

## **Implementation of Convolutional Neural Network (CNN) in Facial Expression Recognition**

### **Implementasi Convolutional Neural Network (CNN) Pada Pengenalan Ekspresi Wajah**

**Augyeris Lioga Seandrio<sup>1</sup>, Awang Hendrianto Pratomo<sup>2</sup>, Mangaras Yanu Florestiyanto<sup>3</sup>**

<sup>1,2,3</sup> Informatika, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

<sup>1</sup>augyerislioga.s@gmail.com, <sup>2\*</sup>awang@upnyk.ac.id, <sup>3</sup>mangaras.yanu@upnyk.ac.id

\*: *Penulis korespondensi (corresponding author)*

#### **Informasi Artikel**

*Received: April 2021*

*Revised: May 2021*

*Accepted: June 2021*

*Published: August 2021*

#### **Abstract**

*Purpose: Helping teachers to monitor student emotions by apply the Convolutional Neural Network as method to the application, as well as knowing the accuracy in recognizing facial expressions.*

*Design/methodology/approach: Using a Convolutional Neural Network to classify processing in the form of images. System development using the prototype method.*

*Findings/result: Based on the results of tests performed using 3589 basic human expression data, best accuracy value is 70.46%, precision value is 71%, and recall value is 70%.*

*Authenticity/state of the art: This study has relatively similar research themes. However, it has differences in the research method, the tools used, and the research output. In previous research, with the same objects, namely faces and facial emotions, the method used, the devices for taking images of emotions and faces, and the steps in the process were different. In this study, emotions on faces are identified through images taken in real-time using a camera and by applying the Convolutional Neural Network method with a visual group geometry (VGG) architecture include 11, 13, 16, and 19 layers which layers will generate expression probability based in 7 basic human expressions and their categories.*

#### **Abstrak**

*Keywords: Facial Expression; Classification; Deep Learning; Convolutional Neural Network*

*Kata kunci: : Ekspresi Wajah; Klasifikasi; Deep Learning; Convolutional Neural Network.*

*Tujuan: Membantu pengajar melakukan monitoring emosi siswa dengan menerapkan metode Convolutional Neural Network pada aplikasi, serta mengetahui akurasi dalam melakukan pengenalan ekspresi wajah.*

*Perancangan/metode/pendekatan: Menggunakan Convolutional Neural Network untuk mengklasifikasi pengolahan berupa citra. Pengembangan sistem menggunakan metode prototype.*

*Hasil: Berdasarkan hasil pengujian yang dilakukan dengan menggunakan 3589 data ekspresi dasar manusia*

---

mendapatkan nilai akurasi sebesar 70,46%, nilai presisi sebesar 71% dan nilai *recall* sebesar 70%.

Keaslian/ *state of the art*: Berdasarkan penelitian sebelumnya, penelitian ini mempunyai karakteristik yang relatif serupa dalam tema penelitian. Namun memiliki perbedaan pada metode penelitian, perangkat yang digunakan, dan hasil keluaran penelitian.

Pada penelitian sebelumnya, dengan objek yang sama yaitu wajah dan emosi wajah, pada metode yang digunakan, perangkat dalam pengambilan citra emosi dan wajah, serta langkah-langkah dalam prosesnya pun berbeda. Pada penelitian ini emosi pada wajah diidentifikasi melalui citra yang diambil secara *real-time* menggunakan kamera dan dengan menerapkan metode *Convolutional Neural Network* dengan arsitektur *visual group geometry* (VGG) dengan 11, 13, 16 dan 19 lapisan yang akan menghasilkan probabilitas ekspresi dalam 7 ekspresi dasar manusia beserta kategorinya.

---

## 1. Pendahuluan

Kemajuan pada bidang pengolahan citra digital membuat banyak inovasi baru dalam suatu sistem, secara tidak langsung teknologi pengolahan citra mulai menggeser sistem-sistem yang masih menggunakan teknologi lama. Tidak hanya teknologi pengolahan citra saja yang berkembang tetapi metode-metode pengenalan juga ikut berkembang seperti contohnya ialah sistem yang menggunakan pengenalan wajah sebagai sistem keamanan dan pemantauan. Banyak aplikasi yang mengembangkan pengenalan wajah yang berkaitan dengan penguatan ketepatan sistem dalam mengenali wajah, sehingga wajah masih menjadi objek yang terus menerus dikembangkan oleh peneliti [1].

Wajah merupakan salah satu instrumen yang dapat digunakan untuk mengidentifikasi seseorang [2]. Dari wajah manusia dapat menunjukkan ekspresi atau emosi mereka dalam berinteraksi sosial. Dalam interaksi tersebut, kondisi emosional yang baik menjadi hal penting dalam menjalin komunikasi yang baik. Menurut seorang psikolog bernama Mehrabian bahwa ekspresi wajah menyumbang sebesar 55% dalam penyampaian pesan [3].

