

IMPLEMENTASI DEEP LEARNING MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK KLASIFIKASI GAYA BELAJAR SISWA BERDASARKAN EKSPRESI WAJAH SISWA SELAMA PROSES PEMBELAJARAN

¹⁾Sherly Dwi Puspita, ²⁾Sutardi, ³⁾L.M. Golok Jaya, ⁴⁾L.M. Bahtiar Aksara, ⁵⁾Andi Tenriawaru, ⁶⁾Isnawaty

^{1,2,3,4,5,6)}Jurusan Informatika, Fakultas Teknik, Universitas Halu Oleo

¹⁾sherlydwipuspita15@gmail.com, ²⁾sutardi@uho.ac.id, ³⁾laodemgj@uho.ac.id,

⁴⁾bahtiar.aksara@uho.ac.id, ⁵⁾andi.tenriawaru@uho.ac.id, ⁶⁾isnawaty@uho.ac.id

INFO ARTIKEL

Riwayat Artikel :

Diterima : 22 Juli 2026

Disetujui : 10 Agustus 2026

Kata Kunci :

Deep Learning, Convolutional Neural Network, Ekspresi Wajah, Gaya Belajar, VARK

ABSTRAK

Perbedaan gaya belajar siswa menjadi tantangan bagi guru dalam menentukan metode pembelajaran yang tepat, terutama pada kelas dengan jumlah siswa yang besar. Penelitian ini bertujuan mengembangkan sistem analisis respons afektif dan estimasi kecenderungan gaya belajar siswa secara otomatis berdasarkan ekspresi wajah selama proses pembelajaran tatap muka menggunakan metode *Convolutional Neural Network* (CNN). Data dikumpulkan melalui perekaman video pembelajaran di kelas XI.A4 SMA Negeri 1 Kendari dan diproses menggunakan Haar Cascade dan CSRT untuk menghasilkan 1.652 citra wajah yang terbagi ke dalam empat kelas ekspresi yaitu Tertarik, Fokus, Bingung, dan Bosan. Proses pelabelan divalidasi oleh dua anotator independen dengan nilai Cohen's Kappa ($\kappa = 0,8138$) dan Gwet's AC1 ($AC1 = 0,8323$) yang termasuk kategori *Almost Perfect Agreement*. Model CNN yang dibangun dengan arsitektur empat blok konvolusi bertingkat menghasilkan akurasi sebesar 94,76% pada data pengujian. Hasil klasifikasi ekspresi selanjutnya dipetakan ke kecenderungan gaya belajar berdasarkan model VARK melalui mekanisme *majority voting* dan diimplementasikan dalam sistem berbasis web bernama EduFace yang memungkinkan guru menganalisis kecenderungan gaya belajar siswa secara otomatis dari video pembelajaran tanpa memerlukan observasi manual.

ARTICLE INFO

Article History :

Received : Jul 22, 2026

Accepted : Aug 10, 2026

Keywords:

Deep Learning, Convolutional Neural Network, Facial Expression, Learning Styles, VARK

ABSTRACT

Differences in student learning styles pose a challenge for teachers in determining appropriate teaching methods, especially in classes with large numbers of students. This study aims to develop a system for automatically analyzing students' affective responses and estimating learning style tendencies based on facial expressions during face-to-face learning using the Convolutional Neural Network (CNN) method. Data were collected through learning video recordings in class XI.A4 of SMA Negeri 1 Kendari and processed using Haar Cascade and CSRT to produce 1,652 facial images divided into four expression classes: Interested, Focused, Confused, and Bored. The labeling process was validated by two independent annotators with Cohen's Kappa ($\kappa = 0.8138$) and Gwet's AC1 ($AC1 = 0.8323$) values, both categorized as Almost Perfect Agreement. The CNN model built with a four-tiered convolutional block architecture

achieved an accuracy of 94.76% on the test data. The expression classification results were then mapped to learning style tendencies based on the VARK model through a majority voting mechanism and implemented in a web-based system called EduFace, which enables teachers to automatically analyze students' learning style tendencies from learning videos without the need for manual observation.

1. PENDAHULUAN

Gaya belajar merupakan pendekatan khas yang digunakan setiap individu dalam menerima dan memproses informasi selama pembelajaran. Perbedaan gaya belajar berpengaruh langsung terhadap efektivitas pembelajaran, karena metode yang tidak sesuai dengan preferensi siswa dapat menurunkan fokus, motivasi, serta tingkat pemahaman (Fleming & Mills 1992). Dalam praktik pembelajaran tatap muka, guru sering menghadapi kesulitan dalam mengidentifikasi gaya belajar siswa secara akurat akibat jumlah siswa yang besar, keterbatasan waktu, dan sifat observasi manual yang subjektif. Kondisi ini dipertegas oleh berbagai penelitian yang menunjukkan bahwa identifikasi kebutuhan belajar siswa masih menjadi tantangan nyata dalam praktik pembelajaran di Indonesia (Musa & Kumalasari 2024; Suryani et al. 2024; Irawan 2025).

Seiring kemajuan teknologi, ekspresi wajah mulai dimanfaatkan sebagai indikator untuk menganalisis kondisi emosional dan kognitif siswa secara lebih objektif. Ekspresi wajah dapat mencerminkan tingkat keterlibatan, kebingungan, serta respons afektif yang berhubungan langsung dengan proses memahami materi. Dengan perkembangan *deep learning*, data ekspresi wajah kini dapat dianalisis secara otomatis menggunakan *Convolutional Neural Network* (CNN) yang mampu mengenali pola visual kompleks melalui ekstraksi fitur bertingkat, mulai dari tepi dan tekstur hingga perubahan kontur alis, mata, dan mulut yang merepresentasikan ekspresi tertentu. Oleh karena itu, diperlukan suatu sistem otomatis yang mampu mengidentifikasi ekspresi wajah siswa selama proses pembelajaran sebagai dasar untuk memetakan kecenderungan gaya belajar berdasarkan respon afektif siswa secara lebih objektif dan konsisten dibandingkan observasi manual.

Salah satu model yang banyak digunakan untuk mengidentifikasi preferensi belajar adalah

VARK (Visual, Auditory, Read/Write, dan Kinesthetic). Model ini memperkirakan kecenderungan siswa berdasarkan preferensi modalitas belajar mereka dalam menerima dan mengolah informasi, sehingga dapat membantu guru menentukan strategi pembelajaran yang lebih sesuai dengan karakteristik masing-masing siswa. Estimasi kecenderungan gaya belajar menggunakan model VARK telah banyak diterapkan karena mampu memberikan gambaran yang sistematis mengenai preferensi belajar siswa serta menjadi acuan dalam pengembangan pembelajaran yang lebih adaptif.

