

**MENDETEKSI TINGKAT KETERLIBATAN SISWA PADA
PEMBELAJARAN DARING BERDASARKAN
EMOSI WAJAH DAN SIKAP SISWA
(STUDI KASUS: INSTITUT PEMERINTAHAN DALAM
NEGERI)**

TESIS

**Karya tulis sebagai salah satu syarat
kelulusan MK IF5099 Metodologi Penelitian/Tesis 1**

**Oleh
ARI APRIANSYAH
NIM: 23523301
(Program Studi Magister Informatika)**



**INSTITUT TEKNOLOGI BANDUNG
Maret 2025**

**MENDETEKSI TINGKAT KETERLIBATAN SISWA PADA
PEMBELAJARAN DARING BERDASARKAN
EMOSI WAJAH DAN SIKAP SISWA
(STUDI KASUS: INSTITUT PEMERINTAHAN DALAM
NEGERI)**

Oleh
Ari Apriansyah
NIM: 23523301
(Program Studi Magister Informatika)

Institut Teknologi Bandung

Menyetujui
Tim Pembimbing

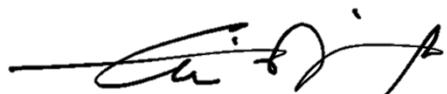
Tanggal 24 Maret 2025

Ketua



Dr. Ir. Kusprasapta Mutijarsa, S.T., M.T.

Anggota



Dr. Fadhil Hidayat, S.Kom., M.T.

DAFTAR ISI

HALAMAN PENGESAHAN	i
DAFTAR ISI.....	ii
DAFTAR LAMPIRAN.....	iii
DAFTAR GAMBAR DAN ILUSTRASI	iv
DAFTAR TABEL.....	v
DAFTAR SINGKATAN DAN LAMBANG	vi
BAB I Pendahuluan.....	1
I.1 Latar Belakang	1
I.2 Rumusan Masalah	6
I.3 Tujuan Penelitian.....	6
I.4 Batasan Penelitian	7
I.5 Kontribusi Penelitian.....	7
I.6 Sistematika Penulisan.....	8
BAB II Tinjauan Pustaka.....	10
II.1 <i>Learning Management System</i>	10
II.2 <i>Affective Computing</i>	11
II.3 <i>Automated Emotion Recognition</i> di Lingkungan Pendidikan	11
II.3.1 Modalitas Teks.....	13
II.3.2 Modalitas Audio	16
II.3.3 Modalitas Sinyal Tubuh.....	19
II.3.4 Modalitas Visual	22
II.3.5 Multimodal.....	26
II.4 <i>Facial Expressions Recognition</i> di Lingkungan Pendidikan Daring	27
II.5 Oklusi pada Pengenalan Ekspresi Wajah	29
II.6 Klasifikasi Emosi dalam Deteksi <i>Engagement</i> Siswa	31
II.6.1 <i>Machine Learning</i> (ML)	32
II.6.2 <i>Deep Learning</i> (DL)	32
II.6.3 Kombinasi <i>Machine Learning</i> Klasik dan <i>Deep Learning</i> ..	34
II.7 <i>Knowledge Discovery in Databases</i>	35
II.8 Peta Literatur	36
BAB III Metodologi Penelitian.....	37
III.1 Identifikasi Masalah dan Motivasi	37
III.2 Definisi Tujuan dan Solusi	39
III.3 Perancangan dan Pengembangan	42
III.4 Demonstrasi.....	44
III.5 Evaluasi	44
III.6 Komunikasi	45
BAB IV Analisis dan Desain Awal.....	46
IV.1 Lingkungan Penerapan	46
IV.2 Analisis Wajah	46
DAFTAR PUSTAKA	50
LAMPIRAN	57

DAFTAR LAMPIRAN

Lampiran A Potongan Paparan Konsep Perubahan RENIP IPDN 2020-2045 57

DAFTAR GAMBAR DAN ILUSTRASI

Gambar I.1	Contoh kondisi wajah yang tertutup sebagian oleh tangan pada saat pembelajaran daring (gambar dibuat dengan AI)	5
Gambar II.1	Siklus penerapan AER pada lingkungan pendidikan (Yu dkk., 2024)	12
Gambar II.2	Deteksi <i>engagement</i> siswa menggunakan FER, <i>eye blink</i> , dan <i>head pose</i> (Gupta dkk., 2023a)	24
Gambar II.3	Alur pengenalan emosi dengan FER digabungkan dengan analisis sentimen (Cui dkk., 2021)	26
Gambar II.5	Proses utama <i>facial expressions recognition</i>	28
Gambar II.6	(a) Keadaan wajah dengan oklusi <i>partial</i> oleh tangan, (b) keadaan wajah dengan oklusi total (Jain & Learned-Miller, 2010)	30
Gambar II.7	Gambaran umum alur <i>machine learning</i>	32
Gambar II.8	Distribusi penggunaan metode <i>machine learning</i> dan <i>deep learning</i> dalam rentang tahun 2010 – 2024 (Yu dkk., 2024)	34
Gambar II.9	Alur metode <i>knowledge discovery in databases</i>	35
Gambar II.10	Peta literatur deteksi emosi siswa pada media pembelajaran daring	36
Gambar II.11	Peta literatur tentang cara mengatasi masalah oklusi pada FER ...	36
Gambar III.1	<i>Design Science Research Methodology</i>	37
Gambar III.2	Gambaran umum ajuan solusi	40
Gambar III.3	Rancangan umum ajuan solusi	40
Gambar III.4	Metode ajuan dalam penanganan masalah oklusi pada wajah pada penerapan FER	41
Gambar IV.1	Alur penerapan mekanisme umpan balik pada LMS	46
Gambar IV.2	Contoh hasil analisis wajah dengan paramater arah pandangan mata, arah kepala, dan emosi wajah.....	47
Gambar IV.3	Contoh kondisi wajah yang arah kepala dan mata tidak menghadap monitor atau menghadap kesamping (gambar dibuat dengan AI)	47
Gambar IV.4	Diagram alir algoritma kondisi penggunaan FER untuk menentukan <i>engagement</i> siswa.....	48

DAFTAR TABEL

Tabel II.1	Studi terkait yang menggunakan metode <i>deep learning</i> pada deteksi emosi dengan modalitas visual pada lingkungan pendidikan	33
Tabel III.1	Tahap identifikasi masalah dan motivasi.....	37
Tabel III.2	Tahap tujuan dan solusi	39
Tabel III.3	Tahap perancangan dan pengembangan	42
Tabel III.4	Tahapan pada metode KDD.....	43
Tabel III.5	Tahapan demonstrasi	44
Tabel III.6	Tahap evaluasi	45
Tabel III.7	Tahap komunikasi.....	45

DAFTAR SINGKATAN DAN LAMBANG

SINGKATAN	Nama	Pemakaian pertama kali pada halaman
2D	Dua Dimensi	5
3D	Tiga Dimensi	5
AER	<i>Automated Emotion Recognition</i>	11
BVP	<i>Blood Volume Pressure</i>	20
CMU	<i>Carnegie Mellon University</i>	17
CNN	<i>Convolutional Neural Network</i>	3
CNN-RF	<i>Convolutional Neural Network - Random Forest</i>	34
COVID-19	<i>Corona Virus Disease 2019</i>	1
Daring	Dalam Jaringan	3
DL	<i>Deep Learning</i>	4
DSRM	<i>Design Science Reseacrh Methodology</i>	8
ECG	<i>Elektrokardiography</i>	20
EDA	<i>Electrodermal Activity</i>	19
EEG	<i>Electroencephalography</i>	16
EKG	<i>Electrocardiography</i>	19
EMG	<i>Elektromiography</i>	20
FER	<i>Facial Expressions Recognition</i>	3
GSR	<i>Galvanic Skin Response</i>	20
IB-BiLSTM	<i>Information Block Bidirectional LSTM</i>	14
IEMOCAP	<i>Interactive Emotional Dyadic Motion Capture</i>	18
IoT	<i>Internet of Things</i>	3
IPDN	Institut Pemerintahan Dalam Negeri	1
KDD	<i>Knowledge Discovery in Databases</i>	35
LBP	<i>Local Binary Pattern</i>	22
LMS	<i>Learning Management System</i>	1
LSTM	<i>Long Short Term Memory</i>	13
MER	<i>Multimodal Emotion Recognition</i>	26
MFCC	<i>Mel-Frequency Cepstral Coefficients</i>	17
ML	<i>Machine Learning</i>	15
MOOC	<i>Massive Online Open Course</i>	3
NLP	<i>Natural Language Processing</i>	27
PANAS	<i>Positive and Negative Affect Schedule</i>	15
PPG	<i>Photoplethysmography</i>	20
RSP	<i>Respiration</i>	20
SAM	<i>Self-Assessment Manikin</i>	15
SER	<i>Speech Emotion Recognition</i>	16
SKT	<i>Skin Temperature</i>	20
SVM	<i>Support Vector Machine</i>	14
TTS	<i>Text-To-Speech</i>	17
WHO	<i>World Health Organization</i>	1

LAMBANG

BAB I Pendahuluan

Bab ini membahas berbagai dasar dan pedoman yang menjadi fokus utama penelitian. Tujuannya adalah untuk memberikan pemahaman menyeluruh mengenai penelitian serta urgensinya dalam mencapai tujuan yang telah ditetapkan.

I.1 Latar Belakang

Institut Pemerintahan Dalam Negeri (IPDN) sebagai lembaga pendidikan tinggi kedinasan di bawah naungan Kementerian Dalam Negeri sedang melakukan transformasi digital dalam dunia pendidikan untuk dapat mengikuti era revolusi industri 4.0. Dalam usaha untuk mencapai hal tersebut, IPDN merencanakan salah satu program jangka panjang yaitu penerapan *smart campus* yang tertuang dalam dokumen rencana strategis IPDN 2020-2024 dan di dukung dengan rencana induk pengembangan IPDN 2020-2045. Salah satu poin yang ditekankan dalam penerapan *smart campus* tersebut adalah pengembangan teknologi pembelajaran. Hal ini juga didukung dengan salah satu poin pada misi IPDN yaitu meningkatkan mutu dan kinerja penyelenggaraan pendidikan yang mengarah pada pemenuhan standar mutu pendidikan tinggi nasional berbasis teknologi informasi dan komunikasi.

Pandemi *Corona Virus Disease 2019* (COVID-19) beberapa tahun silam membawa dampak besar bagi dunia pendidikan, memaksa sistem pembelajaran beradaptasi dengan cepat ke metode daring. Sekolah dan universitas di seluruh dunia termasuk IPDN harus mengandalkan teknologi digital untuk memastikan kelangsungan pembelajaran. Perubahan secara mendadak ini otomatis membuat sejumlah permasalahan dalam pembelajaran, banyak siswa mengalami kesulitan dalam memahami materi tanpa interaksi langsung dengan guru. Walaupun pandemi COVID-19 dinyatakan berakhir oleh *World Health Organization* (WHO) dan diperkuat dengan Keputusan Presiden (Keppres) Nomor 17 Tahun 2023 tentang Penetapan Berakhirnya Status Pandemi COVID-19 di Indonesia, namun pengembangan teknologi pembelajaran di dunia pendidikan masih tetap perlu untuk disempurnakan. Salah satu media yang saat ini tetap digunakan adalah *Learning Management System* (LMS).

LMS memungkinkan institusi pendidikan untuk menyusun materi pembelajaran, memberikan tugas, melakukan evaluasi, serta memantau perkembangan siswa secara sistematis. Dengan fitur seperti kelas virtual, forum diskusi, dan integrasi dengan berbagai alat pembelajaran, LMS mendukung fleksibilitas dalam proses belajar mengajar, baik secara daring maupun *hybrid*. Selain itu, LMS juga membantu meningkatkan efisiensi administrasi pendidikan, mengurangi penggunaan kertas, serta memfasilitasi akses pembelajaran yang lebih inklusif dan personal bagi setiap pengguna. LMS juga adalah bagian integral dari era pendidikan 4.0, dimana teknologi terintegrasi dengan praktik pedagogis untuk meningkatkan hasil belajar (Mahmud dkk., 2022).

IPDN dalam perkembangannya mengikuti era pendidikan 4.0 juga telah menggunakan sebuah aplikasi LMS yang berfungsi untuk menyimpan dan mengorganisir materi dari pengajar. Materi yang diberikan pengajar dalam LMS IPDN bisa berbentuk dokumen maupun video *asynchronous learning*. *Asynchronous learning* dapat memberikan kesempatan yang lebih fleksibel pada siswa dalam hal waktu, walaupun demikian *asynchronous learning* juga memiliki kelemahan yaitu tidak adanya interaksi langsung antara pengajar dan siswa sehingga pengajar tidak dapat menilai keseriusan atau kefokusinan dari siswa pada saat mengakses materi yang diberikan (Peterson, 2023). Diperlukan suatu mekanisme umpan balik untuk mengatasi permasalahan tersebut, salah satunya adalah penerapan teknologi informasi yang dapat menyimpulkan tingkat keseriusan dan kefokusinan siswa memanfaatkan data emosi baik dalam bentuk *text*, *audio*, maupun *visual* yang didapat dari pengamatan pada siswa (Yu dkk., 2024).

Salah satu bentuk mekanisme umpan balik dalam pembelajaran daring adalah berbasis *text* seperti analisis sentimen dari komentar yang diberikan siswa pada forum tertentu, ataupun dari jawaban tugas yang diberikan oleh pengajar pada media pembelajaran daring (Xu, 2023). Beberapa penelitian tentang penerapan analisis sentimen untuk mengetahui emosi dari siswa telah dilakukan oleh beberapa peneliti. Peneliti X. Zhang & Qin, (2022) membuat sistem yang menerapkan analisis sentimen untuk mengevaluasi emosi siswa pada pelajaran Ideologi dan Politik dalam *platform* daring. Pendekatan analisis sentimen juga telah dilakukan

oleh Mao dkk., (2023) untuk mengevaluasi tingkat emosional siswa dalam pelajaran *Internet of Things* (IoT). Walaupun demikian, pendekatan analisis sentimen dinilai kurang tepat untuk dijadikan aspek penilaian utama untuk mendapatkan data emosi siswa, karena dimungkinkan untuk terjadinya emosi palsu dari siswa dalam tulisannya (Khoshnam & Baraani-Dastjerdi, 2022).

