

KLASIFIKASI EMOSI PADA WAJAH DARI REKAMAN APLIKASI VIDEO CONFERENCE DENGAN MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK

Arthur Adhitya Marunduh ¹⁾ Lina ²⁾

^{1,2)} Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara
Jl. Letjen S. Parman No. 1, Jakarta 11440 Indonesia
arthur.535170001@stu.untar.ac.id ¹⁾, lina@untar.ac.id ²⁾

ABSTRACT

The COVID-19 pandemic has now become a very heartbreaking disaster for all inhabitants of the earth. All activities and life are certainly disrupted due to this epidemic. Education is also one of the areas that is hampered, all levels of education from TK, SD, SMP, and SMA, while in college must carry out the teaching and learning process online. By controlling emotions, it can have a good impact on controlling one's emotional intelligence. Humans generally have 2 ways of expressing their emotions, namely verbally and nonverbally. Verbal ability is to express everything that is felt consciously in words, while nonverbal ability is to express what the body feels using media on the body, such as hand movements, shrugging, facial expressions, and others. This system will classify emotions from facial expressions using the CNN method and face detection using the Viola-Jones method. Viola-jones is a method for detecting facial objects on images and CNN is a method for recognizing facial expressions. The viola-jones method has an average accuracy of 85.67% for detecting the location of faces and for CNN, the average accuracy for classifying emotions on faces is 52.46%.

Key words

Covid 19, CNN, Viola-Jones, emotions, classification

1. Pendahuluan

Manusia umumnya mempunyai 2 cara dalam mengekspresikan emosinya yaitu dengan verbal dan nonverbal. Kemampuan verbal adalah mengungkapkan segala sesuatu yang dirasakan secara sadar dengan kata-kata, sedangkan kemampuan non verbal adalah mengekspresikan apa yang dirasakan tubuh menggunakan media yang ada pada tubuh, seperti gerakan tangan, mengangkat bahu, ekspresi wajah, dan lainnya. Wajah merupakan sebuah sistem multi sinyal atau multi pesan. Wajah dapat memperlihatkan satu macam sinyal tetapi memiliki banyak pesan yang terkandung di dalamnya. Hal ini membuat seseorang bisa saja salah mengartikan ekspresi

dari sinyal ataupun wajah seseorang karena banyaknya makna dari satu ekspresi wajah. Sulit untuk mengetahui kesan pada wajah seseorang secara pasti, karena tidak cukup kata untuk menggambarkan wajah seseorang. Tidak mudah menggambarkan ekspresi wajah yang merupakan fenomena visual, sehingga dibutuhkan data-data untuk mendefinisikan ekspresi dari wajah manusia.[1]

Perkembangan teknologi pada saat ini menyebabkan banyak peneliti melakukan penelitian terhadap klasifikasi emosi pada wajah dengan menggunakan metode yang beragam. Klasifikasi emosi pada wajah dalam teknologi disebut sebagai salah satu bagian dari Computer vision dan Artificial intelligence (AI). Artificial intelligence (AI) adalah program yang dapat membuat komputer atau mesin untuk berpikir cerdas seperti halnya manusia. Sebagai contoh dalam perkembangan Artificial intelligence saat ini, facebook dapat melakukan pengenalan terhadap orang-orang yang berada di dalam suatu foto yang diupload, dan juga memberi saran untuk memberikan tag kepada orang yang bersangkutan. Kemampuan tersebut merupakan cabang dari Artificial intelligence itu sendiri yaitu Computer vision. Computer vision adalah suatu teknik dalam ilmu komputer yang dapat melakukan ekstraksi informasi yang berada di dalam sebuah citra (baik statis maupun dinamis). Computer vision merupakan bagaimana cara suatu mesin dapat melihat lalu mengambil informasi-informasi yang terkandung pada suatu citra.[2]

CNN merupakan salah satu jenis dari neural network yang digunakan biasanya untuk data image. CNN dapat juga digunakan untuk melakukan pendeteksian dan mengenali object pada suatu image. Secara keseluruhan CNN tidak jauh berbeda dengan neural network pada umumnya. CNN terdiri atas neuron-neuron yang memiliki weight, bias dan activation function.[3]

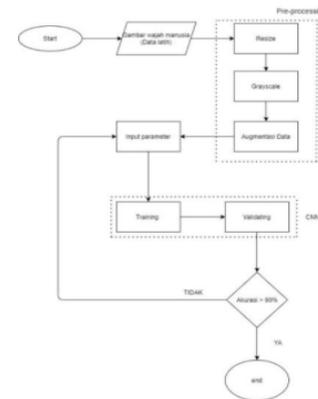
Batasan dari rancangan pada sistem yang akan dibuat adalah hanya dapat mendeteksi 4 emosi yaitu senang, sedih, marah, dan juga netral dan juga Data uji harus berupa input dari rekaman video conference

seperti zoom atau microsoft teams dalam 1 halaman sebanyak 2 orang, 5 orang dan juga 10 orang. Data latih yang akan digunakan adalah dari Facial Expression Recognition (FER)[4]. FER dikenalkan pada international conference on machine learning (ICML) 2013. FER memiliki 35.887 citra wajah yang dibagi menjadi 7 jenis emosi yang berbeda. Untuk perancangan pada sistem klasifikasi emosi pada wajah dari rekaman video conference akan digunakan 4 emosi dasar pada manusia.[5]

