diperoleh penulis, pada tahap ini penulis akan membangun model yang sesuai dengan kebutuhan dan hasil yang ingin dicapai. Perancangan model dimulai dari persiapan dataset, pelatihan dataset, evaluasi proses pelatihan, dan pengujian terhadap model.

3.6. Penulisan Laporan

Pada tahapan terakhir, penulis akan mendokumentasikan semua proses dan hasilnya pada laporan tugas akhir mengikuti format penyusunan yang sudah ditentukan.

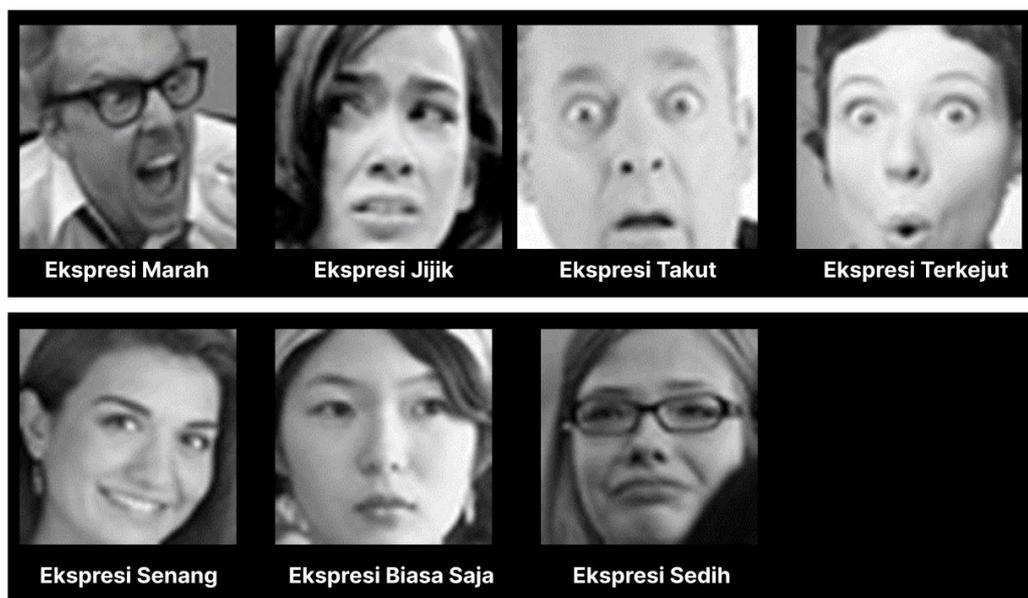
4. Hasil dan Diskusi

4.1 Dataset

Pada penelitian ini penulis mengumpulkan data berupa citra wajah manusia. Data tersebut berformat png dan juga jpeg yang terdiri dari kumpulan ekspresi yang berbeda-beda. Data tersebut diambil dari Flickr-Faces-HQ Dataset (FFHQ) yang menyeduakan sampai 22000 dataset berupa citra wajah manusia dengan ragam ekspresi, usia, jenis kelamin, dan juga warna kulit dari berbagai negara. Dengan resolusi dan dimensi yang berbeda-beda dari segi warna, pencahayaan, dan kontras.

4.1.1. Analisa Dataset

Dari data yang sudah didapatkan, lalu diproses selanjutnya dalam penelitian ini terlihat pada Gambar 1. menunjukkan citra wajah yang diklasifikasi kedalam 7 ekspresi yaitu marah, jijik, terkejut, senang, biasa saja, sedih, dan terkejut. Klasifikasi tersebut berdasarkan penelitian yang sudah dilakukan oleh Ekman.