Ekspresi atau mimik wajah merupakan bentuk komunikasi nonverbal yang merupakan hasil dari satu atau lebih gerakan atau posisi otot pada wajah serta dapat menyampaikan keadaan emosi dari seseorang kepada orang yang mengamatinya [3]. Menurut psikolog Amerika Paul Ekman, emosi pada dasarnya diklasifikasikan ke dalam enam jenis emosi dasar yaitu senang, sedih, terkejut, marah, takut, dan jijik[3]. Pengenalan emosi seseorang dapat membantu dalam menentukan kegiatan yang membangkitkan suasana salah satunya ialah dalam kegiatan belajar mengajar. Kestabilan emosi penting bagi keberhasilan proses belajar mengajar, karena emosi mempengaruhi siswa dalam memusatkan perhatian dan menggunakan pikirannya dengan baik [5]. Kualitas dan kuantitas belajar siswa sebagian besar dipengaruhi oleh emosi siswa tersebut dalam menjalani proses belajar mengajar [19]. Dalam proses belajar mengajar emosi positif dapat mempercepat penangkapan materi sehingga memperoleh hasil belajar yang maksimal, sebaliknya emosi negatif akan mengurangi penangkapan materi dalam proses belajar.

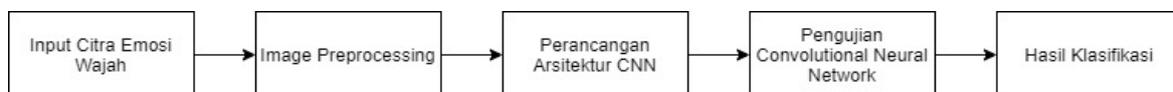
Penelitian yang berkaitan dengan wajah sebagai objek yang dideteksi sangat beragam, deteksi objek wajah dapat dilakukan dengan mencari objek citra di dalam gambar yang memiliki

kemungkinan sebagai sebuah wajah. Salah satu metode deteksi wajah yang pernah dilakukan yaitu dengan melabelkan area kulit serta memisahkan citra kulit pada wajah dan bukan, metode ini mendapatkan tingkat akurasi mencapai 81,58% [7]. Penelitian lain yang juga berkaitan dengan deteksi wajah adalah metode *viola-jones*, penelitian ini dilakukan dengan menguji objek wajah manusia dalam keadaan frontal, sistem mampu mendeteksi dengan akurasi 100% dan waktu deteksi yang didapatkan kurang dari 0.5 detik. Menurut penelitian tersebut jarak wajah yang dapat diidentifikasi ialah 21 sentimeter sampai 134 sentimeter dari camera [8]. Penelitian mengenai ekspresi wajah pernah dilakukan dengan *Local Binary Pattern* menggunakan JAFFE Database dalam mengklasifikasi ekspresi wajah berdasarkan inputan citra berupa histogram yang kemudian direpresentasikan menjadi matriks  $1 \times n$ . Tingkat keberhasilan tertinggi mencapai 84% dengan skema pengujian *person-dependent* dan konfigurasi operator LBP ( $p=8, r=8$ ), serta segmentasi 16 *region* [9]. Penelitian lain yang menggunakan *Convolutional Neural Network* (CNN) dan *K-Nearest Neighbours* (KNN) membuktikan bahwa CNN lebih baik dalam mengenali ekspresi wajah dengan menggunakan dataset JAFFE dan CK+. Akurasi didapatkan dengan menganalisis kinerja dan database yang sama, menunjukkan bahwa algoritma CNN unggul dibandingkan dengan KNN dalam mendeteksi ekspresi wajah dengan akurasi 76,7% dan 65,1% pada JAFFE dataset, serta 80,3% dan 77,2% pada CK+ dataset [10]. Penerapan algoritma *Convolutional Neural Network* (CNN) terbukti berhasil mengklasifikasi gambar wajah dengan tingkat akurasi rata – rata mencapai 87%, dengan menggunakan 7 lapisan model konvolusi pada CNN [11].

Dari permasalahan dan penelitian yang telah dijabarkan mengenai pentingnya pemantauan emosi siswa pada saat belajar mengajar, maka penelitian ini akan mengimplementasikan metode CNN sebagai pengolahan citra untuk identifikasi ekspresi wajah yang akan digunakan sebagai monitoring emosi di kelas. Kamera akan digunakan untuk merekam aktifitas emosi di kelas, serta metode deteksi wajah yang digunakan pada penelitian ini adalah Viola-jones dan tersedia dalam OPENCV dengan ekstraksi citra ekspresi wajah menggunakan CNN, yang berfokus pada 5 ekspresi yaitu frustrasi, terkejut, tidak tertarik, tertarik, dan netral yang berdasarkan 7 ekspresi dasar yaitu marah, jijik, senang, sedih, takut, terkejut, dan biasa. sehingga sistem diharapkan dapat membantu pengajar dalam pemantauan emosi serta dalam penyampaian materi.

