

## KLASIFIKASI EMOSI PADA RAUT WAJAH PELAJAR MENGGUNAKAN EKSTRAKTOR FITUR FACE MESH DAN METODE SUPPORT VECTOR MACHINE

Muhammad Nugraha Delta Revanza<sup>\*1</sup>, Fitra Abdurrachman Bachtiar<sup>2\*</sup>, Budi Darma Setiawan<sup>3</sup>

<sup>1,2,3</sup>Universitas Brawijaya, Malang

Email: <sup>1</sup>drevanza19@student.ub.ac.id, <sup>2</sup>fitra.bachtiar@ub.ac.id, <sup>3</sup>s.budidarma@ub.ac.id

\*Penulis Korespondensi

(Naskah masuk: 22 Febuari 2025, diterima untuk diterbitkan: 27 Agustus 2025)

### Abstrak

Dalam lingkup pembelajaran, rasa emosional dan perhatian memegang peranan penting dalam keterlibatan pelajar terhadap proses pembelajaran yang sedang berlangsung. Emosi pelajar menimbulkan reaksi afektif terhadap proses pembelajaran, seperti *boredom*, *engagement*, *confusion*, dan *frustration*. Reaksi afektif tersebut dapat digunakan sebagai tolok ukur dalam melakukan evaluasi kegiatan pembelajaran. Pengenalan emosi dapat dilakukan dengan pengamatan citra wajah, namun pemrosesan sebuah citra memerlukan sebuah model yang dapat melakukan klasifikasi emosi berdasarkan raut wajah secara tepat. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membangun sebuah sistem pengenalan emosi melalui raut wajah pelajar dengan ekstraktor fitur Mediapipe Face Mesh dan Support Vector Machine (SVM). Proses ekstraksi *frame* dan ekstraksi fitur dilakukan untuk mendapatkan total 1404 titik tiga dimensi *facial landmark* untuk kemudian diklasifikasikan dengan menggunakan SVM. Untuk meningkatkan kinerja klasifikasinya, dilakukan optimasi algoritma SVM melalui *hyperparameter tuning* dan Grid Search Cross Validation untuk menghasilkan kombinasi parameter model dengan kinerja terbaik. Hasil yang diperoleh adalah 53%, mengalami peningkatan 25% dibandingkan dengan model standar tanpa proses *hyperparameter tuning* yang menunjukkan bahwa *hyperparameter tuning* memiliki pengaruh terhadap kinerja model. Selain itu, terdapat titik-titik *facial landmark* yang berperan dalam klasifikasi emosi berdasarkan hasil analisis, yaitu titik yang berada di sekitar mata.

**Kata kunci:** pengenalan emosi, *facial landmark*, *hyperparameter tuning*, SVM, *face mesh*

## EMOTION CLASSIFICATION ON STUDENTS' FACIAL EXPRESSIONS USING FACE MESH FEATURE EXTRACTOR AND SUPPORT VECTOR MACHINE

### Abstract

*In the context of learning, emotion and attention are significant factors influencing the learner's engagement with the ongoing learning process. The affective reactions of learners to the learning process, which may include boredom, engagement, confusion, or frustration, can be influenced by their emotional state. Such affective reactions may be employed as benchmarks for the evaluation of learning activities. The emotions can be recognized by analyzing the image of human face. However, image processing needs a model that can accurately categorize emotions based on facial expressions. This research aims to address these issues through the construction of an emotion recognition system based on student facial expressions using the Mediapipe Face Mesh feature extractor and Support Vector Machine. First, a frame extraction and feature extraction process was conducted to obtain a total of 1,404 three-dimensional facial landmark points as input data. Subsequently, the SVM algorithm was optimized through hyperparameter tuning and Grid Search Cross Validation to produce a combination of model parameters with the best performance. The resulting value was 53%, representing a 25% increase compared to the standard model without hyperparameter tuning, which demonstrates that hyperparameter tuning has a significant impact on model performance. Additionally, the analysis revealed that certain facial landmark points, particularly those around the eyes, play a crucial role in emotion classification.*

**Keywords:** emotion recognition, facial landmark, hyperparameter tuning, SVM, confusion matrix.

### 1. PENDAHULUAN

Pembelajaran pada era modern sekarang dapat dilakukan tidak hanya dengan bertatap muka secara

langsung, tetapi juga dapat dilakukan secara daring. Namun, sebuah penelitian menunjukkan bahwa tingkat kehadiran peserta didik pada tahun 2020 berkisar antara 55% dengan tingkat partisipasi

mahasiswanya sebesar 40% (Qutishat, Obeidallah and Qawasmeh, 2022). Penelitian yang dilakukan oleh Pascarella, Seifert dan Blaich (2010) mengungkapkan bahwa ketertarikan pelajar juga dapat berdampak terhadap partisipasi maupun prestasi yang dapat dicapai oleh pelajar. Oleh karena itu, evaluasi terhadap tingkat ketertarikan pelajar diperlukan untuk memberikan umpan balik bagi pengajar dalam menyusun strategi pembelajaran (De Carolis et al., 2019).