Meskipun demikian, penelitian sebelumnya di bidang ini masih banyak berfokus pada konteks pembelajaran daring dengan kondisi yang lebih terkontrol, sehingga belum sepenuhnya merepresentasikan dinamika pembelajaran tatap muka yang lebih kompleks dan natural (Gambo et al. 2020). Sejalan dengan hal tersebut, Tonguç & Ozaydin (2020) menunjukkan bahwa pengenalan emosi siswa secara otomatis dari ekspresi wajah selama pembelajaran memiliki potensi besar, namun masih menghadapi tantangan pada kondisi lingkungan kelas tatap muka yang tidak terkontrol. Pabba & Kumar (2022) juga mengembangkan sistem monitoring keterlibatan siswa di kelas besar menggunakan ekspresi wajah, namun belum mengintegrasikan hasil analisis tersebut dengan estimasi kecenderungan gaya belajar siswa. Gambo et al. (2020) telah mengusulkan pendekatan serupa dengan memetakan ekspresi wajah langsung ke kategori VARK menggunakan CNN, namun pendekatan tersebut menggunakan dataset yang dikonstruksi secara artifisial tanpa mempertimbangkan konteks gaya mengajar, sehingga akurasi yang diperoleh terbatas pada kisaran 50%-71% dan belum diujikan pada kondisi kelas nyata yang tidak terkontrol. Berdasarkan tinjauan sistematis yang dilakukan Lek & Teo (2023), sebagian besar penelitian klasifikasi emosi akademik masih berfokus pada konteks pembelajaran

daring, sehingga penelitian pada lingkungan tatap muka masih sangat terbatas. Selain itu, integrasi antara hasil klasifikasi ekspresi wajah dengan estimasi kecenderungan gaya belajar berdasarkan model VARK masih jarang dieksplorasi, khususnya pada konteks pembelajaran tatap muka di lingkungan kelas nyata yang tidak terkontrol.

Kondisi nyata di lapangan, seperti yang ditemukan pada pembelajaran fisika di kelas XI.A4 SMA Negeri 1 Kendari, menunjukkan bahwa masih terdapat siswa yang belum mencapai pemahaman optimal, yang mengindikasikan perlunya pendekatan yang lebih adaptif dalam mengenali karakteristik belajar siswa (Sasmita & Hartoyo 2020). Dengan demikian, masih terdapat peluang untuk mengembangkan sistem yang mampu menganalisis respons afektif siswa secara langsung dari rekaman video pembelajaran tatap muka di lingkungan kelas nyata sebagai dasar estimasi kecenderungan gaya belajar siswa.

Berdasarkan celah tersebut, penelitian ini bertujuan untuk mengembangkan sistem klasifikasi ekspresi wajah siswa menggunakan CNN pada rekaman video pembelajaran tatap muka, serta mengintegrasikan hasil klasifikasi dengan analisis kecenderungan gaya belajar berdasarkan model VARK. Perlu ditekankan bahwa dalam konteks penelitian ini, istilah klasifikasi gaya belajar merujuk pada proses estimasi kecenderungan gaya belajar siswa berdasarkan analisis respons afektif yang tercermin melalui ekspresi wajah, dan tidak dimaksudkan sebagai diagnosis psikologis yang bersifat absolut. Pendekatan ini diharapkan dapat membantu guru memahami gaya belajar siswa secara lebih objektif dan mendukung pelaksanaan pembelajaran yang lebih adaptif, personal, dan responsif terhadap kebutuhan siswa, sejalan dengan tujuan SDGs poin 4.5 tentang pendidikan yang inklusif dan berkeadilan.

a. Penelitian terdahulu

Tabel 1. Penelitian Terdahulu

Peneliti	Dataset	Metode	Jumlah Kelas	Performa	Kontribusi

Gambo et al. (2020)	ELSD(5000 sampel/dimensi)	CNN	4 kelas (VARK)	71,03% (Visual)	Deteksi gaya belajar VARK dari ekspresi wajah secara otomatis menggunakan CNN
Tonguç & Ozaydin Ozkara (2020)	Foto Webcam 67 mahasiswa	Microsoft Emotion API	7 emosi	N/A	Analisis perubahan emosi siswa selama kuliah tatap muka
Pabbala & Kumar (2022)	Video kelas + BAUM-1, DAiSEE, YawDD	CNN	6 kelas	76,90%	Monitoring keterlibatan siswa di kelas besar tatap muka
Gupta, Kumar & Tekchandani	FER-2013, CK+, RAF-BD	ResNet-50	2 kelas	92,32%	Deteksi keterlibatan pelajar berbasis ekspresi

(2023)					si wajah pada pembelajaran daring
Hu & Gao (2025a)	Video Pembelajaran Daring	Analisis Korelasi	N/A	N/A	Korelasi ekspresi wajah dengan keterlibatan siswa daring
Penelitian ini	Video Pembelajaran Tatap Muka	Haar Cascade + CSRT + CNN Custom	4 kelas	94,76%	Sistem web EduFace: klasifikasi ekspresi wajah dan estimasi kecenderungan gaya belajar VARK pada pembelajaran tatap muka

2. METODE

Penelitian ini menggunakan pendekatan eksperimen yang bertujuan untuk merancang sistem analisis kecenderungan gaya belajar

berdasarkan respon afektif siswa secara otomatis melalui analisis ekspresi wajah. Sistem yang dibangun mengintegrasikan tiga metode utama, yaitu *Haar Cascade* dan algoritma CSRT (*Channel and Spatial Reliability Tracking*) dari pustaka OpenCV untuk deteksi dan pelacakan wajah, CNN untuk klasifikasi ekspresi wajah, serta model VARK sebagai kerangka pemetaan kecenderungan gaya belajar siswa.

2.1 Pengumpulan Data

Pengumpulan data video pembelajaran telah mendapat izin dari pihak SMA Negeri 1 Kendari dan dilaksanakan dengan sepengetahuan guru serta siswa yang bersangkutan. Penelitian ini dilakukan sesuai dengan prosedur etik penelitian yang berlaku, di mana seluruh data visual siswa yang dikumpulkan hanya digunakan untuk kepentingan penelitian dan tidak disebarluaskan kepada pihak yang tidak berkepentingan.

Tahap awal penelitian diawali dengan kajian pustaka terhadap literatur yang relevan meliputi topik klasifikasi ekspresi wajah, gaya belajar siswa, dan penerapan *deep learning* dalam konteks pendidikan. Data primer diperoleh melalui perekaman langsung aktivitas belajar mengajar mata pelajaran fisika di kelas XI.A4 SMA Negeri 1 Kendari. Perekaman dilakukan menggunakan kamera yang diposisikan di dalam ruang kelas untuk menangkap ekspresi wajah siswa secara alami selama sesi pembelajaran. Perlu dicatat bahwa tidak seluruh siswa dapat terekam secara penuh dikarenakan keterbatasan jangkauan kamera, sehingga hanya sebagian siswa yang posisinya berada dalam area pandang kamera yang dapat diikutsertakan dalam proses pengambilan data.