Bentuk mekanisme umpan balik yang lainnya adalah berbasis *audio*. Berdasarkan referensi penelitian lampau, Bao dkk., (2024) menggabungkan fitur *audio* dan gerakan mata untuk mengevaluasi emosi siswa pada *massive online open course* (MOOC). Mereka merancang sistem yang dapat mengklasifikasikan emosi seperti minat, kebahagiaan, kebingungan, dan kebosanan dengan hasil akurasi sebesar 81,90%. Penelitian lain yang menggunakan fitur *audio* dalam mendeteksi emosi siswa pada pembelajaran daring adalah Bekmanova dkk., (2022), didasari oleh pembelajaran jarak jauh karena COVID-19, mereka membangun sistem yang dapat mendeteksi emosi positif dan negatif dari ucapan lisan siswa dengan bahasa Kazakh dengan hasil akurasi sebesar 79,7%. Berdasarkan hasil penelitian – penelitian di atas, membuktikan bahwa *audio* dapat digunakan sebagai basis penilaian untuk memberikan masukan dari siswa kepada pengajar berupa data emosi siswa. Walaupun demikian, penggunaan *audio* dalam deteksi emosi siswa pada pembelajaran daring memiliki sejumlah masalah, seperti penggunaan fitur *audio* yang menyebabkan penurunan akurasi jika digabungkan dengan fitur lain (Pan dkk., 2024), lalu adanya batasan kuantitas terhadap *database* yang akan digunakan sebagai data latih (Zhu-Zhou dkk., 2022), hingga mahalnya biaya komputasi pengolahan sinyal *audio* (Z. T. Liu dkk., 2021).

Bentuk mekanisme umpan balik yang lainnya adalah berbasis *visual*. Memanfaatkan teknologi *computer vision*, banyak penelitian yang menggunakan teknologi *facial expressions recognition* (FER) untuk mendapatkan data emosi dari pengamatan ekspresi wajah siswa pada lingkungan pembelajaran daring. Bhardwaj dkk., (2021) merancang sistem untuk menghitung tingkat keterlibatan siswa dengan menerapkan FER selama kelas daring dan mendapatkan hasil akurasi sebesar 93,6% dalam klasifikasi emosinya menggunakan algoritma *Convolutional Neural Network* (CNN). Penerapan FER dalam lingkungan pendidikan daring

lainnya juga pernah dilakukan oleh Gupta dkk., (2023). Mereka membangun sistem FER yang dapat mendeteksi tingkat keterlibatan *real-time* dengan menganalisis ekspresi wajah, lalu mengklasifikasikan siswa menjadi “Terlibat” atau “Tidak Terlibat” berdasarkan keadaan emosinya. Algoritma yang mereka gunakan untuk klasifikasi emosi siswa adalah algoritma ResNet-50 dan mendapatkan hasil mencapai akurasi 92,3% dalam skenario *real-time*. Berdasarkan hal tersebut, dibandingkan dengan solusi alternatif seperti analisis sentimen dan pengenalan suara, penerapan FER pada pembelajaran daring baik *synchronous* maupun *asynchronous* bisa dijadikan sebagai solusi yang paling logis untuk menyelesaikan masalah kebutuhan mekanisme umpan balik dari siswa pada lingkungan pembelajaran daring baik dari sisi batasan-batasan tiap solusi yang ada maupun dari biaya dan beban komputasi.

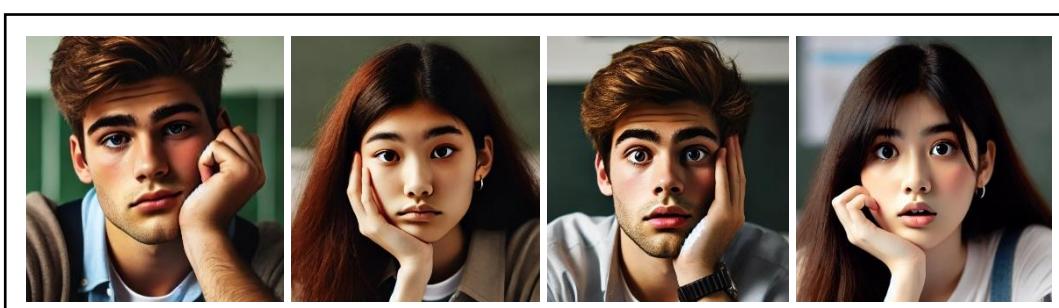
Walaupun demikian, penerapan FER pada pembelajaran daring juga mempunyai beberapa kekurangan, terutama dalam keadaan oklusi, seperti wajah yang tertutup oleh tangan, kacamata, atau objek lain (Li dan Deng, 2022). Kondisi ini dapat mengurangi akurasi dalam mendeteksi ekspresi secara tepat, sehingga sistem mungkin menghasilkan prediksi yang kurang akurat atau bahkan keliru dalam menginterpretasikan emosi pengguna (Mensah dkk., 2024). Selain itu, variasi pencahayaan dan sudut pandang kamera yang tidak seragam juga dapat memengaruhi performa model dalam mengenali ekspresi wajah secara konsisten.

Sudah ada beberapa penelitian yang mencoba untuk mengatasi masalah oklusi pada penerapan FER pada lingkungan nyata. Ada beberapa cara untuk mengatasi oklusi pada penerapan FER yang telah dilakukan peneliti, antara lain dengan menerapkan teknik pemrosesan citra seperti merekonstruksi bagian wajah yang terhalang (Poux dkk., 2022), menggunakan model *deep learning* (DL) yang dilatih dengan data augmentasi oklusi agar lebih *robust* terhadap data oklusi (X. Zhang dkk., 2022), serta memanfaatkan multi-modalitas, seperti kombinasi informasi dari gambar dan data lainnya, misalnya gabungan data wajah dan suara (Freire-obregón dan Castrillón-santana, 2024 dan Mamieva dkk., 2023). Meskipun berbagai metode telah dikembangkan untuk mengatasi oklusi pada penerapan FER, solusi-solusi di

atas merupakan teknik yang memerlukan komputasi yang kompleks, sehingga mengakibatkan kenaikan latensi pada prosesnya.

Ada cara lain yang dilakukan peneliti untuk mengatasi oklusi yaitu pada bagian *pre-processing* data pada sistem FER. Peneliti Cao dkk., (2021) melakukannya dengan cara mengambil setengah wajah dan menggabungkannya dalam bentuk tiga dimensi (3D), kemudian dilakukan penguatan bentuk otot disekitar hidung untuk dapat meningkatkan akurasi pengenalan emosi dari data wajah 3D. Hasilnya mengalami kenaikan akurasi terhadap beberapa emosi dasar setelah melakukan proses penguatan otot muka disekitar hidung. Atas dasar tersebut, penulis mengajukan cara yang mirip dengan apa yang dilakukan Cao dkk. yaitu dengan cara memanfaatkan setengah wajah yang tidak teroklusi, kamudian dilakukan *mirroring* dan digabungkan hingga menjadi sebuah gambar wajah yang utuh dengan posisi menghadap kedepan (*frontal*) dalam bentuk dua dimensi (2D), dengan harapan akurasi yang didapatkan minimal menyamai hasil pemrosesan wajah 3D dengan beban komputasi yang lebih rendah dari 3D.

Berdasarkan uraian di atas, penulis akan membuat sebuah aplikasi media pembelajaran daring yang megacu pada LMS IPDN dengan menambahkan fitur umpan balik menggunakan sistem yang dapat menilai tingkat keterlibatan siswa berdasarkan emosi wajah dan sikap siswa. Sikap siswa yang akan dinilai adalah berdasarkan arah kepala dan arah mata siswa. Hasil penilaian tingkat keterlibatan siswa yang dihasilkan oleh aplikasi diharapkan bisa menjadi data yang bisa digunakan sebagai umpan balik ketika konsep ini diterapkan pada suatu media pembelajaran daring.



Gambar I.1 Contoh kondisi wajah yang tertutup sebagian oleh tangan pada saat pembelajaran daring (gambar dibuat dengan AI)

Penelitian ini akan berfokus kepada bagaimana membangun sistem FER pada media pembelajaran daring yang *robust* terhadap wajah dalam oklusi terutama saat wajah tertutup tangan (contoh pada gambar I.1), dimana keadaan tersebut memang lazim dilakukan oleh siswa dalam menjalani pembelajaran secara daring baik karena masalah kesehatan leher atau kepala (Beekman dkk., 2002), maupun karena faktor kefokusan atau kebosanan. Pengaplikasian FER dalam media pembelajaran daring yang penulis usulkan juga diharapkan dapat beroperasi pada *resource* yang minim sehingga tidak menimbulkan efek keuangan yang dapat membebani lembaga dalam penerapannya.

I.2 Rumusan Masalah

Berdasarkan latar belakang di atas, bahwa LMS pada IPDN masih memiliki kekurangan yaitu belum adanya mekanisme umpan balik untuk menilai tingkat keterlibatan siswa. Oleh karena itu, dirumuskan suatu masalah penelitian sebagai pertanyaan yang akan dijawab pada penelitian ini, yaitu “Bagaimana menerapkan FER yang dipadukan dengan penilaian sikap siswa pada media pembelajaran daring untuk mendapatkan data tingkat keterlibatan siswa ?”.

Dari penjelasan di atas, terdapat tiga pertanyaan penelitian yang akan diajukan untuk membantu dalam menjawab rumusan masalah, yaitu:

1. Bagaimana meyimpulkan tingkat keterlibatan siswa dengan modalitas FER, arah pandangan mata, dan arah kepala?
2. Apakah *pre-processing* gambar 2D yang memanfaatkan data setengah wajah dapat meningkatkan probabilitas emosi wajah siswa?
3. Seberapa besar efek pemrosesan gambar 2D terhadap waktu eksekusi *pre-processing* sebuah gambar wajah yang teroklusi dibandingkan dengan menggunakan pemrosesan gambar 3D?

I.3 Tujuan Penelitian

Tujuan penelitian ini dibagi menjadi tujuan umum dan tujuan khusus. Tujuan umum penelitian ini adalah membuat konsep media pembelajaran daring dengan menambahkan fitur umpan balik menggunakan modalitas FER, arah pandangan, dan arah kepala untuk mendapatkan data tingkat keterlibatan dari siswa saat

melakukan pembelajaran daring. Adapun tujuan khusus dari penelitian ini adalah sebagai berikut:

1. Aplikasi media pembelajaran daring dapat meyimpulkan tingkat keterlibatan siswa menggunakan *multi-modal* yaitu FER, arah pandangan mata, dan arah kepala.
2. Melakukan pengukuran efek *pre-processing* gambar 2D yang memanfaatkan data setengah wajah terhadap probabilitas emosi wajah siswa.
4. Melakukan evaluasi waktu hasil *pre-processing* gambar 2D wajah yang teroklusi dibandingkan dengan menggunakan pemrosesan gambar 3D.

I.4 Batasan Penelitian

Adanya keterbatasan dalam hal waktu, tenaga dan biaya, dirasa perlu untuk menetapkan batasan dalam penelitian ini, Hal ini penting untuk dilakukan agar penelitian tetap fokus pada permasalahan yang diangkat. Berikut batasan-batasan yang ditetapkan untuk penelitian ini.

1. Media pembelajaran daring yang akan dibuat mengacu pada LMS yang telah diterapkan di IPDN.
2. Penerapan sistem FER hanya pada pembelajaran *asynchronous* yang berbentuk video rekaman paparan perkuliahan.
3. Hasil penelitian belum akan diimplementasikan pada LMS IPDN.
4. Bentuk oklusi yang diteliti hanya pada kondisi sebagian wajah tertutup tangan, sebagaimana penjelasan di latar belakang yang menyebutkan bahwa keadaan sebagian wajah tertutup tangan lazim dilakukan oleh siswa dalam menjalani pembelajaran secara daring baik karena masalah kesehatan leher atau kepala (Beekman dkk., 2002).

I.5 Kontribusi Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi baik terhadap penyelesaian masalah maupun dalam bidang ilmu pengetahuan. Adapun kontribusi yang diharapkan adalah:

1. Menghasilkan sebuah konsep mekanisme umpan balik pada pembelajaran daring berdasarkan tingkat keterlibatan siswa menggunakan data emosi wajah dan sikap siswa.
2. Menghasilkan temuan ilmiah baru yang membuktikan bahwa *pre-processing* gambar wajah yang teroklusi menggunakan data setengah wajah dalam 2D memiliki akurasi yang sama baiknya dengan 3D dalam mengenali emosi wajah namun dengan waktu eksekusi yang lebih cepat.

I.6 Sistematika Penulisan

Sistematika penulisan tesis akan dibagi menjadi lima bab yang diawali dengan penjabaran umum di setiap rinciannya.

Bab I Pendahuluan

Memberikan penjelasan mengenai latar belakang dilakukannya penelitian. Dari latar belakang yang telah dijabarkan, selanjutnya disusun suatu rumusan masalah, tujuan penelitian, batasan penelitian, kebaruan dan kontribusi penelitian serta sistematika penulisan.

Bab II Tinjauan Pustaka

Menjelaskan mengenai gambaran tinjauan literatur dan kajian ilmu yang digunakan untuk mendukung suatu penelitian. Tinjauan pustaka ini selanjutnya dimanfaatkan dalam mengidentifikasi suatu permasalahan, konsep, dan ilmu pengetahuan untuk mendapatkan solusi dari penelitian yang akan dilakukan.

Bab III Metodologi Penelitian

Pembahasan terkait dengan metodologi yang akan digunakan di dalam penelitian. *Design Science Research Methodology* (DSRM) dipilih sebagai metode dalam penulisan tesis ini. Terdapat enam tahapan utama di dalam DSRM, yaitu identifikasi masalah & motivasi, menentukan tujuan solusi, perancangan & pengembangan, demonstrasi, evaluasi, dan komunikasi.

Bab IV Pembahasan

Menjelaskan dan menjabarkan hasil dari perancangan pendukung yang berupa suatu aplikasi usulan dan implementasi purwarupa yang telah dibuat. Kemudian

dilanjutkan dengan pembahasan hasil evaluasi pengujian lapangan dan analisis yang menyertainya. Perencanaan mengenai evaluasi dan pembahasan analisisnya terlebih dahulu telah dijabarkan pada Bab III.

Bab V Kesimpulan dan Saran

Memberikan penjelasan mengenai kesimpulan dari penelitian yang telah dilakukan. Selain itu, juga akan disampaikan mengenai beberapa saran yang bisa menjadi pertimbangan untuk dilakukannya penelitian lanjutan.

BAB II Tinjauan Pustaka

Bab ini mencakup tinjauan pustaka yang berperan sebagai acuan penelitian dalam mencapai tujuan penelitian yang telah didefinisikan. Tinjauan pustaka dilakukan terhadap referensi eksternal yang kredibel dan sesuai dengan kebutuhan penggerjaan tesis.

