



## PENERAPAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK DAN HAAR CASCADE UNTUK PRESENSI DENGAN VIDEO REKAMAN ZOOM

Muhammad Verdiansyah <sup>1)</sup>, Achmad Solichin<sup>2)</sup>

<sup>1</sup> Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur

<sup>2</sup> Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi LUHUR

email: <sup>1</sup> verdiansyah1404@gmail.com, <sup>2</sup> achmad.solichin@budiluhur.ac.id

### ARTICLE INFO

#### Article History:

Received : 9 September 2022

Accepted : 17 January 2023

Published : 22 January 2023

#### Keywords:

Student presence,

Face recognition

Haar Cascade

Convolutional Neural

Network

Inception-V3

#### IEEE style in citing this article:

M. Verdiansyah and A. Solichin, "Penerapan Algoritma Convolutional Neural Network dan Haar Cascade Untuk Presensi Dengan Video Rekaman Zoom", *Jurnal.ilmiah.informatika*, vol. 7, no. 2, pp. 117-125, Dec. 2022.

#### Corresponding Author:

Muhammad Verdiansyah

Universitas Budi Luhur

### ABSTRACT

The COVID-19 pandemic that is engulfing the world is one of the situations that urges the use of and even human dependence on technology to become higher. This pandemic that can cause death by spreading through droplets or water droplets, so many countries have implemented social restrictions by prohibiting their citizens from doing many activities outside. The use of technology such as the Google Meet and Zoom applications has become a new habit for the community, especially in learning and teaching. Attendance is important to know and control the presence of students in the teaching and learning process. Currently, attendance at online lectures is still mostly done by calling students one by one or asking students to fill out certain forms. It has weaknesses and is less effective. Therefore, this study proposes presence with face identification using a web-based application system. The method used in this research are Haar Cascade face detection for facial image extraction and Convolutional Neural Network (CNN) algorithm for face identification process. The contribution given from this research in the form of identification can be done using video recordings of zoom and images or screenshots from zoom and displaying student attendance. In this study, the dataset was sourced from taking photos directly on individuals. There are 20 labels used in this study. The identification results of 1119 facial images (test data) on the CNN custom model, obtained 88% accuracy, 88% precision, and 83% recall.

## 1. PENDAHULUAN

Saat ini dunia sedang dalam masa pemulihan dari pengaruh *Corona Virus Disease* 2019 (COVID-19) selama 3 tahun belakangan ini yang telah merubah kebiasaan manusia secara signifikan. Pada Desember 2019 pertama kali pandemi bermula yang dimulai dari kota Wuhan provinsi Hubei Republik Rakyat Tiongkok [1].

Virus COVID-19 memiliki penyebaran infeksi yang cukup tinggi yaitu penyebaran Aerosol yang penyebarannya melalui udara. Partikel ini melayang di udara sebelum mendarat di permukaan benda ataupun makhluk hidup. Penyebaran virus COVID-19 adalah melalui air liur (*droplets*), dalam kontak dekat tanpa pelindung memiliki tingkat penularan yang tinggi. World Health Organization (WHO) menyatakan bahwa langkah pencegahan lain dalam melawan COVID-19 adalah membiasakan jaga jarak dengan anggota masyarakat lain sehingga risiko tertular virus COVID-19 bisa ditekan.

Pencegahan yang dilakukan untuk melawan virus COVID-19 antara lain dengan cara pembatasan interaksi antar individu atau biasa disebut sebagai Pembatasan Sosial Berskala Besar (PSBB) [2]. Dalam pembatasan sosial tersebut masyarakat diharuskan melaksanakan kegiatan sosial seperti bekerja, belajar, dan berkomunikasi dengan teman melalui aplikasi komunikasi jarak jauh seperti *Zoom* dan *Google Meet*.

Pada implementasinya, banyak universitas yang melakukan pencegahan penyebaran virus COVID 19 dengan melarang mahasiswanya untuk belajar dari universitas dan melakukan belajar dan mengajar melalui aplikasi komunikasi jarak jauh seperti *Zoom* dan *Google Meet*.

