

Identification of Fatigue from Facial Expressions Using Transfer Learning

Jefri Manurung, Andika Setiawan*, Meida Cahyo Untoro

Department of Informatics Engineering, Institut Teknologi Sumatera, Lampung, Indonesia
e-mail: jefri14117107@student.itera.ac.id, *andika.setiawan@if.itera.ac.id,
cahyo.untoro@if.itera.ac.id

Abstract

Initially, teaching and learning activities were carried out face-to-face in the provided room, but now they have switched to online. Online learning has an impact on student learning disengagement, which is known through indicators of aspects of emotional exhaustion, physical fatigue, cognitive fatigue, and loss of motivation. Besides, the teacher must provide the material that has been provided. The teacher must also pay attention to all students who are participating in the online learning. This can be overcome by a system that can detect student disengagement using a camera device. The system works by scanning the direction of students' faces and views using OpenCV technology and Transfer Learning methods. Using context, facial expressions, and heart rate can be used to recognize student disengagement. However, with the widespread availability of cameras, it is easier to identify disengagement using facial expressions. The facial expression recognition system in this study will use the FER2013 dataset and Transfer Learning method. Facial expression recognition using the FER-2013 dataset and Transfer Learning method has a reading accuracy rate of 68% in 25 epochs. Then, after being implemented as an impression parameter in the disengagement identification system using 7 scenarios, the accuracy rate is 83.33%, precision is 100%, recall is 75%, and the f1-score is 85.71%.

Keywords—Disengagement, Facial Expressions, Transfer Learning, OpenCV, FER2013

Abstrak

Semula kegiatan belajar mengajar dilakukan dengan tatap muka di ruangan yang sudah disediakan, namun kini beralih secara daring. Pembelajaran daring memberikan dampak terhadap kejenuhan belajar siswa, dimana hal tersebut diketahui melalui indikator-indikator dari aspek kelelahan emosi, kelelahan fisik, kelelahan kognitif, dan kehilangan motivasi. Disamping pengajar harus memberikan materi yang sudah disediakan, pengajar juga harus memperhatikan seluruh siswa yang ada pada pembelajaran daring tersebut. Hal ini dapat diatasi dengan sistem yang dapat mendeteksi kejenuhan siswa menggunakan perangkat kamera. Sistem bekerja dengan memindai arah wajah dan pandangan siswa menggunakan teknologi OpenCV dan metode Transfer Learning. Untuk dapat mengenali kejenuhan siswa, dapat dilakukan dengan menggunakan context performance, ekspresi wajah dan detak jantung. Namun dengan ketersediaan kamera yang sudah tersebar luas, maka identifikasi kejenuhan menggunakan ekspresi wajah menjadi lebih mudah dilakukan. Sistem pengenalan ekspresi wajah pada penelitian kali ini akan menggunakan dataset FER2013 dan metode Transfer Learning. Pengenalan ekspresi wajah dengan dataset FER2013 dan metode Transfer Learning memiliki tingkat akurasi pembacaan 68% dalam 25 epochs. Kemudian setelah diimplementasikan menjadi parameter impresi pada sistem identifikasi kejenuhan menggunakan

7 scenario, didapat tingkat akurasi sebesar 83.33%, presisi sebesar 100%, recall sebesar 75 % dan f1-score sebesar 85.71%..

Kata kunci—Kejenuhan, Ekspresi Wajah, Transfer Learning, OpenCV, FER2013

1. PENDAHULUAN

Begitu banyak sektor di Indonesia yang mengalami dampak dari penyebaran Covid-19, salah satunya adalah sektor Pendidikan. Semula kegiatan belajar mengajar dilakukan dengan tatap muka di ruangan yang sudah disediakan, namun kini beralih secara daring dengan memanfaatkan jejaring komputer, internet dan sebagainya untuk dijadikan media pembelajaran yang baru. Hal ini dilakukan untuk menjaga kesehatan dan keamanan pengajar dan siswa selama masa pandemi Covid-19 [1].

Seiring berjalannya waktu, pembelajaran secara daring menunjukkan dampak pada proses belajar mengajar itu sendiri. Pembelajaran daring memberikan dampak terhadap kejenuhan belajar siswa, dimana hal tersebut diketahui melalui indikator-indikator dari aspek kelelahan emosi, kelelahan fisik, kelelahan kognitif, dan kehilangan motivasi. Efek paling parah yang didapat dari kejenuhan adalah frustrasi akut yang dapat menimbulkan perilaku menyimpang dan gangguan kesehatan [2].

Pengajar dituntut untuk lebih kreatif dalam melakukan kegiatan belajar mengajar secara daring. Disamping pengajar harus memberikan materi yang sudah disediakan, pengajar juga harus dapat mengimbangnya dengan kegiatan yang bisa memunculkan gairah atau semangat selama proses belajar mengajar. Namun, pengajar tentu akan kesulitan ketika ia harus memberikan materi secara lengkap, sekaligus memperhatikan seluruh siswa yang ada pada pembelajaran daring tersebut. Pengajar harus dapat memperhatikan setiap ekspresi yang ada pada para siswa untuk mengetahui jenuh atau tidaknya siswa pada saat menerima materi. Hal ini dapat diatasi dengan sistem yang dapat mendeteksi kejenuhan siswa menggunakan perangkat kamera. Sistem bekerja dengan memindai arah wajah dan pandangan siswa menggunakan teknologi OpenCV dan metode *Transfer Learning*.

Identifikasi kejenuhan dapat dilakukan dengan menggunakan *context performance* [3], ekspresi wajah [4] dan detak jantung [5]. Namun dengan ketersediaan kamera yang sudah tersebar luas, maka identifikasi kejenuhan menggunakan ekspresi wajah menjadi lebih mudah dilakukan.

Ekspresi dengan kata lain mimik wajah merupakan bentuk komunikasi non-verbal yang terdiri dari satu atau lebih gerakan atau posisi otot pada wajah serta dapat mengungkapkan keadaan emosi individu kepada individu yang sedang mengamatinya. Ekspresi wajah dapat mengungkapkan isi hati pada diri individu. Sebagai contoh, mengernyitkan alis mata menunjukkan kemarahan atau ketidaksukaan, mengangkat alis mata menunjukkan ekspresi heran dan terkejut. Semua emosi dan berbagai macam isi hati manusia tergambar pada ekspresi wajah yang berbeda-beda [6]. Ekspresi wajah yang akan digunakan pada penelitian ini antara lain senang, marah, sedih, terkejut, takut, jijik, dan netral.

Deteksi wajah merupakan langkah pertama yang harus dilakukan dalam analisis wajah, termasuk di dalamnya adalah pengenalan ekspresi wajah. Deteksi wajah bertujuan untuk menentukan apakah ada wajah atau tidak di dalam citra, dan jika ada dimana letak wajah tersebut dan ukuran masing-masing wajah pada citra. Dalam deteksi wajah terdapat beberapa tantangan seperti posisi wajah, skala wajah, ekspresi wajah, wajah terhalang objek lain, dan kondisi pencahayaan [7]. Setelah wajah dapat dideteksi, barulah dapat dilakukan pengenalan ekspresi wajah.