II.1 *Learning Management System*

Learning Management System (LMS) adalah sebuah platform perangkat lunak yang dirancang untuk mengelola, mendistribusikan, dan melacak aktivitas pembelajaran secara daring. LMS banyak digunakan di institusi pendidikan, perusahaan, dan organisasi untuk menyelenggarakan pelatihan serta kursus secara efisien. Dengan LMS, pengguna dapat mengakses materi pembelajaran, mengikuti kuis, mengerjakan tugas, dan berinteraksi dengan pemateri atau sesama peserta secara fleksibel dari mana saja dan kapan saja.

Manfaat dari Learning Management System (LMS) sangat beragam dan memberikan dampak positif bagi proses pembelajaran. Pertama, LMS dapat meningkatkan keterlibatan siswa serta hasil pembelajaran, karena memungkinkan pengalaman belajar yang lebih interaktif dan terstruktur (Alotaibi, 2024). Selain itu, sistem ini juga berperan dalam meningkatkan aksesibilitas dan kesetaraan pendidikan, memungkinkan siswa dari berbagai latar belakang untuk memperoleh materi pembelajaran dengan lebih mudah (Alotaibi, 2024).

Selain manfaat tersebut, LMS juga mendukung manajemen dan penilaian kursus yang lebih efisien, membantu pendidik dalam mengorganisir materi, memberikan tugas, serta mengevaluasi kemajuan siswa secara sistematis (Tran dan Meacheam, 2020). Lebih jauh, penggunaan LMS memiliki dampak positif terhadap kepuasan siswa, karena menyediakan pengalaman belajar yang fleksibel, mudah diakses, dan sesuai dengan kebutuhan individu (Zaareer dkk., 2024). Terakhir, LMS dapat memfasilitasi interaksi yang lebih baik antar peserta, baik antara siswa dengan pemateri maupun dengan sesama siswa, melalui fitur diskusi, kolaborasi, dan komunikasi yang tersedia dalam platform (Costley dkk., 2022).

Dengan semakin berkembangnya teknologi, LMS terus mengalami inovasi, termasuk dengan penerapan kecerdasan buatan untuk personalisasi pembelajaran dan analisis data yang lebih mendalam (Alotaibi, 2024). Penggunaan LMS di berbagai sektor menunjukkan bahwa sistem ini menjadi bagian penting dalam transformasi digital pendidikan dan pelatihan, memberikan kemudahan dalam akses ilmu pengetahuan serta meningkatkan efisiensi proses pembelajaran.

II.2 *Affective Computing*

Affective computing adalah bidang interdisipliner dalam kecerdasan buatan yang berfokus pada pengenalan, interpretasi, dan simulasi emosi manusia oleh komputer. Konsep ini pertama kali diperkenalkan oleh Rosalind Picard pada tahun 1995, yang menekankan pentingnya emosi dalam interaksi manusia dan teknologi. Dengan kemajuan teknologi, *affective computing* kini diterapkan dalam berbagai domain, seperti interaksi manusia-mesin, perawatan kesehatan, pendidikan, dan hiburan. Kemampuan sistem untuk mengenali dan merespons emosi pengguna dapat meningkatkan pengalaman pengguna secara signifikan (Afzal dkk., 2024).

Dalam dunia pendidikan *affective computing* memiliki peran, terutama dalam meningkatkan pengalaman belajar yang lebih personal dan adaptif (Villegas-Ch dkk., 2024). Teknologi ini memungkinkan sistem pembelajaran digital untuk mengenali emosi siswa melalui ekspresi wajah, nada suara, atau bahkan respons fisiologis (Yu dkk., 2024). Dengan memanfaatkan *affective computing*, sistem dapat menyesuaikan materi dan metode pembelajaran sesuai dengan kondisi emosional siswa, misalnya dengan memperlambat penjelasan jika siswa terlihat kebingungan atau memberikan motivasi saat mereka merasa frustasi (Shi, 2024). Hal ini dapat meningkatkan efektivitas pembelajaran dan membantu siswa lebih fokus serta termotivasi dalam proses belajar (Lin dkk., 2022).

II.3 *Automated Emotion Recognition* di Lingkungan Pendidikan

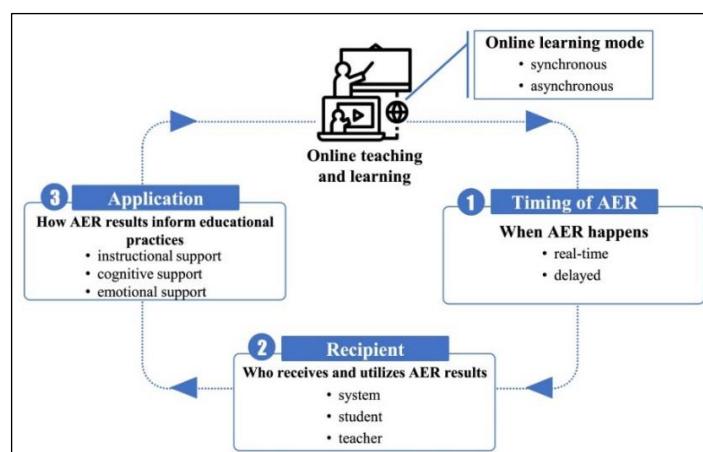
Automated emotion recognition (AER) dalam lingkungan pendidikan merujuk pada teknologi berbasis kecerdasan buatan yang secara otomatis mendeteksi dan menganalisis emosi siswa selama proses pembelajaran. Sistem ini menggunakan berbagai teknik seperti pengenalan wajah, analisis suara, pemantauan fisiologis,

dan pemrosesan bahasa alami untuk memahami kondisi emosional siswa secara *real-time* (Yu dkk., 2024). Dengan memanfaatkan teknologi ini, pendidik dan sistem pembelajaran adaptif dapat menyesuaikan strategi pengajaran agar lebih sesuai dengan keadaan emosional siswa, sehingga meningkatkan keterlibatan mereka dalam proses belajar .

Dalam praktiknya, AER dapat diterapkan baik dalam kelas fisik maupun pembelajaran daring. Di kelas tradisional, kamera dan sensor dapat digunakan untuk menganalisis ekspresi wajah dan bahasa tubuh siswa, sementara dalam pembelajaran daring, sistem dapat mengevaluasi ekspresi wajah dan intonasi suara dari video konferensi. Informasi yang diperoleh dari AER memungkinkan pengajar untuk memberikan intervensi yang lebih tepat, seperti mengubah metode penyampaian materi atau memberikan dukungan tambahan bagi siswa yang mengalami kesulitan. Dengan demikian, teknologi ini berpotensi meningkatkan efektivitas pembelajaran serta menciptakan pengalaman belajar yang lebih adaptif dan responsif terhadap kebutuhan emosional siswa.

Dalam konteks pendidikan, AER bertujuan untuk:

1. Membantu guru dalam memahami kebutuhan siswa dengan memberikan wawasan tentang kondisi emosional mereka (Llurba dkk., 2024).
2. Meningkatkan keterlibatan siswa dengan menyesuaikan strategi pembelajaran berdasarkan emosi mereka (Yu dkk., 2024).
3. Mempersonalisasi pengalaman belajar dan meningkatkan penyesuaian diri siswa dalam tugas-tugas berbasis kelompok (Ngo dkk., 2024).



Gambar II.1 Siklus penerapan AER pada lingkungan pendidikan (Yu dkk., 2024)

Peniliti Yu dkk., (2024) dalam ulasannya mencoba untuk memetakan bagaimana siklus proses pengaplikasian AER dalam lingkungan pendidikan yang dapat dilihat pada gambar II.1. Yu dkk., (2024) menjelaskan bahwa penerapan AER pada pembelajaran daring bisa dalam pembelajaran *synchronous* maupun pembelajaran *asynchronous*. Penerapannya juga bisa dalam keadaan *real-time* maupun *delayed*. Kemudian entitas yang terkait adalah aplikasi sistem pembelajaran daring, guru atau pemateri, dan siswa. Hasil dari penerapan AER bisa berbentuk instruksional, sebagai contoh jika diterapkan secara real-time, sistem bisa memperingatkan siswa jika tidak fokus, lalu bisa juga berbentuk penilaian kognitif, dan juga berbentuk data emosi dari siswa.

Atas dasar tersebut, penulis memilih untuk melakukan penelitian pada pembelajaran *asynchronous* dengan keadaan *delayed* dan menghasilkan data emosi dari siswa dengan tambahan modalitas arah pandangan dan arah kepala yang kemudian disimpulkan untuk menentukan keterlibatan dari siswa pada pembelajaran daring dengan alasan bahwa pengenalan emosi pada keadaan *delayed* memiliki beban *resource* yang lebih ringan dibandingkan dengan sistem *real-time* (Pascual dkk., 2022).

II.3.1 Modalitas Teks

Penggunaan modalitas teks dalam lingkungan pendidikan bisa digunakan untuk penilaian pedagogik. Penilaian pedagogik menggunakan teks tanpa menerapkan AER telah biasa dilakukan seperti menggunakan kuisioner (Herrero-Alvarez dkk., 2024), melakukan survey (Benabbes dkk., 2023), dan penilaian komulatif berbasis teks pada platform *online education* (Xu, 2023). Namun, pendekatan tersebut tidak akurat untuk menilai siswa dari segi emosionalnya. Analisis sentimen adalah salah satu pendekatan yang dapat diterapkan untuk membangun sebuah sistem penilaian pedagogik dengan menerapkan AER melalui data modalitas teks.

Peneliti Zhang & Qin, (2022) membuat sistem yang menerapkan analisis sentimen untuk mengevaluasi emosi siswa pada pelajaran Ideologi dan Politik dalam platform online. Mereka mengombinasikan *Convolutional Neural Network* (CNN) dan *Long Short-Term Memory* (LSTM) untuk menganalisis komentar mengenai kursus ideologis dan politik secara daring. Pendekatan ini bertujuan untuk

meningkatkan akurasi dalam mengenali sentimen pengguna dibandingkan dengan metode tradisional. Hasil penelitiannya menunjukkan bahwa algoritma CNN+LSTM berhasil mengungguli metode konvensional seperti SVM dan LSTM. Kinerja terbaik dicapai dengan menggunakan metode vektorisasi teks Word2vec, yang mampu menangkap representasi semantik kata dengan lebih baik. Dalam analisis terhadap 12.000 komentar, ditemukan bahwa 86,36% dari komentar tersebut mengandung sentimen positif, sementara 13,64% lainnya bersifat negatif. Temuan ini memberikan wawasan berharga bagi pengelola kursus untuk melakukan perbaikan berdasarkan umpan balik negatif yang diberikan oleh peserta. Pendekatan analisis sentimen juga dilakukan oleh Mao dkk., (2023) dalam evaluasi terhadap tingkat emosional siswa dalam pelajaran *Internet of Things* (IoT). Mao dkk., (2023) menggunakan algoritma *Information Block Bidirectional Long-Short Term Memory* (IB-BiLSTM) untuk meningkatkan analisis sentimen dalam teks pada pendidikan animasi secara *daring*. Algoritma ini memanfaatkan teknologi IoT untuk mengumpulkan data multimodal. Model yang dikembangkan mencapai akurasi 93,92% dan skor F1 90,34%, menunjukkan keefektifannya dalam mendeteksi perubahan emosi selama proses pembelajaran. Penelitian ini berkontribusi pada pengembangan pendidikan yang lebih personal dan meningkatkan keterlibatan emosional siswa.

Walaupun demikian, analisis sentimen masih memiliki kelemahan hanya mendeteksi emosi positif dan negatif dan memiliki keterbatasan dalam menangkap nuansa emosi yang lebih kompleks. Pendekatan biner ini tidak dapat membedakan emosi spesifik seperti marah, sedih, takut, atau senang, yang dapat memberikan wawasan lebih mendalam tentang perasaan siswa (Gamage dkk., 2024). Berdasarkan hal tersebut, penulis menyimpulkan bahwa modalitas teks kurang cocok untuk digunakan dalam mendeteksi emosi siswa pada penelitian ini.

Data modalitas teks dapat berasal dari dokumen tertulis, transkrip percakapan, media sosial, artikel berita, hingga korpus linguistik yang dikurasi. Metode koleksi dapat dilakukan melalui *web scraping*, ekstraksi dari basis data publik, penggunaan API dari *platform* pembelajaran daring, serta melalui anotasi manual oleh pakar bahasa. Format penyimpanan data teks dapat berupa TXT, JSON, CSV, atau

database berbasis teks seperti Elasticsearch. Data yang telah dikumpulkan kemudian digunakan untuk berbagai keperluan, seperti analisis sentimen, pemrosesan bahasa alami (*Natural Language Processing*), dan pengembangan model kecerdasan buatan berbasis teks.

Peneliti Benabbes dkk., (2023) mengumpulkan data siswa langsung dari *platform e-learning*. Pengumpulan data tersebut dilakukan dengan mendapatkan umpan balik melalui survei dari siswa dan guru untuk menilai fungsionalitas dan kemudahan dalam penggunaan sistem pembelajaran. Data umpan balik yang didapatkan selanjutnya dianalisis untuk menilai keterlibatan siswa dalam pembelajaran.

Metode serupa juga dilakukan oleh Herrero-Alvarez dkk., (2024), mereka mengumpulkan data tes menggunakan dua kuesioner yang diberikan kepada siswa di akhir kegiatan. Proses ini menghasilkan total 10 tanggapan per siswa. Tanggapan dikumpulkan menggunakan *platform Google Forms*, yang memfasilitasi pengumpulan data emosional dari peserta. Selain itu, keadaan emosional siswa dinilai melalui instrumen *Self-Assessment Manikin* (SAM) dan *Positive and Negative Affect Schedule* (PANAS), yang dirancang khusus untuk mengukur berbagai dimensi emosi.

Jenis data modalitas teks lainnya yang bisa digunakan untuk menilai keterlibatan siswa adalah melalui data kuantitatif. Peneliti Tao dkk., (2022) menganalisis data yang diambil dari data seberapa banyak jumlah *post* yang dibuat siswa dalam media pembelajaran daring menggunakan metode *machine learning* (ML) dengan tambahan parameter lainnya. Walaupun tidak dapat menilai emosi dari siswa, namun metode tersebut cukup efisien jika hanya untuk menilai keterlibatan atau keaktifan siswa dalam pembelajaran online.

Dalam analisis sentimen, Zhang dan Qin, (2022) memanfaatkan komentar siswa pada platform *Massive Open Online Course* (MOOC). Dari data komentar tersebut dikumpulkan dan dilakukan analisis dengan menggunakan *deep learning* untuk menilai emosi dari siswa berdasarkan teks komentar yang ditulisnya.