Rekaman video yang diperoleh selanjutnya diproses melalui serangkaian tahapan otomatis meliputi ekstraksi frame per satuan waktu, identifikasi wajah menggunakan Haar Cascade, serta pelacakan wajah antar frame menggunakan algoritma *Channel and Spatial Reliability Tracker* (CSRT). Setelah melalui proses filtrasi kualitas yang meliputi pembuangan frame redundan, *frame blur*, dan frame dengan oklusi wajah, serta *augmentasi* data untuk mengatasi ketidakseimbangan distribusi kelas, diperoleh total dataset sebanyak 1.652 citra wajah berukuran 224×224 piksel.

Dataset tersebut kemudian dibagi menjadi tiga subset menggunakan *random split* yang tersedia pada *platform* Roboflow dengan proporsi 70% (1.156 citra) untuk pelatihan, 15% (248 citra) untuk validasi, dan 15% (248 citra) untuk pengujian model. Pembagian dilakukan pada tingkat citra individual secara acak tanpa mempertimbangkan identitas siswa, sehingga terdapat kemungkinan frame dari siswa yang sama terdistribusi ke dalam subset yang berbeda. Pembagian secara *subject-independent* tidak dapat diterapkan dalam penelitian ini dikarenakan keterbatasan jumlah subjek dalam dataset serta tidak semua siswa menampilkan keempat kelas ekspresi secara lengkap selama sesi perekaman, sehingga tidak memungkinkan untuk membagi data berdasarkan identitas siswa secara proporsional. Kondisi ini berpotensi menyebabkan akurasi evaluasi bersifat sedikit optimistis dan perlu diperhatikan dalam menginterpretasikan hasil penelitian ini.

2.2 Pelabelan Data

Setiap citra wajah yang dihasilkan diklasifikasikan ke dalam empat kategori ekspresi, yaitu Tertarik, Fokus, Bingung, dan Bosan, mengacu pada indikator visual yang dikembangkan berdasarkan *Facial Action Coding System* (FACS) (Ekman & Friesen 1978). Untuk menjamin objektivitas proses pelabelan, validasi dilakukan oleh dua penilai independen yakni guru mata pelajaran fisika dan guru Bimbingan Konseling SMA Negeri 1 Kendari. Konsistensi antar penilai diukur menggunakan dua metrik statistik yaitu Cohen's Kappa ($\kappa = 0,8138$) dan Gwet's AC1 ($AC1 = 0,8323$), keduanya berada pada kategori *Almost Perfect Agreement* (Cohen 1960; Wongpakaran et al. 2013).

2.3 Perancangan Model CNN

Convolutional Neural Network (CNN) merupakan salah satu arsitektur deep learning yang dirancang khusus untuk mengolah data berbentuk citra dengan cara mengekstraksi fitur secara hierarkis melalui operasi konvolusi (Govindwar et al. 2024). CNN terbukti efektif dalam berbagai tugas pengenalan pola visual termasuk klasifikasi ekspresi wajah karena kemampuannya dalam mempelajari fitur lokal seperti tepi, tekstur, hingga pola kompleks

secara otomatis tanpa memerlukan rekayasa fitur secara manual.

Arsitektur CNN yang dikembangkan dalam penelitian ini tersusun atas empat blok konvolusi bertingkat yang masing-masing dilengkapi *Batch Normalization* dan *Dropout* sebagai mekanisme regularisasi untuk mencegah *overfitting*. Representasi fitur spasial dari blok konvolusi selanjutnya diringkas melalui *Global Average Pooling* sebelum diteruskan ke lapisan klasifikasi dengan fungsi aktivasi *softmax* untuk menghasilkan probabilitas keempat kelas ekspresi. Proses pelatihan dikonfigurasi menggunakan *optimizer Adam* dengan *learning rate awal* 0,001 dan fungsi kerugian *categorical crossentropy*. Untuk menangani ketidakseimbangan distribusi kelas, diterapkan teknik augmentasi citra meliputi rotasi, pembalikan horizontal, dan variasi kecerahan.

2.4 Deteksi dan Pelacakan Wajah

Deteksi dan pelacakan wajah merupakan tahap krusial dalam pipeline sistem EduFace yang berfungsi sebagai jembatan antara input video mentah dengan proses klasifikasi ekspresi oleh model CNN. Setiap frame yang diekstraksi dari video terlebih dahulu dikonversi ke dalam ruang warna *grayscale* dan diproses menggunakan *Contrast Limited Adaptive Histogram Equalization* (CLAHE) untuk meningkatkan kontras citra secara adaptif, khususnya pada kondisi pencahayaan ruang kelas yang tidak merata.

Deteksi wajah dilakukan menggunakan algoritma *Haar Cascade Classifier* dengan parameter *scaleFactor* 1,4, *minNeighbors* 6, dan ukuran minimum wajah 80×80 piksel (Viola & Jones 2001). Setiap wajah yang berhasil dideteksi kemudian dipotong (*cropping*) dari frame asli dan diberikan identitas unik untuk membedakan wajah satu siswa dengan siswa lainnya. Untuk menghindari deteksi berulang pada wajah yang sama, diterapkan perhitungan *Intersection over Union* (IoU) dengan ambang batas 0,3 sebagai filter duplikasi antar wajah yang terdeteksi.

Pelacakan wajah antar frame dilakukan menggunakan algoritma CSRT (*Channel and Spatial Reliability Tracking*) yang diinisialisasi setiap kali wajah baru terdeteksi oleh Haar Cascade. Deteksi ulang dilakukan setiap 15

frame untuk memperbarui posisi wajah yang bergerak selama sesi pembelajaran berlangsung. Hasil crop wajah per siswa yang diperoleh dari proses pelacakan selanjutnya divalidasi menggunakan sistem skor berlapis yang mempertimbangkan variasi tekstur piksel, tingkat kecerahan, deteksi mata, dan rasio warna kulit sebelum diteruskan ke model CNN untuk proses klasifikasi ekspresi.

2.5 Pemetaan Gaya Belajar VARK

Model VARK merupakan kerangka klasifikasi gaya belajar yang dikembangkan oleh Fleming & Mills (1992) yang membagi kecenderungan gaya belajar individu ke dalam empat kategori berdasarkan modalitas sensorik yang paling dominan digunakan dalam menerima dan memproses informasi.

1. Kategori pertama adalah Visual, yaitu kecenderungan belajar melalui representasi grafis seperti diagram, grafik, peta konsep, dan media visual lainnya.
2. Kategori kedua adalah Auditory, yaitu kecenderungan belajar melalui pendengaran seperti penjelasan lisan, diskusi, dan tanya jawab.
3. Kategori ketiga adalah Read/Write, yaitu kecenderungan belajar melalui aktivitas membaca dan menulis seperti mencatat, membaca buku teks, dan merangkum materi.
4. Kategori keempat adalah Kinesthetic, yaitu kecenderungan belajar melalui pengalaman langsung dan praktik seperti eksperimen, simulasi, dan aktivitas fisik yang berkaitan dengan materi pembelajaran.