Hasil survei yang dilakukan Sukawati dkk. menunjukkan efektivitas *zoom meeting*

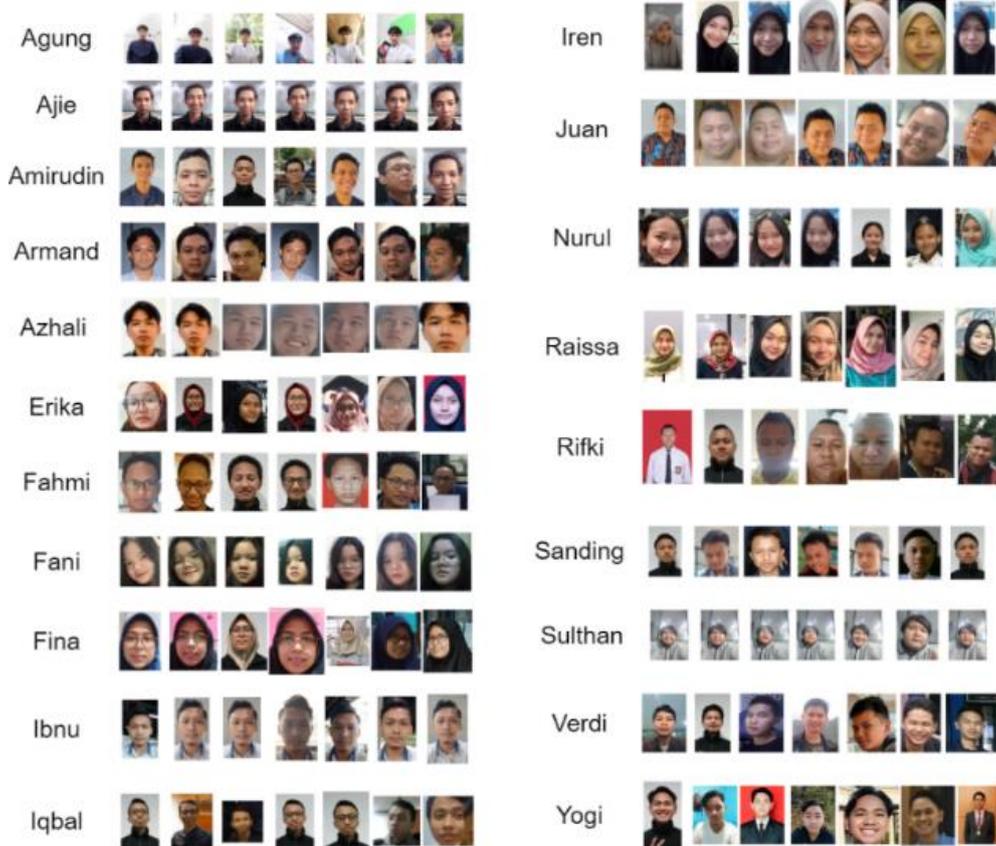
dan *Google Classroom* (GCR) berada di kategori cukup baik. Sebanyak 71% mahasiswa menyatakan *zoom meeting* dan GCR bermanfaat dalam perkuliahan [3]. Namun demikian, permasalahan yang sering muncul dalam pemanfaatan media Zoom adalah pencatatan kehadiran yang dilakukan oleh dosen. Pada saat melakukan presensi, dosen harus melakukan presensi dengan memanggil nama mahasiswa satu per satu. Selain itu, kehadiran dan partisipasi mahasiswa pada saat kuliah daring seringkali sulit diverifikasi karena banyak mahasiswa yang tidak menyalakan kamera.

Pada penelitian ini dikembangkan sistem presensi yang dapat mengidentifikasi wajah mahasiswa saat wajah mahasiswa muncul di dalam video rekaman zoom atau tangkapan layar saat perkuliahan berlangsung. Aplikasi dapat digunakan untuk mengidentifikasi wajah mahasiswa berdasarkan video rekaman maupun tangkapan layar dari perkuliahan pada zoom meeting. Pada sistem aplikasi yang dirancang oleh penulis, sistem aplikasi dapat mengenali beberapa wajah mahasiswa secara bersamaan pada sebuah video dan menyimpan hasil identifikasi wajah mahasiswa yang digunakan untuk presensi perkuliahan. Dengan sistem aplikasi ini penulis berharap akan bermanfaat dalam melakukan presensi terhadap pengajar saat perkuliahan daring.