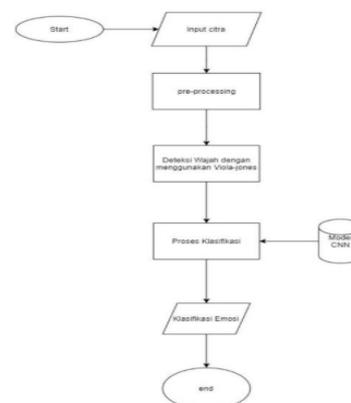
2. Metode

2.1 Design

Program pengklasifikasi emosi ini dirancang dengan menggunakan metode convolutional neural network (CNN) yaitu salah satu dari jenis neural network yang digunakan untuk data image. Metode ini cocok untuk melakukan pengklasifikasian berdasarkan kategori yang telah ditentukan karena CNN dapat digunakan untuk mengenali citra dan pemrosesan secara khusus yang didesain memproses data piksel. Tahap awal dari system ini adalah dari citra masukan(input image), dilakukan preprocessing pada citra yaitu berupa resize, grayscaling dan augmentasi data. Untuk tahap augmentasi data dilakukan beberapa proses yaitu zoom in secara acak dan juga melakukan rotasi pada citra latih dengan maksimal 90°. Tahap preprocessing ini dilakukan dengan tujuan agar meningkatkan kualitas citra sebelum melalui proses training agar informasi dari citra dapat dilatih dengan menghasilkan hasil yang maksimal. Setelah dilakukan preprocessing pada data citra selanjutnya adalah hasil dari pre-processing tersebut dimasukan kedalam layer dari pembelajaran mesin dengan menggunakan metode CNN. Setelah itu dilakukan proses testing untuk menguji berapa besar tingkat keberhasilan dari pembelajaran mesin dengan menggunakan metode CNN. Untuk perhitungan tingkat akurasi akan dihitung secara manual dan dibandingkan dengan hasil dari klasifikasi oleh sistem yang akan dirancang. Skema Pelatihan klasifikasi emosi dapat dilihat pada Gambar 1 dan Skema perancangan pengujian klasifikasi emosi dapat dilihat pada gambar 2.



Gambar 1 skema perancangan pelatihan Klasifikasi emosi



Gambar 2 skema perancangan pengujian Klasifikasi emosi

2.2 Data Yang digunakan

Data yang digunakan oleh aplikasi berupa data dalam bentuk citra dua dimensi. Data tersebut dibagi menjadi 2 yaitu citra latih dan citra validasi. Citra latih adalah citra yang digunakan untuk melatih model CNN. Citra validasi adalah citra yang digunakan untuk melakukan evaluasi model CNN yang telah dilatih. Citra latih dan validasi untuk metode CNN berupa citra ekspresi wajah manusia yang terdiri dari 4 kelas dan setiap kelas memiliki jumlah yang berbeda beda. Data citra didapatkan dari website Kaggle yang bernama fer2013. Format bobot hasil pelatihan dari model CNN berformat .h5. Detail Jumlah data latih untuk proses pelatihan dapat dilihat pada Tabel 1 dan Detail Jumlah data validasi untuk proses pelatihan dapat dilihat pada Tabel 2.

Tabel 1 Detail Jumlah data latih untuk proses pelatihan.

Kelas	Jumlah Data
Marah	3993
Sedih	4938
Senang	7164
Netral	4982

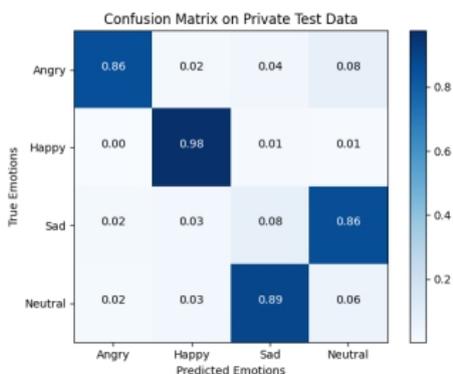
TOTAL	21077
-------	-------

Tabel 2 Detail Jumlah data validasi untuk proses pelatihan.