Berbeda dengan peneliti di atas, Mao dkk., (2023) mendapatkan data modalitas teks dari konversi data sensor yaitu *touchscreen* dan *keyborad* untuk menilai feed back

berupa teks, lalu untuk menangkap emosi menggunakan beberapa perangkat seperti sensor *body motion*, perangkat *eye tracker*, dan sensor EEG. Mereka juga menggunakan *smart pen* atau *handwriting tablet* untuk mendapatkan data tulisan siswa, dan yang terakhir menggunakan *speech recognition* untuk mendapatkan ekspresi verbal siswa. Semua data yang dikumpulkan berupa teks yang kemudian diproses dengan metode IB-BiLSTM yang hasil keluarannya adalah klasifikasi emosi. Apa yang dilakukan oleh Mao dkk., (2023) menunjukkan bahwa penggunaan modalitas teks dapat diambil dari bentuk apapun, baik visual, *audio*, ataupun sinyal walaupun hasilnya akan lebih spesifik dalam mengenali emosi, namun prosesnya menjadi tidak efisien untuk diterapkan dalam lingkungan pendidikan khususnya dalam pembelajaran daring.

Berdasarkan uraian di atas, pendekatan deteksi emosi dengan modalitas teks dinilai kurang tepat untuk diterapkan pada media pembelajaran daring, selain karena hanya bisa mendeteksi emosi biner (*positif* atau *negatif*) dan tidak dapat membedakan emosi spesifik seperti marah, sedih, takut, atau senang, yang dapat memberikan wawasan lebih mendalam tentang perasaan siswa (Gamage dkk., 2024), alasan lainnya adalah dimungkinkan untuk terjadinya emosi palsu dari siswa dalam tulisannya (Khoshnam & Baraani-Dastjerdi, 2022), seperti kata “tidak” jika diungkapkan dengan nada rendah bukan berarti emosinya negatif, sehingga diperlukan modalitas pendukung lain, seperti audio ataupun visual (Das dkk., 2024).

II.3.2 Modalitas *Audio*

Dalam lingkungan pendidikan, modalitas audio dapat digunakan untuk menentukan emosi siswa dengan teknologi *speech emotion recognition* (SER). Beberapa literatur penelitian terdahulu, ada beberapa peneliti yang menggunakan modalitas audio sebagai parameter penentu emosi di bidang pendidikan. Tanko dkk., (2022) pada penelitian membuat sebuah pendekatan SER yang bertujuan untuk mendeteksi emosi pemateri saat mengajar berdasarkan ucapan mereka. Untuk mencapai tujuan ini, dua kumpulan data pidato dalam bahasa Turki dan Inggris dikumpulkan. Model pembelajaran baru tersebut dinamai *ShoePat23* dan berhasil mencapai akurasi klasifikasi sebesar 94,97% untuk bahasa Inggris dan 96,41% untuk bahasa Turki.

Hasil penelitian menunjukkan bahwa *ShoePat23* merupakan teknik yang efektif dan efisien untuk pengenalan emosi dalam ucapan, terutama dalam aplikasi dunia nyata. Selain itu, studi ini mengungkapkan bahwa model ini dapat digunakan untuk mengembangkan alat pengendalian diri baru yang dapat diintegrasikan ke dalam platform pendidikan jarak jauh guna mengevaluasi kinerja pemateri.

Pendekatan lain bernama *Esernet* yang merupakan improvisasi dari sistem SER juga dapat digunakan dalam deteksi emosi siswa (T. Liu dkk., 2025). Ada dua temuan utama dalam penelitian yang dilakukan oleh T. Liu dkk, yaitu adanya isyarat emosional yang signifikan serta perbedaan kecil antara emosi yang berbeda, yang keduanya berperan penting dalam meningkatkan akurasi SER. Model yang diperkenalkan menggunakan strategi dua jalur. Jalur pertama berfokus pada ekstraksi isyarat ucapan penting dengan teknik masking, sementara jalur kedua menggabungkan informasi multiskala dari sinyal suara. Arsitektur *Transformer* digunakan untuk menangkap hubungan semantik jarak jauh dalam sinyal ucapan. Hasil eksperimen menunjukkan bahwa *Esernet* mampu mengenali isyarat penting dan hubungan jarak jauh dengan baik. Namun, penelitian ini juga mengakui bahwa model ini memiliki keterbatasan, terutama dalam hal kebutuhan komputasi yang tinggi akibat strategi multi-tugas yang digunakan.

Dalam modalitas audio khususnya dalam pendekatan SER, beberapa literatur terdahulu menunjukkan bahwa mereka menggunakan dua jenis data, yaitu data publik, dan data *private*. Data modalitas *audio* dapat diperoleh dari berbagai sumber, termasuk menggunakan mikrofon secara langsung, sensor ultrasonik, atau perangkat rekaman lainnya, dengan format yang beragam seperti sinyal mentah, spektrogram, atau fitur yang diekstraksi seperti *Mel-Frequency Cepstral Coefficients* (MFCC). Dari berbagai literatur disebutkan bahwa beberapa peneliti menggunakan dua jenis dataset modalitas audio, yaitu dataset *publik* dan *private*.

Beberapa peneliti yang menggunakan dataset publik adalah Guan, (2024). Dia menggunakan dataset CMU ARCTIC, dimana CMU ARCTIC adalah kumpulan dataset ucapan (*speech corpus*) yang dikembangkan oleh Carnegie Mellon University (CMU). Dataset ini dirancang untuk penelitian dan pengembangan dalam bidang *speech synthesis* atau *text-to-speech* (TTS). CMU ARCTIC berisi

rekaman suara dari berbagai penutur dengan aksen yang berbeda, masing-masing membaca sekitar 1.132 kalimat yang dipilih secara khusus untuk mencakup berbagai fonema dalam bahasa Inggris. Selain dari situs resminya, dataset tersebut juga tersedia pada situs *Kaggle* maupun *Hugging Face*.

Peneliti T. Liu dkk., (2025) juga menggunakan dataset publik pada penelitiannya. Mereka menggunakan gabungan dari dataset IEMOCAP dan EmoDB. IEMOCAP adalah kumpulan data yang terdiri dari rekaman video aktor yang melakukan dialog naskah dengan berbagai ekspresi emosional. Dataset mencakup data multimodal, seperti bingkai video, sinyal audio, dan subtitle teks, sehingga cocok untuk mengevaluasi metode pengenalan emosi multimodal (Le dkk., 2023). Sedangkan Berlin Database of Emotional Speech atau yang lebih dikenal dengan nama EmoDB adalah kumpulan data yang terdiri dari rekaman suara aktor yang mengekspresikan emosi yang berbeda. Rekaman ini dirancang untuk menangkap berbagai keadaan emosional, termasuk kemarahan, ketakutan, kebahagiaan, kesedihan, jijik, kebosanan, dan keadaan netral. Basis data disusun untuk menyediakan serangkaian ekspresi emosional yang komprehensif untuk analisis dan pelatihan model (Kanwal & Asghar, 2021). Dataset IEMOCAP dan EmoDB bisa didapatkan pada situs *Kaggle* maupun *Hugging Face*.

Peneliti Tanko dkk., (2022), menggunakan dataset *private* dengan cara mengumpulkan dataset *English Speech Emotion Recognition* dari total 45 pembicara, yang mencakup 15 wanita dan 30 pria, berusia antara 24 hingga 42 tahun. File *audio* direkam ketika pembicara menyampaikan kuliah tentang forensik digital menggunakan slide kuliah yang disiapkan. Rekaman dibuat dengan ponsel, dan audio ditangkap dalam lima format berbeda, yaitu MP4, MPEG, AAC, OGG, dan M4A. Dataset yang terkumpul terdiri dari berbagai kelas emosional, dengan 44 suara dengan label “menarik”, 47 dengan label “netral”, dan 43 dengan label “membosankan”.

Berdasarkan uraian di atas, meskipun modalitas audio dapat mendeteksi emosi dengan baik, ada beberapa masalah saat menggunakan saat mengidentifikasi emosi siswa dalam pembelajaran daring. Jika fitur *audio* digabungkan dengan fitur lain, ini dapat menurunkan akurasi (Pan dkk., 2024). Selain itu, pemrosesan sinyal

audio membutuhkan sumber daya komputasi yang besar (Z. T. Liu dkk., 2021), yang membuatnya tidak efisien dalam penerapan skala luas. Kemudian masalah lainnya yaitu jumlah data yang tersedia untuk pelatihan sering kali terbatas (Zhu-Zhou dkk., 2022). Atas pertimbangan tersebut, penulis tidak menggunakan modalitas *audio* dalam menentukan *engagement* siswa berdasarkan emosi, mengingat *resource* yang dimiliki pada lokus penelitian tidak terlalu besar.

II.3.3 Modalitas Sinyal Tubuh

Dalam lingkungan yang terkontrol, deteksi emosi siswa kadang juga dilakukan dengan memanfaatkan signal dari tubuh yang ditangkap melalui *tools* atau *sensor* yang terpasang pada bagian tubuh siswa. Penggunaan perangkat IoT juga menjadi pilihan *tools* untuk mendapatkan data signal tubuh siswa (Awais dkk., 2021). Peneliti Awais dkk., (2021) mengusulkan kerangka kerja berbasis IoT untuk menganalisis emosi manusia dengan menggunakan sinyal fisiologis yang diproses melalui model pembelajaran mendalam berbasis LSTM. Model ini mampu mengenali berbagai emosi, seperti lucu, membosankan, relaksasi, dan menakutkan, dengan akurasi lebih dari 95%. Kerangka IoT yang dikembangkan memastikan komunikasi data yang cepat dan andal antara perangkat sensor dan hub IoT, dengan latensi kurang dari 1 milidetik. Sistem ini diharapkan dapat mendukung siswa, institusi pendidikan, dan layanan kesehatan, terutama dalam pembelajaran jarak jauh selama pandemi COVID-19 dan potensi wabah di masa depan.

Pendekatan lainnya adalah dengan memanfaatkan sinyal EEG (Fernandez dkk., 2024). Studi yang dilakukan Fernandez dkk., (2024) menyoroti meningkatnya minat dalam deteksi stres menggunakan sinyal fisiologis, terutama data EEG, untuk meningkatkan pemantauan kesehatan mental dan kesejahteraan. Data dikumpulkan melalui protokol khusus yang dirancang untuk menginduksi stres pada peserta, dengan biosinyal yang direkam dari EEG, EDA, dan EKG. Berbagai teknik ekstraksi fitur dan model pembelajaran mesin diuji untuk mengidentifikasi stres secara akurat, dengan hasil yang menunjukkan bahwa pemilihan fitur, model, dan ukuran jendela yang tepat sangat berpengaruh terhadap kinerja sistem. LightGBM terbukti sebagai metode paling efektif, mencapai akurasi tertinggi di berbagai kondisi, sementara data EEG memberikan hasil yang lebih baik dibandingkan EKG

dan EDA. Integrasi semua sinyal fisiologis menghasilkan performa terbaik, meskipun efektivitasnya tetap bergantung pada konteks pengukuran stres.

Penggunaan sinyal EEG dalam deteksi emosi siswa juga dilakukan oleh peneliti Moise dkk., (2024). Studi yang mereka lakukan bertujuan untuk mengembangkan model pengenalan emosi yang berkinerja tinggi dengan memanfaatkan data psikologis yang relevan dalam proses pendidikan. Tujuh emosi yang dianalisis meliputi kebosanan, kebingungan, frustrasi, rasa ingin tahu, kegembiraan, konsentrasi, dan kecemasan. Model ini menggunakan arsitektur 1D-CNN yang dioptimalkan untuk lima saluran EEG, dengan hasil akurasi tinggi, di mana pengenalan emosi kegembiraan mencapai 99,91%. Untuk menjelaskan prediksi model, studi ini menerapkan LIME, yang menunjukkan bahwa analisis frekuensi fitur dapat membantu memperjelas hasil klasifikasi. Namun, terdapat beberapa keterbatasan, seperti data yang tidak dikumpulkan langsung dalam skenario pembelajaran serta tantangan dalam mengumpulkan data EEG dari peserta pendidikan.

Data modalitas sinyal dapat berasal dari berbagai sumber, seperti EEG, ECG, EMG, atau sinyal dari sensor IoT. Proses koleksi data umumnya dimulai dengan pemilihan perangkat akuisisi yang sesuai, seperti biosensor, perangkat medis, atau sensor industri, yang mampu menangkap sinyal dengan resolusi dan kecepatan sampling yang memadai. Setelah itu, data dikumpulkan dalam kondisi yang telah ditentukan, seperti lingkungan laboratorium yang terkendali atau melalui perangkat wearable yang digunakan dalam aktivitas sehari-hari. Selama proses akuisisi, penting untuk memastikan kualitas sinyal dengan mengurangi *noise* melalui teknik penyaringan (*filtering*) atau *artifact removal*. Data yang telah dikumpulkan kemudian disimpan dalam format yang sesuai, seperti CSV, EDF, atau HDF5, untuk selanjutnya diproses dan dianalisis dalam penelitian atau pengembangan sistem berbasis sinyal.

Peneliti Awais dkk., (2021) mengumpulkan data uji menggunakan sensor fisiologis, yang mencakup berbagai modalitas, seperti EEG, EKG, EMG, GSR, RSP, SKT, BVP, dan PPG. Sebanyak tiga puluh peserta terlibat dalam eksperimen akuisisi data, di mana mereka diperlihatkan delapan video yang dirancang untuk membangkitkan empat kategori emosi yang berbeda yaitu, relaksasi, kebosanan,

kesenangan, dan ketakutan. Sinyal fisiologis direkam secara kontinu dengan frekuensi pengambilan sampel 1.000 Hz untuk menangkap respons emosional peserta secara objektif.

Lalu peneliti Fernandez dkk., (2024) mengumpulkan data menggunakan pengaturan komprehensif yang mencakup akuisisi sinyal fisiologis khususnya EEG, EKG, dan EDA dari subjek yang sama. Proses ini dilakukan dengan mengikuti protokol standar untuk memastikan konsistensi dalam seluruh data yang dikumpulkan. Konfigurasi elektroda yang digunakan untuk menangkap sinyal ini dirancang agar memungkinkan akuisisi yang akurat dan simultan, sehingga memfasilitasi analisis holistik terhadap respons stres subjek. Selain itu, wawancara pasca-eksperimen dan kuesioner dilakukan untuk mengumpulkan data kualitatif, yang kemudian dibandingkan dengan data fisiologis yang dikumpulkan selama pengujian.