Sistem pengenalan ekspresi wajah membutuhkan dataset yang cukup besar. Dalam hal ini dataset FER2013 akan digunakan dalam penelitian ini karena memiliki enam dasar emosi manusia, yaitu bahagia, terkejut, sedih, marah, jijik, dan takut. Namun, jumlah citra pada dataset

FER2013 masih belum cukup besar jika dibandingkan dengan dataset ImageNet yang digunakan dalam melatih model VGG16 dengan jumlah lebih dari 14 juta citra. Untuk mengatasi hal tersebut, dibutuhkan suatu metode yang dapat menggunakan dataset yang tidak besar. Oleh karena itu, sistem pengenalan ekspresi wajah pada penelitian kali ini akan menggunakan metode *Transfer Learning*.

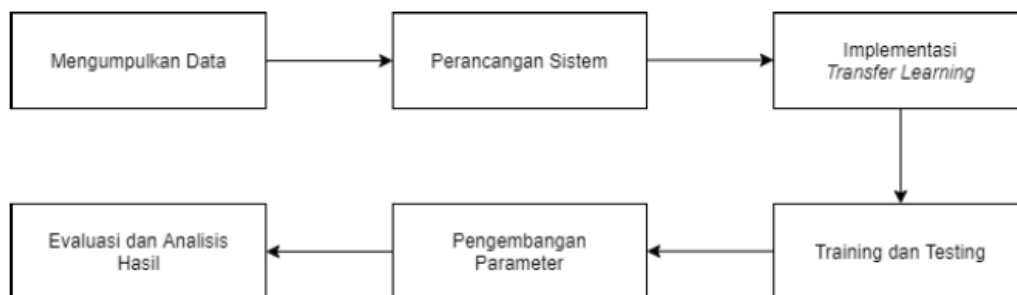
Transfer learning adalah suatu metode yang memanfaatkan model *deep learning* yang telah dilatih (*pre-trained model*) terhadap suatu dataset untuk menyelesaikan permasalahan lain yang serupa dengan cara menggunakannya sebagai *starting point*, memodifikasi dan memperbaharui parameternya sehingga sesuai dengan dataset yang baru [8]. Dataset dan masukkan yang digunakan pada sistem pengenalan ekspresi wajah ini berupa citra wajah, sehingga dibutuhkan media yang dapat menghasilkan keluaran berupa citra wajah, salah satunya yaitu kamera. Dengan bantuan media audiovisual seperti kamera dan pengenalan ekspresi wajah menggunakan *Transfer Learning* dapat dibuat sistem untuk menentukan tingkat kejenuhan siswa.

2. METODE PENELITIAN

Pada bagian ini akan dijelaskan tentang analisa dan perancangan sistem dari penelitian ini. Terdapat beberapa tahapan, yaitu tahapan penelitian yang dilakukan, kebutuhan sistem yang akan dibuat dan penyelesaian masalah identifikasi kejenuhan menggunakan *transfer learning*.

2.1 Alur Penelitian

Perihal pelaksanaan penelitian ini, untuk mempermudahnya maka dijabarkan langkah-langkah apa saja yang akan diambil dalam melakukan penelitian ini. Alur dari penelitian ini direpresentasikan pada Gambar 1 berikut ini.

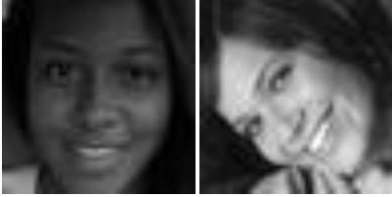








Gambar 1 Alur penelitian

2.1.1 Pengumpulan Data

Pada penelitian kali ini, untuk keperluan pengenalan ekspresi wajah maka diperlukan objek berupa foto ekspresi wajah untuk diidentifikasi. Pada penelitian kali ini studi kasus yang akan digunakan adalah dataset *Facial Expression Recognition 2013* (FER2013). Dataset *Facial Expression Recognition 2013* (FER 2013) adalah set data yang disediakan oleh Kaggle, diperkenalkan pada *International Conference on Machine Learning* (ICML) pada tahun 2013 [18] diperkenalkan oleh Pierre-Luc Carrier dan Aaron Courvill. Kumpulan data dibuat menggunakan *Google Image Search API* dan mencari gambar wajah yang cocok dengan kumpulan 184 kata kunci emosi seperti “bahagia” dan “kesal”. Kata kunci ini dikombinasikan dengan kata-kata yang berhubungan dengan jenis kelamin, usia, atau etnis. Dataset yang dihasilkan berisi 35.887 gambar, meliputi 4.953 gambar “marah”, 547 gambar “jijik”, 5.121 gambar “takut”, 8.989 gambar “senang”, 6.077 gambar “sedih”, 4.002 gambar “terkejut” dan 6.198 gambar “netral”. Oleh karena itu, untuk ekspresi wajah yang akan diidentifikasi pada penelitian ini dibagi menjadi 7 kategori ekspresi [19], yaitu:

Tabel 1 Kategori ekspresi yang akan diidentifikasi dalam penelitian ini

Senang		Marah	
			
<ul style="list-style-type: none"> • Posisi alis mata rileks • Posisi mulut terbuka dan sedikit terlihat gigi atas dan bawah • Ujung mulut tertarik ke arah telinga 		<ul style="list-style-type: none"> • Alis kebawah dan menyatu • Mata terbuka • Bibir menyempit 	
Sedih		Terkejut	
			
<ul style="list-style-type: none"> • Kelopak mata atas menurun • Mata tidak focus • Ujung mulut tertarik kebawah 		<ul style="list-style-type: none"> • Alis mata naik • Mata terbuka lebar • Mulut terbuka lebar 	
Jijik		Netral	
			
<ul style="list-style-type: none"> • Mata sedikit mengecil • Hidung agak melengkung • Bibir agak terbuka 		<ul style="list-style-type: none"> • Seluruh otot wajah dalam kondisi rileks • Kelopak mata bersinggungan dengan retina • Bibir atas dan bawah saling bersentuhan • Mulut tertutup 	
Takut			
			<ul style="list-style-type: none"> • Alis mata naik dan menyatu bersamaan • Kelopak mata naik • Kelopak mata bawah menegang • Bibir sedikit horizontal ke arah telinga

2.1.2 Desain Sistem

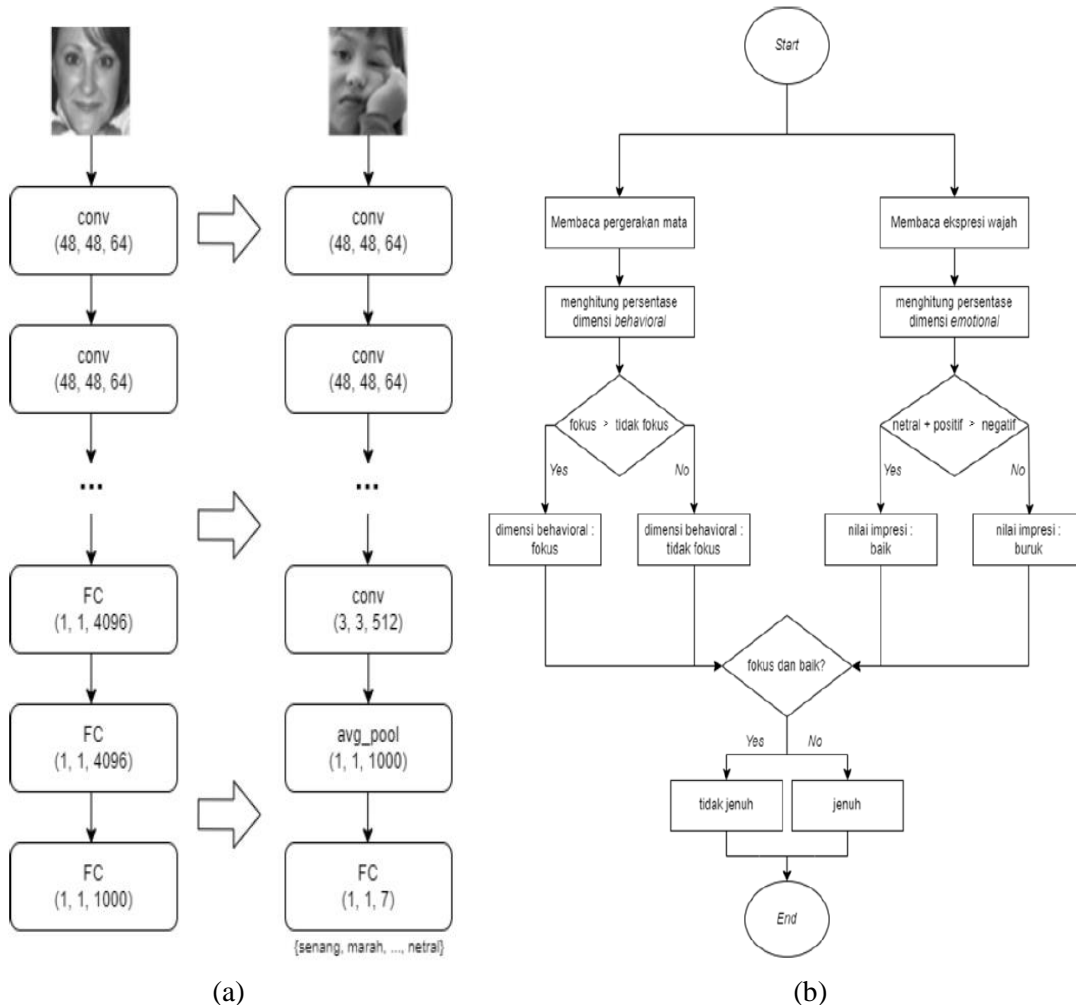
2.1.2.1 Input Dataset

2.1.2.2 Pre-processing data

Dalam penelitian ini, sebelum citra dijadikan masukan untuk pelatihan, citra diolah terlebih dahulu agar memudahkan metode *transfer learning* untuk melakukan pelatihan dan menemukan ciri dari citra yang dimasukkan. Ada 2 tahap preprocessing citra yang dilakukan sebelum citra tersebut diolah oleh metode *transfer learning*:

2.1.2.3 Implementasi Transfer Learning

Metode *transfer learning* diimplementasikan untuk membuat model pengenalan ekspresi wajah, *pre-trained* model yang telah dilatih akan digunakan pada penelitian kali ini. Namun, meskipun model sudah dilatih, nantinya *pre-trained* model ini akan dilatih lagi untuk kebutuhan yang baru dengan bobot *pre-trained* model sebagai bobot awal pelatihan atau bisa disebut *Fine Tuning*. Setelah model dilatih, arsitektur VGG16 yang sudah di-*tuning* tersebut nantinya disimpan untuk kemudian dipanggil pada tahap *testing* pengenalan ekspresi wajah. Pada Gambar 2 di bawah ini dapat diperhatikan hubungan arsitektur VGG16 dan model pengenalan ekspresi wajah.



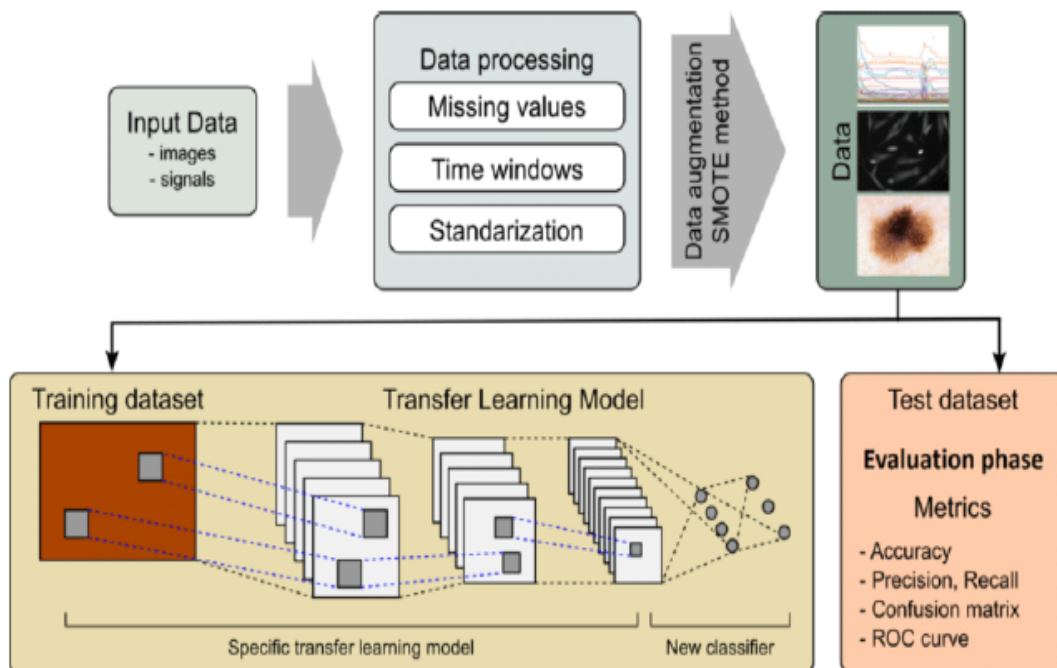
Gambar 2 Hubungan Arsitektur VGG16 (kiri) dengan Model Pengenalan Ekspresi Wajah (kanan) dan Sistem Identifikasi Kejenuhan

Program sistem identifikasi kejenuhan terdiri dari dua sistem pendeteksi yang berjalan bersamaan: Deteksi Pergerakan Mata dan Pengenalan Ekspresi Wajah. Deteksi pergerakan mata menggunakan library `shape_predictor_68_face_landmarks.dat` dari `dlib` untuk mendeteksi dimensi behavioral siswa berdasarkan posisi bola mata terhadap layar. Pengenalan ekspresi wajah menggunakan metode transfer learning dengan arsitektur VGG16 untuk mendeteksi dimensi emosional siswa berdasarkan ekspresi wajah, mengonversi output menjadi tiga dimensi emosional: Netral, Positif, dan Negatif. Sistem ini juga menganalisis dimensi behavioral dengan mendeteksi keberadaan wajah siswa untuk mengetahui partisipasi dalam pembelajaran. Berdasarkan Gambar 2, proses identifikasi kejenuhan dimulai dengan input berupa citra wajah,

seperti wajah sedih dan mata menatap layar. Sistem kemudian menjalankan dua program: deteksi pergerakan mata dan pengenalan ekspresi wajah. Misalnya, wajah sedih dan mata menatap layar meningkatkan persentase variabel "Fokus". Pada pengenalan ekspresi wajah, input dibaca oleh model VGG16 untuk menentukan ekspresi wajah siswa, seperti wajah sedih yang meningkatkan persentase variabel "Reaksi Negatif". Dari nilai persentase ini, didapat dimensi behavioral (fokus atau tidak fokus) dan dimensi emosional (negatif, positif, netral) untuk klasifikasi kejenuhan. Misalnya, wajah sedih dan mata menatap layar menunjukkan bahwa siswa tersebut "Jenuh".

2.1.2.3.1 Proses Training

Dalam sistem identifikasi kejenuhan, model yang telah dilatih sebelumnya perlu dilatih dengan data pelatihan untuk mencapai akurasi pengenalan objek yang tinggi. Tujuan dari pelatihan model ini adalah untuk menemukan properti setiap gambar dan mengidentifikasi neuron yang diaktifkan saat gambar diklasifikasikan. Oleh karena itu, perlu dibuat skema untuk melatih model pre-trained tersebut. Gambar 3 menunjukkan skema pembangunan model menggunakan metode transfer learning. Sebelum mengolah gambar dalam transfer learning, data pelatihan yang diperlukan harus diperoleh dan diproses. Selanjutnya, parameter pelatihan seperti epoch, yang merupakan jumlah putaran dari awal hingga akhir, perlu diinisialisasi. Metode transfer learning kemudian diterapkan untuk pembelajaran fitur dan klasifikasi citra, dengan mendesain ulang arsitektur VGG16 sesuai kebutuhan penelitian ini menggunakan library Keras. Proses pelatihan ini, yang dikenal sebagai model fitting, dijalankan pada arsitektur baru dan berjalan selama sejumlah epoch tertentu, dengan sistem secara otomatis menampilkan nilai akurasi dan loss dari data pelatihan. Setelah selesai, model yang terbentuk disimpan dan dipanggil kembali selama proses pengujian.

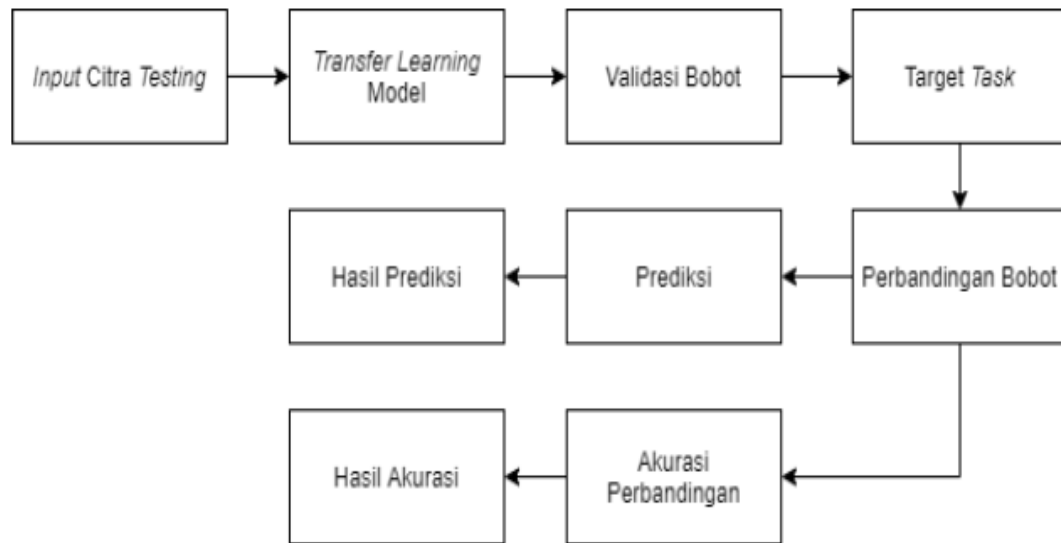


Gambar 3 Skema *Training* dan *Testing*

2.1.2.3.2 Proses Testing

Setelah model pembelajaran transfer learning melalui proses pelatihan, model tersebut diuji untuk mengevaluasi seberapa baik kinerja model yang telah dilatih sebelumnya. Pengujian ini dilakukan dengan menguji citra baru yang belum pernah dilatih oleh model, sehingga dapat mengidentifikasi kinerja model. Langkah pertama adalah mengingat kembali model yang telah terbentuk selama pelatihan. Selanjutnya, sistem akan memunculkan output prediksi berdasarkan citra yang diuji dengan membandingkan bobot antara citra yang akan diprediksi dengan bobot

dari contoh yang telah dipanggil sebelumnya. Setelah bobot citra tervalidasi, citra tersebut akan diklasifikasikan ke hasil berdasarkan bobot terdekat. Proses pengujian ini dapat dilihat pada Gambar 4.



Gambar 4 Output Klasifikasi

2.1.2.3.3 Pengujian Performansi

Pengujian performansi dalam identifikasi kejenuhan dan ekspresi wajah pada video bertujuan untuk mengevaluasi seberapa baik sebuah sistem atau model dapat mendeteksi dan mengklasifikasikan kejenuhan dan berbagai ekspresi wajah siswa. Pengujian ini dilakukan menggunakan berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1 score untuk memberikan gambaran menyeluruh tentang kinerja sistem berbasis Transfer Learning ini. True Positive (TP) adalah jumlah prediksi positif yang benar, di mana model berhasil mengidentifikasi kasus positif dengan benar. True Negative (TN) adalah jumlah prediksi negatif yang benar, di mana model dengan benar mengidentifikasi kasus negatif. False Positive (FP) adalah jumlah prediksi positif yang salah, di mana model salah mengidentifikasi kasus negatif sebagai positif. False Negative (FN) adalah jumlah prediksi negatif yang salah, di mana model salah mengidentifikasi kasus positif sebagai negatif, Persamaan 1-4 menunjukkan rumus untuk menghitung akurasi, presisi, recall dan F-1 score dari sistem yang diajukan.

$$\text{Akurasi} = (TP + TN) / (TP + FP + FN + TN) \quad (1)$$

$$\text{Presisi} = TP / (TP + FP) \quad (2)$$

$$\text{Recall} = TP / (TP + FN) \quad (3)$$

$$\text{F1-score} = (2 \times \text{presisi} \times \text{recall}) / (\text{presisi} + \text{recall}) \quad (4)$$

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian Pengenalan Ekspresi Wajah dengan Transfer Learning

Pengenalan Ekspresi Wajah menggunakan VGG16 sebagai arsitektur utama yang kemudian disesuaikan dengan kebutuhan penelitian ini dengan *Transfer Learning*. *Video input* akan di *capture* setiap *frame*-nya menggunakan fungsi *ret, frame = cap.read()* pada program sistem identifikasi kejenuhan. Setiap *frame* yang telah di *capture* tersebut akan diolah menggunakan metode *haar cascade* yang digunakan untuk mendeteksi wajah dalam setiap *frame* dari *video input*. Area wajah akan di *crop* dan dilakukan *resizing* dan *grayscale* menjadi ukuran 48x48 piksel kemudian diolah dan dijadikan *input* untuk model pengenalan ekspresi wajah.