Kelas	Jumlah Data
Marah	960
Sedih	1139
Senang	1825
Netral	1216
TOTAL	5140

3. Hasil Percobaan

Pengujian untuk model CNN dilakukan terhadap model yang sudah dilatih dengan jumlah epoch sebanyak 200 epoch. Pelatihan dilakukan menggunakan jumlah data latih yang berupa citra ekspresi wajah manusia yang didapatkan dari website Kaggle dengan nama fer2013. Detail jumlah data pada setiap kelas untuk proses pelatihan dapat dilihat pada Tabel 1. Detail jumlah data validasi pada setiap kelas untuk proses pelatihan dapat dilihat pada Tabel 2. Setelah dilakukan proses pengujian pada model CNN dengan menggunakan arsitektur VGG-16 yang sudah dilatih didapatkan akurasi yang tinggi untuk klasifikasi emosi pada setiap citra yaitu sebesar 88.39% dan akurasi validasi sebesar 72%. Untuk nilai hasil pengujian confusion matrix dapat dilihat pada Gambar 3.



Gambar 3 hasil pengujian confusion matrix

Pengujian terhadap hasil keluaran program dilakukan untuk mengetahui akurasi klasifikasi emosi pada wajah manusia dari metode-metode yang digunakan program terhadap dua data pengujian yaitu dataset pengujian 1 dan dataset pengujian 2 dimana dataset pengujian 1 adalah sebanyak 68 citra uji dari 17 rekaman video conference yang telah dibagi menjadi 4 kelas dan dataset pengujian 2 adalah hasil dari deteksi langsung dari kamera (live detection) sebanyak 64 citra set dari 16 orang. Sedangkan untuk mengetahui apakah sebuah klasifikasi emosi pada wajah manusia ini berhasil

dideteksi dengan baik atau tidak, digunakan patokan saat pengambilan data set pengujian, yaitu dimana pada saat pengambilan data set pengujian instruktur menyebutkan emosi yang harus ditunjukkan dan semua peserta menunjukkan emosi tersebut itulah yang menjadi patokan untuk menentukan baik atau tidaknya hasil yang terdeteksi oleh program.

Hasil persentase keberhasilan untuk setiap data set pengujian akan dijabarkan berdasarkan keberhasilan pengklasifikasian emosi pada wajah manusia untuk seluruh data set pengujian dari rekaman video conference dengan jumlah data sebanyak 68 citra dari 17 rekaman video conference dan dibagi menjadi 4 kelas. Hasil dari klasifikasi dari dataset pengujian pertama dapat dilihat pada Tabel 3.

Tabel 3 Hasil klasifikasi pada Dataset pengujian pertama.

Skenario	Akurasi
Berdasarkan Jumlah Area Terdeteksi	31.01%
Berdasarkan Jumlah Citra Uji	39.65%

Selanjutnya untuk pengujian dengan menggunakan Dataset pengujian kedua yaitu Dataset pengujian dari hasil pengklasifikasian melalui kamera secara langsung (live detection). Jumlah Data Set pengujian dari hasil klasifikasi melalui kamera adalah sebanyak 64 citra wajah dari 16 orang yang telah terbagi dalam 4 kelas yaitu senang, sedih, marah dan juga netral. Hasil dari klasifikasi dari Dataset pengujian kedua dapat dilihat pada Tabel 4.

Tabel 4 Hasil klasifikasi pada Dataset pengujian kedua.

Skenario	Akurasi
Berdasarkan Jumlah Area Terdeteksi	72.06%
Berdasarkan Jumlah Citra Uji	67.12%

Pengujian juga dilakukan berdasarkan jenis emosi pada Dataset pengujian pertama. Jenis emosi pada wajah yang digunakan adalah senang, marah, netral dan marah. Pada keseluruhan citra yang dideteksi yaitu berjumlah 68 citra rekaman video conference, terdapat 147 emosi marah, 147 emosi senang, 147 emosi sedih dan 146 emosi netral. Perhitungan akurasi pada emosi marah yang terdeteksi pada citra dengan jumlah citra yang terdapat emosi marah adalah 11/147 sebesar 7.48%. Begitu juga dengan perhitungan akurasi pada emosi senang yaitu 84/147 sebesar 57.14%, netral 38/147 sebesar 25.85% dan sedih 49/146 sebesar 33.56%.

Tabel 5 Hasil klasifikasi pada Dataset pengujian pertama berdasarkan jenis emosi..