## 2. Metode/Perancangan

Penelitian ini dilakukan dalam beberapa tahap. Tahapan-tahapan tersebut dapat dilihat pada Gambar 1.



**Gambar 1.** Tahapan Penelitian

Pada tahap pertama adalah proses memasukkan citra emosi wajah yang direkam menggunakan kamera. Kemudian tahapan *image preprocessing* dilakukan pada citra untuk mempersiapkan citra agar selanjutnya ekstraksi menggunakan *convolutional neural network*. Sebelum dilakukan klasifikasi tentu terdapat perancangan arsitektur *convolutional neural network*, pada penelitian ini arsitektur *convolutional neural network* yang digunakan ialah arsitektur *visual group geometri* (VGG) dengan 4 jenis yaitu 11 lapisan, 13 lapisan, 16 lapisan dan 19 lapisan. Tahap selanjutnya ialah pengujian *convolutional neural network* dengan menguji arsitektur CNN. Dari hasil pengujian, data akan disimpan sebagai model yang telah dilatih menggunakan *Convolutional Neural Network* yang akan melakukan klasifikasi pada citra

## 2.1 Input Citra Emosi Wajah

Tahap input citra emosi wajah dilakukan dengan merekam wajah menggunakan kamera. Citra yang di *input* akan direkam secara *real-time*, citra *input* berasal dari kamera yang memiliki minimal kualitas HD agar citra yang akan diidentifikasi dapat dikenali dengan jelas oleh mesin. Citra yang di *input* yaitu tujuh ekspresi dasar manusia yaitu marah, sedih, bahagia, terkejut, marah, jijik, dan biasa.

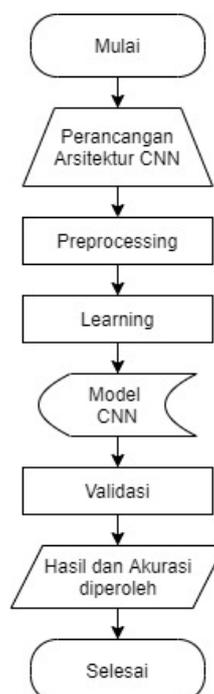
## 2.2 Image Preprocessing

Pada tahap *image preprocessing*, citra akan dilakukan dengan beberapa proses, diantaranya sebagai berikut:

1. *Grayscale* merupakan proses untuk mengubah citra masukan menjadi citra keabuan. Nilai dari tiap piksel citra masukan awal mula memiliki nilai RGB kemudian dihitung nilai rata-ratanya sehingga nilai piksel warna berubah dan menjadi nilai keabuan.
2. *Resize* merupakan proses yang berfungsi untuk meminimalisir penggunaan memori pada saat proses klasifikasi, sehingga dapat mempercepat penangkapan dan klasifikasi citra.
3. Normalisasi Citra diperlukan karena citra yang terbaca pada sistem memiliki ukuran yang cukup besar, sehingga pada proses ini akan dilakukan perubahan nilai pada piksel citra dengan membagikan nilai maksimal pada RGB untuk mempercepat proses klasifikasi citra emosi wajah.
4. Augmentasi Citra adalah teknik memanipulasi sebuah data tanpa kehilangan inti dari data tersebut. Augmentasi data diperlukan untuk training data secara signifikan, tanpa benar-benar mengumpulkan data baru. menggunakan data augmentasi dapat meningkatkan kinerja model dan memperluas limitasi sebuah dataset [10].

## 2.3 Perancangan Arsitektur CNN

Dalam membuat model untuk dapat mengenali citra yang akan diklasifikasi maupun identifikasi harus melakukan beberapa percobaan, agar bisa menyesuaikan lapisan dan parameter yang cocok dengan data yang dimiliki. Ilustrasi tahapan pembuatan model CNN dapat dilihat pada **Gambar 2**.



**Gambar 2.** Tahapan Pembuatan Model

Pada penelitian ini arsitektur *convolutional neural network* yang akan digunakan ialah arsitektur *Visual Group Geometri* (VGG). Proses klasifikasi ekspresi wajah menggunakan *convolutional neural network* dengan varian model *Visual Group Geometri* (VGG) yang terdiri dari 11, 13, 16, dan 19 lapisan[11]. Rancangan arsitektur *convolutional neural network* dengan menggunakan model VGG dapat dilihat pada **Tabel 1**.