Gambar 1. Ekspresi Wajah Manusia

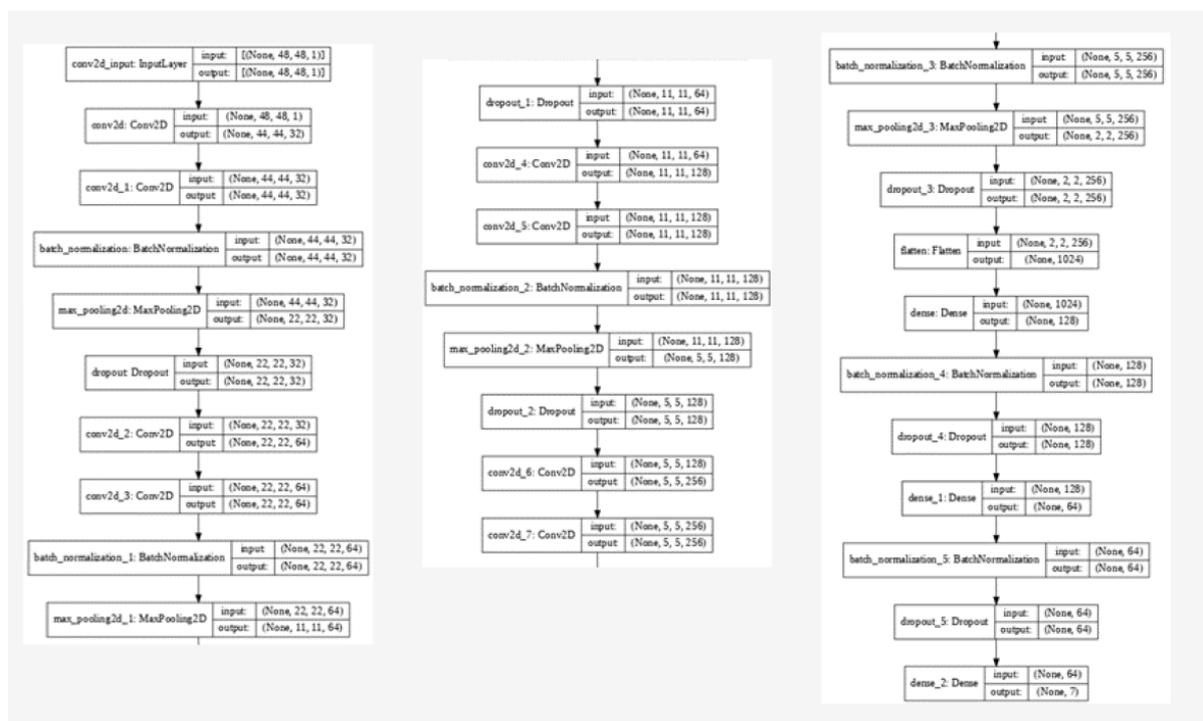
Pada penelitian ini masing-masing klasifikasi diberi label untuk dijadikan sebagai penanda dalam proses pelatihan. Pelabelan masing-masing klasifikasi dijabarkan dala Tabel 1.

Tabel 1. Rincian Pelabelan Masing-masing Kelas

Label / Kelas	Keterangan
<i>Happy</i>	Citra wajah seseorang sedang senang
<i>Angry</i>	Citra wajah seseorang sedang marah
<i>Disgust</i>	Citra wajah seseorang sedang jijik
<i>Surprise</i>	Citra wajah seseorang sedang terkejut
<i>Sad</i>	Citra wajah seseorang sedang sedih
<i>Fear</i>	Citra wajah seseorang sedang takut
<i>Neutral</i>	Citra wajah seseorang sedang biasa saja

4.2. Pengembangan Model

Pada penelitian impelentasi deep learning pada pengenalan pola ekspresi wajah menggunakan metode CNN. CNN termasuk dalam jenis Deep Neural Network karena kedalaman jaringan yang tinggi dan banyak diaplikasikan pada data citra[7] CNN digunakan untuk menganalisis gambar visual, mendeteksi dan mengenali objek pada gambar, yang merupakan vector berdimensi tinggi yang akan melibatkan banyak parameter untuk mencirikan jaringan. Secara garis besar, CNN tidak terlalu jauh berbeda dengan neuron yang memiliki weight, bias dan activation function. Arsitektur CNN ditunjukkan pada Gambar 2. dimana terdapat layer input dan dilanjutkan pada layer selanjutnya layer convolutional dan pooling, yang akan diulang berdasarkan kedalaman jaringannya atau *hidden layer*. [8]



Gambar 2. Diagram Model CNN

4.2. Pelatihan dan Evaluasi Model

Hasil uji dapat dilihat pada table 2. Nilai tertinggi pada batch size 64 dan learning rate 0.001 adalah pada epoch 200 dengan akurasi 0,82. Namun, validasi akurasi yang memiliki selisih yang jauh yang hanya 0,67 dengan selisih 15. Sehingga dapat cenderung overfitting. Oleh karena itu dipilih epoch 100 dengan hasil akurasi 0,71 dan validasi 0,65.

Tabel 2. Tingkat Akurasi Menurut Data Training.

No	Batch Size	Epoch	Accuary	Validasi Akurasi
1	64	50	0.6166	0.6296
2	64	100	0.7170	0.6571
3	64	200	0.8201	0.6725

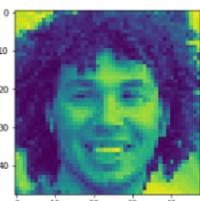
4.3. Pengujian Model

Pengujian model menggunakan data citra diluar dari pelatihan dan data validasi. Dan ditemukan hasil sebagai berikut dengan yang tertera pada Gambar 4. Model yang telah dibuat berhasil mengenali citra ekspresi wajah sesuai dengan yang diharapkan yaitu pada gambar tersebut diperlihatkan bahwa hasil prediksi dan gambar sesuai dengan yang ditentukan pada label.

```
In [17]: img = image.load_img("00508.png", target_size = (48,48), color_mode = "grayscale")
img = np.array(img)
plt.imshow(img)
print(img.shape)

label_dict = {0: 'Angry', 1: 'Disgust', 2: 'Fear', 3: 'Happy', 4: 'Neutral', 5: 'Sad', 6: 'Surprise'}
(48, 48)

0
10
20
30
40
0 10 20 30 40
```



```
In [18]: img = np.expand_dims(img, axis = 0) #makes image shape (1,48,48)
img = img.reshape(1,48,48,1)
result = model.predict(img)
result = list(result[0])
print(result)

[1.5442117e-29, 0.0, 3.717496e-28, 1.0, 0.0, 0.0, 6.920339e-35]
```

```
In [19]: img_index = result.index(max(result))
print(label_dict[img_index])
plt.show()

model.save_weights('model_weights_epoch100.h5')

Happy
```

Gambar 4. Hasil Pengujian

5. Pengujian Sistem

Berdasarkan hasil pengujian, website dapat digunakan dengan baik dan website melalui kamera yang sudah disambungkan dapat mengenali wajah secara real-time. Dalam prosesnya ekspresi yang terdeteksi juga mampu ditampung ke dalam database yang sudah dibuat.