Dalam penelitian ini, pemetaan kecenderungan gaya belajar dilakukan berdasarkan kombinasi antara ekspresi dominan siswa dan konteks gaya mengajar yang diterapkan selama sesi pembelajaran. Ekspresi dominan setiap siswa ditentukan melalui mekanisme *majority voting* terhadap seluruh hasil klasifikasi ekspresi yang diperoleh sepanjang sesi, di mana ekspresi yang paling sering muncul ditetapkan sebagai ekspresi representatif siswa tersebut. Interpretasi ekspresi wajah terhadap keterlibatan siswa dalam pembelajaran dirangkum pada Tabel 2, sedangkan pemetaan kecenderungan gaya

belajar berdasarkan kombinasi gaya mengajar dan ekspresi negatif siswa ditunjukkan pada Tabel 3.

Tabel 2. Interpretasi Ekspresi Wajah

No	Ekspresi Wajah	Kategori Respon	Interpretasi Keterlibatan
1	Tertarik	Positif	Menunjukkan minat dan penerimaan terhadap metode pembelajaran.
2	Fokus	Positif	Menunjukkan perhatian dan konsentrasi tinggi.
3	Bingung	Negatif	Mengindikasikan kesulitan memahami materi.
4	Bosan	Negatif	Mengindikasikan rendahnya keterlibatan.

Berdasarkan kategori respon pada Tabel 2, apabila siswa secara dominan menunjukkan ekspresi negatif, kecenderungan gaya belajarnya dipetakan ke kategori VARK yang berbeda dari gaya mengajar yang sedang diterapkan sebagaimana ditunjukkan pada Tabel 3.

Tabel 3. Pemetaan Gaya Belajar Alternatif

No	Gaya Mengajar	Ekspresi Wajah	Indikasi VARK
1	Visual	Bingung	Auditory
2	Visual	Bosan	Kinesthetic
3	Auditory	Bingung	Visual
4	Auditory	Bosan	Kinesthetic
5	Read/Write	Bingung	Auditory
6	Read/Write	Bosan	Kinesthetic
7	Kinesthetic	Bingung	Auditory
8	Kinesthetic	Bosan	Read/Write

Berdasarkan Tabel 3, apabila ekspresi dominan siswa termasuk dalam kategori negatif, kecenderungan gaya belajar siswa dipetakan ke kategori VARK yang berbeda dari gaya mengajar yang diterapkan. Sebagai contoh, siswa yang secara dominan menunjukkan ekspresi Bingung saat guru menggunakan pendekatan Visual diindikasikan memiliki kecenderungan gaya belajar Auditory, yang berarti siswa tersebut lebih mudah memahami materi melalui penjelasan lisan dibandingkan melalui representasi visual. Perlu ditekankan

bahwa pemetaan ini tidak dimaksudkan sebagai klasifikasi psikologi yang bersifat absolut, melainkan sebagai aturan inferensi operasional yang dirancang berdasarkan sintesis konsep gaya belajar VARK dan indikator afektif yang tercermin melalui ekspresi wajah siswa. Penelitian di bidang *affective computing* menunjukkan bahwa ekspresi wajah dapat merepresentasikan tingkat keterlibatan, kebingungan, maupun kebosanan siswa selama proses belajar (Whitehill et al. 2014; Hu & Gao 2025b), sehingga informasi afektif tersebut dapat dimanfaatkan sebagai dasar untuk memperkirakan kecenderungan gaya belajar siswa. Hasil pemetaan ini selanjutnya digunakan sebagai dasar penyajian informasi kecenderungan gaya belajar per siswa pada antarmuka sistem EduFace.

2.6 Implementasi Sistem

Sistem EduFace diwujudkan dalam bentuk aplikasi web menggunakan *framework* Flask berbasis Python yang memungkinkan guru mengunggah video pembelajaran dan memperoleh hasil analisis kecenderungan gaya belajar siswa secara otomatis. Sistem dirancang tanpa mekanisme penyimpanan data permanen, sehingga video yang diunggah hanya diproses sementara di memori sistem dan dihapus secara otomatis setelah analisis selesai guna menjaga kerahasiaan data visual siswa.

Alur kerja sistem dimulai ketika pengguna mengunggah video pembelajaran melalui antarmuka web dan memilih konteks gaya mengajar yang diterapkan selama sesi tersebut. Video kemudian diproses secara otomatis di *background thread* agar antarmuka web tetap responsif selama proses berlangsung. Tahapan pemrosesan meliputi ekstraksi frame dengan mekanisme skip frame untuk efisiensi komputasi, deteksi dan pelacakan wajah menggunakan Haar Cascade dan CSRT, klasifikasi ekspresi menggunakan model CNN yang telah dilatih, rekapitulasi ekspresi per siswa melalui *majority voting*, dan pemetaan kecenderungan gaya belajar berdasarkan model VARK. Selama proses berlangsung, sistem menampilkan progress bar secara *real-time* yang diperbarui setiap dua detik melalui mekanisme polling untuk memberikan umpan balik visual kepada pengguna mengenai progres analisis.

2.7 Evaluasi Model

Evaluasi performa model CNN dilakukan menggunakan data pengujian sebesar 15% dari total dataset yang tidak pernah digunakan selama proses pelatihan. Metrik evaluasi yang digunakan meliputi akurasi, *precision*, *recall*, dan *F1-score* untuk mengukur kemampuan model dalam mengklasifikasikan keempat kelas ekspresi secara menyeluruh. Selain itu, *confusion matrix* digunakan untuk menganalisis distribusi hasil prediksi model terhadap label sebenarnya sehingga dapat diidentifikasi kelas ekspresi mana yang paling sering mengalami kesalahan klasifikasi.

3. HASIL DAN PEMBAHASAN

Berikut adalah hasil implementasi sistem klasifikasi kecenderungan gaya belajar siswa berbasis ekspresi wajah menggunakan metode *Convolutional Neural Network* (CNN) yang dikombinasikan dengan model VARK, yang dikembangkan berdasarkan tahapan-tahapan yang telah dijelaskan pada metode penelitian.

3.1 Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari citra wajah siswa yang merepresentasikan empat kelas ekspresi selama proses pembelajaran berlangsung. Setiap kelas memuat sejumlah citra wajah dengan karakteristik visual yang bervariasi, mencakup perbedaan individu siswa, sudut pengambilan gambar, dan kondisi pencahayaan ruang kelas. Penggunaan empat kelas ekspresi ini bertujuan untuk mengukur kemampuan model CNN dalam mengenali variasi ekspresi wajah yang relevan dengan kondisi afektif siswa selama pembelajaran. Adapun dataset pada penelitian ini terbagi atas empat kelas yaitu Tertarik, Fokus, Bingung, dan Bosan.