**Tabel 1.** Rancangan Arsitektur

VGG11	VGG13	VGG16	VGG19
Input 48x48 Grayscale image			
Conv 64	Conv 64 Conv 64	Conv 64 Conv 64	Conv 64 Conv 64
Maxpooling			
Conv 128	Conv 128 Conv 128	Conv 128 Conv 128	Conv 128 Conv 128
Maxpooling			
Conv 256 Conv 256	Conv 256 Conv 256	Conv 256 Conv 256 Conv 256	Conv 256 Conv 256 Conv 256 Conv 256
Maxpooling			
Conv 512 Conv 512	Conv 512 Conv 512	Conv 512 Conv 512 Conv 512	Conv 512 Conv 512 Conv 512 Conv 512
Maxpooling			
Conv 512 Conv 512	Conv 512 Conv 512	Conv 512 Conv 512 Conv 512	Conv 512 Conv 512 Conv 512 Conv 512
Maxpooling			
FC 4096			
FC 4096			
FC 7			

Pada **Tabel 1** percobaan pertama *convolutional neural network* dengan arsitektur VGG dengan 11 lapisan menggunakan 8 lapisan konvolusi dan 3 lapisan *fully connected*. Lapisan konvolusi yang digunakan pada VGG 11 ialah 64 filter dilakukan satu kali, 128 filter dilakukan satu kali 256 filter dilakukan 2 kali, 512 filter dilakukan 4 kali.

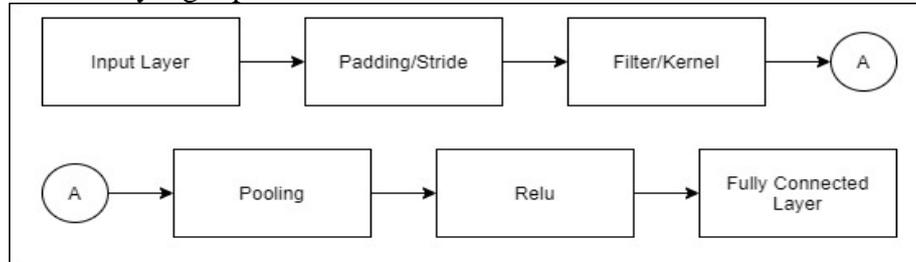
Pada percobaan kedua *convolutional neural network* dengan arsitektur VGG dengan 13 lapisan menggunakan 10 lapisan konvolusi dan 3 lapisan *fully connected*. Lapisan konvolusi yang digunakan pada VGG 13 ialah 64 filter dilakukan 2 kali, 128 filter dilakukan 2 kali, 256 filter dilakukan 2 kali, 512 filter dilakukan 4 kali.

Pada percobaan ketiga *convolutional neural network* dengan arsitektur VGG dengan 16 lapisan menggunakan 13 lapisan konvolusi dan 3 lapisan *fully connected*. Lapisan konvolusi yang digunakan pada VGG 16 ialah 64 filter dilakukan 2 kali, 128 filter dilakukan 2 kali, 256 filter dilakukan 3 kali, 512 filter dilakukan 6 kali.

Pada percobaan keempat *convolutional neural network* dengan arsitektur VGG dengan 19 lapisan menggunakan 16 lapisan konvolusi dan 3 lapisan *fully connected*. Lapisan konvolusi yang digunakan pada VGG 19 ialah 64 filter dilakukan 2 kali, 128 filter dilakukan 2 kali, 256 filter dilakukan 4 kali, 512 filter dilakukan 8 kali.

Pada **Gambar 2** tahap dimulai dari perancangan arsitektur *convolutional neural network*, untuk mengatur jumlah lapisan konvolusi, *pooling*, dan *fully connected* yang digunakan. Persiapan

dilakukan setelah dilakukan *preprocessing* berupa citra augmentasi, lalu citra hasil augmentasi akan menjadi data masukkan pada proses pembelajaran data oleh rancangan *convolutional neural network*. Pada saat dilakukan proses pembelajaran maka pola-pola setiap pixel yang ada akan diolah menjadi dua tahapan yaitu *feed-forward* dalam pengenalan setiap fitur yang terdapat pada citra, dan *backpropagation* digunakan untuk mengatasi pengecilan nilai *error*, dan *loss* yang didapatkan. Tahap selanjutnya ialah menghasilkan model *convolutional neural network* dari proses pembelajaran, dan model akan dilakukan validasi untuk mengetahui hasil dari akurasi dan *loss* yang diperoleh.



**Gambar 3.** Tahapan *Convolutional Neural Network*

Pada **Gambar 3** tahapan ekstraksi fitur menggunakan metode *Convolutional Neural Network* (CNN) memiliki 6 proses utama, yaitu:

1. *Input layer*  
*Input layer* merupakan lapisan yang menerima citra 7 emosi dasar manusia yaitu marah, senang, jijik, sedih, terkejut, dan takut, serta non ekspresi yang telah dilakukan proses *preprocessing* sebelumnya. Kemudian citra akan dilakukan proses konvolusi pada *convolution layer*.
2. *Padding/stride*  
*Padding/stride* merupakan bagian dari *convolution layer* dimana *padding* digunakan untuk mengontrol lebar dan ukuran pada kernel untuk menghasilkan output yang diinginkan, hal ini digunakan agar dimensi dari output dari *feature map* tidak banyak terbuang. Sedangkan *stride* merupakan pergeseran yang akan dilakukan oleh *filter/kernel* pada citra.[13]
3. *Filter/kernel*  
*Filter/kernel* merupakan fitur *detector* pada *convolution layer*. Kernel yang biasa digunakan pada *convolution layer* ialah *array* dua dimensi bisa berukuran 5x5, 3x3, dan 1x1. Pada proses *filtering convolutional neural network* menggunakan operasi fungsi *cross-correlation* **Persamaan 1** [13].
 