## 2. METODE PENELITIAN

Dataset atau data yang digunakan dalam penelitian ini bersumber dari data foto mahasiswa yang diperoleh mulai tanggal 20 April 2022 -25 Mei 2022 sejumlah 389 citra foto berekstensi .jpg, .png, dan .jpeg. Data tersebut diperoleh melalui pengambilan foto terhadap individu secara langsung. Dataset disimpan ke dalam 20 folder terpisah

sesuai dengan masing-masing kelas atau individu. Contoh dataset citra yang dapat



Gambar 1. Cuplikan Dataset Penelitian

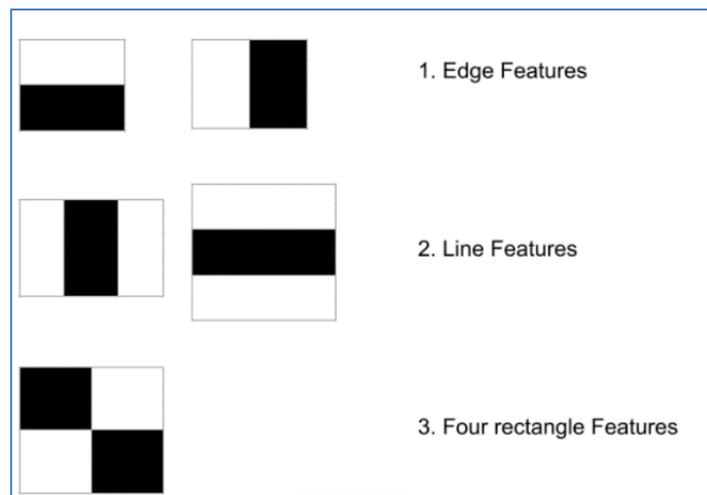
Jumlah total dataset yang terkumpul sejumlah 389 citra wajah. Lalu dilakukan tahap augmentation flip left-right citra dan rotation citra 10, 25, serta 35 derajat ke kiri dan ke kanan sehingga menghasilkan citra yang lebih variasi dan jumlah citra setelah diaugmentasi sebanyak 5554 buah.

Haar cascade adalah sebuah metode deteksi objek yang dibuat oleh Paul Viola dan Michael Jones. Pada tahun 2001, dari hasil penelitian "Rapid Object Detection using a Boosted Cascade of Simple Features" mereka dapat mendeteksi masker dari citra yang berasal dari foto atau video dari webcam dengan baik,

dilihat pada Gambar 1.

dengan nilai akurasi tertinggi 88,7% dan terendah 44,9% [4], [5].

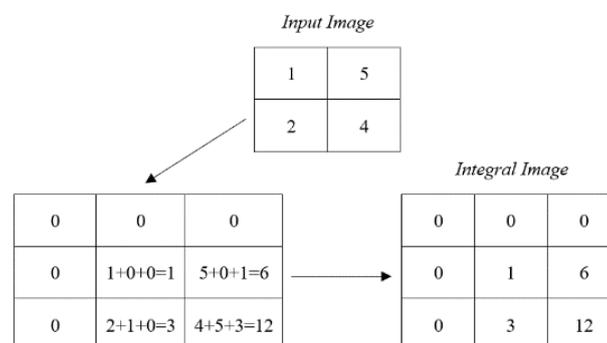
Algoritma Haar Cascade menggunakan algoritma tepi atau deteksi garis dimana algoritma tersebut dilatih dengan menerima banyak gambar berupa gambar positif yang berisi wajah dan gambar negatif yang tidak berisi wajah untuk dilatih. Dibawah ini adalah konsep fitur yang diajukan oleh Paul Viola dan Michael Jones metode haar cascade menggunakan teknik deteksi citra dengan membandingkan intensitas piksel pada citra. Pada gambar 2 adalah contoh dari fitur haar cascade untuk deteksi wajah [6].



Gambar 2. Fitur Haar

Edge Features adalah fitur yang membandingkan 2 nilai yang berbeda dari penjumlahan matriks piksel yang ditempatkan pada 2 bagian persegi secara vertikal maupun horizontal dengan ukuran dan bentuk matriks yang sama antar perbandingan. Fitur ini dapat digunakan untuk menentukan posisi alis mata pada citra. Line Features adalah fitur yang mengkalkulasi jumlah dari nilai fitur yang berada di tengah line Features, fitur ini digunakan untuk menentukan posisi hidung dan mulut pada citra. Kemudian adalah four rectangle features yang bertujuan untuk melakukan kalkulasi terhadap perbedaan intensitas piksel secara diagonal [7]. Dari beberapa jenis fitur dengan bagian tertentu yang berbeda-

beda ukuran menghasilkan matriks yang cukup rumit. Sehingga untuk mendapatkan multiple filter yang akan diaplikasikan pada fitur diperlukan algoritma tambahan, algoritma tambahan yang digunakan pada haar cascade adalah AdaBoost. Dalam hal ini nilai fitur yang telah didapatkan sebelumnya akan menjadi sebuah input pada algoritma AdaBoost. Tetapi kalkulasi dengan menggunakan Teknik ini akan membutuhkan komputasi yang tinggi. Sehingga untuk mengurangi tingginya tingkat komputasi ini. Digunakan pendekatan integral image. Ilustrasi proses integral image dapat dilihat pada Gambar 3.