Ketertarikan emosi sendiri bisa dilihat melalui *micro-expression* yang dapat merepresentasikan emosi yang dirasakan oleh pelajar secara refleksi (Thuseethan, Rajasegarar and Yearwood, 2019). Mengetahui emosi pada pelajar sendiri dibutuhkan untuk membuat suasana pembelajaran menjadi lebih interaktif dan menghilangkan kebosanan yang biasanya terjadi di tengah proses pembelajaran (Krithika L.B and Lakshmi Priya GG, 2016). Penelitian Whitehill et al., (2014) mengidentifikasi beberapa metode yang digunakan untuk melakukan evaluasi terhadap ketertarikan pelajar, mulai dari *self-report* hingga pengukuran fitur, misalnya emosi wajah, secara otomatis.

Pengukuran terhadap tingkat ketertarikan pelajar terhadap proses belajar menjadi hal penting sebagai evaluasi proses pembelajaran. Paidja dan Bachtiar (2022) berhasil mendapatkan akurasi 97,51% melalui model Convolutional Neural Network (CNN) dan ekstraktor fitur Dlib pada deteksi emosi pada pelajar. Selain itu, Bhatia, Tomar dan Jain (2021) melakukan eksperimen menggunakan ekstraktor fitur HOG dan membandingkan hasil model SVM dan CNN yang menghasilkan akurasi hingga 98%. Metode *machine learning* konvensional dengan ekstraktor fitur Face Mesh dan Principal Component Analysis (PCA) menghasilkan akurasi hingga 97% dalam mendeteksi emosi pada wajah manusia (Siam et al., 2022). Multi-level CNN dan fitur ekstraktor OpenFace dan Mediapipe Pose mendapatkan akurasi hingga 82% pada salah satu kelas, yaitu *confusion* (Solanki and Mandal, 2022).

Berdasarkan kekurangan yang ada pada penelitian sebelumnya, penelitian ini mengusulkan sebuah metode untuk melakukan klasifikasi emosi menggunakan ekstraktor fitur Mediapipe Face Mesh dan Support Vector Machine (SVM). Penelitian ini bertujuan untuk meneliti penerapan ekstraktor fitur Face Mesh dan SVM terhadap klasifikasi emosi pelajar menggunakan dataset dari DAiSEE. Secara garis besar, tahapan penelitian meliputi pengumpulan data, ekstraksi gambar dari video, ekstraksi fitur, normalisasi data, pelatihan, dan pengujian. Hasil dari penelitian ini diharapkan dapat dijadikan sebagai referensi dalam penyusunan strategi dalam kegiatan pembelajaran serta pengembangan penelitian *affective computing* pada pengenalan emosi wajah.

## 2. KAJIAN PUSTAKA

Paidja dan Bachtiar (2022) membuat sebuah model klasifikasi emosi menggunakan metode CNN, ekstraktor fitur Dlib, dan dataset yang DAiSEE. *Facial landmark* yang berhasil diekstrak lalu dihitung jarak antar titiknya menggunakan *euclidean distance*. Penelitiannya menghasilkan model dengan akurasi 97,80% dengan waktu komputasi 411,45 detik.

Penelitian yang dilakukan oleh Siam et al. (2022) melakukan klasifikasi emosi manusia menggunakan beberapa metode *machine learning*. Dataset Japanese Female Facial Expression (JAFPE), RAF-DB dan the Extended Cohn-Kanade (CK+) dilakukan ekstraksi fitur menggunakan Mediapipe Face Mesh sebelum direduksi fiturnya menggunakan PCA. Setelah itu, dataset dilakukan pelatihan dan pengujian menggunakan berbagai metode *machine learning*, yaitu Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naive Bayes, Regresi Logistik, dan Random Forest. Metode K-NN berhasil menghasilkan akurasi hingga 97% dengan waktu komputasi yang singkat dibandingkan dengan penggunaan metode berbasis *deep learning*. Penelitian ini menyarankan adanya penambahan fitur untuk deteksi emosi, seperti suara, tulisan, atau video.