Tabel 4. Sebaran Dataset

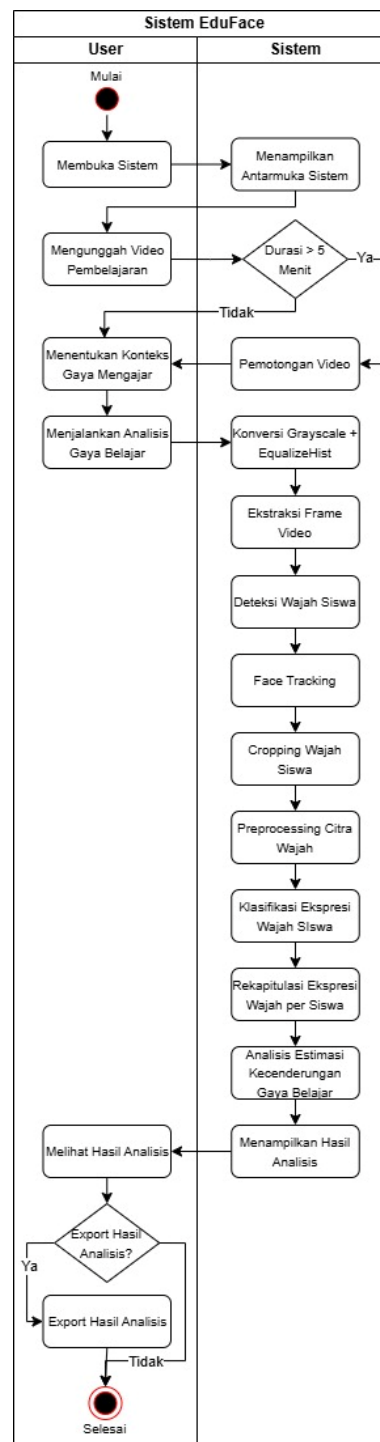
No	Ekspresi	Gambar
1	Tertarik	

2	Fokus	
3	Bingung	
4	Bosan	

3.2 Unified Modelling Language (UML)

Unified Modeling Language (UML) merupakan bahasa pemodelan standar yang digunakan untuk memvisualisasikan, merancang, dan mendokumentasikan sistem perangkat lunak secara terstruktur. Salah satu diagram dalam UML yang digunakan dalam penelitian ini adalah *activity diagram*, yaitu diagram yang menggambarkan alur aktivitas dan interaksi antara pengguna dan sistem secara sekuensial untuk merepresentasikan proses bisnis atau alur kerja suatu sistem.

Dalam penelitian ini, *activity diagram* digunakan untuk mengilustrasikan alur kerja sistem EduFace secara menyeluruh, mulai dari tahap pengguna membuka sistem, mengunggah video pembelajaran, menentukan konteks gaya mengajar, hingga sistem menghasilkan dan menampilkan hasil analisis kecenderungan gaya belajar siswa. Berikut ditampilkan *activity diagram* sistem klasifikasi gaya belajar yang menjadi acuan implementasi dalam penelitian ini.



Gambar 1. Activity Diagram Sistem

Diagram aktivitas pada gambar menunjukkan alur kerja sistem klasifikasi gaya belajar siswa berdasarkan analisis ekspresi wajah dari video pembelajaran, yang melibatkan interaksi antara *user* (pengguna/guru) dan *sistem*. Proses dimulai ketika pengguna membuka sistem, kemudian sistem menampilkan antarmuka sistem kepada pengguna.

Selanjutnya, pengguna mengunggah video pembelajaran ke dalam sistem. Sistem kemudian

memeriksa durasi video tersebut. Jika durasi video lebih dari 5 menit, sistem akan melakukan pemotongan video terlebih dahulu sebelum melanjutkan proses. Namun, jika durasi video tidak lebih dari 5 menit, proses langsung dilanjutkan tanpa pemotongan.

Setelah durasi video sesuai, pengguna menentukan konteks gaya mengajar yang digunakan, kemudian menjalankan proses analisis gaya belajar. Sistem selanjutnya memproses video melalui beberapa tahapan berurutan, yaitu konversi citra ke grayscale sekaligus *equalize histogram*, ekstraksi frame dari video, deteksi wajah siswa, *face tracking*, serta *cropping* area wajah siswa yang terdeteksi.

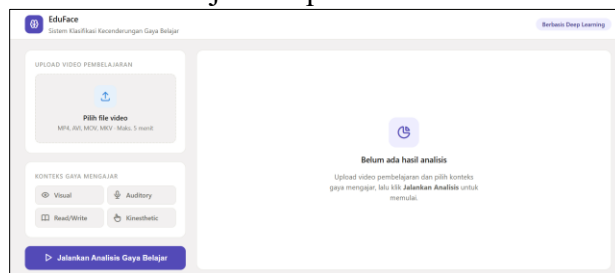
Citra wajah yang telah dipotong kemudian melalui tahap *preprocessing* citra wajah, dilanjutkan dengan klasifikasi ekspresi wajah siswa. Hasil klasifikasi ekspresi wajah dari setiap siswa kemudian direkapitulasi, dan sistem melakukan analisis kecenderungan gaya belajar berdasarkan rekapitulasi tersebut.

Setelah proses analisis selesai, sistem menampilkan hasil analisis kepada pengguna, yang kemudian dapat melihat hasil analisis tersebut. Pengguna selanjutnya diberikan pilihan untuk melakukan *export* hasil analisis. Jika pengguna memilih Ya, sistem akan melakukan proses *export* hasil analisis, dan alur kembali ke tahap pilihan *export* (memungkinkan pengguna melakukan *export* berulang bila diperlukan). Jika pengguna memilih Tidak, maka proses langsung berakhir. Alur diagram ini diakhiri setelah tahap *export* (atau tidak) selesai dilakukan.

3.3 Implementasi Sistem

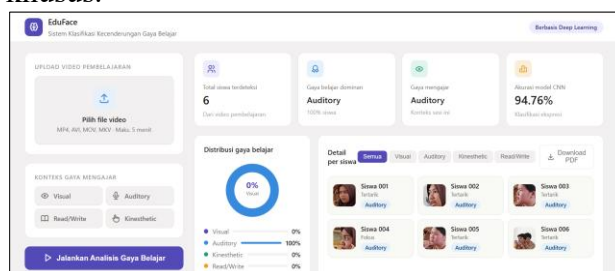
Pada penelitian ini telah dirancang dan dibangun sebuah sistem EduFace berbasis web statis. Sistem ini digunakan untuk menganalisis kecenderungan gaya belajar siswa berdasarkan ekspresi wajah selama proses pembelajaran menggunakan metode *Convolutional Neural Network* sebagai model klasifikasi ekspresi wajah dan model VARK sebagai dasar pemetaan gaya belajar. Sistem dirancang tanpa basis data permanen, sehingga video yang diunggah hanya diproses sementara di memori sistem dan dihapus secara otomatis setelah analisis selesai guna meminimalkan risiko kebocoran data

visual siswa. Tampilan antarmuka yang dihasilkan ditunjukkan pada Gambar berikut.