Kelas	Jumlah	Terdeteksi	Akurasi
Marah	147	11	7.48%
Senang	147	84	57.14%
Netral	147	38	25.85%
Sedih	147	39	33.56%

Pengujian dilakukan berdasarkan jenis emosi pada Dataset pengujian kedua yaitu deteksi langsung melalui kamera. Jenis emosi pada wajah yang digunakan adalah senang, marah, netral dan marah. Pada keseluruhan citra yang dideteksi yaitu berjumlah 64 citra tangkapan kamera, terdapat 17 emosi marah, 17 emosi senang, 17 emosi sedih dan 17 emosi netral. Perhitungan akurasi pada emosi marah yang terdeteksi pada citra dengan jumlah citra yang terdapat emosi marah adalah 8/17 sebesar 47.06%. Begitu juga dengan perhitungan akurasi pada emosi senang yaitu 17/17 sebesar 100%, netral 13/17 sebesar 76.47% dan sedih 11/17 sebesar 64.71%.

Tabel 7. Hasil klasifikasi pada Dataset pengujian kedua berdasarkan jenis emosi.

Kelas	Jumlah	Terdeteksi	Akurasi
Marah	17	8	47.06%
Senang	17	17	100%
Netral	17	13	76.47%
Sedih	17	11	64.71%

Confusion matrix juga akan diimplementasikan kepada dataset pengujian 1 yaitu rekaman video conference dan data set pengujian 2 yaitu tangkapan langsung dari kamera. Hasil klasifikasi dapat dilihat pada Tabel 8 untuk dataset pengujian 1 dan Tabel 9 untuk dataset pengujian 2.

Tabel 7. klasifikasi confusion matrix pada data set pengujian 1.

Kelas	Marah	Senang	Netral	Sedih
Marah	11	59	18	25
Senang	4	84	13	15
Netral	0	54	38	20
Sedih	3	44	14	49

Tabel 8. klasifikasi confusion matrix pada data set pengujian 2.

Kelas	Marah	Senang	Netral	Sedih
Marah	8	7	1	3
Senang	0	17	1	0
Netral	0	3	13	1
Sedih	0	5	2	1

Berdasarkan nilai yang telah didapatkan pada Tabel 7 dan Tabel 8 dapat diperoleh nilai akurasi, presisi, recall dan juga F1-Score. Untuk lebih rinci nya dapat dilihat pada Tabel 9.

Tabel 9. Nilai Akurasi, Presisi, Recall, dan F1-Score.

	Akurasi	Presisi	Recall	F1-Score
Data set Pengujian 1	39.65%	46.68%	40.16 %	43.17%
Data set Pengujian 2	67.12%	75.73%	68.53 %	71.95%

4. Kesimpulan

1. Model CNN dengan menggunakan arsitektur VGG-16 yang sudah dilakukan pelatihan menggunakan data latih FER 2013 dan dengan menggunakan 200 epoch memiliki performa akurasi yang cukup baik yaitu model memiliki akurasi sebesar 88.39% dan validasi akurasi sebesar 72%.
2. Secara keseluruhan sistem rata-rata dari Data Set pengujian pertama dan Data Set pengujian kedua mendapatkan akurasi 85.67% untuk melakukan pendeteksian wajah dan 52.46% untuk melakukan klasifikasi emosi pada wajah manusia.
3. Hasil pengujian semua modul yang terdapat pada program klasifikasi emosi dengan menggunakan metode blackbox testing dapat berjalan dengan baik. Fungsi dan fitur yang terdapat dalam setiap modul sudah berjalan sesuai dengan spesifikasi rancangan.

REFERENSI

[1] Asitiningrum, Nian and Prawitasari, Johana Endang. Hubungan Antara Minat Terhadap Komik Jepang (Manga), Jurnal Psikologi VOLUME 34.

[2] Abdurrohman, Harits. COMPUTER VISION. <https://medium.com/@Otabekku/computer-vision-versi-gue-a98774cc7522>. 4 september 2020.

[3] Connie Tee, Al-Shabi Mundher, Cheah Wooi Ping, Goh Michael. Facial Expression Recognition Using a Hybrid CNN-SIFT Aggregator. MIWAI 2017: Multi- disciplinary Trends in Artificial Intelligence pp 139-149, 2017.

[4] Giannopoulos Panagiotis, Perikos Isidoros, Hatzilygeroudis Ioannis. Deep Learning Approaches for Facial Emotion Recognition: A Case Study on FER-2013. Advances in Hybridization of Intelligent Methods pp 1-16. oktober 2017.

[5] Novyantika, Rizky Dwi. Deteksi Tanda Nomor Kendaraan Bermotor Pada Media Streaming Dengan Algoritma Convolutional Neural Network Menggunakan Tensorflow. Yogyakarta: Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia. 2018.

Arthur Adhitya Marunduh, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara.

Lina S.T., M.Kom., Ph.D., memperoleh gelar Sarjana dari Universitas Tarumanagara, Indonesia tahun 2001 dan gelar Magister dari Universitas Indonesia, Indonesia tahun 2004. Kemudian tahun 2009 memperoleh gelar Ph.D. dari Nagoya University, Jepang. Saat ini sebagai Dosen Tetap Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara.