

Klasifikasi Ekspresi Wajah Menggunakan Metode *Convolutional Neural Network* (Studi Kasus Kuliah Daring)

Dwiki Darmawan*, Handoko Supeno**

Jurusan Teknik Informatika, Fakultas Teknik, Universitas Pasundan
Jln. Dr. Setiabudhi no. 193 Bandung, Jawa Barat
* dwiki_173040069@mail.unpas.ac.id, ** handoko@unpas.ac.id

Abstrak - Perubahan cara belajar pada tahun 2020 menjadikan media daring sangat dibutuhkan karena dengan adanya pandemi covid-19 maka semua instansi Pendidikan dan pemerintah hampir di seluruh dunia mulai menuntut agar pembelajaran menggunakan media daring yang disediakan instansi masing-masing maupun media pihak ketiga. cara belajar daring ini memiliki beberapa kekurangan seperti tingkat antusias belajar yang dirasa tidak diketahui, maka dari itu untuk mengetahui tingkat antusias suatu pembelajaran daring tadi diperlukan media untuk mengenali antusiasme dari pengguna video conference guna mendapatkan feedback yang dapat diukur apakah antusiasme pembelajaran sesuai dengan apa yang diharapkan oleh pengajar, maka pembuatan model deep learning untuk klasifikasi ekspresi wajah menggunakan metode Convolutional neural network dilakukan untuk mengetahui antusiasme dari kuliah daring dengan cara melihat real-time webcam yang diklasifikasikan oleh ekspresi yang sudah ditentukan atau dengan meninjau ulang dari rekaman perkuliahan yang sudah terjadi. Hasil akurasi menjadi tolok ukur seberapa akurat pembelajaran yang diterima untuk mengklasifikasikan ekspresi wajah guna mengetahui antusias pada model sebanyak 36.000 dataset dengan 64 % akurasi untuk melakukan deteksi ekspresi wajah pada sebuah frame video, dimana pada studi kasus ini tingkat tidak antusiasnya belajar secara daring lebih tinggi 1% dibandingkan dengan tingkat antusias belajar siswa di masa pandemi covid-19.

Kata Kunci - *Deep Learning, Convolutional Neural Network, Klasifikasi Ekspresi Wajah, Kuliah Daring*

1. PENDAHULUAN

Perubahan cara belajar pada tahun 2020 menjadikan media daring sangat dibutuhkan karena dengan adanya pandemi covid-19 maka semua instansi Pendidikan dan pemerintah hampir di seluruh dunia mulai menuntut agar pembelajaran menggunakan media daring yang disediakan instansi masing-masing maupun media pihak ketiga lainnya seperti google classroom, google meet, moodle elearning, zoom meeting dll (Adhetya, 2020) Sehingga cara belajar daring ini memiliki beberapa kekurangan seperti tingkat antusias belajar yang dirasa tidak diketahui, maka dari itu untuk mengetahui tingkat antusias suatu pembelajaran daring tadi diperlukan media untuk mengenali antusiasme dari pengguna video conference guna mendapatkan feedback yang dapat diukur apakah antusiasme pembelajaran sesuai dengan apa yang diharapkan oleh pengajar(Adhetya, 2020).

Salah satu bagian dari *Artificial Intelligence* adalah *Machine Learning*. *Machine Learning* merupakan bagian dari *Artificial Intelligence* yang memungkinkan mesin dapat “belajar” untuk meningkatkan kemampuannya berdasarkan pengalaman.(Mahmudi, 2019) *Machine Learning* ini memiliki banyak algoritma yang salah satunya adalah *Artificial Neural Network* (ANN) atau Jaringan Syaraf Tiruan (JST) yang digunakan untuk pengklasifikasian. Algoritma ANN ini meniru cara kerja otak manusia yang terdiri dari banyak sekali neuron, dimana neuron-neuron ini saling berkomunikasi satu sama lain untuk mengolah data yang ditangkap otak kita menjadi sebuah informasi yang dapat dimanfaatkan. Inilah konsep yang dipakai algoritma *Artificial Neural Network*. ANN ini memiliki model yang dinamakan *Multi-Layer Perceptron* (MLP) yang memiliki banyak lapisan untuk menghubungkan antar neuron, MLP ini memiliki kemampuan pengklasifikasian untuk permasalahan seperti pengenalan pola, pengenalan suara, pengenalan karakter untuk pembacaan dokumen, pengenalan sinyal, penentuan pola gizi, dan pengolahan citra maupun permasalahan yang lainnya. Namun pengklasifikasian menggunakan Model MLP ini memiliki kelemahan jika digunakan untuk input dalam bentuk gambar digital. Jauh lebih dalam lagi dari *Machine Learning* kita akan mengetahui yang namanya *Deep Learning* yang tentu saja lebih kompleks dari *Machine Learning*. *Deep Learning* memiliki metode yang menjadi solusi dari permasalahan yang dimiliki ANN yaitu *Convolutional Neural Network*.

Convolutional neural network (CNN) adalah metode yang digunakan untuk mendeteksi dan mengenali objek dalam bentuk citra digital. Dengan kesimpulan tersebut bahwa penggunaan metode *Convolutional neural network* CNN tersebut bisa digunakan untuk klasifikasi ekspresi wajah dalam pembelajaran daring. Hal yang menunjang metode pengujian terhadap implementasi klasifikasi ekspresi wajah ini membedakan jenis ekspresi yang nantinya bisa dikelompokkan sebagai tingkat antusiasme dari belajar daring, tahap implementasi membedakan jenis ekspresi dengan metode *Convolutional neural network* CNN ini dilakukan pada training jaringan, kemudian jaringan tersebut diuji dengan data validasi. Jaringan yang telah diuji dengan data validasi kemudian diterapkan untuk memproses data tes. Sehingga diperoleh efisiensi dalam melakukan klasifikasi ekspresi wajah dengan *video conference*. (Septian, 2019)

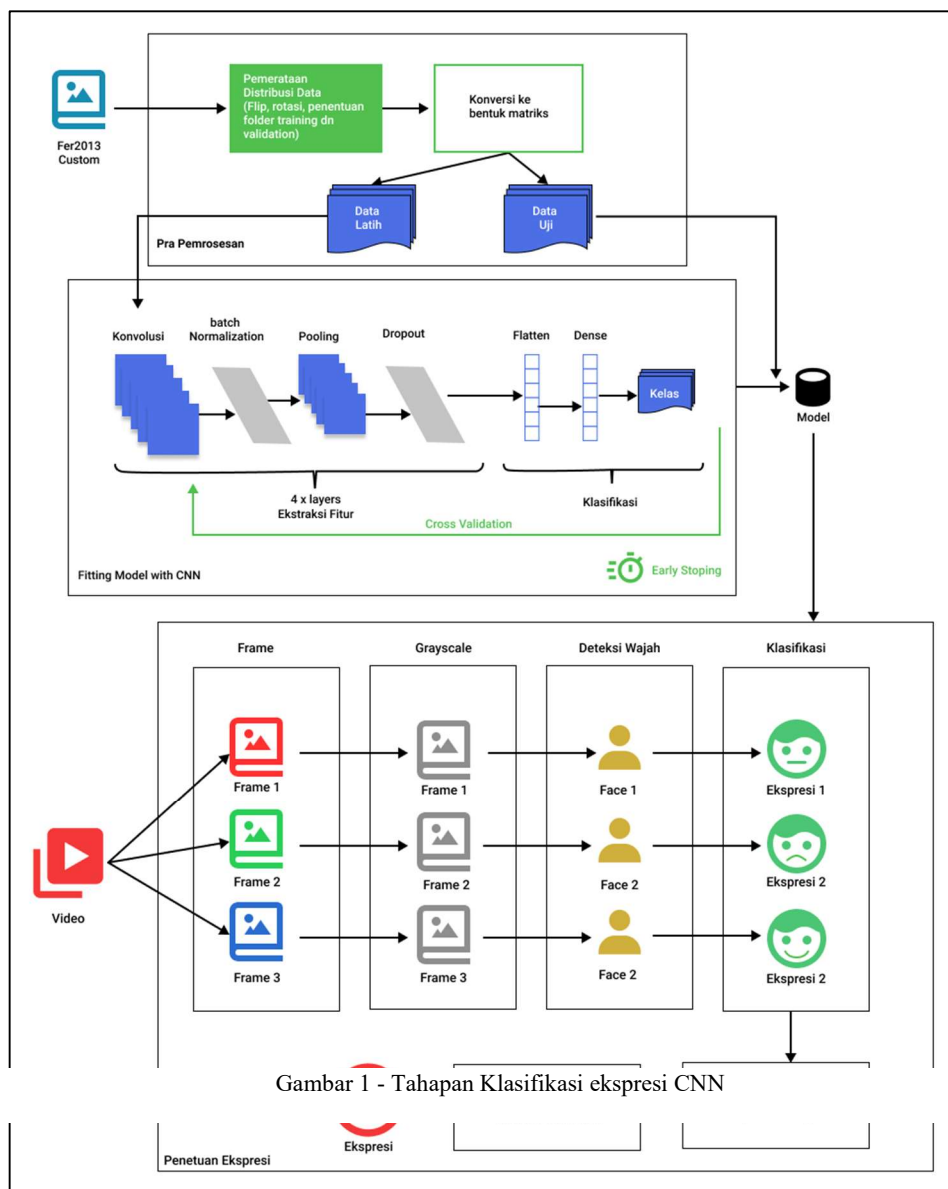
2. METODE PENELITIAN

Penelitian ini menggunakan data berupa gambar dengan 3 ekspresi dasar untuk menentukan ekspresi antusias diantaranya adalah *Neutral*, *Interested*, dan *Disappointed*. Untuk testing dari keluaran penelitian ini menggunakan video dengan objek siswa/mahasiswa berupa video offline/playback. Video tersebut merepresentasikan siswa/mahasiswa ketika mengikuti pembelajaran secara online/daring. Eksperimen yang dilakukan pada penelitian ini adalah dengan menggunakan dataset FER13custom serta *tools keras* untuk training, *Haarcascade* dan *OpenCV* untuk deteksi wajah yang nantinya akan digunakan sebagai parameter utama dalam klasifikasi ekspresi wajah. Berikut tahapan klasifikasi ekspresi pada gambar 2 - Tahapan Klasifikasi ekspresi CNN.

Dalam deep learning, umumnya penentuan ekspresi pada video dilakukan melalui tiga tahap, yaitu prapemrosesan(preprocessing), pemodelan(*fitting model with CNN*), dan penentuan ekspresi. Pada penelitian kali ini yaitu prapemrosesan akan mengubah struktur folder dari dataset FER13 Custom untuk memenuhi kebutuhan fitting dan training model dengan memindahkan file yang bernama *private_test* dan *public_test* disimpan pada folder baru yaitu *validation* dan file sebelumnya yang tidak dipindahkan disimpan pada folder *training* untuk memisahkan datasets sesuai dengan file *training* atau *validation*.

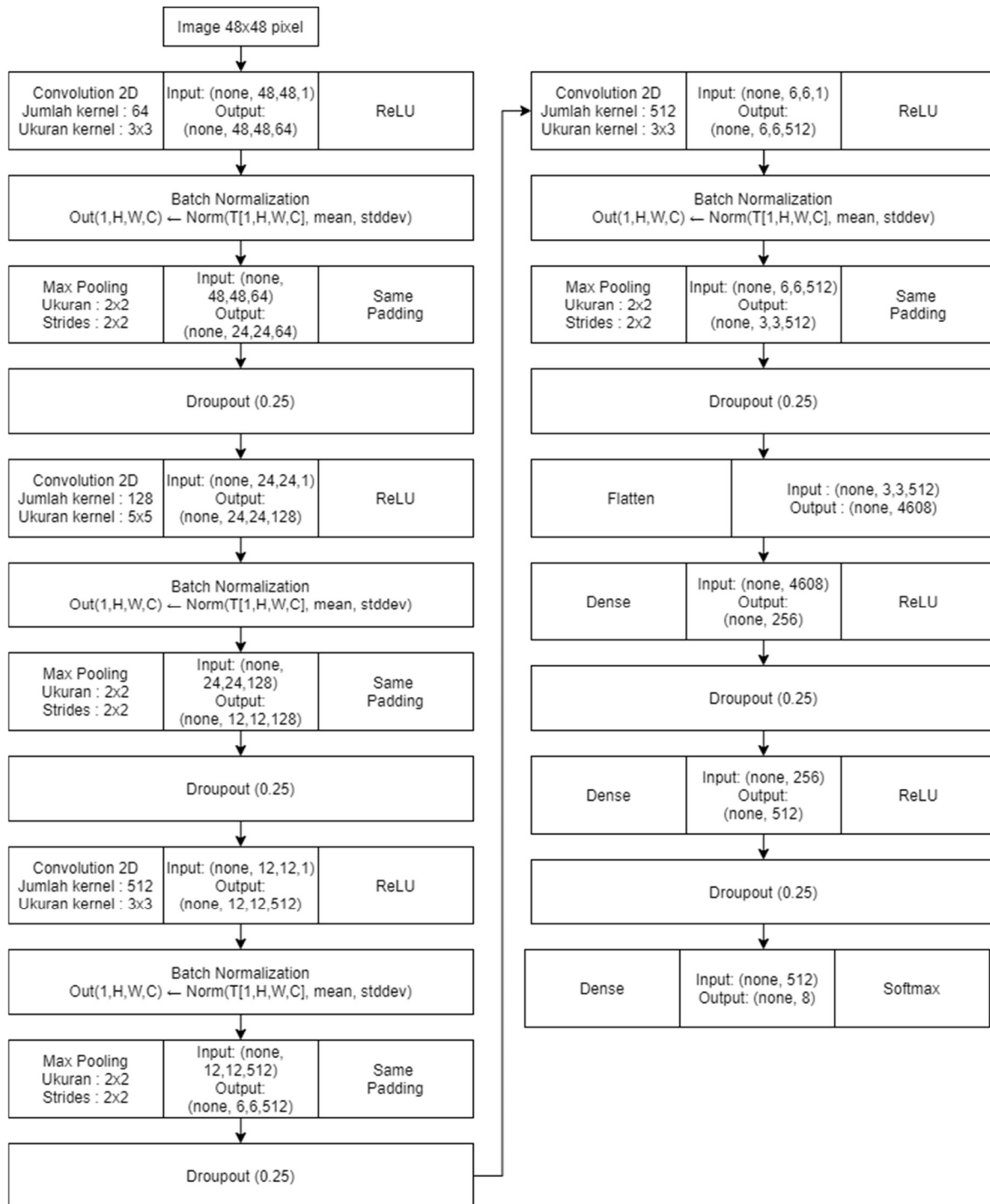
2.1 Arsitektur Convolutional Neural Network

Pembuatan Model CNN adalah proses fitting dataset yang dilakukan dengan memanfaatkan Pustaka *Tensorflow keras* dari Google untuk membantu dalam proses pemodelan Convolutional neural network. tahap pertama dalam pemodelan adalah membuat arsitektur CNN. Arsitektur yang dibangun sangat sederhana dengan tujuan untuk mendapatkan model ringan karena akan digunakan untuk video yang didalamnya terdapat tiga ekspresi wajah dan mengalami perubahan wajah yang sangat cepat.(Praseyawan, 2020)



Gambar 1 - Tahapan Klasifikasi ekspresi CNN

Arsitektur yang digunakan pada penelitian ini sendiri adalah memiliki beberapa layer konvolusi dan dense untuk memaksimalkan fitting model agar dapat berjalan dengan baik saat proses testing dilakukan, berikut gambar ..



Gambar 2 - Arsitektur CNN

2.2 Confusion Matrix

Untuk mengevaluasi metode-metode yang telah diterapkan digunakan suatu metode tersendiri yaitu metode confusion matrix. Setelah model diuji akan dihasilkan suatu persentase yang menunjukkan tingkat keakuratan. Tingkat keakuratan ini dapat diukur dengan membandingkan ekspresi wajah yang berhasil dikenali dengan benar dengan jumlah keseluruhan data pengujian. Confusion

matrix menghitung akurasi dengan menghitung kumulatif dari jumlah prediksi yang benar dibagi dengan jumlah dari prediksi yang benar ditambah dengan prediksi yang salah. Perhitungan dirumuskan sebagai berikut :

$$\text{Akurasi} = \frac{\sum \text{Data uji yang dikenali}}{\sum \text{Data uji}} * 100\%$$

Atau

$$\text{Accuracy} = \frac{TP+T}{TP+TN+FP+F} * 100\%$$

Setelah beberapa kali dilakukan pengujian dan percobaan pada pengklasifikasian ekspresi dengan deteksi objek wajah. Hasil berupa data presentase. Proses testing untuk perhitungan sebuah pengujian hasil data testing dengan menggunakan data uji sebanyak 5.663, untuk setiap kelas jenis ekspresi sebanyak 1.166 data untuk Neutral, 2.228 data untuk Interested, dan 2.269 data untuk Disappointed. Hasil Confusion Matrix adalah sebagai berikut :

Tabel 1 - Confusion Matrix

Matrix		Predict class		
		Disappointed	Interested	Neutral
Actual class	Disappointed	880	91	1300
	Interested	240	1700	250
	Neutral	83	56	1000

Berdasarkan tabel 1 – Confusion Matrix hasil prediksi dari model terhadap data testing data baru menunjukkan hasil yang baik. Prediksi terhadap ekspresi Disappointed di klasifikasikan sebagai benar dan missing data dari Interested sebanyak 91 data dan neutral sebanyak 1300 data. Prediksi terhadap ekspresi Interested di klasifikasikan sebagai benar dan missing data dari Disappointed sebanyak 240 data dan neutral sebanyak 250 data. Prediksi terhadap ekspresi Neutral di klasifikasikan sebagai benar dan missing data dari Interested sebanyak 56 data dan Disappointed sebanyak 83 data.

Hasil dari accuracy, recall dan precision yang disajikan pada gambar 5.4 – Precision recall accuracy yang menunjukkan hasil nilai recall, accuracy dan precision dengan parameter yang diberikan oleh confusion matrix.

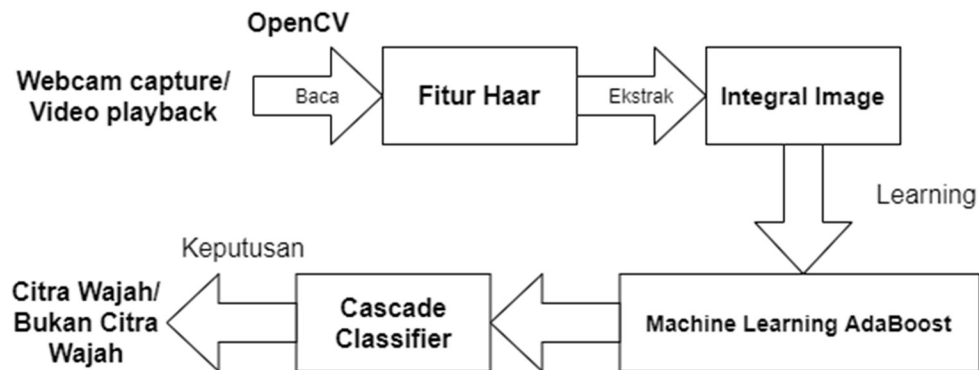
	precision	recall	f1-score	support
Disappointed	0.73	0.39	0.50	2269
interested	0.92	0.78	0.84	2228
neutral	0.40	0.88	0.55	1166
accuracy			0.64	5663
macro avg	0.68	0.68	0.63	5663
weighted avg	0.74	0.64	0.65	5663

Gambar 3 - Accuracy Precision Recall

2.3 Deteksi Wajah

Deteksi wajah bertujuan untuk menghubungkan antara model dengan program deteksi yang dibuat dengan menggunakan beberapa library seperti *OpenCV*, *HaarCascade* dan *Keras from tensorflow* agar model dapat dijalankan dan diuji dengan data baru.

Berikut merupakan gambaran untuk proses klasifikasi ekspresi melalui deteksi wajah dengan menggunakan library HaarCascade sebagai object detection. HaarCascade yang digunakan adalah “haarcascade_frontalface_default.xml” yang memiliki fungsi utama sebagai pendeteksi objek khusus wajah. Berikut gambar 4.1 - Deteksi wajah menggunakan haarcascade



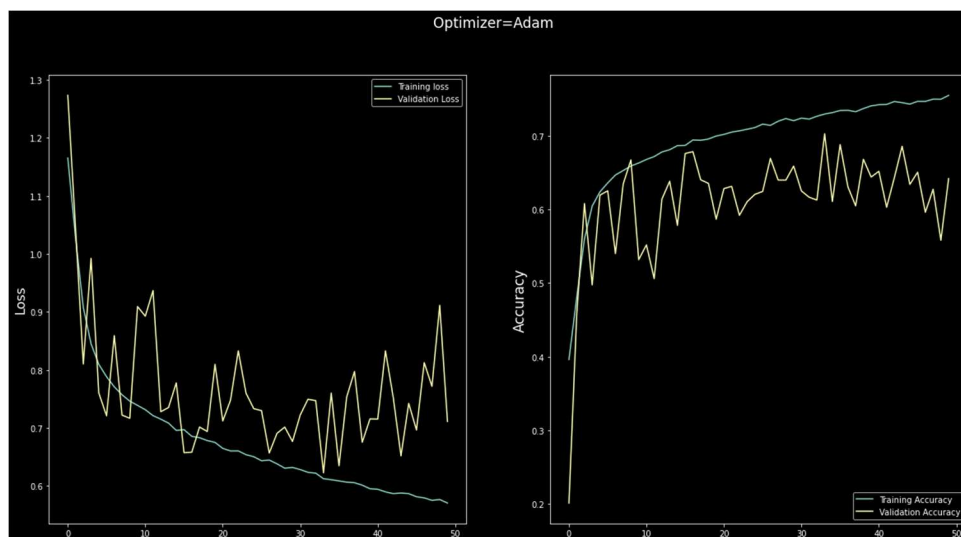
Gambar 4 - Alur Deteksi wajah

3. HASIL DAN PEMBAHASAN

Hasil dan pembahasan pengujian dilakukan menggunakan video playback atau rekaman kuliah daring dengan banyak wajah yang diujikan. Pengujian akan menghasilkan wajah terdeteksi dengan benar dan menunjukkan ekspresi objek wajah saat itu.

3.1 Data latih

Data yang digunakan pada *training* yaitu 80% dari keseluruhan data. Proses training menggunakan GPU dari Kaggle dan parameter seperti learning rate yaitu 0,001, batch size 128, dan Epoch sebanyak 34. Sehingga memberikan akurasi yang baik dan ditunjukkan dengan grafik dari akurasi dan kesalahan proses training tersebut pada gambar 6 – grafik akurasi kesalahan *training* dibawah ini.

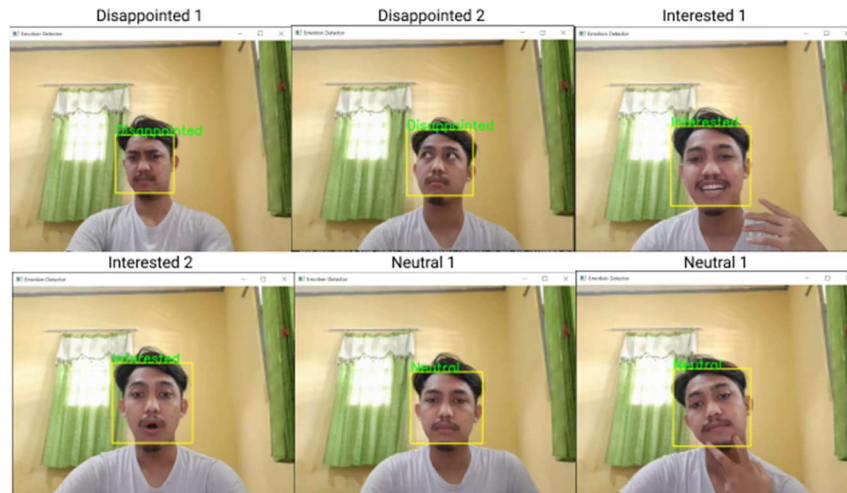


Gambar 5 – Grafik akurasi kesalahan *training*

3.2 Percobaan

A. Percobaan dengan webcam

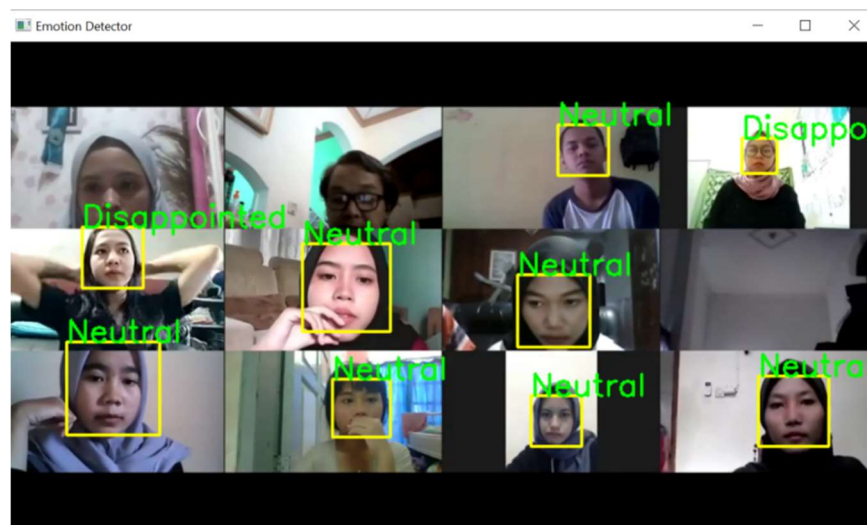
Tampilan saat deteksi untuk klasifikasi ekspresi wajah secara real-time dengan metode CNN. Pada tampilan deteksi peneliti menggunakan webcam laptop yang digunakan untuk menampilkan wajah didalam layar dan menunjukkan beberapa ekspresi berbeda dalam satu ekspresi namun menyajikannya juga berbeda seperti *neutral* karna sedang menyimak pembelajaran dan *neutral* saat sedang diam tanpa berpikir, *interested* karena sedang memperhatikan hal yang disukai dari pelajaran atau dilihatnya ke layar dan sedang tersenyum, dan *disappointed* saat merasa teralihkan dan merasa tidak nyaman mengikuti kelas berlangsung. Dimana deteksi untuk klasifikasi objek wajah secara real-time seperti pada gambar 7 - Klasifikasi ekspresi wajah secara real-time



Gambar 6 - Klasifikasi ekspresi wajah secara real-time

B. Percobaan dengan video playback

Tampilan saat deteksi untuk klasifikasi ekspresi wajah via video playback dengan memanfaatkan rekaman kuliah daring yang telah berlangsung. Pada kasus kedua ini (video dengan objek banyak) pada sampel video kuliah daring yang diambil dari youtube yang tengah membahas materi silabus kuliah. Kuliah tersebut menggambarkan suasana antusias dan tidak antusiasnya dengan kebijakan perkuliahan yang sedang dijelaskan oleh dosen bersangkutan. Dimana deteksi untuk klasifikasi objek wajah via video playback seperti pada gambar 8 - Klasifikasi ekspresi wajah via video playback



Gambar 7 - Klasifikasi ekspresi wajah via video playback

3.3 Pembahasan

Bagian terpenting dari berhasilnya algoritma *Convolutional neural network* untuk melakukan klasifikasi ekspresi wajah terutama dalam kasus penelitian ini adalah klasifikasi terhadap suatu objek wajah dengan banyaknya jumlah data ekspresi. Semakin banyak jumlah data yang dilatih pada proses training, maka hasil yang didapatkannya pun akan semakin tinggi.

Berdasarkan hasil dari penelitian ini telah melalui beberapa percobaan dalam melakukan *training* model CNN untuk mendapatkan hasil yang maksimal dan bisa digunakan untuk studi kasus penelitian ini, pada beberapa percobaan yang telah dilakukan diantaranya adalah percobaan menggunakan *notebook* dari *jupyter notebook* untuk percobaan pertama namun berganti karena proses training menggunakan sistem *localhost* yang digunakan bukan berbasis *cloud computing* dan *cloud storage* yang dimana sistem *localhost* sendiri memberatkan sehingga beralih ke *notebook* lain, *google colab* untuk percobaan kedua yang menggunakan sistem *cloud computing* untuk menggunakan GPU dari *colab* itu sendiri dan penyimpanan dengan *cloud storage* menggunakan *google drive* namun pada case ini peneliti menemukan kendala dengan penggunaan dataset yang tidak memiliki sistem *integrated* dengan *notebook project* yang berjalan dan *library OpenCV* tidak bisa digunakan secara maksimal, maka dari itu peneliti beralih mencari *notebook* lain sehingga menemukan *notebook* yang terintegrasi dengan *datasets* yang dibutuhkan dengan penelitian ini yaitu *Kaggle notebook* dan peneliti pun mencoba mencari formula untuk arsitektur CNN yang bisa digunakan dalam proses training CNN berjalan maksimal.

Percobaan final dilakukan dengan data validasi dan *testing* dari model CNN yang telah ditraining, pada validasi dan *testing* ternyata memiliki hal yang timpang tindih seperti dimana data dari validasi menunjukkan nilai model CNN yang telah ditraining memiliki kekurangan atau cacat pada prosesnya sendiri yang mengakibatkan ekspresi yang terbaca sama dengan ekspresi lain, namun pada proses testing yang dilakukan dengan perhitungan manual ekspresi yang muncul, semua ekspresi menunjukkan deteksi yang 64% akurat dengan skala kesalahan atau cacat sangatlah minim jika dibandingkan dengan nilai dari *confusion matrix* dari hasil validasi. Di dalam video dengan objek wajah banyak, jumlah wajah dengan ekspresi tertentu dapat ditentukan pada setiap frame. Pada testing dengan dataset video testing dengan skenario kuliah daring pada scene video dosen sedang mengadakan sesi tanya jawab dengan mahasiswa untuk mata kuliah yang sedang diajarkan yang berlangsung 10 detik terdeteksi keseluruhan ekspresi dalam satu video per orang yaitu 250 ekspresi dan dalam satu frame terdapat 12 orang dan jika ditotal maka menjadi 3000 total ekspresi yang akan muncul dengan persentase ekspresi yang paling dominan adalah ekspresi neutral (2910 atau 97%) kemudian Interested (30 atau 1%) dan Disappointed (60 atau 2%). Jumlah ekspresi yang menunjukkan ekspresi tidak antusias atau kecewa (disappointed) terdapat 1% lebih banyak dari total keseluruhan ekspresi semua frame dibandingkan dengan ekspresi antusias atau tertarik (interested) pada studi kasus ini.

4. KESIMPULAN

Dalam penelitian klasifikasi ekspresi menggunakan *deep learning* dengan metode CNN didapatkan Model CNN pada penelitian ini menggunakan input shape berukuran 48x48, nilai learning rate 0,001, ukuran filter 3x3, jumlah epoch 340, Data training lebih dari 30.000 dan data validation lebih dari 6000. Menghasilkan tingkat akurasi dan testing dalam melakukan klasifikasi ekspresi wajah untuk mengetahui antusiasme siswa/mahasiswa dalam pembelajaran kuliah daring sebesar 1% lebih besar tidak antusias dibandingkan dengan yang antusias. Dari percobaan yang telah dilakukan didapatkan hasil precision sebesar 68% recall sebesar 68%, dan akurasi sebesar 64% dan dapat dinilai bekerja dengan baik untuk klasifikasi ekspresi wajah dengan tingkat *poor classification*.

Dalam proses mengetahui antusiasme perkuliahan daring dapat diketahui dengan ekspresi yang sudah terklasifikasi dalam tiga ekspresi diantaranya *Interested*, *neutral*, dan *disappointed* yang dimana ketiga ekspresi tersebut mewakili tingkat antusias dari ekspresi yang dilakukan oleh siswa/mahasiswa saat sedang melaksanakan kuliah daring yaitu ekspresi Interested yang mewakili rasa antusias dalam mengikuti pembelajaran, neutral yang mewakili rasa biasa atau bisa dibilang normal dalam antusias mengikuti pembelajaran, dan disappointed yang mewakili rasa tidak antusias atau kecewa terhadap pembelajaran yang disajikan ataupun masalah dengan faktor lain sehingga berpengaruh tidak antusiasnya dalam belajar daring itu sendiri.

Daftar Pustaka

- Adhetya Cahyani, Lin Diah Listiana, Sari Puteri Deta Larasati. 2020. "Motivasi Belajar Siswa SMA Pada Pembelajaran Daring di masa pandemi Covid-19." *IQ (Ilmu Al-qur'an): Jurnal Pendidikan Islam* III (01): 123-140.
- Arriaga, Octavio, Paul G. Ploger, and Matias Valdenegro. 2017. "Real-time Convolutional Neural Networks for Emotion and Gender Classification." *Hochschule Bonn-Rhein-Sieg* 5.
- Arrofiqoh, E. N., & Harintaka, H. 2018. "Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi." *Geomatika* 61.
- Falah, Risal. 2018. "Intent Recognition Bahasa Indonesia Pada Chat Bot Portal Praktikum Menggunakan Convolutional Neural Network." *Sumedang, Jawa Barat: Universitas Padjadjaran*.

- Fibrianda, Mercury Fluorida, and Adhitya Bhawiyuga. 2018. "Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM)." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 3112-3123.
- Halprin Abhirawa, Jondri, Anditya Arfianto. 2017. "Pengenalan Wajah Menggunakan Convolutional Neural Network." *e-Proceeding of Engineering* 4907.
- I Wayan Suartika E. P, Arya Yudhi Wijaya, dan Rully Soelaiman. 2016. "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101." *JURNAL TEKNIK ITS V* (1): 2301-9271.
- Ioffe, S., & Szegedy, C. 2015. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *32nd International Conference on Machine Learning, ICML 2015 I*: 448–456.
- Lino, A. F., B. C. Silva, D. P. Rocha, and G. P. Furriel and P. W. Calixto. 2017. "Performance of haar and LBP Feature in cascade classifiers to whiteflies and counting." *Chil. Conf. Electr. Electron. Eng. Inf. Commun. Technol. CHILECON 2017- Proc* 1-6.
- Mahmudi, Muhammad Fatahillah Z, and Kusri. 2019. "IMPLEMENTASI METODE VIOLA JONES UNTUK MENDETEKSI WAJAH MANUSIA." *Jurnal INFORMA Politeknik Indonusa Surakarta V* (1).
- Nugroho, Pulung Adi, Indah Fenriana, and Rudy Arijanto. 2020. "Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Pada Ekspresi Manusia." *JURNAL ALGOR* 2715-0577.
- Prasetyawan, Daru, and Shofwatul Uyun. 2020. "Penentuan Emosi pada Video dengan Convolutional Neural Network." *JISKA* 23-35.
- Rere, L. M. Rasdi, Sunarto Usna, and Soegijanto. 2019. "Studi Pengenalan Ekspresi Wajah Berbasis Convolutional Neural Network." *Seminar Nasional Teknologi Informasi dan Komunikasi STI&K* 8.
- Septian, Ripan, Dede Irawan Saputra, and Susanto Sambasri. 2019. "Klasifikasi Emosi Menggunakan Convolutional Neural Network." *SENTER 2019: Seminar Nasional Teknik Elektro 2019* 53-62.
- Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R Salakhutdinov. 2015. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning* ." *Research* 1929–1958.
- SUBRAMANIAN, PADMAVATHY. 2021. *A Trusted Companion: AI Software Keeps Drivers Safe and Focused on the Road Ahead*. NVIDIA. January 19. Accessed September 14, 2021. <https://blogs.nvidia.com/blog/2021/01/19/drive-ix-ai-software-drivers-safe/>.
- Yusuf, Achmad, Randy Cahya Wihandika, and Candra Dewi. 2019. "Klasifikasi Emosi Berdasarkan Ciri Wajah Menggunakan Convolutional Neural Network." *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 10595-10604.
- Zisserman, Karen Simonyan & Andrew. 2015. "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION." *Visual Geometry Group, Department of Engineering Science, University of Oxford* 14.