Gambar 3. Integral Image

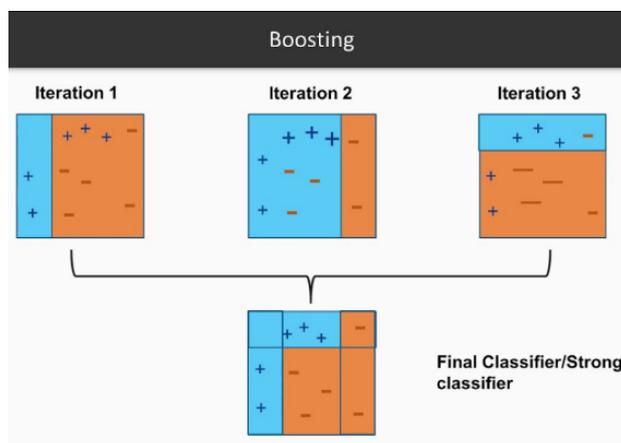
Proses integral dilakukan pada gambar dengan menambahkan 1 kolom dan 1

baris baru di sebelah kiri dan atas matriks, dengan perhitungan pada matriks orisinal

(1,1) ditambahkan dengan nilai di atasnya yaitu 0 dan di kirinya 0 menjadi  $1+0+0 = 1$  dan pada matriks orisinil (1,2) kalkulasi yang dilakukan tetap sama yaitu menambahkan dengan nilai di atasnya 0 dan nilai disebelah kirinya 1 sehingga dapat menghasilkan  $5+0+1=6$ . Untuk matriks orisinil (2,2) perlakuan tetap sama dengan menambahkan nilai di atasnya 5 (sebelum penjumlahan) dan nilai 3 (setelah penjumlahan) disebelah kiri menjadi  $4+5+3 = 12$  [8].

Penerapan boosting pada haar cascade dapat memberikan bobot pada tingkat

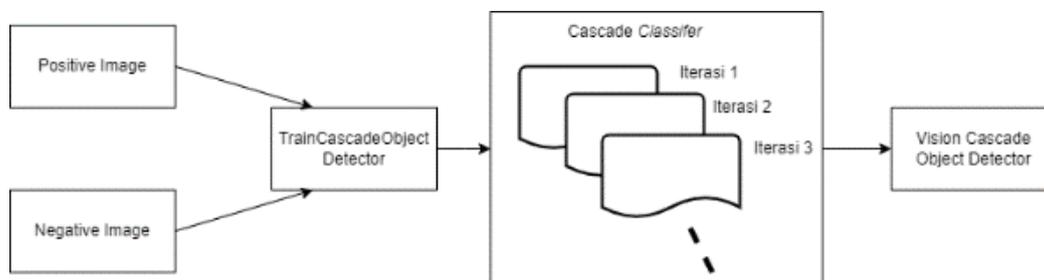
error klasifikasi yang dapat mempengaruhi distribusi data. Dengan memilih fitur yang terbaik dan melatih classifiers untuk menggunakan fitur yang telah terpilih. Serta menggunakan kombinasi weak classifier untuk membuat strong classifier sehingga algoritma haar cascade dapat mendeteksi objek. Algoritma Adaboost akan melakukan iterasi pertama untuk membuat weak classifier lalu dari hasil iterasi tersebut akan digabungkan menjadi strong classifier [9]. Ilustrasi dari penjelasan dapat dilihat pada Gambar 4.



Gambar 4. Adaboost

Setelah proses boosting dilakukan, selanjutnya akan menerapkan cascade classifier yang terdiri dari beberapa tingkatan, dimana setiap tingkatan berisi weak learners. Weak learners akan dilatih

menggunakan boosting, sehingga bisa mendapatkan akurasi tinggi classifier dari rata-rata seluruh weak learners. Pada Gambar 5 merupakan ilustrasi alur cascade classifier.



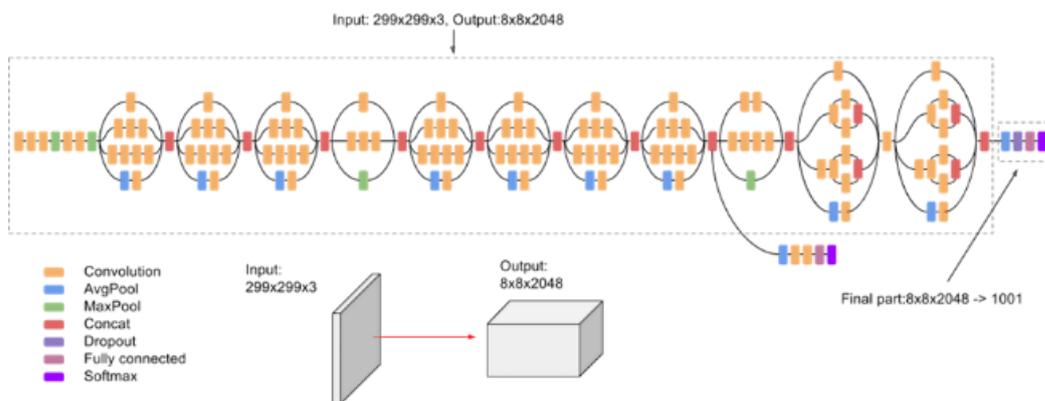
Gambar 5. Cascade Classifier

Dari hasil cascade classifier tersebut akan menjadi sebuah model yang dapat melakukan pendeteksian citra, dalam hal ini adalah model pendeteksi wajah.

*Convolutional Neural Network* adalah proses ekstraksi fitur dari citra wajah yang telah dilakukan preprocessing sebelumnya. Dalam penelitian ini modeling

menggunakan arsitektur Convolutional Neural Network (CNN). Penulis dalam menerapkan convolutional neural network akan menggunakan arsitektur model inception-v3 sebagai feature model atau base model dengan melakukan modifikasi

pada layer tertentu terutama pada layer terakhir model. Setelah itu akan melakukan training dengan dataset citra wajah, maka dapat custom model CNN. Pada Gambar 6 merupakan ilustrasi arsitektur inception-v3.



Gambar 6. Arsitektur Inception-V3

Inception-v3 pada layer-layer nya melakukan operasi yang sama dengan arsitektur CNN dasar. Namun jumlahnya saja yang lebih banyak dan kompleksitas yang lebih tinggi. Pada arsitektur terdapat beberapa bagian utama yang membedakan dengan versi inception sebelumnya, antara lain: RMSProp Optimizer, Factorized 7x7 convolutions, BatchNormalization pada Auxillary Classifiers, dan Label Smoothing [10].

Inception-v3 merupakan salah satu model deep convolutional network. Arsitektur ini diperkenalkan oleh Szegedy. Dikembangkan pada tahun 2014 dengan judul paper “Going Deeper with Convolutions”. Model Inception-V3 ini pernah dilatih dengan data pada ImageNet dan mampu untuk mengidentifikasi 1000 kelas dalam ImageNet dengan error rate terendah yaitu 3,5% dan error rate tertinggi yaitu 17.3 [11].

Pada prinsipnya, inception berfokus

kepada modul yang meningkatkan jumlah layer Neural Network secara paralel untuk mengekstraksi fitur yang ada pada fitur sebelumnya, untuk meningkatkan kapabilitas pembelajaran sistem. Tujuan inception adalah ekstraksi harus dilakukan secara bersama-sama atau paralel. modul Inception akan melakukan ekstraksi fitur dari satu layer ke beberapa layer yang berbeda secara bersamaan dan setelahnya akan disatukan Kembali dengan filter concatenation. Arsitektur inception-v3 memiliki 3 modul utama yang digunakan untuk mereduksi dimensi citra.

Pada Penelitian ini, penulis menerapkan Inception-V3 sebagai base model dengan perubahan pada layer-layer akhir dan melakukan training dengan dataset citra wajah, sehingga menghasilkan custom model CNN. Berikut pada tabel 1 merupakan susunan layer arsitektur yang digunakan pada penelitian ini.