$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n) \quad (1)$$

Operasi ini dilakukan dengan melakukan kombinasi linear yang menghasilkan *feature map* yang baru dari perkalian *dot product* antara *input layer* dengan *kernel/filter*. Operasi kombinasi linear dapat dilihat pada **Persamaan 2**.

$$a . b = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_n b_n \quad (2)$$
4. *Pooling*  
*Pooling* merupakan lapisan yang terdapat pada *pooling layer* pada lapisan ini *feature map* yang dihasilkan dari *convolution layer* dilakukan *downsampling*. Hal ini dilakukan agar citra yang dihasilkan dapat merangkum informasi dan mengurangi dimensi. Pada penelitian ini proses *pooling* menggunakan *maxpooling*, dimana *downsampling* dilakukan dengan mengambil nilai terbesar pada setiap piksel.[14]
5. Aktivasi ReLu

Aktivasi ReLu merupakan aktivasi yang digunakan untuk meminimalisir *vanishing gradient*. *Vanishing gradient* merupakan problem umum yang terjadi pada *neural network* yang dalam (*deep*)[15]. Operasi ReLu akan meminimalisir penggunaan komputasi dan *vanishing gradient* pada proses pembelajaran, karena nilai negatif akan diubah menjadi 0 dan nilai positif dikalikan dengan 1. Operasi ReLu dapat dilihat pada **Persamaan 3** [16].

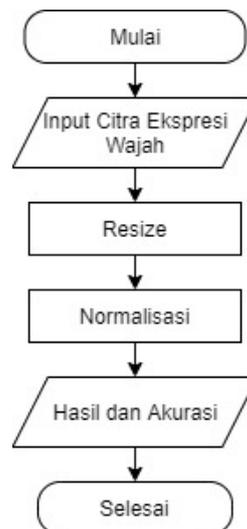
$$f(x) = \max(x,0) \tag{3}$$

6. *Fully connected layer*

Setelah didapatkan nilai pada *maxpooling* maka akan disejajarkan menggunakan *Flatten* atau lebih tepatnya akan diubah dimensinya dari 2 dimensi menjadi 1 dimensi linear, karena *neural network* hanya dapat melakukan klasifikasi secara *linear*. Proses klasifikasi pada *fully connected layer* menggunakan aktivasi *softmax*. Operasi aktivasi *softmax* dapat dilihat pada **Persamaan 4** [12].

$$y_j = \frac{e^{z_i}}{\sum_k e^{z_k}} \tag{4}$$

## 2.4 Pengujian Convolutional Neural Network



**Gambar 4.** Pengujian Model

Pada **Gambar 4**, arsitektur yang telah dirancang akan divalidasi agar mengetahui apakah susunan arsitektur dan parameter yang diproses oleh model dapat mengetahui pola yang terdapat pada emosi wajah. Dengan mengetahui nilai terbaik dari 4 arsitektur yang di uji, dengan nilai *loss* yang rendah dan proses pelatihan yang tidak memakan waktu. Akurasi terbaik akan disimpan dalam model yang akan diterapkan pada aplikasi.

## 2.5 Confusion Matrix

*Confusion matrix* merupakan informasi perbandingan hasil klasifikasi yang dilakukan oleh sistem (model) dengan hasil klasifikasi sebenarnya. Untuk menghitung performa suatu model diperlukan parameter pengukuran yaitu tingkat akurasi, recall, dan presisi. Persamaan untuk mengukur ketiga parameter tersebut disajikan pada persamaan (5), persamaan (6), dan persamaan (7)[12].

$$\text{Akurasi} = \frac{TP+TN}{TP+FP+FN+TN} \tag{5}$$

$$\text{Presisi} = \frac{\sum_{i=1}^n \left( \frac{TP_i}{FP_i+TP_i} \right)}{n} \tag{6}$$