Moise dkk., (2024) juga melakukan metode serupa dengan menggunakan alat sensor dalam pengumpulan datanya. Mereka mengumpulkan data uji menggunakan pendekatan multimodal, dengan fokus pada sinyal EEG dan berbagai sinyal fisiologis perifer. Dataset DEAP, yang digunakan dalam penelitian ini, mencakup sinyal EEG yang direkam dari 32 subjek saat mereka menonton 40 video musik. Sinyal EEG ditangkap menggunakan elektroda AgCl, yang diposisikan sesuai dengan sistem 10/20 internasional, dengan laju pengambilan sampel 512 Hz. Selain itu, sinyal fisiologis *perifer* seperti *galvanic skin response* (GSR), amplitudo respirasi, suhu kulit, EKG, dan EMG, juga direkam menggunakan sensor yang ditempatkan di wajah, leher, dan tangan kiri, serta sabuk respirasi untuk mengukur pola pernapasan. Metode pengumpulan data yang komprehensif ini bertujuan untuk meningkatkan keakuratan pengenalan emosi dalam proses pembelajaran.

Walaupun pengenalan emosi menggunakan signal asli tubuh manusia seperti EEG lebih efektif dibanding metode tradisional (Li dan Deng, 2022a), namun penggunaan sensor atau tools dalam menangkap sinyal tubuh manusia pasti memerlukan perangkat yang lebih banyak daripada pengenalan emosi lainnya seperti FER, SER, dan SA (Li & Deng, 2022a). Atas dasar tersebut, penulis juga tidak

memilih pendekatan menggunakan modalitas sinyal tubuh untuk diterapkan dalam penentuan *engagement* siswa berdasarkan emosinya.

II.3.4 Modalitas Visual

Dalam beberapa tahun ke belakang, modalitas visual merupakan modalitas terbanyak yang digunakan oleh beberapa peneliti dalam mendeteksi emosi pada lingkungan pendidikan (Yu dkk., 2024), terutama dengan pendekatan FER. Ketersediaan dataset publik yang relatif banyak dan masih banyaknya peluang penelitian tentang FER, menjadikan FER sebagai topik penelitian yang masih menarik untuk terus dikembangkan terutama di bidang pendidikan seperti yang dilakukan oleh Wang, (2021). Wang menyimpulkan bahwa perilaku belajar daring dapat dianalisis secara efektif menggunakan teknik pengenalan emosi berbasis gambar. Proses ini mencakup ekstraksi *keyframe* dari ekspresi wajah dengan menerapkan *enhanced local binary pattern* (LBP) dan *wavelet transform*. *Keyframe* yang diperoleh digunakan untuk mengekstrak fitur ekspresi rata-rata untuk mengoptimalkan akurasi dalam *emotion recognition*. Sistem analisis perilaku pembelajaran daring dikembangkan oleh Wang, (2021) adalah dengan mengintegrasikan *face-based emotion recognition methods*. Model klasifikasi emosi pada *image-based online learning* dirancang menggunakan *attention mechanism*, yang terbukti meningkatkan performa klasifikasi. Hasil penelitiannya menunjukkan bahwa model ini dapat diimplementasikan untuk *online learning behavior analysis*.

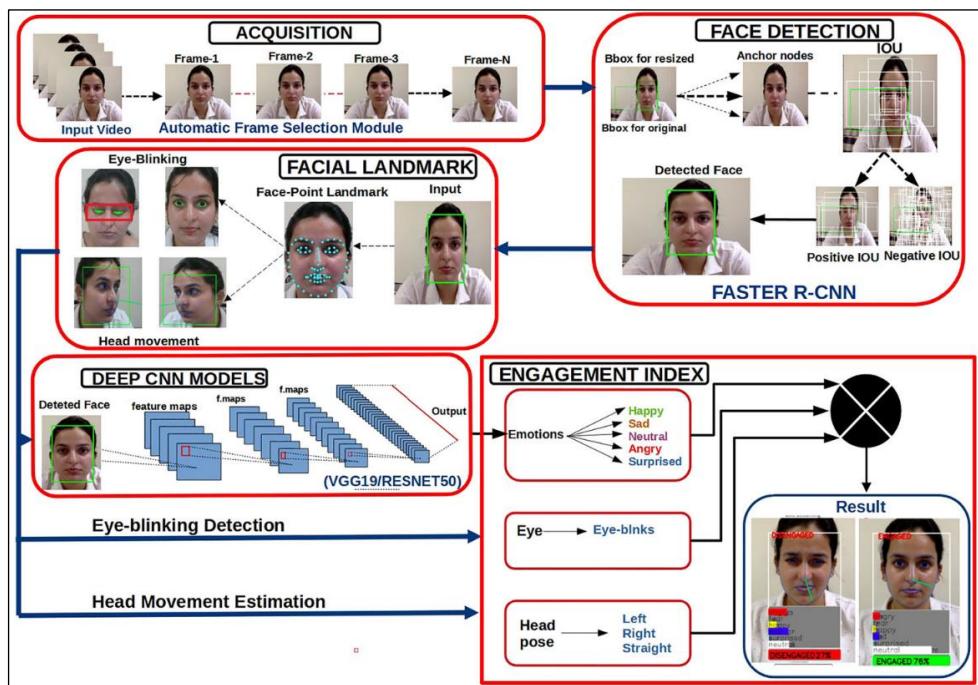
Pendekatan menggunakan modalitas *visual* juga bisa digunakan dalam mem membuat sebuah *visual analytics system* (Zeng dkk., 2021). Peneliti Zeng dkk., (2021) memperkenalkan EmotionCues, yaitu sebuah *visual analytics system* yang dirancang untuk merangkum dan menganalisis emosi siswa dalam suatu video kelas secara luring. Dengan fitur visualisasi interaktif, EmotionCues memungkinkan *end-users*, seperti guru dan para orang tua untuk mendapatkan *insights* mengenai *engagement* siswa dan emosinya.

Penerapan FER memiliki tantangan tersendiri terhadap emosi anak (Rathod dkk., 2022). Peneliti Rathod dkk., (2022) menyimpulkan bahwa model *deep learning* efektif untuk deteksi emosi pada anak-anak dalam media pembelajaran daring

dengan menggunakan dataset LIRIS serta dataset khusus yang dikembangkan untuk anak usia 7 hingga 10 tahun. Rathod dkk., (2022) menyatakan bahwa *facial expressions* pada anak-anak berbeda dari orang dewasa, sehingga digunakan 3D 468 *key geometric points* untuk meningkatkan *emotion recognition accuracy* untuk mendeteksi emosi pada anak-anak. Hasil akurasinya mencapai 90.98% melalui *comprehensive comparative analysis* dengan tujuh model arsitektur CNN.

Penerapan FER juga bisa dalam keadaan real-time pada video saat pembelajaran daring (Mehta dkk., 2022). Peneliti Mehta dkk., (2022) menyimpulkan bahwa level konsentrasi siswa dapat dideteksi secara efektif melalui *real-time facial expression analysis*. Tiga tingkat konsentrasi yang diidentifikasi dan diperoleh dari *facial emotions* adalah “*highly concentrated*”, “*nominally concentrated*”, dan “*not concentrated at all*”. Hasil penelitian ini menunjukkan efektivitas model *deep learning*, khususnya VGG16, dalam *facial emotion classification*, dengan akurasi tinggi dalam mendeteksi tingkat konsentrasi. Secara keseluruhan, hasil penelitian mengindikasikan bahwa mengintegrasikan emotion recognition ke dalam *e-learning systems* dapat meningkatkan pengalaman belajar dengan menyediakan umpan balik secara langsung terkait *engagement* siswa.

Selain dengan FER, deteksi *engagement* siswa juga bisa menggunakan *head pose* (Gupta dkk., 2023b) dan *eye gaze* (Tian dkk., 2024). Dengan teknik tersebut, sistem bisa menyimpulkan keadaan siswa apakah fokus terhadap pelajaran atau tidak. Peneliti Gupta dkk., (2023a), menggabungkan *head pose* dan *eye blink* (gambar II.2) dalam mendeteksi *engagement* siswa. Dilatarbelakangi atas meningkatnya permintaan pembelajaran daring selama pandemi COVID-19, Gupta dkk., (2023a) mengusulkan sistem *student engagement detection* untuk memberikan umpan balik. Sistem ini dirancang untuk mendukung *e-learning platforms* dengan mendeteksi tingkat keterlibatan siswa secara *real-time*. Pendekatan multimodal diterapkan dengan mengombinasikan data dari *facial expressions*, *eye movements*, dan *head movements* untuk mengklasifikasikan siswa sebagai *engaged* atau *disengaged*. Model *deep learning* digunakan untuk *facial emotion recognition*, *blink detection*, dan *head movement tracking*, yang kemudian dihitung menjadi *Engagement Index* (EI). Hasil eksperimen menunjukkan bahwa sistem ini mencapai akurasi sebesar 92.58%.



Gambar II.2 Deteksi *engagement* siswa menggunakan FER, *eye blink*, dan *head pose* (Gupta dkk., 2023a)

Data modalitas visual dapat berasal dari berbagai sumber, seperti kamera RGB, sensor, kamera termal, atau kombinasi dari beberapa perangkat. Beberapa literatur menunjukkan banyak peneliti yang menggunakan dataset publik untuk membuat model latih, seperti Affectnet (Trabelsi dkk., 2023), FER-2013 (Pordoy dkk., 2024), CK+ (Harb dkk., 2023). Selain itu, ada pula yang menggabungkan beberapa dataset publik, seperti yang dilakukan oleh Gupta dkk., (2023c), Meriem dkk., (2022), dan Aly dkk., (2023).

Untuk mendapatkan dataset publik pada modalitas visual, langkah pertama yang dapat dilakukan adalah mencari sumber dataset yang sesuai dengan kebutuhan penelitian, seperti Kaggle, ImageNet, AffectNet, atau Google Dataset Search. Setelah menemukan dataset yang relevan, periksa lisensi dan syarat penggunaannya untuk memastikan kepatuhan terhadap hak cipta. Selanjutnya, dataset dapat diunduh melalui API atau tautan resmi yang disediakan oleh penyedia dataset. Selain itu, jika dataset yang tersedia belum mencukupi, data tambahan dapat dikumpulkan dari sumber terbuka lainnya atau melalui teknik *web scraping* dengan tetap memperhatikan etika pengambilan data.

Selain dataset publik, beberapa peneliti juga menggunakan data private yang terdiri dari kumpulan wajah siswa yang diambil langsung dalam pengamatan selama beberapa waktu. Pada modalitas visual, peneliti lebih cendrung menggunakan kamera, baik webcam yang tertanam pada perangkat komputer maupun kamera profesional seperti yang dilakukan oleh Rathod dkk., (2022) yang menggunakan kamera DSLR untuk mengambil data wajah siswa yang merupakan anak-anak untuk dijadikan model latih.

Pengumpulan data menggunakan *webcam* pada lingkungan pembelajaran daring lazim dilakukan oleh beberapa peniliti. Gupta dkk., (2023a) pada penelitiannya mengumpulkan data uji dengan menangkap video *streaming* secara *real-time* dari *webcam* yang digunakan oleh siswa di lingkungan *E-learning*. Video yang ditangkap kemudian diekstrak menjadi *frame* wajah setiap 20 detik. Kumpulan *frame* wajah tersebut dianalisis untuk mendeteksi titik-titik kunci pada wajah, kedipan mata, dan gerakan kepala. Cara yang sama juga dilakukan oleh Gupta dkk., (2023b) pada penelitiannya yang serupa.

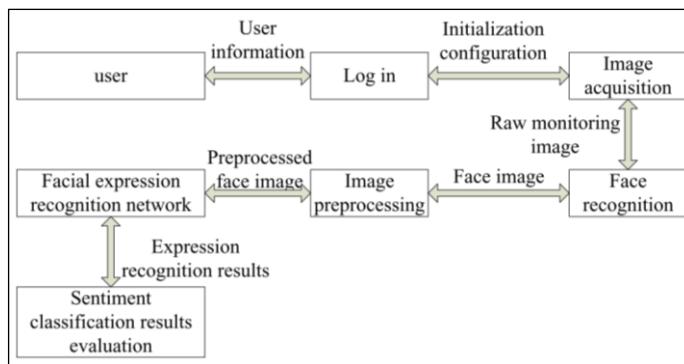
Penggunaan *webcam* juga dilakukan oleh Shobana dan Kumar, (2021), yaitu dengan cara *webcam* diaktifkan segera setelah pelajar masuk ke sistem kuis, merekam video penilaian berdasarkan pertanyaan, dan mengubah video menjadi *frame* gambar secara berkala untuk analisis lebih lanjut. Peneliti seperti Aly dkk., (2023), Trabelsi dkk., (2023), Sassi dkk., (2023), dan Li, (2025), juga menggunakan *webcam* dalam pengumpulan data baik untuk data latih, maupun data uji.

Cara pengumpulan data yang berbeda dilakukan oleh Bao dkk., (2024b). Mereka mengumpulkan data uji menggunakan perangkat pelacak mata, yaitu TobiiTX300 yang mencakup modul gerakan mata dan modul tampilan 23 inci. Perangkat beroperasi pada frekuensi sampling 60 Hz/s dan terhubung ke komputer desktop yang mengontrol prosedur eksperimental dan merekam data. Pengumpulan data melibatkan 68 mahasiswa berusia 20-23 tahun, dengan rasio pria dan wanita seimbang 1:1. Percobaan dilakukan pada lingkungan laboratorium terkontrol dengan kecerahan konstan. Apa yang dilakukan oleh Bao dkk., (2024b) menunjukkan bahwa ada cara lain untuk mengumpulkan data dari modalitas visual selain menggunakan kamera.

Berdasarkan uraian di atas, modalitas visual merupakan pilihan yang sangat logis untuk diterapkan pada sebuah sistem pembelajaran daring. Hal ini didasari pada banyaknya data publik yang bisa digunakan dalam membuat model FER sehingga proses pengumpulan data penelitian bisa lebih cepat dan efisien. Namun, penggunaan modalitas visual pada pembelajaran daring juga mempunyai beberapa kekurangan, terutama pada penerapan FER dalam keadaan oklusi, seperti wajah yang tertutup oleh tangan, kacamata, atau objek lain (Li dan Deng, 2022). Atas dasar tersebut, penelitian tentang cara mengatasi keadaan oklusi dalam penerapan FER masih dipandang perlu untuk dilakukan untuk menyempurnakan penerapan FER di berbagai bidang.

II.3.5 Multimodal

Beberapa peneliti menggunakan lebih dari satu elemen dalam pendekripsi emosi siswa atau yang lebih dikenal dengan sebutan *multimodal emotion recognition* (MER). Cui dkk., (2021) pada penelitiannya menggabungkan FER (visual) dan analisis sentimen (teks) untuk mendekripsi emosi siswa pada *Business English Class* dengan alur pada gambar II.3.