Tabel 3.

Ekspresi	Hasil yang diharapkan	Keluaran
<i>Neutral</i>	Mesin dapat mengenali ekspresi wajah <i>neutral</i>	Berhasil
<i>Happy</i>	Mesin dapat mengenali ekspresi wajah <i>happy</i>	Berhasil
<i>Angry</i>	Mesin dapat mengenali ekspresi wajah <i>angry</i>	Berhasil
<i>Fear</i>	Mesin dapat mengenali ekspresi wajah <i>fear</i>	Berhasil
<i>Sad</i>	Mesin dapat mengenali ekspresi wajah <i>sad</i>	Berhasil
<i>Disgust</i>	Mesin dapat mengenali ekspresi wajah <i>disgust</i>	Gagal
<i>Surprise</i>	Mesin dapat mengenali ekspresi wajah <i>surprise</i>	Berhasil

5. Kesimpulan dan Saran

Berdasarkan hasil dari penelitian ini, didapati tiga kesimpulan. Pertama, implementasi dengan model CNN dengan perpustakaan tensorflow. Dengan percobaan dilakukan sebanyak tiga kali dan mengambil hasil pelatihan yang memiliki tingkat akurasi yang terbaik. Kedua, tingkat akurasi yang terbaik dari model yang telah dibangun merupakan hasil dari percobaan pelatihan sebanyak tiga kali. Model CNN yang dibuat menghasilkan training accuracy sebesar 0,71 dan besar validation accuracy sebesar 0,65. Ketiga, implementasi model dibuat pada website dengan video-realtime. Dalam membangun video real-time penulis menggunakan framework Django dan beberapa library dari python. Dan mendapatkan hasil yang cukup memuaskan meskipun ada beberapa kekurangan seperti tidak bisa mendeteksi 1 klasifikasi yaitu disgust.

Dari hasil yang sudah didapatkan, dari membuat model sampai implementasi . didapatkan saran pengembangan lebih lanjut untuk model *machine learning* yang telah dibuat yaitu dengan meningkatkan lagi akurasi yang telah diperoleh dengan menambahkan regulasi lain agar tidak terjadi overfitting dan penambahan data yang baik agar tingkat dapat meningkat.

Referensi

- [1] D. Goleman, "Kecerdasan Emosional," 2002.
- [2] P. Ekman, "Universals and Cultural Differences in Facial Expressions of Emotion BT - Nebraska Symposium on Motivation," *Nebraska Symposium on Motivation*, vol. 19. pp. 207–282, 1972, [Online]. Available: papers3://publication/uuid/FDC5E29A-0E28-4DDF-B1A4-F53FEE0B4F70.
- [3] I. Cholissodin and A. A. Soebroto, "AI , *MACHINE LEARNING* & DEEP LEARNING (Teori & Implementasi)," no. December, 2021.
- [4] P. A. Nugroho, I. Fenriana, and R. Arijanto, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (Cnn) Pada Ekspresi Manusia," *Algor*, vol. 2, no. 1, pp. 12–21, 2020.
- [5] A. Riadi and R. Sulaehani, "Analisis Implementasi Preprocessing Dengan Otsu-Gaussian Pada Pengenalan Wajah," *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 200–205, 2019, doi: 10.33096/ilkom.v11i3.457.200-205.
- [6] I. P. A. E. D. Udayana and I. K. D. G. Supartha, "Implementasi Kombinasi Metode Mean Denoising dan Convolutional Neural Network pada Facial Landmark Detection," *J. Nas. Pendidik. Tek. Inform.*, vol. 10, no. 1, p. 1, 2021, doi: 10.23887/janapati.v10i1.29779.
- [7] U. Tensorflow and A. F. Gad, *Practical Computer Vision Applications Using Deep Learning with CNNs Applications Using Deep Learning with CNNs*. .
- [8] Google, "TensorFlow Guide." <https://developers.google.com>.
- [9] A. N. Link and J. T. Scott, "Performance Evaluation Metrics," *Public Account.*, no. 1995, pp. 17–21, 1998, doi: 10.1007/978-1-4615-5639-8_4.
- [10] P. Payal and M. M. Goyani, "A comprehensive study on face recognition: methods and challenges," *Imaging Sci. J.*, vol. 68, no. 2, pp. 114–127, 2020, doi: 10.1080/13682199.2020.1738741.