Gambar 2. Tampilan Dashboard Sistem

Gambar 2 menampilkan antarmuka utama sistem EduFace yang menggunakan konsep *single-page interface* dengan tema *light mode* untuk mendukung keterbacaan hasil analisis. Antarmuka ini menggabungkan seluruh modul fungsional dalam satu layar, terdiri dari header berisi identitas sistem, panel input di sisi kiri untuk mengunggah video pembelajaran dan memilih konteks gaya mengajar berdasarkan empat kategori model VARK (*Visual*, *Auditory*, *Read/Write*, *Kinesthetic*), serta panel output di sisi kanan yang menampilkan *empty state* sebelum analisis dijalankan. Susunan antarmuka yang hierarkis dan intuitif memungkinkan pengguna menjalankan proses analisis secara mandiri tanpa memerlukan pelatihan teknis khusus.



Gambar 3. Tampilan Hasil Analisis

Gambar 3 menampilkan antarmuka hasil analisis gaya belajar setelah proses klasifikasi selesai dijalankan. Panel ringkasan di bagian atas menyajikan empat indikator utama, yaitu total siswa yang terdeteksi, gaya belajar dominan pada kelas, konteks gaya mengajar yang digunakan, serta akurasi model CNN dalam melakukan klasifikasi ekspresi wajah. Di bawahnya, sistem menampilkan diagram distribusi gaya belajar dalam bentuk *donut chart* yang merepresentasikan persentase masing-masing kategori model VARK (*Visual*, *Auditory*, *Read/Write*, *Kinesthetic*) pada keseluruhan siswa.

Pada sisi kanan, terdapat panel Detail per Siswa yang menampilkan hasil klasifikasi gaya belajar untuk setiap siswa secara individual, lengkap dengan foto wajah, label ekspresi dominan, dan kategori gaya belajar yang terdeteksi. Panel ini dilengkapi filter berdasarkan kategori VARK untuk memudahkan penelusuran data, serta fitur unduh hasil analisis dalam format PDF sebagai dokumentasi bagi guru. Penyajian informasi secara ringkas dan visual ini memudahkan guru dalam menginterpretasikan kecenderungan gaya belajar siswa secara cepat dan menyeluruh.

Berdasarkan hasil analisis yang ditampilkan pada Gambar 3, dari total 11 siswa yang berhasil terdeteksi, sebesar 82% menunjukkan kecenderungan gaya belajar Kinesthetic, diikuti Auditory dan Read/Write masing-masing sebesar 9%, serta Visual sebesar 0%. Hasil ini diperoleh dari sesi pembelajaran dengan konteks gaya mengajar Kinesthetic, sehingga tingginya persentase gaya belajar Kinesthetic mengindikasikan bahwa sebagian besar siswa menunjukkan respons positif berupa ekspresi Tertarik atau Fokus terhadap pendekatan pembelajaran praktik yang diterapkan, yang sesuai dengan pemetaan gaya belajar berdasarkan model VARK.

3.4 Hasil Evaluasi Klasifikasi

Gambar berikut menunjukkan hasil evaluasi klasifikasi dari model CNN yang digunakan untuk mengklasifikasikan ekspresi wajah siswa ke dalam empat kelas berdasarkan metrik akurasi *precision*, *recall*, *f1-score* dan *support*.

```

==== Classification Report ====

```

	precision	recall	f1-score	support
Tertarik	0.7805	1.0000	0.8767	32
Fokus	1.0000	0.9178	0.9571	146
Bingung	0.9756	1.0000	0.9877	40
Bosan	0.9062	0.9667	0.9355	30
accuracy			0.9476	248
macro avg	0.9156	0.9711	0.9392	248
weighted avg	0.9564	0.9476	0.9491	248

Gambar 4. Hasil Evaluasi Klasifikasi

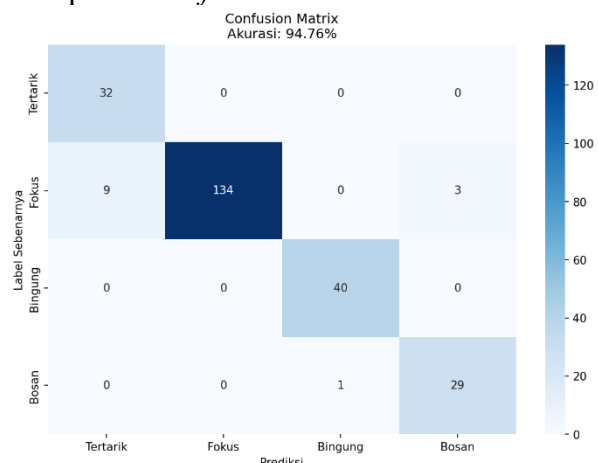
Berdasarkan hasil evaluasi yang ditampilkan pada Gambar 4 model menghasilkan akurasi sebesar 94,76%, dengan *precision* rata-rata (macro avg) 91,56%, *recall* rata-rata 97,11%, dan *F1-score* rata-rata 93,92% pada data uji sebanyak 248 sampel. Secara rinci, kelas Bingung memperoleh performa terbaik dengan *precision* 97,56%, *recall* 100%, dan *F1-*

score 98,77%, menunjukkan model sangat akurat dalam mengenali ekspresi ini. Kelas Fokus juga menunjukkan hasil yang tinggi dengan *precision* 100%, *recall* 91,78%, dan *F1-score* 95,71%. Kelas Bosan memperoleh *precision* 90,62%, *recall* 96,67%, dan *F1-score* 93,55%.

Sementara itu, kelas Tertarik memiliki *recall* sempurna (100%) namun *precision* paling rendah di antara kelas lain (78,05%), yang mengindikasikan model cenderung salah mengklasifikasikan ekspresi lain sebagai "Tertarik" (false positive), meskipun model berhasil mengenali seluruh sampel Tertarik yang sebenarnya. Secara keseluruhan, nilai-nilai ini menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan ekspresi wajah siswa, dengan performa terbaik pada kelas Bingung dan ruang perbaikan pada kelas Tertarik.

3.5 Confusion Matrix

Confusion matrix merupakan metode yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi model terhadap label kelas sebenarnya pada data uji. Melalui matrix ini, dapat diketahui jumlah prediksi yang benar maupun yang keliru untuk setiap kelas, sehingga kesalahan klasifikasi antar kelas dapat diidentifikasi secara lebih spesifik. Untuk memvisualisasikan distribusi prediksi pada masing-masing kelas, ditampilkan *confusion matrix* berikut:



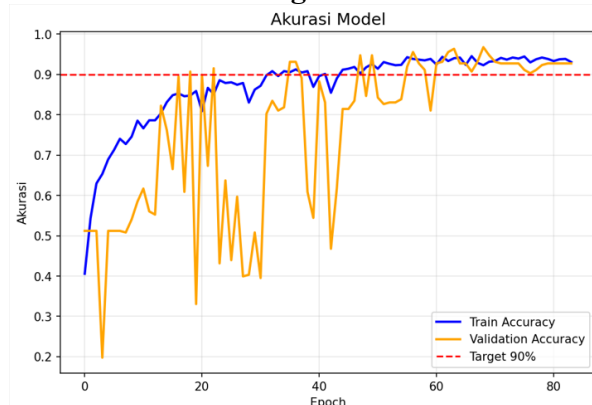
Gambar 5. Hasil *Confusion Matrix*

Confusion matrix pada Gambar 5 menunjukkan hasil evaluasi model klasifikasi ekspresi wajah dengan akurasi keseluruhan sebesar 94,76%. Pada kelas Tertarik, seluruh 32

sampel berhasil diklasifikasikan dengan benar tanpa ada kesalahan. Kelas Fokus memiliki 134 sampel yang diklasifikasikan dengan tepat, namun 9 sampel salah diprediksi sebagai Tertarik dan 3 sampel salah diprediksi sebagai Bosan. Kelas Bingung menunjukkan hasil sempurna dengan seluruh 40 sampel diklasifikasikan secara benar. Sementara itu, kelas Bosan memiliki 29 sampel yang diprediksi dengan tepat, dengan 1 sampel salah diklasifikasikan sebagai Bingung.

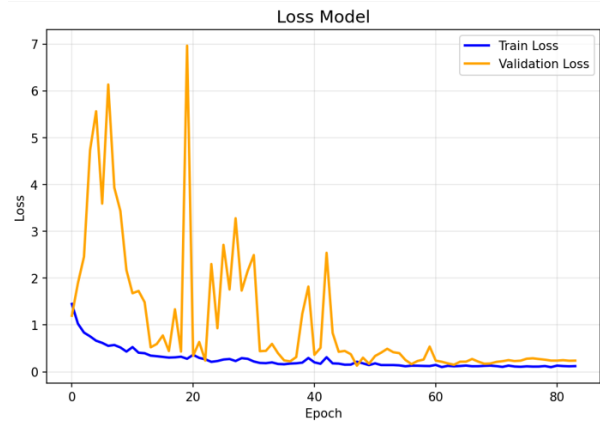
Secara keseluruhan, distribusi nilai pada diagonal utama matrix yang dominan menunjukkan bahwa model mampu mengklasifikasikan sebagian besar ekspresi wajah dengan tepat pada masing-masing kelas, dengan kesalahan klasifikasi (*misclassification*) yang relatif kecil dan terkonsentrasi pada kelas Fokus terhadap Tertarik dan Bosan.

3.6 Grafik Training



Gambar 6. Grafik Akurasi Model

Gambar 6 menunjukkan kurva akurasi model selama proses pelatihan (*training*) dan validasi. *Train accuracy* meningkat secara stabil dan konsisten hingga mencapai kisaran 93-95%, sementara *validation accuracy* mengalami fluktuasi cukup tinggi pada epoch awal (0-50) sebelum akhirnya konvergen mendekati nilai *train accuracy* di atas epoch ke-60. Konvergensi antara kedua kurva tanpa *gap* yang melebar ini mengindikasikan bahwa model mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat sebelumnya. Fluktuasi tajam pada *validation accuracy* di epoch-epoch awal merupakan hal yang wajar terjadi, terutama disebabkan oleh ukuran data validasi yang relatif kecil serta *learning rate* yang masih tinggi pada tahap awal pelatihan.



Gambar 7. Grafik Loss Model

Gambar 7 menunjukkan kurva *loss* model selama proses pelatihan dan validasi. *Train loss* menurun secara konsisten, sedangkan *validation loss* menunjukkan lonjakan tajam pada beberapa epoch awal sebelum akhirnya stabil dan konvergen bersama *train loss* menuju nilai yang rendah ($<0,2$) pada epoch akhir. Konvergensi antara *train* dan *validation loss* menuju nilai yang sama tanpa *gap* yang melebar ini mengindikasikan bahwa model tidak mengalami *overfitting*, karena ciri utama *overfitting* adalah *training loss* yang terus menurun sementara *validation loss* menstabil atau justru meningkat. Sebaliknya, pada model yang baik (*just-right*), fungsi *loss* pada *training* dan validasi cenderung mengikuti satu sama lain secara berdekatan dan konvergen menuju nilai yang kecil, sebagaimana ditunjukkan pada grafik ini setelah epoch ke-60. Lonjakan sesekali pada *validation loss* di awal training dapat terjadi akibat data yang *noisy* atau *learning rate* yang tinggi, dan hal ini merupakan karakteristik yang umum ditemui pada proses pelatihan model *deep learning*.

Untuk memvalidasi kecenderungan gaya belajar yang dihasilkan sistem, hasil analisis EduFace dibandingkan dengan pengelompokan gaya belajar siswa yang telah ditetapkan sebelumnya oleh guru mata pelajaran serta hasil kuesioner yang telah diisi oleh siswa. Instrumen kuesioner dalam penelitian ini diadaptasi dari konsep VARK yang dikembangkan oleh Fleming (2001) dan disesuaikan dengan konteks pembelajaran yang diamati. Hasil perbandingan menunjukkan tingkat kesesuaian sekitar 80% atau 4 dari 5 siswa yang terdeteksi memiliki kesesuaian kecenderungan gaya belajar antara estimasi sistem dengan hasil kuesioner dan pengelompokan dari guru mata pelajaran. Hal ini

mengindikasikan bahwa sistem EduFace mampu memperkirakan kecenderungan gaya belajar siswa dengan cukup baik melalui analisis ekspresi wajah selama proses pembelajaran berlangsung. Berbeda dengan Gambo et al. (2020) yang mengklasifikasikan ekspresi wajah secara langsung ke kategori VARK dengan akurasi berkisar 50%-71% menggunakan dataset yang dikonstruksi secara artifisial, penelitian ini memisahkan tahapan klasifikasi ekspresi wajah dengan tahapan pemetaan ke model VARK yang mempertimbangkan konteks gaya mengajar, sehingga menghasilkan pendekatan yang lebih kontekstual dan terukur.