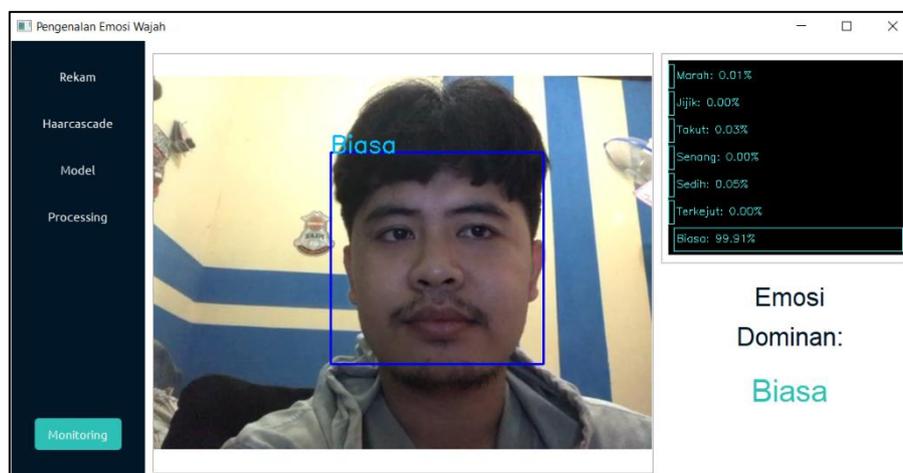
$$Recall = \frac{\sum_{i=1}^n \left( \frac{TP_i}{FN_i + TP_i} \right)}{n} \quad (7)$$

Terdapat empat istilah untuk merepresentasikan hasil proses klasifikasi yaitu TP (*True Positive*) jumlah data positif yang terklasifikasi dengan benar oleh sistem, TN (*True Negative*) jumlah data negative yang terklasifikasi dengan benar oleh sistem, FN (*False Negative*) jumlah data negative yang terklasifikasi salah oleh sistem, FP (*False Positive*) jumlah data positif yang terklasifikasi salah oleh sistem.

### 3. Hasil dan Pembahasan

#### 3.1 Hasil Interface

Hasil dari perancangan dan implementasi pada penelitian ini berupa tampilan aplikasi pada saat digunakan yaitu seperti **Gambar 5**. Halaman tersebut merupakan halaman monitoring emosi wajah dan probabilitas emosi.



**Gambar 5.** Proses Monitoring

Pada **Gambar 5** terdapat 5 tombol utama diantaranya tombol rekam, tombol haarcascade, tombol model, tombol *processing*, dan tombol monitoring. Tombol rekam berfungsi untuk mengakses kamera, tombol haarcascade digunakan untuk memanggil *library* haar-cascade untuk memproses wajah, tombol model untuk memanggil model yang telah dilatih, dan tombol *processing* digunakan untuk melihat hasil *preprocessing*. Selanjutnya, tombol monitoring yang digunakan untuk melakukan klasifikasi dan monitoring emosi wajah pada citra emosi.

#### 3.2 Hasil Pengujian

Hasil percobaan dari masing-masing arsitektur pada **Tabel 1** dapat dilihat sebagai berikut.

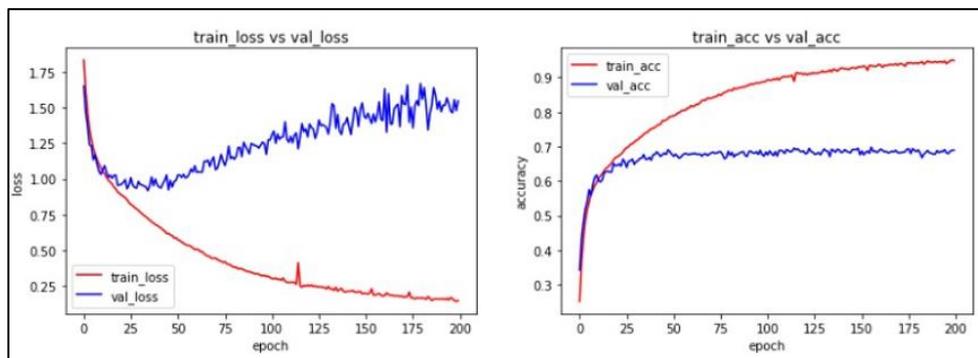
**Tabel 2.** Hasil percobaan CNN

Prediksi		Data Aktual							Total Prediksi
		Marah	Jijik	Takut	Senang	Sedih	Kaget	Biasa	
VGG19	Benar	294	39	270	762	366	338	402	2471
	Salah	197	16	258	117	228	78	224	1118
VGG16	Benar	321	40	290	776	320	331	451	2529
	Salah	170	15	238	103	274	85	175	1060
VGG13	Benar	302	40	268	758	280	352	467	2467
	Salah	189	15	260	121	314	64	159	1122
VGG11	Benar	318	44	246	753	267	348	465	2441
	Salah	173	11	282	126	327	68	161	1148

Masing-masing model menggunakan *maxpooling* untuk melakukan *down sampling* pada citra, menggunakan aktivasi ReLu, serta menggunakan *neurons* 4096 sebanyak 2 kali. Dengan bantuan 4096 neuron membantu menguraikan dan menghubungkan menjadi keluaran yang lebih sesuai, dengan menggunakan aktivasi *softmax* untuk melakukan *multiclassification*

Dilihat pada **Tabel 2** susunan arsitektur model *Visual Group Geometry* (VGG) yang menghasilkan total prediksi benar paling banyak adalah arsitektur *Visual Group Geometry* (VGG) dengan 16 lapisan. Susunan lapisan yang terdapat pada arsitektur VGG16 diantaranya 8 lapisan konvolusi, 5 lapisan *pooling*, dan 3 lapisan *fully connected layer*. Pada lapisan *fully connected layer* terdapat bantuan 4096 *neurons* dan dilakukan sebanyak 2 kali. Penggunaan *neurons* untuk membantu menguraikan dan menghubungkan menjadi keluaran yang lebih sesuai. Penggunaan 4096 *neurons* didasarkan pada arsitektur asli yang dikembangkan oleh peneliti *Visual Group Geometry*[11].

Berikut ini merupakan gambar kurva hasil selama proses pembelajaran data pada arsitektur *visual group geometry* (VGG) dengan 16 lapisan.

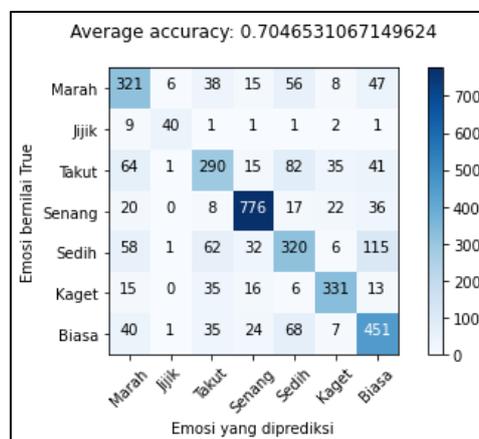


**Gambar 6.** Hasil Kurva Pembelajaran

Pada **Gambar 6** meskipun hasil kurva validasi (berwarna biru) dan data latih (berwarna oren) sedikit melebar, namun tidak menjadi masalah yang besar, karena rentang nilainya tidak jauh. Jadi kesalahan yang dihasilkan juga kecil, serta kurva validasi mengikuti trend dengan data latih.

Setelah melakukan pembuatan model serta proses validasi, maka selanjutnya ialah pengukuran proses pengujian. Pada penelitian ini pengujian tersebut dapat diketahui setiap citra yang benar dan salah pada kelasnya masing-masing. Data pengujian yang digunakan sebanyak 3589 dataset, terdapat 491 data marah, 55 data jijik, 528 data takut, 879 data senang, 594 data sedih, 416 data kaget, dan 626 data biasa.

Hasil pada pengujian menggunakan *confusion matrix* yang dilakukan tampak pada gambar berikut.



**Gambar 7.** Visualisasi Hasil Confusion Matrix VGG 16

Berdasarkan **Tabel 2**, nilai akurasi didapatkan dengan cara melakukan perhitungan akurasi seperti pada persamaan 5, presisi pada persamaan 6, dan recall pada persamaan 7. Hasil pengujian yang diperoleh dari *confusion matrix* yang telah dilakukan sebelumnya adalah sebagai berikut.

**Tabel 3.** Hasil Pelatihan Model CNN

Model	Akurasi	Presisi	Recall
VGG11	68.00%	65.86%	68.71%
VGG13	68.74%	67.43%	68.57%
VGG16	70.47%	70.86%	69.57%
VGG19	68.90%	69.00%	68.00%

Sehingga pada pengujian menggunakan *confusion matrix* tersebut model VGG dengan 16 lapisan merupakan model dengan akurasi, presisi dan recall terbaik. Akurasi yang diperoleh model VGG dengan 16 lapisan ialah akurasi 70,47%, presisi 70,86%, dan recall 69,57%.

#### 4. Kesimpulan dan Saran

Berdasarkan penelitian yang telah dilakukan, banyaknya lapisan pada lapisan konvolusi tidak menjadi jaminan baik dan buruknya hasil akurasi yang diperoleh. Arsitektur *Convolutional Neural Network* menggunakan model *Visual Group Geometri* (VGG) dengan 16 lapisan, terbukti dapat menangani data emosi wajah yang diklasifikasi. Hal itu dapat dibuktikan dengan pengujian *confusion matrix* yang menghasilkan nilai rata-rata akurasi sebesar 70,47%, presisi rata-rata sebesar 70,86%, dan *recall* rata-rata sebesar 69,57%. Pengujian dilakukan dengan menggunakan data uji sebanyak 3589 dataset, yang terdiri dari 491 data marah, 55 data jijik, 528 data takut, 879 data senang, 594 data sedih, 416 data kaget, dan 626 data biasa. Hasil klasifikasi yang kurang tepat pada penelitian ini dipengaruhi oleh faktor tertentu seperti buram pada kamera dan kemiripan antar emosi.