Gambar II.3 Alur pengenalan emosi dengan FER digabungkan dengan analisis sentimen (Cui dkk., 2021)

Penggunaan multimodal dalam deteksi di lingkungan pendidikan pernah dilakukan peneliti. Orosoo dkk., (2024) menggunakan dua kumpulan data utama dalam penelitiannya yaitu *Multimodal Emotion Lines Dataset* (MELD) dan *EnterFace'05*. MELD terdiri dari percakapan dalam serial televisi *Friends*, yang mencakup modalitas audio, visual, dan teks, memungkinkan analisis emosi yang komprehensif. Dataset *EnterFace'05* berisi urutan video yang menyediakan

informasi penting untuk pengenalan emosi, menunjukkan bahwa data video dikumpulkan untuk analisis emosi. Studi ini juga menekankan penggunaan *Natural Language Processing* (NLP) untuk mengonversi suara menjadi teks, yang menunjukkan bahwa data audio dikumpulkan dan diproses untuk meningkatkan akurasi pengenalan emosi. Secara keseluruhan, pengumpulan data melibatkan modalitas audio dan visual, kemungkinan melalui rekaman video dan input audio.

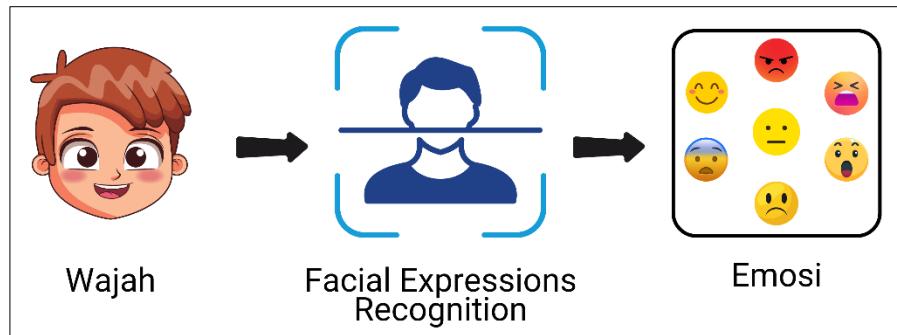
Hal berbeda dilakukan oleh peneliti Roy dkk., (2023). Mereka mengekstrak video menjadi text yang merupakan hasil konversi dari audio, lalu audio itu sendiri diekstrak terpisah dan dijadikan fitur MFCC kemudian diproses dengan *deep learning*, lalu *frame* dari video diproses dengan model *yolo*, yang pada akhirnya menggabungkan semua skor dari sisi teks, audio, dan video menjadi nilai akhir untuk menetukan keterlibatan dalam media pembelajaran daring. Roy dkk., (2023) mengumpulkan data uji dengan menerapkan model di server *cloud*, di mana pengguna diperlihatkan video dari kuliah video sumber terbuka berdasarkan minat mereka. Setelah menyelesaikan video, pengguna diminta untuk memberikan peringkat untuk keterlibatan mereka pada skala 1 hingga 10. Metode ini memungkinkan pengumpulan data keterlibatan pengguna yang terkait langsung dengan konten video yang dilihat.

Walaupun pengenalan emosi dengan multimodal berpotensi meningkatkan akurasi pengenalan emosi (Hunukumbure dkk., 2021), namun pengenalan emosi dengan multimodal memiliki beberapa tantangan, salah satunya adalah masalah sinkronisasi data dari berbagai sumber dan pengembangan model yang harus mampu menangani data yang heterogen (Das dkk., 2024), dan juga pasti memiliki beban komputasi yang lebih tinggi daripada pengenalan emosi dengan *unimodal* (Khan dkk., 2024).

II.4 *Facial Expressions Recognition* di Lingkungan Pendidikan Daring

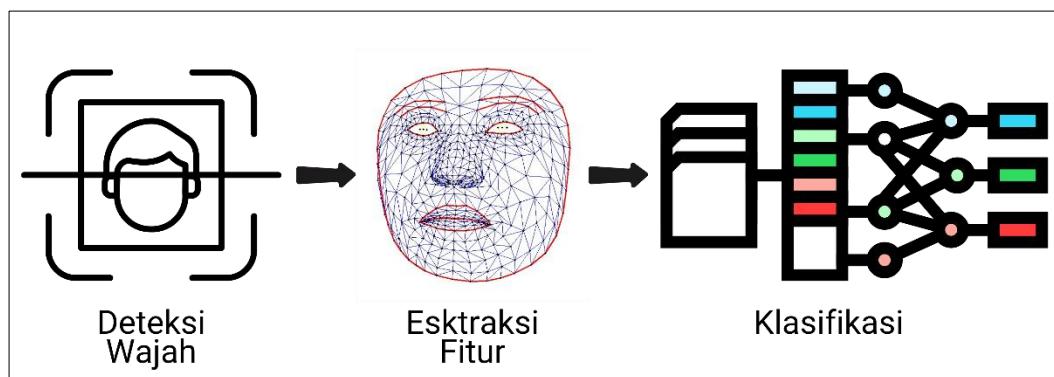
Facial Expression Recognition (FER) adalah bidang penelitian dalam computer vision dan kecerdasan buatan yang berfokus pada identifikasi dan klasifikasi ekspresi wajah manusia berdasarkan fitur visual. Teknologi ini memanfaatkan algoritma *machine learning* dan *deep learning* untuk mengenali berbagai emosi seperti bahagia, sedih, marah, terkejut, takut, jijik, dan netral. Proses utama dalam

sistem FER mencakup beberapa tahap, yaitu deteksi wajah, ekstraksi fitur, dan klasifikasi ekspresi (gambar II.5).



Gambar II.4 Sistem pengenalan emosi wajah

Deteksi wajah dilakukan dengan metode seperti *Haarcascade* (Sukumaran dan Manoharan, 2024) atau MTCNN (Zeng dkk., 2021) untuk menemukan posisi wajah dalam gambar. Setelah wajah terdeteksi, fitur-fitur penting seperti bentuk mata, mulut, dan alis diekstraksi menggunakan teknik seperti *Local Binary Pattern* (LBP) dan *Convolutional Neural Networks* (Sowjanya & Krithiga, 2024). Tahap akhir adalah klasifikasi ekspresi, yang biasanya dilakukan menggunakan model seperti *Random Forest* (Shobana & Kumar, 2021), atau jaringan saraf dalam seperti ResNet50 (Aly dkk., 2023) dan VGG19 (Gupta dkk., 2023c).



Gambar II.5 Proses utama *facial expressions recognition*

Dalam lingkungan pendidikan daring, FER dapat digunakan untuk menganalisis ekspresi wajah siswa guna memahami tingkat keterlibatan, kebingungan, atau kejemuhan mereka selama pembelajaran. Dengan menggunakan kamera dan algoritma berbasis *deep learning*, sistem FER dapat mendeteksi emosi siswa secara *real-time*, memungkinkan pengajar untuk menyesuaikan metode pengajaran agar lebih interaktif dan efektif. Teknologi ini juga dapat diterapkan dalam sistem *e-*

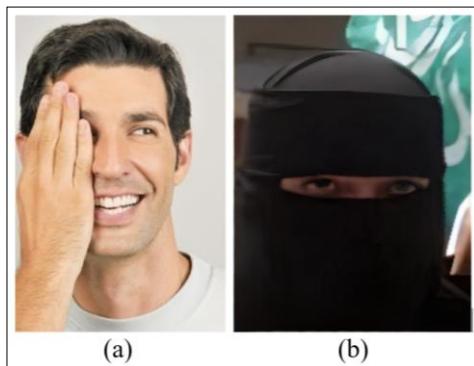
learning adaptif, di mana materi pelajaran secara otomatis disesuaikan berdasarkan respons emosional siswa. Selain itu, FER dapat membantu dalam evaluasi pengalaman belajar, dan mendukung pengajar dalam memberikan umpan balik yang lebih personal, serta meningkatkan kualitas interaksi dalam pembelajaran jarak jauh.

Meskipun banyak kemajuan telah dicapai, FER masih menghadapi beberapa tantangan, seperti perbedaan budaya dalam ekspresi emosi (Mattioli & Cabitza, 2024), variasi pencahayaan, serta oklusi atau halangan pada wajah. Salah satu tantangan terbesar adalah menangani wajah yang terhalang oleh tangan, kacamata, atau masker, yang dapat mengurangi akurasi pengenalan emosi wajah (Li dan Deng, 2022).

II.5 Oklusi pada Pengenalan Ekspresi Wajah

Oklusi adalah kondisi di mana suatu objek atau bagian tertentu tertutupi oleh objek lain sehingga tidak terlihat sepenuhnya. Dalam konteks pengenalan ekspresi wajah, oklusi terjadi ketika bagian wajah tertutup oleh elemen seperti kacamata, masker, tangan, rambut, atau benda lainnya. Dalam banyak kasus, metode pengenalan ekspresi wajah berbasis *deep learning* mengalami penurunan akurasi ketika menghadapi kondisi oklusi yang signifikan (Wickline dkk., 2025).

Oklusi dapat bersifat parsial atau total, tergantung pada seberapa besar bagian wajah yang tertutup. Oklusi parsial terjadi ketika hanya sebagian kecil wajah yang tertutup, seperti mata yang tertutup kacamata hitam atau mulut yang tertutup masker (Wickline dkk., 2025). Sementara itu, oklusi total dapat terjadi ketika hampir seluruh wajah tertutupi (C. Zhu dkk., 2022). Untuk mengatasi tantangan ini, berbagai teknik seperti augmentasi data, deteksi oklusi, dan rekonstruksi fitur wajah digunakan dalam sistem pengenalan ekspresi berbasis kecerdasan buatan.



Gambar II.6 (a) Keadaan wajah dengan oklusi *partial* oleh tangan, (b) keadaan wajah dengan oklusi total (Jain & Learned-Miller, 2010)

Untuk mengatasi permasalahan ini, berbagai pendekatan telah dikembangkan, seperti augmentasi data (Malakar dkk., 2024), penggunaan model yang lebih robust terhadap oklusi (Kim & Lee, 2023), serta teknik deteksi dan rekonstruksi wajah (N. Zhang dkk., 2023). Penelitian tentang rekonstruksi wajah lainnya juga pernah dilakukan, seperti yang dilakukan oleh K. Zhu dkk., (2023). Mereka melakukan rekonstruksi wajah dalam bentuk 3D. Menghapus objek oklusi di wajah dan menggunakan *template* (database bagian-bagian wajah) dalam bentuk 3D, berbagai jenis path wajah sebagai media untuk mengenerate (mengisi) bagian oklusi yang dihapus. Penelitian ini berhasil menghapus dan menggenerate wajah oklusi menjadi wajah tanpa oklusi, namun sayangnya tidak ada pengujian lebih lanjut terhadap nilai akurasi klasifikasi emosi wajah. Walaupun demikian metode yang dilakukan Kaifeng Zhu dkk, bisa dijadikan referensi dalam hal pendekatan yang tepat dalam menghapus oklusi pada wajah.

Ada cara lain untuk mengatasi oklusi pada wajah yaitu pada bagian *pre-processing* Peneliti Cao dkk., (2021) mencoba untuk mengenali emosi dari wajah yang *non-frontal* dengan berbasis gambar 3D. Mereka melakukannya dengan mengambil setengah bagian wajah dan menggabungkannya dalam bentuk 3D kemudian melakukan penguatan bentuk otot disekitar hidung untuk dapat meningkatkan akurasi pengenalan emosi dari wajah 3D. Hasilnya ada kenaikan akurasi terhadap beberapa emosi dasar setelah melakukan proses penguatan otot muka disekitar hidung.

Walaupun bentuk data wajah dalam 3D dapat mengandung lebih banyak informasi dan meningkatkan akurasi pengenalan emosi wajah (Cao dkk., 2021), namun mengubah gambar 2D ke bentuk 3D dapat merubah informasi asli dari data wajahnya seperti bentuk otot muka yang tidak sesuai dengan data aslinya. Data gambar dalam bentuk 3D juga memiliki kompleksitas komputasi yang lebih tinggi gambar 2D sehingga memerlukan waktu proses yang lebih lama dari gambar 2D (Morar dkk., 2017). Apa yang dilakukan oleh Cao dkk., (2021) tersebut menginspirasi penulis untuk mengusulkan sebuah metode yang sama namun dalam bentuk gambar 2D dengan harapan hasil akurasinya bisa menyamai dengan bentuk gambar 3D namun dengan kompleksitas pemrosesan yang lebih sederhana sehingga bisa diterapkan pada resource dengan spesifikasi rendah.

II.6 Klasifikasi Emosi dalam Deteksi *Engagement* Siswa

Berdasarkan beberapa literatur, dalam modalitas visual para peneliti melakukan berbagai macam cara untuk menilai *engagement* siswa dalam pembelajaran, ada yang langsung dengan menilai terhadap *eye gaze* dan *head pose* (Gupta dkk., 2023a), namun tidak sedikit juga yang menggunakan data emosi siswa yang didapat dari proses *emotion recognition*. *Emotion Recognition* adalah proses dimana sistem menganalisis ekspresi wajah untuk mengklasifikasikan emosi seseorang. Beberapa peneliti pada studi ini menggunakan nilai probabilitas setiap emosi pada siswa untuk menentukan tingkat engagement dari siswa dalam pembelajaran yang dilakukan. Untuk mendapatkan nilai emosi tersebut beberapa peneliti menggunakan teknik *machine learning* konvensional, seperti *Random Forest* (Shobana dan Kumar, 2021), atau bahkan metode *deep learning*, seperti CNN (Bao dkk., 2024b) dan ResNet-50 (Aly dkk., 2023). Model ini dilatih pada dataset ekspresi wajah seperti FER2013, AffectNet, atau CK+ yang mengandung berbagai kategori emosi seperti marah, bahagia, sedih, netral, takut, jijik, dan terkejut. Setelah wajah terdeteksi, fitur-fitur penting diekstraksi dan digunakan untuk melakukan klasifikasi emosi. Keberhasilan Emotion Detection sangat bergantung pada kualitas data, teknik augmentasi, serta arsitektur model yang digunakan untuk menangkap perbedaan ekspresi wajah secara akurat.

II.6.1 Machine Learning (ML)

Kelebihan ML dalam *automatic emotion recognition* (AER) terletak pada efisiensi komputasi, interpretabilitas, dan fleksibilitas dalam berbagai jenis data. Dibandingkan dengan *deep learning* (DL) yang membutuhkan jumlah data besar dan daya komputasi tinggi, ML dapat bekerja dengan baik bahkan pada dataset terbatas dan dengan *hardware* yang lebih sederhana (Ülgen Sönmez dan Varol, 2024). Adapun gambaran umum pada tugas klasifikasi menggunakan *machine learning* dapat dilihat pada gambar II.7.