Solanki dan Mandal (2022) melakukan penelitian terhadap klasifikasi ekspresi wajah dengan *engagement analysis* pada DAiSEE dataset. Penelitian ini melakukan ekstraksi fitur wajah menggunakan OpenFace dan ekstraksi pose menggunakan Mediapipe pada 2.157.890 dataset hasil *frame extraction*. Dengan metode multi-level model berbasis CNN, didapatkan hasil klasifikasi 63,8% (*boredom*), 77,2% (*engagement*), 82% (*confusion*), dan 86,6% (*frustration*). Penelitian ini menyarankan adanya proses *hyperparameter tuning* untuk meningkatkan kinerja dari model.

Penelitian terakhir melakukan perbandingan model terhadap sistem klasifikasi emosi akademik. Penelitian ini menerapkan berbagai macam kondisi terhadap model, mulai dari penggunaan ekstraktor fitur, deteksi secara *real-time*, hingga kombinasi algoritma klasifikasi yang digunakan. Hasil terbaik yang didapatkan dari penelitian ini merupakan algoritma *long-term recurrent CNN* tanpa ekstraktor fitur dengan akurasi sebesar 94,6%. Penelitian ini menyarankan adanya standarisasi label pada dataset dan melakukan pengamatan terhadap siswa lebih lama dengan tambahan sensor atau rekaman video (Lek and Teo, 2023).

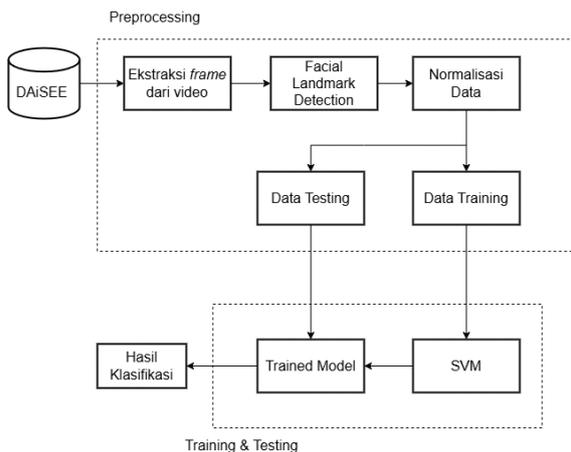
Berdasarkan penelitian yang dilakukan oleh Paidja dan Bachtiar (2022), penggunaan ekstraktor fitur, dalam penelitian ini Dlib, membantu kinerja model dalam memproses data citra menjadi data latih. Namun, Dlib memiliki kekurangan dalam mendeteksi bentuk wajah yang tidak menghadap ke arah kamera. Penelitian lain yang dilakukan oleh Bhatia, Tomar dan Jain (2021) menunjukkan bahwa jumlah data akan mempengaruhi kinerja model, terutama model yang berbasis Neural Network agar model tidak mengalami *overfitting*. Penelitian yang dilakukan

Solanki dan Mandal (2022) dilakukan tanpa proses *hyperparameter tuning*, sehingga hasil prediksi terhadap *affective state* pada data DAiSEE kurang baik. Lek dan Teo (2023) menjelaskan bahwa dataset DAiSEE memiliki kekurangan pada standarisasi label pada dataset video yang disediakan sehingga mengurangi keakuratan prediksi emosi citra dibandingkan dengan emosi asli.

Penelitian ini mengusulkan sistem klasifikasi emosi menggunakan ekstraktor fitur Mediapipe Face Mesh dan SVM. Ekstraktor fitur digunakan berdasarkan metode yang diusulkan pada penelitian Paidja dan Bachtiar (2022) dan Solanki dan Mandal (2022). Pada penelitian ini, model SVM dengan *hyperparameter tuning* untuk meningkatkan kinerja model. Klasifikasi dilakukan berdasarkan *affective state* dari pelajar, yaitu *boredom*, *engagement*, *confusion*, dan *frustration* tanpa mengikutsertakan intensitas emosi.

### 3. METODE YANG DIUSULKAN

Metodologi yang diusulkan pada penelitian ini direpresentasikan pada Gambar 1. Pertama, data berupa video dari DAiSEE diekstrak menjadi sebuah *frame*. Kemudian, dilakukan ekstraksi fitur menggunakan Face Mesh dan normalisasi data berupa koordinat tiga dimensi menggunakan Min-Max Scaler. Data yang telah diproses kemudian dimasukkan ke dalam model *hyperparameter tuning* berdasarkan *hyperparameter* yang telah ditentukan. Kombinasi *hyperparameter* terbaik digunakan sebagai model pelatihan sebelum dilakukan pengujian terhadap kinerja model serta analisis titik *facial landmark* terhadap kinerja model.



Gambar 1. Gambaran Umum Sistem

#### 3.1. Dataset

Data yang digunakan dalam penelitian ini adalah Dataset for Affective States in E-Environments (DAiSEE), sebuah dataset publik dikenalkan oleh Gupta et al., (2016). DAiSEE berisi 9.068 klip video dari 112 pelajar, dengan rentang umur 18-30 tahun, saat mengikuti pembelajaran daring. Dataset ini telah dilabeli dengan 4 kategori,

yaitu *boredom*, *engagement*, *confusion*, dan *frustration*. Masing-masing label memiliki intensitas dengan rentang 0 (*very low*), 1 (*low*), 2 (*high*), dan 3 (*very high*). DAiSEE berusaha menjadi solusi dalam pengenalan tingkat *engagement* seseorang yang dapat diterapkan dalam berbagai bidang, mulai dari kesehatan hingga *e-learning* (Gupta et al., 2016).

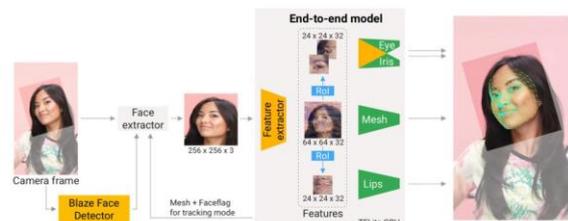
#### 3.2. Preprocessing

Citra berupa video diekstraksi terlebih dahulu menggunakan *library* FFmpeg berdasarkan penelitian Li, (2018). Video berdurasi 10 detik diekstrak menjadi sebuah *frame* dengan interval 15 fps (*frame-per-second*). Dataset yang digunakan dalam penelitian berjumlah masing-masing 20 video pada label *boredom*, *engagement*, dan *confused*, serta 17 video pada label *frustration*. Label dari tiap video ditentukan berdasarkan skor intensitas emosi tertinggi yang diberikan pada penelitian (Gupta et al., 2016). *Frame* yang diekstrak melalui proses ekstraksi berjumlah 101.724 *frame* video.

Ekstraksi fitur dilakukan dengan Mediapipe Face Mesh, sebuah model *machine learning* pendeteksi topologi dan bentuk wajah dalam bentuk koordinat tiga dimensi. Arsitektur Face Mesh terdiri atas kombinasi model BlazeFace pada deteksi citra atau video menggunakan sebuah *bounding box* (Nuralif et al., 2023). Setelah *bounding box* berhasil diidentifikasi, Face Mesh akan melakukan estimasi 468 titik koordinat tiga dimensi *facial landmark* tiap *frame*. Gambaran dari *facial landmark* pada wajah dapat dilihat pada Gambar 3.



Gambar 2. Gambaran 468 titik landmark menggunakan Face Mesh



Gambar 3. Arsitektur Face Mesh

Mediapipe Face Mesh akan mendeteksi posisi wajah dan menjadikannya sebuah *input* dengan ukuran 256×256. Citra masukan kemudian dideteksi

posisi wajahnya menjadi citra berukuran 64×64 dan membagi model menjadi dua, yaitu pendeteksi *bounding box* dan pendeteksi area spesifik seperti mulut dan hidung. Proses ini menghasilkan citra berukuran 6×6 sehingga proses deteksi wajah lebih ringan dan cepat (Nuralif et al., 2023). Gambaran proses dari Face Mesh dapat dilihat pada Gambar 2.

### 3.3. Training dan Testing

Data yang digunakan pada penelitian merupakan *subset* dari data citra yang sudah terekstrak secara acak, berjumlah 1.000 data per kelas, sehingga total data yang digunakan adalah 4.000. Data kemudian dibagi menjadi data latih dan data tes dengan rasio 75:25, menghasilkan 3.000 data latih dan 1.000 data uji. Data latih selanjutnya dijadikan masukan untuk model pencarian *hyperparameter tuning* dengan *hyperparameter* SVM yang telah ditentukan, yaitu *C*, *kernel*, dan *tolerance (tol)* menggunakan Grid Search Cross Validation dengan jumlah dengan nilai *K* = 10. Skenario *hyperparameter tuning* direpresentasikan pada **Error! Reference source not found.**

Tabel 1. Skenario *Hyperparameter Tuning*

| Parameter     | Nilai  |     |            |
|---------------|--------|-----|------------|
| <i>C</i>      | 0,1    | 1   | 10         |
| <i>kernel</i> | linear | rbf | polynomial |
| <i>Tol</i>    | 0,001  | 0,1 | 1          |

*Hyperparameter* yang menghasilkan model terbaik dijadikan nilai *hyperparameter* pada proses pelatihan. Proses pelatihan dilakukan dengan Support Vector Machine (SVM) yang didasarkan pada *library* Scikit-Learn (Pedregosa et al., n.d.). Model hasil pelatihan selanjutnya akan diuji menggunakan data tes untuk dievaluasi kinerjanya berdasarkan *f1-score*. Hasil pengujian akan dijadikan sebagai acuan dalam melakukan analisis terhadap kinerja model dan pengaruh titik *facial landmark* terhadap akurasi model klasifikasi emosi pada raut wajah pelajar menggunakan Face Mesh dan SVM.

### 3.4. Evaluasi

Evaluasi digunakan untuk membandingkan label prediksi dengan label asli atau menafsirkan hasil klasifikasi berdasarkan data yang telah diberikan (Sun, Wong and Kamel, 2009). Model yang baik merupakan model yang hanya memiliki nilai pada diagonalnya dan distribusi nilainya tidak merata (Yun, 2021). Metrik evaluasi yang digunakan pada penelitian ini adalah akurasi, *recall*, *precision*, dan *f1-score*. Metrik tersebut dihitung berdasarkan nilai *confusion matrix*, yaitu *true positive (TP)*, *true negative (TN)*, *false positive (FP)*, dan *false negative (FN)*. Perhitungan nilai akurasi dapat dilihat pada (1). (1)(2) menunjukkan perhitungan dari nilai *precision*. Persamaan (3) (4) menunjukkan masing-masing perhitungan dari nilai *recall* dan *f1-score*.

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

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN} \quad (3)$$

$$f1 - score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \quad (4)$$

## 4. HASIL EKSPERIMEN DAN ANALISIS

### 4.1. Pengujian *Hyperparameter Tuning*

Pengujian *hyperparameter tuning* dilakukan dengan metode Grid Search Cross Validation. Metode ini melakukan uji coba terhadap kombinasi nilai *hyperparameter* yang terdapat pada **Error! Reference source not found.** untuk mencari kombinasi *hyperparameter* terbaik. Kemudian, hasil pencarian *hyperparameter* dibandingkan dengan model tanpa penyetelan. Hasil analisis dijadikan tolok ukur dalam meningkatkan kinerja model SVM. Berikut adalah kombinasi nilai *hyperparameter* yang diuji.

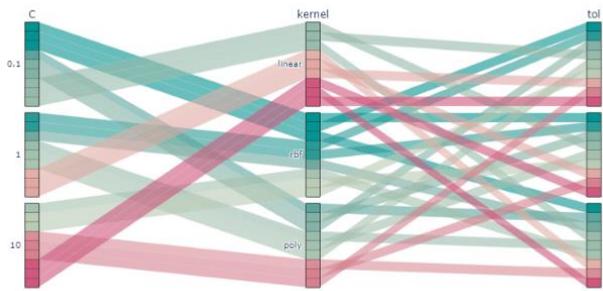
Tabel 2. Hasil Pengujian *Hyperparameter Tuning*

| No  | <i>C</i> | <i>kernel</i> | <i>tol</i> | <i>f1-score</i> |
|-----|----------|---------------|------------|-----------------|
| 1   | 1        | poly          | 1          | 0,312871        |
| 2   | 10       | linear        | 0,001      | 0,536227        |
| 3   | 10       | linear        | 0,1        | 0,534972        |
| ... | ...      | ...           | ...        | ...             |
| 25  | 10       | poly          | 0,001      | 0,493177        |
| 26  | 10       | poly          | 0,1        | 0,492529        |
| 27  | 10       | poly          | 1          | 0,479409        |

Riwayat pencarian dari kombinasi *hyperparameter* terbaik dapat dilihat pada Gambar 4 (warna merah menunjukkan kinerja model yang lebih baik). Kombinasi *hyperparameter* terbaik dapat dilihat pada Tabel 3. Kemudian, dilakukan analisis terhadap masing-masing parameter SVM terhadap kinerja model. **Error! Reference source not found.** menunjukkan bahwa parameter *C* dan *kernel* memiliki pengaruh besar terhadap kinerja dari model SVM. Sementara itu, parameter *tol* menunjukkan bahwa tidak ada perbedaan besar antara ketiga parameter yang diujikan.

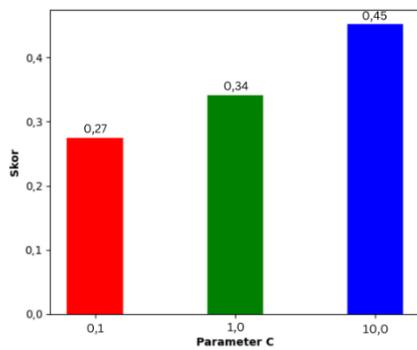
Tabel 3. Kombinasi *Hyperparameter* Terbaik

| Parameter     | Nilai  |
|---------------|--------|
| <i>C</i>      | 10     |
| <i>kernel</i> | linear |
| <i>tol</i>    | 0,001  |

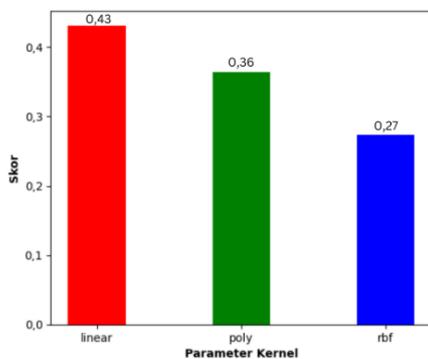


Gambar 4. Riwayat Pencarian *Hyperparameter*

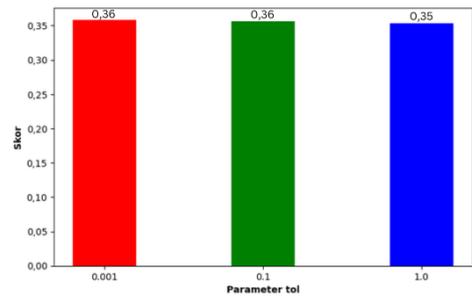
Gambar 5 menunjukkan bahwa nilai *C* yang tinggi cenderung menghasilkan kinerja model yang lebih baik. Hal ini dibuktikan dengan *C* yang bernilai 10 memiliki perbedaan hingga 18% dibandingkan dengan nilai 0,1 dan 1. Gambar 6 menunjukkan bahwa *kernel* berjenis *linear* memiliki pengaruh terbesar terhadap kinerja model. Pengaruh pemilihan *kernel* dibuktikan dengan adanya perbedaan hingga 16% terhadap rata-rata nilai skor selama pengujian. Hal ini juga menunjukkan bahwa data masukan merupakan data yang dapat dipisahkan secara linear. Gambar 7 menunjukkan bahwa tidak ada perbedaan yang signifikan pada variasi nilai *tol*. Oleh karena itu, variasi nilai *tol* tidak memiliki pengaruh besar terhadap kinerja model saat melakukan prediksi label data uji.



Gambar 5. Pengaruh Parameter *C* terhadap Model



Gambar 6. Pengaruh Parameter *Kernel* terhadap Model



Gambar 7. Pengaruh Parameter *tol* terhadap Model

Berdasarkan *hyperparameter* terbaik, dilakukan pengujian terhadap skor model menggunakan *confusion matrix*. Melalui pengujian ini, didapatkan hasil seperti pada Tabel 4. Melalui Tabel 4, didapatkan bahwa akurasi yang didapatkan model adalah 53% dengan kelas *boredom* memiliki nilai *precision*, *recall*, dan *f1-score* paling besar dibandingkan dengan kelas yang lain. Hal ini menunjukkan bahwa model lebih mudah dalam mendeteksi kelas *boredom* dibandingkan dengan lainnya.

Tabel 4 juga menunjukkan bahwa model yang diusulkan pada penelitian tidak lebih baik dari model-model yang sudah ada pada penelitian sebelumnya. Kurangnya data citra masukan sebagai data latihan menjadi salah satu faktor kurang baiknya kinerja model. Penelitian yang dilakukan oleh Paidja and Bachtiar (2022) juga menunjukkan bahwa kinerja CNN lebih baik dibandingkan dengan metode klasifikasi menggunakan *machine learning* dan ekstraktor fitur Face Mesh. Namun, *hyperparameter tuning* berhasil meningkatkan kinerja model. Perbedaan hasil evaluasi dari model dengan *hyperparameter tuning* dengan model tanpa *hyperparameter tuning* dapat dilihat pada Tabel 4 dan Tabel 5.

Tabel 4. Evaluasi Pelatihan dengan Parameter Terbaik

| Kelas             | <i>precision</i> | <i>recall</i> | <i>f1-score</i> | Akurasi |
|-------------------|------------------|---------------|-----------------|---------|
| <i>Boredom</i>    | 0,86             | 0,75          | 0,80            | 0,53    |
| <i>Engaged</i>    | 0,48             | 0,35          | 0,40            |         |
| <i>Confused</i>   | 0,37             | 0,38          | 0,38            |         |
| <i>Frustrated</i> | 0,46             | 0,64          | 0,54            |         |

Tabel 5. Evaluasi Pelatihan tanpa Parameter Terbaik

| Kelas             | <i>precision</i> | <i>recall</i> | <i>f1-score</i> | Akurasi |
|-------------------|------------------|---------------|-----------------|---------|
| <i>Boredom</i>    | 0,28             | 0,31          | 0,29            | 0,28    |
| <i>Engaged</i>    | 0,40             | 0,06          | 0,10            |         |
| <i>Confused</i>   | 0,24             | 0,11          | 0,15            |         |
| <i>Frustrated</i> | 0,28             | 0,64          | 0,39            |         |

#### 4.2. Analisis SVM *Coef\_*

SVM *coef\_* merupakan salah satu properti SVM yang menunjukkan bobot data latihan, dalam hal ini bobot titik *facial landmark*, terhadap klasifikasi label kelas saat melakukan prediksi. Variabel ini diambil lalu direpresentasikan dalam bentuk *heatmap* seperti yang terlihat pada Gambar 8. Titik koordinat yang

berwarna gelap memiliki pengaruh yang besar terhadap penentuan label uji. Sebaliknya, titik yang berwarna terang tidak berpengaruh terhadap penentuan label dari data uji saat proses prediksi.

Gambar 8, yang merepresentasikan kelas *boredom*, menunjukkan bahwa hampir seluruh titik memiliki warna terang, bahkan pada bagian sekitar mata. Berdasarkan hasil pengujian, dapat disimpulkan bahwa kelas *boredom* tidak bergantung pada fitur tertentu dalam melakukan klasifikasi emosi. Oleh karena itu, pada kasus ini dapat disimpulkan bahwa model lebih mudah dalam mengidentifikasi fitur-fitur yang terdapat di dalam kelas *boredom* tanpa adanya fitur pembeda.



Gambar 8. SVM Coef pada Kelas *Boredom*



Gambar 9. SVM Coef pada Kelas *Engaged*

Berbeda dengan kelas *boredom*, kelas *engaged*, *confused*, dan *frustration* memiliki fitur-fitur pembeda yang berada di sekitar mata, terutama bagian kelopak mata. Hal ini dapat dilihat pada Gambar 9, Gambar 10, dan Gambar 11. Apabila dibandingkan dengan hasil evaluasi model pada Tabel 4, kelas *engaged* memiliki kinerja yang kurang baik tetapi hasilnya dekat dengan hasil dari kelas *confused*. Hasil tersebut menunjukkan bahwa kelas *engaged* memiliki beberapa kesamaan dalam nilai fitur dengan kelas *confused* dan *frustration*. Kelas *frustration* sendiri memiliki matriks nilai pengujian yang lebih baik dibandingkan dengan kelas *engaged* dan *confused*. Berdasarkan itu, penambahan titik-titik

*facial landmark* yang signifikan juga memiliki pengaruh terhadap kinerja model dalam melakukan klasifikasi.



Gambar 10. SVM Coef pada Kelas *Confused*



Gambar 11. SVM Coef pada Kelas *Frustration*

## 5. KESIMPULAN

Berdasarkan implementasi dan hasil pengujian terhadap klasifikasi emosi pelajar menggunakan Mediapipe Face Mesh dan SVM, dapat ditarik kesimpulan sebagai berikut: Penerapan *hyperparameter tuning* menggunakan Grid Search Cross Validation terbukti meningkatkan kinerja model. Kinerja model yang telah dilakukan penyetalan *hyperparameter* menggunakan parameter terbaik mengalami peningkatan sebesar 25% dibandingkan dengan model SVM bawaan. Setelah dilakukan analisis lebih mendalam, ditemukan bahwa tidak semua parameter berpengaruh ke kinerja model. Parameter C dan kernel merupakan parameter pengaruh yang besar terhadap kinerja model pada saat tahap pengujian. Sementara itu, nilai dari tol sendiri tidak memiliki pengaruh signifikan terhadap kinerja model secara keseluruhan.

SVM *coef\_* memberikan tambahan informasi terkait signifikansi sebuah fitur dalam proses klasifikasi emosi melalui raut wajah. Setelah melakukan tahap pengujian, dapat dilihat bahwa terdapat beberapa titik landmark yang memiliki pengaruh besar terhadap hasil klasifikasi pada kelas *engaged*, *confused*, dan *frustration*, yaitu *landmark*

yang berada di daerah mata. Namun, hal ini tidak berlaku pada kelas boredom karena tidak adanya titik-titik *landmark* khusus yang digunakan sebagai fitur utama dalam melakukan prediksi. Secara umum, model tetap mengambil titik-titik *landmark* pada daerah sekitar mata untuk digunakan sebagai fitur utama dalam melakukan klasifikasi.

Pada penelitian selanjutnya, peneliti ingin melakukan penambahan fitur lain, seperti *blendshape*, untuk meningkatkan kinerja model. Selain itu, penggunaan metode *deep learning* sebagai metode klasifikasi juga dapat digunakan peneliti kedepannya untuk menghasilkan kinerja model yang lebih baik.