Tabel 1. Custom Model CNN

Type	Size/Stride	Input Size
Conv	3 x 3 / 2	299 x 299 x 3
Conv	3 x 3 / 1	149 x 149 x 32

Type	Size/Stride	Input Size
<b>Conv Padded</b>	3 x 3 / 1	147 x 147 x 32
<b>Pool</b>	3 x 3 / 2	147 x 147 x 64
<b>Conv</b>	3 x 2 / 1	73 x 73 x 64
<b>Conv</b>	3 x 3 / 2	71 x 71 x 80
<b>Conv</b>	3 x 3 / 1	35 x 35 x 192
<b>3 x Inception</b>	Modul 1	35 x 35 x 288
<b>5 x Inception</b>	Modul 2	17 x 17 x 768
<b>2 x Inception</b>	Modul 3	8 x 8 x 1280
<b>Pool</b>	8 x 8	8 x 8 x 2048
<b>Linear</b>	Logits	1 x 1 2048
<b>Dropout</b>	FC layer	0.2
<b>Softmax</b>	Classifier	20

Pada Tabel 1, layer terakhir dihapus, dengan menggunakan perintah `include_top = false`. Sehingga layer terakhir bisa ditambahkan, Dalam penelitian ini, penulis menambahkan 2 layer baru yaitu, Dropout, dan dense layer (20) untuk melakukan klasifikasi berjumlah 20 label.

Sebelum melakukan modeling, dataset sebelumnya dilakukan proses rescale = 1./255 terlebih dahulu yang akan mengkonversi rentang citra dari [0,255] menjadi [0,1]. dimana orisinil citra RGB memiliki rentang 0-255, nilai tersebut dapat mempengaruhi proses hasil dari

modeling sehingga penulis mengkonversi rentang tersebut menjadi [0,1] menggunakan `rescale = 1./255`.

### 3. HASIL DAN PEMBAHASAN

Hasil dari Modeling yang dilakukan dengan dataset yang sebelumnya sudah dibagi (data latih: 80% dan data uji: 20%), dengan model yang berbeda menggunakan program yang dikembangkan oleh penulis dalam penelitian ini dan menghasilkan informasi sebagai berikut:

Tabel 2. Hasil Training

Model	Loss	Accuracy	Val Loss	Val Accuracy	Epoch	Durasi Waktu
Custom Model Inception-V3	0.1557	0.9712	0.2055	0.9559	100	145 Menit
Mobilenet-V2	0.1424	0.9882	1.0589	0.7120	44	16 Menit

Dari hasil modeling mobilenet-v2 yang bersumber dari tabel 2 berhenti di epoch ke-44 dikarenakan tidak adanya perkembangan dari val loss saat proses modeling. model mobilenet-v2 dapat digunakan dan memiliki kecepatan proses training yang lebih cepat tetapi memiliki akurasi yang lebih rendah daripada inception-v3 sehingga penulis

menggunakan inception-v3 sebagai custom model CNN.

Pengujian merupakan salah satu hal yang perlu dilakukan dalam setiap pengembangan sistem yang mengevaluasi, menganalisa dan mengetahui keakuratan sistem yang telah dirancang. Pada penelitian ini, dilakukan pengujian dari sisi akurasi, presisi dan recall pada implementasi CNN dalam

memprediksikan atau mengidentifikasi label untuk data uji. Pada penelitian ini juga menguji tingkat presisi dan recall dari

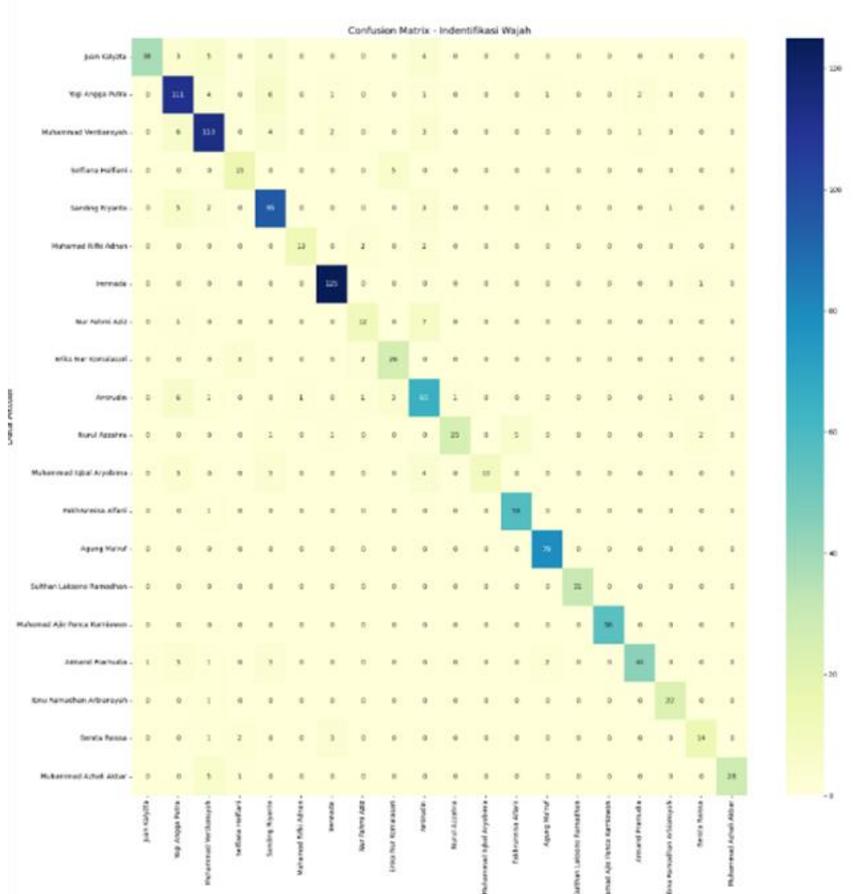
masing-masing label atau kelas. Pada Tabel 3. Berikut adalah sampel data prediksi.