Meskipun tingkat kesesuaian sistem mencapai 80%, terdapat kecenderungan model untuk salah mengklasifikasikan ekspresi Fokus sebagai Tertarik, sebagaimana terlihat pada hasil confusion matrix. Hal ini disebabkan oleh kemiripan visual yang tinggi antara kedua ekspresi tersebut, di mana keduanya sama-sama menampilkan fitur wajah positif seperti mata terbuka terarah dan postur kepala yang stabil, sehingga model CNN kesulitan membedakan keduanya secara konsisten. Dalam konteks pemetaan VARK, kesalahan klasifikasi antara Fokus dan Tertarik tidak berdampak signifikan terhadap hasil estimasi kecenderungan gaya belajar karena keduanya termasuk dalam kategori respons positif yang menghasilkan pemetaan VARK yang sama sesuai dengan konteks gaya mengajar yang diterapkan. Temuan dari kuesioner juga menunjukkan bahwa hampir seluruh siswa menyatakan bahwa cara guru menjelaskan materi sangat mempengaruhi tingkat pemahaman mereka, yang mengindikasikan bahwa kecenderungan gaya belajar siswa tidak dapat dipisahkan dari gaya mengajar yang diterapkan. Dengan demikian, kesesuaian antara gaya belajar siswa dan gaya mengajar guru merupakan faktor penting dalam menentukan efektivitas proses pembelajaran, dan tidak semata-mata ditentukan oleh preferensi belajar siswa secara individual.

4. PENUTUP

4.1. Kesimpulan

Penelitian ini berhasil mengembangkan sistem analisis respons afektif dan estimasi kecenderungan gaya belajar siswa berbasis ekspresi wajah menggunakan metode

Convolutional Neural Network (CNN) yang diimplementasikan dalam bentuk aplikasi web EduFace dengan akurasi 94,76% pada data pengujian. Sistem memproses video pembelajaran secara otomatis melalui tahapan deteksi dan pelacakan wajah, klasifikasi ekspresi, rekapitulasi ekspresi dominan melalui mekanisme *majority voting*, serta pemetaan kecenderungan gaya belajar berdasarkan model VARK, sehingga guru dapat memperkirakan kecenderungan gaya belajar siswa secara lebih objektif tanpa memerlukan observasi manual yang memakan waktu.

4.2. Saran

Komponen deteksi wajah menggunakan algoritma Haar Cascade masih memiliki keterbatasan dalam kondisi kelas nyata yang tidak terkontrol sehingga tidak seluruh siswa dapat terdeteksi secara konsisten. Selain itu, pembagian dataset yang dilakukan pada tingkat citra individual berpotensi menyebabkan akurasi evaluasi bersifat sedikit optimistis. Untuk penelitian selanjutnya disarankan penggunaan algoritma deteksi wajah berbasis *deep learning* seperti MTCNN atau RetinaFace untuk meningkatkan cakupan deteksi wajah dalam kondisi lingkungan kelas yang beragam, perluasan dataset dengan variasi ekspresi dan jumlah subjek yang lebih beragam guna meningkatkan kemampuan generalisasi model, penerapan pembagian dataset berbasis identitas subjek (*subject-independent split*) untuk memperoleh evaluasi yang lebih ketat, serta pengembangan mekanisme validasi kecenderungan gaya belajar yang lebih komprehensif melalui integrasi dengan instrumen psikometrik yang terstandarisasi.

5. DAFTAR PUSTAKA

- Cohen, J., 1960, 'A Coefficient of Agreement for Nominal Scales', *Educational and Psychological Measurement*, 20(1), 37–46.
- Ekman, P. & Friesen, W. V., 1978, *Facial Action Coding System: Manual*.
- Fleming, N.D. & Mills, C., 1992, 'Not Another Inventory, Rather a Catalyst for Reflection', *To Improve the Academy*, 11, 137–155.
- Gambo, F.L., Wajiga, G.M., Garba, E.J. & Abdullahi, A.A., 2020, 'Prediction of

- Learning Style Using Facial Expressions with Convolutional Neural Network Special Issue on Computing and Communication Technologies Prediction of Learning Style Using Facial Expressions with Convolutional Neural Network', *International Journal of Information Processing and Communication (IJIPC)*, 9(2), 389–399.
- Govindwar, V., Aman Akbani, Aachal Nandeshwar, Aishwarya Wanjari, Prachi Agashe & Charan Pote, 2024, 'Facial Expression Recognition Using Convolutional Neural Network', *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(2), 786–794.
- Gupta, S., Kumar, P. & Tekchandani, R.K., 2023, 'Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models', *Multimedia Tools and Applications*, 82(8), 11365–11394.
- Hu, X. & Gao, J., 2025, *Facial expression recognition reveals students' engagement in online class: Correlations with six engagement measurements*, *PLOS One*, 1–22, Public Library of Science.
- Irawan, Y., 2025, 'Analisis Kesulitan Guru dalam Menerapkan Pembelajaran Berdiferensiasi di Kelas Inklusif', *Makarimul Ilmi: Jurnal Pengembangan Madrasah Ibtidayah*, 02(02), 97–108.
- Lek, J.X.-Y. & Teo, J., 2023, 'Academic Emotion Classification Using FER: A Systematic Review', *Human Behavior and Emerging Technologies*, 2023, 1–27.
- Musa, L. & Kumalasari, D., 2024, 'Implementasi Pembelajaran Berdiferensiasi dalam Upaya Pemenuhan Kebutuhan Belajar Murid pada Mata Pelajaran IPS', *Ideguru: Jurnal Karya Ilmiah Guru*, 9(2), 809–814.
- Pabba, C. & Kumar, P., 2022, 'An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition', *Expert Systems*, 39(1).
- Sasmita, P.R. & Hartoyo, Z., 2020, 'Pengaruh Pendekatan Pembelajaran STEM Project Based Learning terhadap Pemahaman Konsep Fisika Siswa', *SILAMPARI Jurnal Pendidikan Ilmu Fisika*, 2(2), 136–148.
- Suryani, E.A., Siswanto, D.H., Pisiwati, S.A. & History, A., 2024, 'Strengthening Teacher Competence through Differentiated Instruction Training as an Implementation of the Merdeka Curriculum Article Info ABSTRACT', *JOELI: Journal of Educational and Learning Innovation*, 1(2), 3063–3966.
- Tonguç, G. & Ozaydın Ozkara, B., 2020, 'Automatic recognition of student emotions from facial expressions during a lecture', *Computers & Education*, 148, 103797.
- Viola, P. & Jones, M., 2001, *Rapid object detection using a boosted cascade of simple features*, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 1-511-1-518, IEEE Comput. Soc.
- Whitehill, J., Serpell, Z., Lin, Y.-C., Foster, A. & Movellan, J.R., 2014, 'The Faces of Engagement Automatic Recognition of Student Engagement from Facial Expressions', *IEEE Transaction on Affective Computing*, 5(1), 86–98.
- Wongpakaran, N., Wongpakaran, T., Wedding, D. & Gwet, K.L., 2013, 'A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples', *BMC Medical Research Methodology*, 13(1), 61.