Beberapa hal yang dapat diberikan untuk pengembangan aplikasi ini yaitu aplikasi ini dapat dikembangkan untuk mengklasifikasi profil wajah disertai emosi, dan menyesuaikan penggunaan *pre-trained* model *convolutional neural network* agar model yang digunakan dapat meningkatkan kestabilan dalam pembelajaran oleh mesin.

#### Daftar Pustaka

- [1] M. Y. Florestiyanto, A. H. Pratomo, And N. I. Sari, "Penguatan Ketepatan Pengenalan Wajah Viola-Jones Dengan Pelacakan," *Teknika*, Vol. 9, No. 1, Pp. 31–37, 2020.
- [2] H. Abhirawan, Jondri, And A. Arifianto, "Pengenalan Wajah Menggunakan Convolutional Neural Networks (Cnn)," *Univ. Telkom*, Vol. 4, No. 3, Pp. 4907–4916, 2017.
- [3] Z. Abidin, "Pengembangan Sistem Pengenalan Ekspresi Wajah Menggunakan Jaringan Syaraf Tiruan Backpropagation (Studi Kasus Pada Database Mug)," *J. Mat. Murni Dan Terap.*, Vol. 5, No. 1, Pp. 21–30, 2011.
- [4] C. Chotimah, P. Studi, D. I. V Kebidanan, F. Kedokteran, And U. S. Maret, "Hubungan Kestabilan Emosi Dengan Prestasi Belajar Pada Siswa Kelas X Di Sma Negeri 1 Karangnom Klaten," Pp. 1–58, 2010.
- [5] R. Wijanarko And N. Eko, "Deteksi Wajah Berbasis Segmentasi Warna Kulit Menggunakan Ruang Warna Ycber & Template Matching," *Cendekia Eksata*, Vol. 2, No. 1, Pp. 1–6, 2017.

- 
- [6] I. N. Dedi Ary Prasetya, “Deteksi Wajah Metode Viola Jones Pada Opencv Menggunakan Pemrograman Python,” *Simp. Nas. Rapi Xi Ft Ums*, Pp. 18–23, 2012.
- [7] A. Saputra And T. Agung Budi W, “Pengenalan Ekspresi Wajah Menggunakan Local Binary Pattern ( Lbp ),” *Repos. Telkom Univ.*, No. 1103081084, 2015.
- [8] K. Shan, J. Guo, W. You, D. Lu, And R. Bie, “Automatic Facial Expression Recognition Based On A Deep Convolutional-Neural-Network Structure,” *Proc. - 2017 15th Ieee/Acis Int. Conf. Softw. Eng. Res. Manag. Appl. Sera 2017*, Pp. 123–128, 2017.
- [9] M. Zufar And B. Setiyono, “Convolutional Neural Networks Untuk Pengenalan Wajah Secara Real-Time,” *J. Sains Dan Seni Its*, Vol. 5, No. 2, Pp. 72–77, 2016.
- [10] C. Shorten And T. M. Khoshgoftaar, “A Survey On Image Data Augmentation For Deep Learning,” *J. Big Data*, 2019.
- [11] K. Simonyan And A. Zisserman, “Very Deep Convolutional Networks For Large-Scale Image Recognition,” *3rd Int. Conf. Learn. Represent. Iclr 2015 - Conf. Track Proc.*, Pp. 1–14, 2015.
- [12] K. Aprilianta, H. C. Rustamaji, Y. Fauziah, P. Studi, T. Informatika, And K. Sleman, “Convolutional Neural Network Dalam,” Vol. Xx, No. Xx, Pp. 1–10, 2019.
- [13] Ian Goodfellow, Y. Bengio, And A. Courville, *Deep Learning*. 2019.
- [14] H. J. Jie And P. Wanda, “Runpool: A Dynamic Pooling Layer For Convolution Neural Network,” *Int. J. Comput. Intell. Syst.*, Vol. 13, No. 1, Pp. 66–76, 2020.
- [15] J. W. G. Putra, “Pengenalan Konsep Pembelajaran Mesin Dan Deep Learning,” Vol. 4, Pp. 1–235, 2019.
- [16] S. R. Dewi, “Deep Learning Object Detection Pada Video,” *Deep Learn. Object Detect. Pada Video Menggunakan Tensorflow Dan Convolutional Neural Netw.*, 2018.
- [17] S. Ioffe And C. Szegedy, “Batch Normalization: Accelerating Deep Network Training By Reducing Internal Covariate Shift,” *32nd Int. Conf. Mach. Learn. Icml 2015*, Vol. 1, Pp. 448–456, 2015.
- [18] B. Yuwono, “Image Smoothing Menggunakan Mean Filtering, Median Filtering, Modus Filtering Dan Gaussian Filtering,” *Telematika*, Vol. 7, No. 1, 2015.
- [19] Khodijah, N. (2014). Psikologi Pendidikan. *Jakarta: Rajawali Pers*.