Tabel 3. Contoh hasil pengujian (prediksi sistem dan data sebenarnya)

Gambar	Data Prediksi	Data Aktual
	Juan Kalyzta	Juan Kalyzta
	Nur Fahmi Aziz	Nur Fahmi Aziz
	Yogi Angga Putra	Yogi Angga Putra
	Agung Ma'Ruf	Yogi Angga Putra
	Selfiana Halfiani	Selfiana Halfiani
...	...	...
	Benita Raissa	Benita Raissa

Keseluruhan hasil prediksi data uji (1119 data uji) kemudian

direpresentasikan ke dalam confusion matrix dapat dilihat pada Gambar 7.



Gambar 7. Confusion Matrix

Pada Tabel 4 Adalah hasil pengukuran akurasi keseluruhan, presisi, dan recall

pada masing-masing label atau kelas dengan molel custom model CNN.

Tabel 4. Hasil Pengujian

Label	Precision	Recall	F1 Score	Support
Juan Kalyzta	0,97	0,7	0,81	54
Yogi Angga Putra	0,8	0,88	0,84	126
Muhammad Verdiansyah	0,84	0,88	0,86	129
Selfiana Halfiani	0,71	0,75	0,73	20
Sanding Riyanto	0,82	0,89	0,85	107
Muhamad Rifki Adnan	0,93	0,76	0,84	17
Irennada	0,95	0,99	0,97	126
Nur Fahmi Aziz	0,71	0,6	0,65	20
Erika Nur Komalasari	0,76	0,84	0,80	31
Amirudin	0,73	0,82	0,77	79
Nurul Azzahra	0,96	0,74	0,84	34
Muhammad Iqbal Aryabima	1	0,5	0,67	20
Fakhrunnisa Alfani	0,92	0,98	0,95	59
Agung Ma'ruf	0,95	1	0,97	79
Sulthan Laksono Ramadhan	1	1	1,00	31

Muhammad Ajie	1	1	1,00	56
Armand	0,94	0,81	0,87	34
Ibnu Ramadhan	0,92	0,96	0,94	23
Benita Raissa	0,82	0,7	0,76	20
Muhammad Azhali Akbar	1	0,82	0,90	34
<i>Accuracy</i>			0,88	1119
<i>Macro Avg</i>			0,89	0,83
<i>Weighted Avg</i>			0,88	0,87

Perolehan nilai akurasi, presisi dan recall menggunakan rumus yang dijabarkan berikut:

a. Akurasi

$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$akurasi = \frac{980}{1119}$$

$$akurasi = 0.87578$$

b. Presisi

$$Presisi_{Keseluruhan} = \frac{\text{Total Presisi seluruh label}}{\text{total label}} \quad (2)$$

$$Presisi_{Keseluruhan} = \frac{17,73}{20}$$

$$Presisi_{Keseluruhan} = 0.8865$$

c. Recall

$$Recall_{Keseluruhan} = \frac{\text{Total recall seluruh label}}{\text{total label}} \quad (3)$$

$$Recall_{Keseluruhan} = \frac{17,73}{20}$$

$$Recall_{Keseluruhan} = 0.831$$

Keterangan:

a. *True Positive* (TP) merupakan data positif yang diprediksi benar. Misalnya: citra 1 berlabel positif dan dari model latihan yang dibuat memprediksi citra 1 bernilai positif juga.