Proses pada pengenalan ekspresi wajah ini terjadi setelah posisi wajah dideteksi menggunakan *haarcascade frontal face* seperti terlampir pada Gambar 5.

```
ret, frame = cap.read()

emotion = {0: "Negatif - Marah", 1: "Negatif - Jijik", 2: "Negatif - Takut", 3: "Positif - Senang",
          4: "Positif - Senang", 5: "Positif - Senang", 6: "Positif - Senang"}

face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades + 'haarcascade_frontalface_default.xml')

gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)

faces = face_cascade.detectMultiScale(
    gray,
    scaleFactor=1.05,
    minNeighbors=7,
    minSize=(224, 224),
    flags=cv2.CASCADE_SCALE_IMAGE
)

framecount += 1

for x, y, w, h in faces:
    roi_gray = gray[y:y+h, x:x+w]
    roi_color = frame[y:y+h, x:x+w]
    cv2.rectangle(frame, (x, y), (x+w, y+h), (255, 0, 0), 2)
    faces = face_cascade.detectMultiScale(roi_gray)

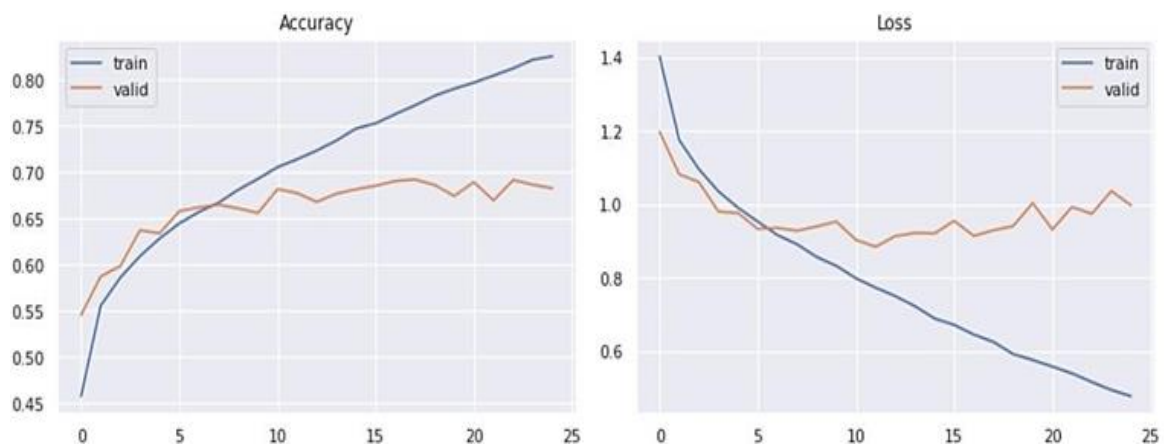
    if len(faces) == 0:
        print("Face not detected")
    else:
        for (ex, ey, ew, eh) in faces:
            face_roi = roi_color[ey:ey+eh, ex:ex+ew]

            final_image = cv2.resize(face_roi, (48, 48))
            final_image = np.expand_dims(final_image, axis=0) #need fourth dimension
            final_image = final_image/255.0 #normalizing

            prediction = model.predict(final_image)
            maxindex = int(np.argmax(prediction))
```

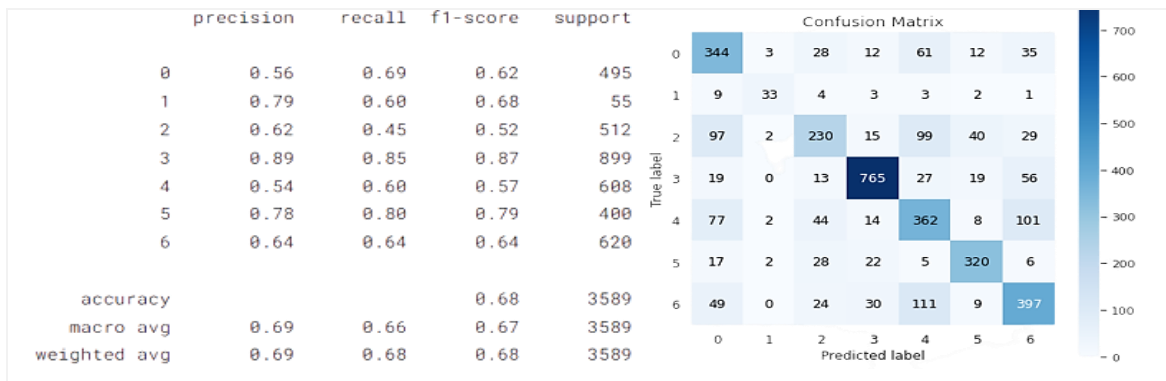
Gambar 5 Script Pengenalan Ekspresi Wajah

Gambar 5 menunjukkan proses pengenalan ekspresi wajah terjadi pada perintah `prediction = model.predict(final_image)`. Keluaran dari proses tersebut akan memberikan *output* berupa *list* berisi *softmax score* dari 7 kelas emosi berdasarkan FER2013 dataset. Sistem akan menyimpan dan menampilkan emosi dengan *score maximum*. Fungsi `int(np.argmax(prediction))` akan memberikan *output* berupa emosi dengan nilai tertinggi.



Gambar 6 Grafik Akurasi (kiri) dan *Loss* (kanan) Model

Gambar 6 menunjukkan grafik loss dan grafik akurasi dari hasil *training* dan evaluasi. Hasil evaluasi menunjukkan bahwa akurasi dari model yang telah dites dengan menggunakan 3589 data adalah 68%, dengan nilai *confusion matrix* seperti pada Gambar 7.



Gambar 7 Confusion Matrix

Pada Gambar 7, sumbu Y menunjukkan bagian *true* dan sumbu X menunjukkan bagian *predicted*. Elemen-elemen diagonal mewakili jumlah titik di mana label yang diprediksi sama dengan label yang sebenarnya (*true positive*). Dari *confusion matrix* ini didapat *performance metrics* yang akan menentukan apakah program yang telah dibuat ini bagus atau tidak. *Performance metrics* yang didapatkan model ini adalah akurasi rata-rata sebesar 0.68, presisi rata-rata sebesar 0.69, *recall* rata-rata bernilai 0.66 dan rata-rata *f1-score* bernilai 0.67. Nilai untuk *performance metrics* setiap masing-masing output dapat dilihat juga pada Gambar 7.

3.2 Hasil Uji Coba Sistem dengan 7 Skenario

Pengujian dilakukan dengan melakukan tes pada sistem dengan 7 input video dengan kondisi berbeda untuk dilakukan analisis. Pengujian dan pengambilan data dilakukan dengan menggunakan perangkat laptop Acer Swift 3 bersistem operasi Windows 10 yang telah memiliki virtual environment *python*. Peneliti melakukan 2 pengujian berupa analisis kejenuhan 7 *scenario* oleh sistem identifikasi kejenuhan dan analisis kejenuhan oleh 2 *annotator*. Analisis kejenuhan menggunakan sistem identifikasi kejenuhan dilakukan pada perangkat dengan spesifikasi yang sudah dijelaskan pada bagian sebelumnya. Sedangkan, analisis kejenuhan oleh 2 *annotator* dilakukan dengan mengisi kuesioner berupa ceklis parameter yang sama dengan sistem identifikasi kejenuhan. *Annotator* menganalisis video dan mengisi ceklis parameter kejenuhan setiap 10 detik video. Tabel 2 menunjukkan penjelasan dan juga hasil dari tiap-tiap 7 *scenario* pengujian yang telah dilakukan pada video untuk menguji sistem ini

Tabel 2 Hasil pengujian berdasarkan 7 *scenario*

<p>Scenario 1</p> <p>Hasil Scenario 1</p>	<p>Siswa selalu melihat layar, siswa menggunakan laptop dalam keadaan setengah berbaring, posisi badan statis, menunjukkan ekspresi netral.</p> 
<p>Scenario 2</p>	<p>Siswa selalu melihat layar, posisi badan duduk dan terkadang bergerak,</p>

	menunjukkan ekspresi tertarik, tertawa, dan tersenyum.
<p>Hasil Scenario 2</p>	
<p>Scenario 3</p>	<p>Siswa selalu melihat layar, siswa menggunakan laptop dalam posisi tengkurap, posisi badan statis, menunjukkan ekspresi netral.</p>
<p>Hasil Scenario 3</p>	
<p>Scenario 4</p>	<p>Siswa hadir, namun terkadang tidak melihat layar, posisi badan duduk dan statis, banyak menengok, siswa terlihat bingung dan mengernyitkan dahi beberapa kali.</p>
<p>Hasil Scenario 4</p>	
<p>Scenario 5</p>	<p>Siswa hadir, namun hampir tidak sama sekali menatap layar, kurang lebih hanya menatap layar selama 3 detik, siswa fokus bermain handphone, posisi siswa berubah: duduk, dan berbaring, tampak ekspresi netral.</p>
<p>Hasil Scenario 5</p>	
<p>Scenario 6</p>	<p>Siswa hadir, namun siswa sama sekali tidak melihat layar, selalu melihat ke bawah, posisi badan duduk, statis, tampak ekspresi netral.</p>

Hasil Scenario 6	
Scenario 7	Tidak ada siswa, hanya ruangan.
Hasil Scenario 7	

Berdasarkan hasil running pada 7 video menggunakan sistem identifikasi kejenuhan yang sudah ditampilkan pada Tabel 2, maka diperoleh data *output* sistem identifikasi kejenuhan pada video menggunakan *Transfer Learning* yang dapat dilihat pada Tabel 3. Persentase pada sistem identifikasi kejenuhan didapat dari jumlah frame dengan reaksi dibagi dengan total frame yang terdeteksi. Dalam scenario 3, didapat jumlah frame dengan mata memperhatikan sebanyak 899 frame, dengan total frame yang terdeteksi sebanyak 1658 frame, sehingga didapat persentase sebesar 54.19%. Kemudian masih pada scenario 3, didapat jumlah frame dengan reaksi netral sebanyak 836 frame, dengan total frame yang terdeteksi sebanyak 1811 frame, sehingga didapat persentase sebesar 46.16%.

Tabel 3 Data *Output* Sistem Identifikasi Kejenuhan

<i>Scenario</i>	<i>On Screen (%)</i>	<i>Off Screen (%)</i>	MataTak Terdeteksi (%)	Respon Netral (%)	Respon Positif (%)	Respon Negatif (%)
1	89.24	10.76	5.82	88.3	0.0	11.7
2	95.63	4.37	2.48	73.42	14.0	12.58
3	54.19	45.81	0.96	46.16	0.17	53.67
4	90.6	9.4	9.29	66.42	0.23	33.35
5	13.52	86.48	86.2	3.54	21.24	75.22
6	0.28	99.72	99.34	34.33	3.99	61.68
7	0.0	100.0	100.0	0.0	0.0	0.0
MIN	0.0	4.37	2.48	0.0	0.0	0.0
MAX	95.63	100.0	100.0	88.30	21.24	99.46
AVG (%)	57.24	42.75	34.01	52.02	6.61	41.37

Selain itu, pada Tabel 3 didapat nilai minimal sebesar 0.00% pada parameter *on screen*, respon netral, respon positif dan respon negatif. Kemudian didapat nilai maksimal sebesar 100.00% pada parameter *off screen* dan mata tak terdeteksi. Hal ini dikarenakan pada *scenario 7* tidak terdeteksi adanya wajah yang ada hanya ruangan. Lalu, didapat nilai rata-rata sebesar 57.24% pada parameter *on screen* dan nilai rata-rata sebesar 52.02% pada parameter respon netral. Hal ini sesuai dengan *scenario 1,2* dan *3* dimana siswa selalu melihat layar dengan respon netral, *scenario 5,6* dan *7* dimana siswa tidak melihat layar, serta *scenario 4* dimana siswa terkadang tidak melihat layar, banyak menengok, siswa terlihat bingung dan

mengernyitkan dahi beberapa kali. Pada Tabel 4, didapat kondisi *null-null*, hal ini dikarenakan pada *scenario 7* tidak terdeteksi adanya wajah yang ada hanya ruangan.

Tabel 4 Analisis Data *Output* Sistem Identifikasi Kejenuhan

<i>Scenario</i>	Presensi	<i>Overall Respon</i>	<i>Attendance</i>	Impresi	Status Kejenuhan		
1	Ya	1	Netral	Fokus	Baik	Tidak Jenuh	1
2	Ya	1	Netral	Fokus	Baik	Tidak Jenuh	1
3	Ya	1	Negatif	Fokus	Buruk	Jenuh	0
4	Ya	1	Netral	Fokus	Baik	Tidak Jenuh	1
5	Tidak	0	Negatif	Tidak Fokus	Buruk	Jenuh	0
6	Ya	1	Negatif	Tidak Fokus	Buruk	Jenuh	0
7	Tidak	0	null	null	null	null	-

Selanjutnya, untuk validasi sistem identifikasi kejenuhan dibutuhkan hasil analisis 7 *scenario* oleh 2 *annotator*, maka didapatkan hasil yang dapat dilihat pada Tabel 5. Persentase pada data kejenuhan oleh *annotator* didapat dari jumlah detik dengan reaksi dibagi dengan total detik yaitu 120 detik. Dalam *scenario 3* oleh *annotator 1*, didapat jumlah detik dengan mata tidak memperhatikan sebanyak 10 detik, dengan total detik yaitu 120 detik, sehingga didapat persentase sebesar 8.33%. Kemudian pada *scenario 3* oleh *annotator 2*, didapat jumlah detik dengan mata tidak memperhatikan sebanyak 20 detik, dengan total detik yaitu 120 detik, sehingga didapat persentase sebesar 16.67%. Setelah itu data dari dua *annotator* akan di rata-ratakan, sehingga didapat nilai persentase mata tidak memperhatikan pada *scenario 3* sebesar 12.5%.

Tabel 5 Data Kejenuhan oleh *Annotator*

<i>Scenario</i>	<i>On Screen (%)</i>	<i>Off Screen (%)</i>	MataTak Terdeteksi (%)	Respon Netral (%)	Respon Positif (%)	Respon Negatif (%)
1	91.67	8.33	0.0	100.0	0.0	0.0
2	70.83	8.33	12.5	16.67	83.33	0.0
3	75.0	12.5	12.5	62.5	33.33	4.16
4	29.16	54.16	16.67	54.16	4.16	41.67
5	0.0	66.67	33.33	50.0	0.0	50.0
6	0.0	50.0	50.0	50.0	0.0	50.0
7	0.0	0.0	100.0	0.0	0.0	0.0
MIN	0.0	0.0	0.0	0.0	0.0	0.0
MAX	91.67	66.67	100.00	100.0	83.33	50.0
AVG (%)	38.1	28.57	32.14	55.55	20.13	24.30

Pada Tabel 5 juga didapati nilai minimal sebesar 0.00% pada semua parameter. Kemudian didapat nilai maksimal sebesar 100.00% pada parameter *off screen* dan mata tak terdeteksi. Hal ini dikarenakan pada *scenario 7* tidak terdeteksi adanya wajah yang ada hanya ruangan. Lalu, didapat nilai rata-rata sebesar 38.1% pada parameter *on screen* dan nilai rata-rata sebesar 55.55% pada parameter respon netral. Hal ini sesuai dengan *scenario 1,2 dan 3* dimana siswa selalu melihat layar dengan respon netral, *scenario 5,6 dan 7* dimana siswa tidak melihat layar, serta *scenario* terkadang 4 dimana siswa tidak melihat layar, banyak menengok, siswa terlihat bingung dan mengernyitkan dahi beberapa kali.

Tabel 6 Analisis Data Kejenuhan oleh *Annotator*

<i>Scenario</i>	Presensi	<i>Overall Respon</i>	<i>Attendance</i>	Impresi	Status Kejenuhan		
1	Ya	1	Netral	Fokus	Baik	Tidak Jenuh	1
2	Ya	1	Positif	Fokus	Baik	Tidak Jenuh	1
3	Ya	1	Netral	Fokus	Baik	Tidak Jenuh	1
4	Ya	1	Netral	Tidak Fokus	Baik	Jenuh	0
5	Ya	1	Negatif	Tidak Fokus	Buruk	Jenuh	0

Scenario	Presensi	Overall Respon	Attendance	Impresi	Status Kejenuhan		
6	Ya	1	Negatif	Tidak Fokus	Buruk	Jenuh	0
7	Tidak	0	null	null	null	null	-

Sementara itu, Tabel 6 menunjukkan bahwa didapati kondisi *null-null*, hal ini dikarenakan pada *scenario 7* tidak terdeteksi adanya wajah dan yang nampak ada hanyalah ruangan tanpa kehadiran manusia

4.3 Hasil Uji Performan Sistem

Pengujian performansi dilakukan dengan menerapkan fitur-fitur yang terdapat pada sistem ini. Pada penelitian ini dilakukan pengujian performansi dalam hal akurasi, presisi, *recall*, dan *f1-score* berdasarkan parameter, presensi, *attendance*, impresi, dan status kejenuhan. Analisis dan pengujian ini akan dilakukan sesuai dengan parameter skenario di atas hasil sistem identifikasi kejenuhan dibandingkan dengan hasil analisis *annotator*, hasil dari pengujian ini akan ditampilkan dalam bentuk tabel dan hasil penghitungan rumusnya. Secara garis besar pada sub-bab ini akan dibagi menjadi 5 bagian yaitu pengujian metrik performa akurasi, presisi, *recall*, dan *f1-score* berdasarkan parameter presensi, *attendance*, impresi, dan status kejenuhan, serta evaluasi metrik performa sistem identifikasi kejenuhan. Tabel 7 menunjukkan hasil evaluasi performa dari sistem identifikasi kejenuhan menggunakan *Transfer Learning* ini.

Tabel 7 Evaluasi Metrik Performa Sistem Identifikasi Kejenuhan

Evaluasi Metrik Performa Sistem Identifikasi Kejenuhan				
Parameter	Akurasi	Presisi	Recall	F1-score
Presensi	85.71	100	83.33	90.91
Attendance	83.33	75.00	100	85.71
Impresi	83.33	100	75.00	85.71
Status Kejenuhan	66.67	66.67	66.67	66.67

Tabel 7 menunjukkan nilai akurasi, presisi, *recall*, dan *f1-score* dari setiap parameter. Tabel menunjukkan nilai akurasi, presisi, *recall*, dan *f1-score* lebih dari 50% pada parameter presensi, *attendance*, impresi dan status kejenuhan. Sedangkan, parameter status kejenuhan memiliki nilai akurasi, presisi, *recall*, dan *f1-score* terendah yaitu masing-masing sebesar 66.67%. Nilai ini tidak hanya bergantung pada pengenalan ekspresi wajah dan deteksi pergerakan mata, melainkan juga jumlah skenario dan juga jumlah *annotator*. Dengan semakin banyaknya jumlah skenario dan jumlah *annotator*, maka persebaran nilainya pun akan semakin besar. Kemudian sistem pengenalan ekspresi wajah setelah diimplementasikan menjadi parameter impresi pada sistem identifikasi kejenuhan menggunakan 7 *scenario*, didapat tingkat akurasi sebesar 83.33%, presisi sebesar 100%, *recall* sebesar 75 % dan *f1-score* sebesar 85.71%. Meskipun pada saat *training* dan *validation* menggunakan dataset FER2013, model pengenalan ekspresi wajah memiliki tingkat akurasi pembacaan 68% dalam 25 *epochs*.



Gambar 8 Kesalahan Pembacaan Arah Bola Mata pada *Scenario 4*

Selanjutnya, sistem ini perlu dioptimalkan untuk meningkatkan presisi dalam deteksi attendance melalui pembacaan arah bola mata. Saat ini, sistem sering salah membaca arah mata sebagai “memperhatikan” meskipun siswa tidak melihat *webcam* seperti yang terlihat pada Gambar 8. Selain itu, deteksi emosi siswa melalui ekspresi wajah juga perlu diperbaiki. Kesalahan dalam membaca ekspresi sering terjadi, seperti ketika ekspresi netral dianggap negatif atau sedih karena posisi wajah yang kurang ideal saat *video conferencing*.

Sistem juga sering mendeteksi ekspresi yang tidak relevan dengan materi pelajaran karena menangkap interaksi siswa dengan lingkungan sekitar. Menggunakan dataset FER-2013 dengan akurasi 68%, sistem ini masih dapat ditingkatkan dengan model *machine learning* yang lebih akurat. Ketepatan pembacaan ekspresi penting karena mempengaruhi penilaian kejuhan siswa. Selain itu, deteksi pergerakan mata juga mempengaruhi nilai fokus, terutama jika siswa sering memalingkan wajah atau mengedip. Berdasarkan penelitian, frekuensi kedipan mata berkaitan dengan tingkat fokus siswa, sehingga deteksi “mata tidak terdeteksi” perlu diperhitungkan secara akurat. Dari uji coba sistem identifikasi kejuhan, dapat disimpulkan bahwa sistem identifikasi kejuhan cukup efektif untuk mengukur kejuhan siswa pada metode pembelajaran daring. Sistem identifikasi kejuhan dapat mengukur presensi, status *attendance* siswa memperhatikan layar, emosi, impresi dan status kejuhan siswa.