Gambar II.7 Gambaran umum alur *machine learning*

Pendekatan ML Konvensional pernah dilakukan peneliti untuk mengklasifikasikan tingkat pemahaman siswa dalam sebuah tes dalam platform digital. Salah satunya adalah peneliti Shobana dan Kumar, (2021) yang memanfaatkan ML untuk menganalisis perilaku non-verbal dalam menilai tingkat akuisisi pengetahuan peserta didik selama proses pengerjaan kuis dalam platform digital (*I-Quiz*). Model yang digunakan adalah *Random Forest Classifier* yang dilatih menggunakan dataset berisi perilaku non-verbal serta skor pertanyaan-jawaban untuk memprediksi tingkat pemahaman peserta secara akurat. Sistem *I-Quiz* dirancang untuk menangkap perilaku non-verbal melalui kamera menghadap ke depan yang merekam ekspresi wajah, pergerakan mata, gestur, serta postur tubuh. Dari hasil rekaman tersebut, sistem menghasilkan fitur yang mencakup berbagai parameter seperti emosi, tingkat kedipan, skor kepercayaan diri, serta waktu yang dibutuhkan dalam menjawab pertanyaan. Dengan menggunakan fitur tersebut Shobana dan Kumar, (2021) melakukan klasifikasi nilai tingkat akuisisi pengetahuan peserta didik.

II.6.2 Deep Learning (DL)

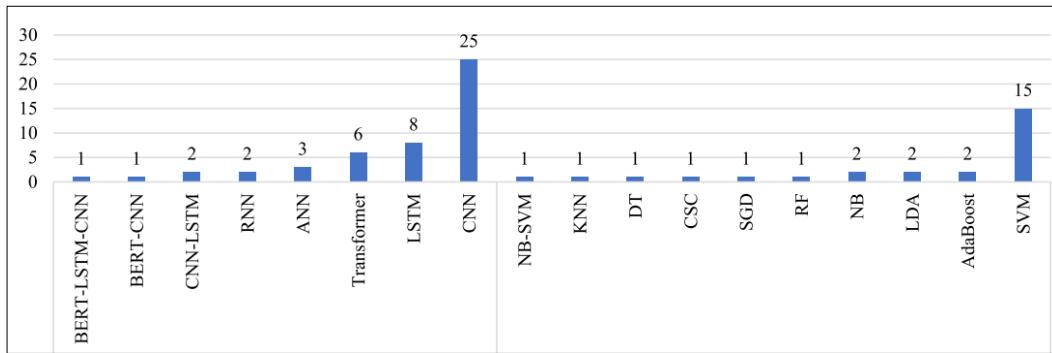
Kelebihan *Deep Learning* (DL) dalam *automatic emotion recognition* (AER) terletak pada kemampuannya mengekstrak fitur secara otomatis, menangkap pola kompleks dalam data, dan menghasilkan akurasi tinggi dibandingkan metode ML

klasik. Dalam penerapan FER, DL dapat menganalisis ekspresi wajah secara lebih mendalam, tanpa memerlukan proses ekstraksi fitur manual (S. Zhang dkk., 2024). Meskipun membutuhkan daya komputasi yang lebih besar daripada ML klasik, DL menawarkan fleksibilitas, skalabilitas, dan kemampuan generalisasi yang lebih baik, menjadikannya pilihan utama dalam banyak aplikasi AER berbasis gambar, audio, dan multimodal (Trigka & Dritsas, 2025).

Tabel II.1 Studi terkait yang menggunakan metode *deep learning* pada deteksi emosi dengan modalitas visual pada lingkungan pendidikan

Metode	Referensi	Matrik Evaluasi
CNN	(Bao dkk., 2024b)	Accuracy (81.9 %)
	(Bhardwaj dkk., 2021)	Accuracy (93.6 %)
	(Sowjanya dan Krithiga, 2024)	Accuracy (94.8%)
	(Rathod dkk., 2022)	Accuracy (90.98%)
	(Sukumaran dan Manoharan, 2024)	Accuracy (>80%)
DenseAttNet	(Mehta dkk., 2022)	Accuracy (>80%)
ResNet50	(Aly dkk., 2023)	Accuracy (97.72%)
VGG19	(Gupta dkk., 2023c)	Accuracy (92.58%)
YoloV5	(Trabelsi dkk., 2023)	Accuracy (76%)
Neural Architecture Search (NAS)	(Harb dkk., 2023)	Accuracy (86.29%)
Dual Path Stacked Attention Network (DPSAN)	(H. Zhu dkk., 2022)	Accuracy (93.26%)
Multi-Frame Transfer Learning (MFTL)	(Pordoy dkk., 2024)	Accuracy (87.8%)

Hasil rangkuman studi pada tabel II.1 menunjukkan bahwa metode CNN adalah metode yang paling banyak digunakan dalam klasifikasi emosi khususnya dalam penerapannya pada bidang pendidikan daring. Hal ini juga di dukung oleh data hasil penelitian yang dilakukan oleh Yu dkk., (2024) yang dapat dilihat pada gambar II.8.



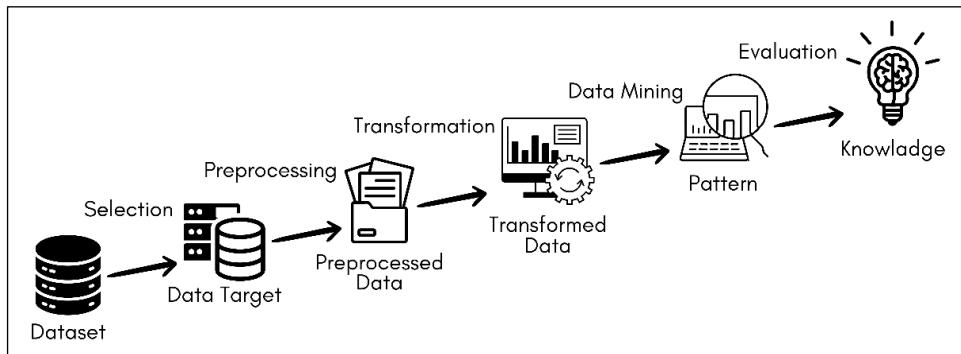
Gambar II.8 Distribusi penggunaan metode *machine learning* dan *deep learning* dalam rentang tahun 2010 – 2024 (Yu dkk., 2024)

II.6.3 Kombinasi *Machine Learning* Klasik dan *Deep Learning*

Penggabungan metode ML Klasik dan DL juga mungkin dilakukan untuk mendeteksi emosi wajah siswa. Sassi dkk., (2023) menggabungkan CNN dan Random Forest untuk menganalisis secara *real-time* emosi siswa saat pembelajaran daring. Dalam implementasinya, Sassi dkk., (2023) menggunakan model CNN-RF untuk pengenalan emosi, yang menunjukkan kinerja lebih tinggi serta kompleksitas yang lebih rendah dibandingkan dengan metode konvensional. Model CNN-RF yang diusulkan menunjukkan kinerja yang unggul dengan mencapai akurasi sebesar 71,86%, presisi 70,56%, *recall* 74,35%, dan skor F1 sebesar 72,09%. Hasil ini lebih tinggi dibandingkan dengan tolok ukur yang ada, menegaskan efektivitas pendekatan yang digunakan. Akurasi yang tinggi menunjukkan bahwa model mampu mengklasifikasikan sampel dengan baik, sementara presisi yang tinggi mengindikasikan rendahnya tingkat prediksi positif palsu. Keunggulan lain dari model ini terletak pada arsitektur yang efisien, yang memungkinkan peningkatan akurasi tanpa mengorbankan kinerja komputasi. Dengan hanya 5,17 juta parameter, model ini tetap ringan dan optimal untuk aplikasi dunia nyata, terutama dalam lingkungan dengan sumber daya yang terbatas. Evaluasi model dilakukan menggunakan dataset FER2013, yang berisi gambar wajah dengan berbagai ekspresi emosi. Selama proses pelatihan, teknik augmentasi data diterapkan untuk memperkaya kumpulan data, sehingga meningkatkan generalisasi model dalam mengenali emosi dari berbagai variasi gambar.

II.7 Knowledge Discovery in Databases

Knowledge Discovery in Databases (KDD) merupakan proses mengidentifikasi pola serta informasi yang bernilai dari data dalam jumlah besar (Gaikwad, 2017). KDD mencakup berbagai teknik dan tahapan untuk mengidentifikasi pola, hubungan, atau informasi tersembunyi dalam data.



Gambar II.9 Alur metode *knowladge discovery in databases*

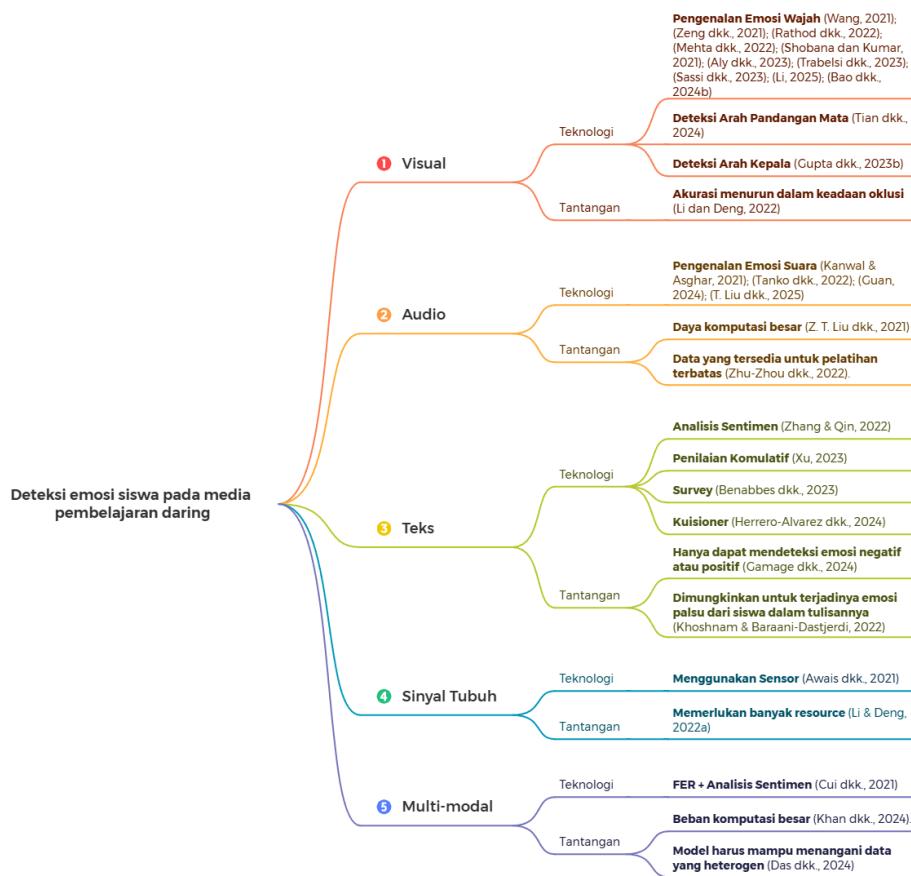
Menurut Gaikwad, (2017), proses ini terdiri dari beberapa langkah utama sesuai dengan gambar II.9, antara lain:

1. Seleksi Data : Memilih dan mengumpulkan data yang relevan untuk analisis.
2. *Pre-processing* (Pembersihan Data) : Membersihkan data dari noise, menangani nilai yang hilang, dan memastikan data dalam format yang sesuai.
3. Transformasi: Mengubah data ke dalam bentuk yang diinginkan untuk selanjutnya dilakukan proses *mining*.
4. Data *Mining* : Menggunakan teknik seperti *machine learning*, statistik, kecerdasan buatan, ataupun algoritma lainnya untuk menemukan pola atau hubungan dalam data.
5. Interpretasi dan Evaluasi : Menganalisis hasil setelah dari pengaruh proses transformasi terhadap dataset dan mengekstrak informasi untuk dijadikan pengetahuan baru.

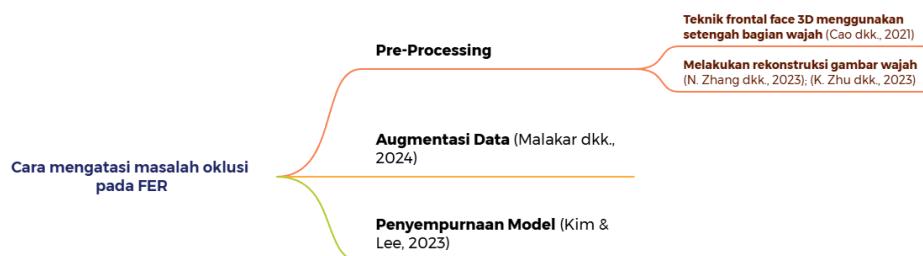
Dari penjelasan di atas, dapat disimpulkan bahwa KDD cocok untuk menguji efek dari sebuah transformasi yang dilakukan pada dataset tertentu, dalam hal ini adalah dataset *facial expresssions recognition* dengan cara membandingkan metrik evaluasi dalam keadaan *baseline* dan metrik evaluasi setelah tranformasi dilakukan.

II.8 Peta Literatur

Berdasarkan beberapa literatur yang telah dibahas, penulis melakukan pemetaan tentang referensi penelitian terdahulu beserta tantangannya yang berkaitan dengan pilihan-pilihan yang ada untuk mendeteksi emosi siswa pada pembelajaran daring, hasilnya dapat dilihat pada gambar II.10. Selain itu, penulis juga membuat peta literatur tentang penelitian yang sudah dilakukan untuk mengatasi masalah wajah yang teroklusi dalam penerapan FER yang dapat dilihat pada gambar II.11.



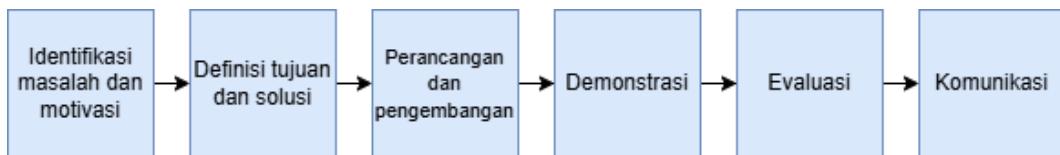
Gambar II.10 Peta literatur deteksi emosi siswa pada media pembelajaran daring



Gambar II.11 Peta literatur tentang cara mengatasi masalah oklusi pada FER

BAB III Metodologi Penelitian

Semua metodologi yang akan digunakan dalam penelitian dibahas dalam bab ini. Gambar III.1 menunjukkan tahapan penelitian *Design Science Research Methodology* (DSRM) yang akan diterapkan dalam penelitian ini.