#### DAFTAR PUSTAKA

- BHATIA, S., TOMAR, U. AND JAIN, A.V., 2021. Comparing SVM and Neural Networks' performance in Face Detection. In: *2021 International Conference on Intelligent Technologies (CONIT)*. [online] 2021 International Conference on Intelligent Technologies (CONIT). Hubli, India: IEEE. pp.1–7. <https://doi.org/10.1109/CONIT51480.2021.9498383>.
- GUPTA, A., D'CUNHA, A., AWASTHI, K. AND BALASUBRAMANIAN, V., 2016. *DAiSEE: Towards User Engagement Recognition in the Wild*. <https://doi.org/10.48550/arXiv.1609.01885>.
- KRITHIKA L.B AND LAKSHMI PRIYA GG, 2016. Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric. *Procedia Computer Science*, 85, pp.767–776. <https://doi.org/10.1016/j.procs.2016.05.264>.
- LEK, J.X.-Y. AND TEO, J., 2023. Academic Emotion Classification Using FER: A Systematic Review. *Human Behavior and Emerging Technologies*, 2023, pp.1–27. <https://doi.org/10.1155/2023/9790005>.
- LI, G., 2018. Special Treatment of Video Image Based on FFmpeg. In: *Proceedings of the 2018 Joint International Advanced Engineering and Technology Research Conference (JIAET 2018)*. [online] 2018 Joint International Advanced Engineering and Technology Research Conference (JIAET 2018). Xi'an, China: Atlantis Press. <https://doi.org/10.2991/jiaet-18.2018.47>.
- NURALIF, I., YUNIARNO, E.M., SUPRAPTO, Y.K. AND WICAKSONO, A.A., 2023. Driver Fatigue Detection Based On Face Mesh Features Using Deep Learning. In: *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. [online] 2023 International Seminar on Intelligent Technology and Its Applications (ISITIA). Surabaya, Indonesia: IEEE. pp.1–5. <https://doi.org/10.1109/ISITIA59021.2023.10221053>.
- NURRAHMA ROSANTI PAIDJA, A. AND BACHTIAR, F.A., 2022. Engagement Emotion Classification through Facial Landmark Using Convolutional Neural Network. In: *2022 2nd International Conference on Information Technology and Education (ICIT&E)*. [online] 2022 2nd International Conference on Information Technology and Education (ICIT&E). Malang, Indonesia: IEEE. pp.234–239. <https://doi.org/10.1109/ICITE54466.2022.9759546>.
- PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A. AND COURNAPEAU, D., n.d. Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.
- SIAM, A.I., SOLIMAN, N.F., ALGARNI, A.D., ABD EL-SAMIE, F.E. AND SEDIK, A., 2022. Deploying Machine Learning Techniques for Human Emotion Detection. *Computational Intelligence and Neuroscience*, 2022, pp.1–16. <https://doi.org/10.1155/2022/8032673>.
- SOLANKI, N. AND MANDAL, S., 2022. Engagement Analysis Using DAiSEE Dataset. In: *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. [online] 2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV). Singapore, Singapore: IEEE. pp.223–228. <https://doi.org/10.1109/ICARCV57592.2022.10004250>.
- SUN, Y., WONG, A.K.C. AND KAMEL, M.S., 2009. CLASSIFICATION OF IMBALANCED DATA: A REVIEW. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04), pp.687–719. <https://doi.org/10.1142/S0218001409007326>.
- THUSEETHAN, S., RAJASEGARAR, S. AND YEARWOOD, J., 2019. Detecting Micro-expression Intensity Changes from Videos Based on Hybrid Deep CNN. In: Q. Yang, Z.-H. Zhou, Z. Gong, M.-L. Zhang and S.-J. Huang, eds. *Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science*. [online] Cham: Springer International Publishing. pp.387–

399. [https://doi.org/10.1007/978-3-030-16142-2\\_30](https://doi.org/10.1007/978-3-030-16142-2_30).
- WHITEHILL, J., SERPELL, Z., LIN, Y.-C., FOSTER, A. AND MOVELLAN, J.R., 2014. The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions. *IEEE Transactions on Affective Computing*, 5(1), pp.86–98.
- <https://doi.org/10.1109/TAFFC.2014.2316163>.
- YUN, H., 2021. Prediction model of algal blooms using logistic regression and confusion matrix. *International Journal of Electrical and Computer Engineering (IJECE)*, 11(3), p.2407. <https://doi.org/10.11591/ijece.v11i3.pp2407-2413>.