4. KESIMPULAN

Dari hasil penelitian yang dilakukan untuk mengidentifikasi kejuhan siswa, dapat diperoleh kesimpulan sebagai berikut:

1. Metode *Transfer Learning* diimplementasikan untuk membuat model pengenalan ekspresi wajah. *Pre-trained model* yang telah dilatih sebelumnya digunakan sebagai bobot awal pelatihan atau disebut *Fine Tuning*. Setelah model dilatih, arsitektur VGG16 yang telah di-tuning disimpan untuk digunakan pada tahap pengujian sistem identifikasi kejuhan. Pengenalan ekspresi wajah dengan dataset FER2013 dan metode *transfer learning* memiliki tingkat akurasi pembacaan 68% dalam 25 *epochs*. Setelah diimplementasikan menjadi parameter impresi pada sistem identifikasi kejuhan menggunakan 7 skenario, didapat tingkat akurasi sebesar 83.33%, presisi sebesar 100%, *recall* sebesar 75%, dan *f1-score* sebesar 85.71%.
2. Pengenalan ekspresi wajah dikombinasikan dengan deteksi pergerakan mata dan dianalisis berdasarkan dimensi *behavioral* (apakah seorang siswa fokus atau tidak fokus) dan dimensi *emotional* (apakah seorang siswa menunjukkan emosi netral, positif, atau negatif).
3. Setelah pengujian sistem identifikasi kejuhan dilakukan menggunakan 7 skenario, didapat tingkat akurasi sebesar 66.67%, presisi sebesar 66.67%, *recall* sebesar 66.67%, dan *f1-score* sebesar 66.67%. Nilai tersebut dipengaruhi oleh kesalahan deteksi pergerakan mata pada sistem yang diakibatkan oleh wajah yang menghadap agak ke kanan atau ke kiri dengan posisi pupil berada di tengah-tengah bola mata meskipun mata sedang tidak menatap *webcam*, seperti pada skenario 4. Selain itu, posisi wajah terhadap layar serta posisi dan kemiringan badan siswa terhadap layar juga mempengaruhi pengenalan ekspresi wajah, seperti pada skenario 3 dan skenario 6.

4. SARAN

Dalam pengembangan pengenalan ekspresi wajah berbasis citra menggunakan metode *transfer learning* ini diperlukan beberapa perbaikan untuk mencapai hasil yang lebih maksimal, maka perlu dilakukan perbaikan pada beberapa hal diantaranya:

1. Memperbaiki sistem pengenalan ekspresi wajah sehingga dapat mendeteksi pergerakan arah kepala (melihat ke kanan, kiri, atas dan bawah) lebih akurat dan tidak dianggap tidak ada wajah.
2. Sistem identifikasi kejenuhan dapat men-train dengan dataset baru, baik yang sudah ada maupun buatan sendiri atau menggunakan model *machine learning* yang lebih akurat.
3. Sistem identifikasi kejenuhan dapat di implementasikan pada *video conference app* yang sudah ada atau yang baru dibuat, sehingga dapat dilakukan evaluasi pada pembelajaran secara langsung.
4. Dipertimbangkan untuk menggunakan pakar ekspresi sebagai anotator, sehingga hasil anotasi menjadi lebih akurat.

DAFTAR PUSTAKA

- [1] A. H. Rustaman, "Efektivitas penggunaan aplikasi daring, video conference dan sosial media pada mata kuliah komputer grafis 1 di masa pandemi covid-19," *JISIP (Jurnal Ilmu Sosial dan Pendidikan)*, vol. 4, no. 3, 2020.
- [2] R. Pawicara and M. Conilie, "Analisis pembelajaran daring terhadap kejenuhan belajar mahasiswa Tadris Biologi IAIN Jember di tengah pandemi Covid-19," *ALVEOLI: Jurnal Pendidikan Biologi*, vol. 1, no. 1, pp. 29–38, 2020.
- [3] N. Alyuz *et al.*, "Semi-supervised model personalization for improved detection of learner's emotional engagement," in *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, 2016, pp. 100–107.
- [4] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, "The faces of engagement: Automatic recognition of student engagement from facial expressions," *IEEE Transactions on Affective Computing*, vol. 5, no. 1, pp. 86–98, 2014.
- [5] H. Monkaresi, N. Bosch, R. A. Calvo, and S. K. D'Mello, "Automated detection of engagement using video-based estimation of facial expressions and heart rate," *IEEE Transactions on Affective Computing*, vol. 8, no. 1, pp. 15–28, 2016.
- [6] D. L. Z. Astuti and S. Samsuryadi, "Kajian Pengenalan Ekspresi Wajah menggunakan Metode PCA dan CNN," in *Annual Research Seminar (ARS)*, 2019, vol. 4, no. 1, pp. 293–297.
- [7] D. Prasetyawan, "Penentuan Emosi pada Video dengan Convolutional Neural Network," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 5, no. 1, pp. 23–35, 2020.
- [8] J. Chen, F. Lécué, J. Z. Pan, I. Horrocks, and H. Chen, "Knowledge-based transfer learning explanation," *arXiv preprint arXiv:1807.08372*, 2018.
- [9] O. M. Nezami, M. Dras, L. Hamey, D. Richards, S. Wan, and C. Paris, "Automatic recognition of student engagement using deep learning and facial expression," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 2019, pp. 273–289.
- [10] A. Zein, "Pendeteksian Kantuk Secara Real Time Menggunakan Pustaka Opencv Dan Dlib Python," *Sainstech: Jurnal Penelitian dan Pengkajian Sains dan Teknologi*, vol. 28, no. 2, 2018.
- [11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [12] D. Triasanti, "Konsep Dasar Python," *Surabaya: Sulita Jaya*, 2002.
- [13] I. Taufiq, "Deep Learning Untuk Deteksi Tanda Nomor Kendaraan Bermotor Menggunakan Algoritma Convolutional Neural Network Dengan Python Dan Tensorflow." Skripsi. Program Studi Sistem Informasi Sekolah Tinggi Manajemen Informatika ..., 2018.
- [14] F. Chollet, "Keras documentation," *keras.io*, vol. 33, 2015.
- [15] K. D. P. Novianti, N. A. Setiawan, and S. S. Kusumawardani, "Peningkatan Nilai Recall dan Precision pada Penelusuran Informasi Pustaka Berbasis Semantik (Studi Kasus:

- Sistem Informasi Ruang Referensi Jurusan Teknik Elektro dan Teknologi Informasi UGM),” *Proceedings Konferensi Nasional Sistem dan Informatika (KNS&I)*, 2015.
- [16] J. Grafsgaard, J. B. Wiggins, K. E. Boyer, E. N. Wiebe, and J. Lester, “Automatically recognizing facial expression: Predicting engagement and frustration,” 2013.
- [17] S. Aslan, N. Alyuz, E. Okur, E. M. Sinem, E. Oktay, and A. A. Esme, “Effect of emotion-aware interventions on students’ behavioral and emotional states,” *Educational Technology, Research and Development*, vol. 66, no. 6, pp. 1399–1413, 2018.
- [18] I. J. Goodfellow et al., “Challenges in representation learning: A report on three machine learning contests,” in *International conference on neural information processing*, 2013, pp. 117–124.
- [19] M. Wardhana and M. Hariadi, “Ekspresi emosi pada model wajah tiga dimensi menggunakan naive bayes dan logika fuzzy,” 2019.
- [20] S. A. H. Alrubaie and A. H. Hameed, “Dynamic weights equations for converting grayscale image to RGB image,” *Journal of University of Babylon for Pure and Applied Sciences*, vol. 26, no. 8, pp. 122–129, 2018.
- [21] T. Nakano, M. Kato, Y. Morito, S. Itoi, and S. Kitazawa, “Blink-related momentary activation of the default mode network while viewing videos,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 2, pp. 702–706, 2013.