b. *True Negative* (TN) merupakan data negatif yang diprediksi benar. Misalnya: citra 1 berlabel negatif dan dari model latihan yang dibuat memprediksi citra 1 bernilai negatif juga.

c. *False Positive* (FP) merupakan data negatif namun diprediksi sebagai data positif. Misalnya: citra 1 berlabel negatif namun dari model latihan yang dibuat memprediksi citra 1 bernilai positif.

d. *False Negative* (FN) merupakan data positif namun diprediksi sebagai data negatif. Misalnya: citra 1 berlabel positif namun dari model latihan yang dibuat memprediksi citra 1 bernilai negatif

#### 4. KESIMPULAN

Berdasarkan hasil penelitian dan pengujian dataset menggunakan metode serta algoritme yang telah dilakukan maka dapat disimpulkan. Pendeteksian wajah dengan menggunakan metode haar cascade berhasil dilakukan dengan parameter scalefactor=1.1 atau mereduksi ukuran gambar sebesar 10% dan minNeighbors=5 atau mengambil bagian/matriks dari gambar sedikitnya berukuran 5 x 5 piksel. Algoritma custom model CNN dapat diimplementasikan serta dapat mengidentifikasi citra wajah

dengan 2 media inputan yaitu gambar dan video rekaman. Proses Identifikasi citra wajah data uji dilakukan menggunakan model custom model CNN based model Inception-V3 dengan data uji sejumlah 1119 buah diperoleh akurasi sebesar 88%, presisi 88%, dan recall 83%. Algoritma CNN dalam melakukan modeling berhasil dilakukan dengan model custom model CNN Inception-V3 yang berhasil dibuat dengan durasi waktu (ETA: 145 menit), dengan Val Loss 0.2055 dan Val Accuracy 0.9559

Adapun saran yang dapat peneliti berikan sebagai pengembangan lebih lanjut untuk aplikasi ini, agar dapat berjalan lebih baik lagi adalah pengembangan sistem dapat dipadukan dengan metode lain, agar hasil yang diberikan dapat lebih baik, dapat melakukan *training* ulang model CNN dengan mengambil hasil prediksi dan mengecek kebenaran dari hasil prediksi (menentukan secara manual data aktual gambar). Sehingga dapat dijadikan dataset baru dengan variasi yang lebih banyak, dan *custom model* CNN dapat dibuat *application programming interface* (API) agar model bisa digunakan pada beberapa sistem aplikasi berbeda.

## 5. REFERENSI

- [1] World Health Organization, "Coronavirus disease (COVID-19)," 2022. .
- [2] W. H. L. Stevany Afrizal1, Septi Kuntari, Rizki Setiawan, "Perubahan Sosial Pada Budaya Digital Dalam Pendidikan Karakter Anak," Pros. Semin. Nas. Pendidik. FKIP, vol. 3, no. 1, pp. 429–436, 2020.
- [3] S. Sukawati, "Pemanfaatan Zoom Meeting Dan Google Classroom Dalam Mata Kuliah Inovasi," Semantik, vol. 10, no. 1, pp. 45–54, 2021.
- [4] G. Aprilian Anarki, K. Auliasari, and M. Orisa, "Penerapan Metode Haar Cascade Pada Aplikasi Deteksi Masker," JATI (Jurnal Mhs. Tek. Inform., vol. 5, no. 1, pp. 179–186, 2021.
- [5] A. Solichin and A. Harjoko, "Deteksi Pejalan Kaki pada Video dengan Metode Fastest Pedestrian Detector in The West (FPDW)," TICOM, vol. 2, no. 1, pp. 202–205, 2013.
- [6] G. S. Behera, "Face Detection with Haar Cascade," 2020. .
- [7] Open CV, "Cascade Classifier," 2022.
- [8] Mathworks.com, "integrateImage," 2022.
- [9] A. Desarda, "Understanding AdaBoost," Towards Data Science, 2019. .
- [10] G. Cloud, "Advanced Guide to Inception v3," 2022.
- [11] M. Baihaqy, A. T. Wibowo, and D. Q. Utama, "Klasifikasi Tanaman Anggrek jenis Phalaenopsis berdasarkan Citra Labellum Bunga Menggunakan Metode Convolutinal Neural Network (CNN)," *e-Proceeding Eng.*, vol. 9, no. 3, pp. 1942–1951, 2022.