Gambar III.1 *Design Science Research Methodology*

Gambar III.1 menunjukkan bahwa DSRM terdiri dari enam tahapan penelitian: identifikasi masalah, definisi tujuan dan solusi, perancangan dan pengembangan, demonstrasi, evaluasi, dan komunikasi. Subbab berikutnya memberikan penjelasan tentang setiap langkah yang harus dilakukan.

III.1 Identifikasi Masalah dan Motivasi

Pada tahap identifikasi masalah dan motivasi, fokus akan berada pada identifikasi masalah dan motivasi untuk melakukan penelitian. Tabel III.1 menunjukkan langkah-langkah *input*, proses, dan *output* pada tahap identifikasi masalah dan motivasi.

Tabel III.1 Tahap identifikasi masalah dan motivasi

<i>Input</i>	<ul style="list-style-type: none">- Dokumen RENSTRA dan RENIP IPDN- Riset penelitian- Dokumen ilmiah
Proses	Kajian literatur
<i>Output</i>	<ul style="list-style-type: none">- Informasi tentang arah pengembangan teknologi informasi di IPDN- Informasi mekanisme umpan balik pada pembelajaran daring- Alasan pemilihan FER sebagai pendekatan umpan balik pada pembelajaran daring

Berdasarkan tabel III.1, *input* pada tahap ini diperoleh dengan melakukan pengkajian literatur terkait topik yang akan dibahas, yaitu penerapan mekanisme umpan balik pada LMS IPDN. Pada penelitian ini masalah diperoleh dari telaahan dokumen rencana strategis IPDN yang salah satunya butirnya disebutkan yaitu penerapan *smart campus* IPDN. Salah satu bentuk penerapan *smart campus* di IPDN adalah penggunaan LMS sebagai pendukung pembelajaran. Berdasarkan hasil wawancara langsung kepada dosen IPDN dan digabungkan dengan studi literatur, ditemukan bahwa LMS IPDN memiliki kekurangan yaitu tidak adanya mekanisme umpan balik saat pembelajaran daring untuk menilai *engagement* dan emosi siswa dalam menjalani pembelajaran khususnya pada pembelajaran *asynchronous*.

Berdasarkan beberapa literatur menyebutkan bahwa penerapan deteksi emosi siswa pada pembelajaran daring bisa melalui pendekatan *facial expression recognition* (FER), *sound emotion recognition* (SER), *analysis sentiment*, atau gabungan beberapa pendekatan tersebut. Memperhitungkan kelebihan dan kelemahan masing-masing pendekatan tersebut, disimpulkan bahwa pendekatan *facial expression recognition* adalah solusi yang paling logis untuk diterapkan di LMS IPDN. Namun penerapan FER pada lingkungan pendidikan khususnya pada pembelajaran daring masing memiliki kekurangan pada kondisi wajah yang teroklusi, dimana kondisi tersebut lazim terjadi saat pembelajaran daring. Contohnya adalah saat siswa menyangga kepalanya dengan tangan karena faktor kelelahan atau terlalu fokus, maupun faktor lainnya. Kondisi wajah yang teroklusi tersebut dapat menurunkan tingkat probabilitas dari suatu emosi, sehingga diperlukan penelitian lebih lanjut untuk dapat mengatasi masalah wajah teroklusi pada penerapan FER di LMS IPDN.

Berdasarkan pengkajian data pada tahapan *input*, didapatkan *output* berupa permasalahan yang dapat diidentifikasi, yaitu diperlukannya pengembangan LMS IPDN yang memiliki fitur untuk dapat menilai *engagement* dan emosi siswa yang *robust* terhadap wajah yang teroklusi, khususnya wajah yang tertutupi oleh tangan sebagian, sehingga didapatkan emosi siswa yang lebih akurat, dan hasilnya dapat membantu pengajar dalam mengevaluasi materi maupun siswa yang bersangkutan

dengan lebih tepat. Hal ini sejalan dengan arah rencana strategis IPDN yaitu menerapkan teknologi untuk menjadi *smart campus*.

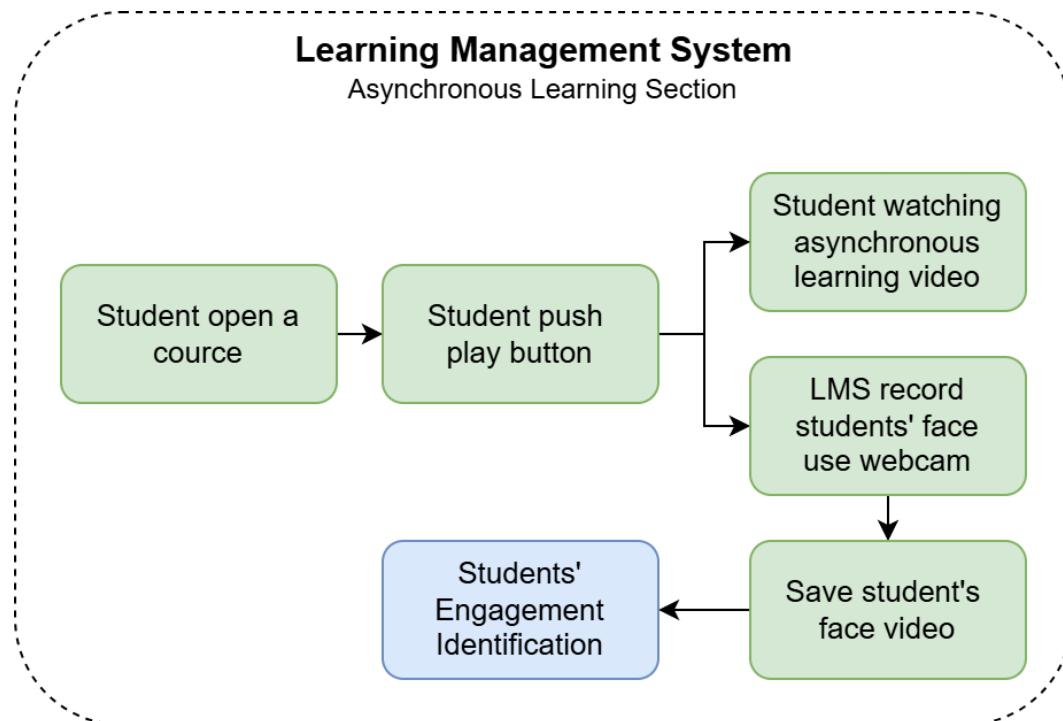
III.2 Definisi Tujuan dan Solusi

Fokus tahap kedua DSRM adalah mendefinisikan tujuan dan solusi berdasarkan masalah yang telah diidentifikasi sebelumnya. Identifikasi tujuan penelitian kemudian memberikan solusi untuk mengatasi masalah yang ada. Pada tahap tujuan dan solusi, Tabel III.2 menunjukkan tahapan *input*, proses, dan *output*.

Tabel III.2 Tahap tujuan dan solusi

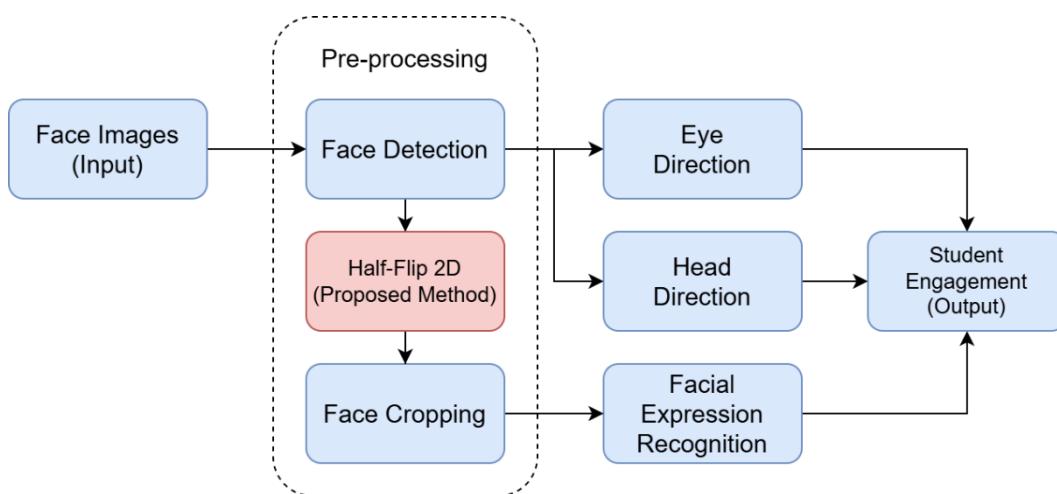
<i>Input</i>	<ul style="list-style-type: none">- Informasi tentang arah pengembangan teknologi informasi di IPDN- Informasi mekanisme umpan balik pada pembelajaran daring- Alasan pemilihan FER sebagai pendekatan umpan balik pada pembelajaran daring
Proses	<ul style="list-style-type: none">- Eksplorasi cara untuk menangani keadaan oklusi pada FER- Eksplorasi model FER yang tersedia dan sesuai
<i>Output</i>	Rancangan solusi

Berdasarkan hasil pada tahap identifikasi masalah dan motivasi, hal selanjutnya yang dilakukan adalah melakukan eksplorasi cara untuk menangani keadaan oklusi pada FER, dan eksplorasi model FER yang tersedia dan sesuai. Dari studi literatur yang telah dilakukan dan menghasilkan beberapa kemungkinan cara yang bisa diterapkan dalam menangani oklusi pada wajah, penulis memilih cara pada bagian *pre-processing* dengan harapan improvisasi pada bagian ini tidak akan menambah beban komputasi yang lebih banyak jika dibandingkan dengan melakukan improvisasi pada bagian pembuatan atau fine-tuning pada model klasifikasinya. Adapun gambaran umum solusi yang penulis ajukan dapat dilihat pada gambar III.2.



Gambar III.2 Gambaran umum ajuan solusi

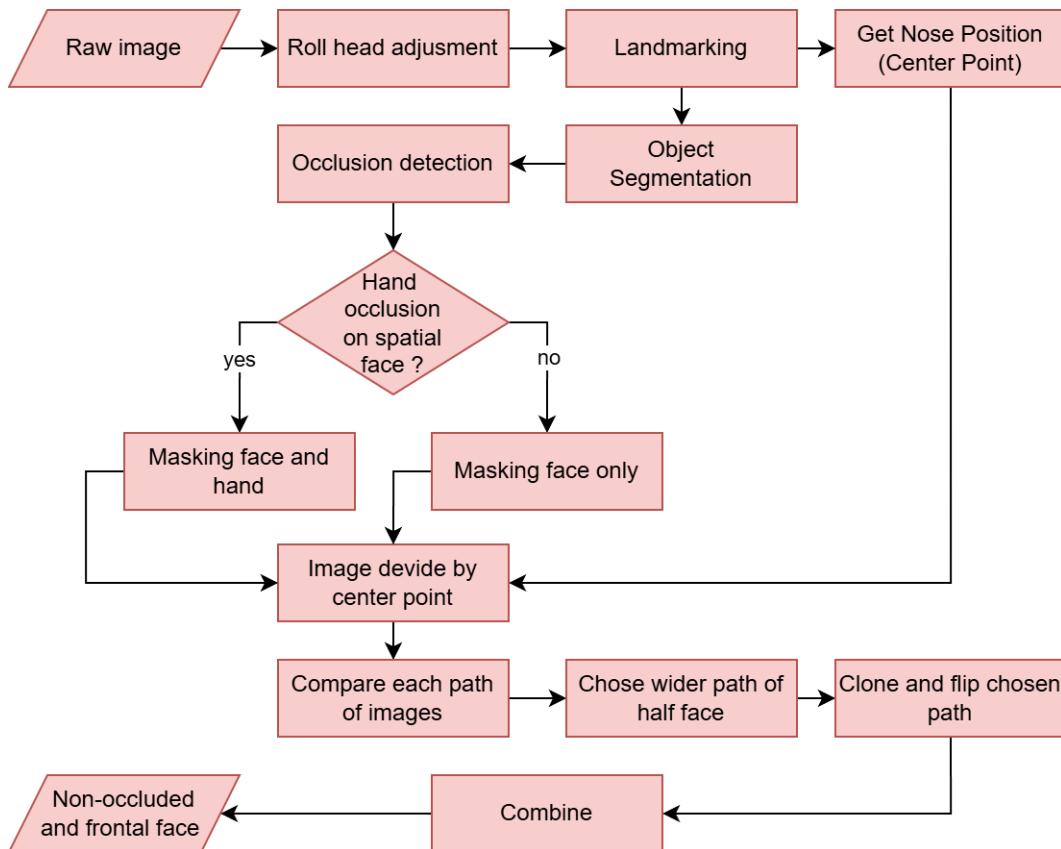
Berdasarkan gambar III.2, akan diperjelas kembali pada bagian *student's engagement identification*, dan menghasilkan sebuah rancangan penerapan FER pada gambar III.3.



Gambar III.3 Rancangan umum ajuan solusi

Pada gambar III.3 terdapat proses *eye direction*, yaitu proses untuk mendapatkan arah pandangan mata, dan juga ada proses *head direction*, yaitu proses untuk mendapatkan data arah kepala menghadap. Terdapat juga proses *Half-flip 2D* yang

merupakan metode usulan pada bagian *pre-processing* untuk menangani masalah oklusi pada wajah saat penerapan FER di LMS IPDN. Penulis mencoba membuat rancangan metode sesuai dengan gambar III.4.



Gambar III.4 Metode ajuan dalam penanganan masalah oklusi pada wajah pada penerapan FER

Penjelasan pada gambar III.4 adalah sebagai berikut:

1. Gambar *input* akan dilakukan penyesuaian arah *roll* kepala, sehingga kepala yang miring akan dibuat tegak lurus.
2. Setelah itu, gambar akan dilakukan *landmarking* untuk mendapatkan titik pada hidung yang akan dijadikan titik dengan suatu wajah.
3. Hasil dari *landmarking* juga selanjutnya dilakukan proses *object segmentation*, untuk memisahkan wajah dan bagian tubuh lain seperti rambut.
4. Selanjutnya hasil dari *segmentation* dilakukan pengecekan apakah ada tangan yang menutupi sebagian wajah.

5. Jika ada maka akan dilakukan *masking* pada wajah dan tangan, namun jika tidak ada oklusi tangan, maka hanya dilakukan masking pada wajah.
6. Memanfaatkan titik tengah wajah pada poin 2, gambar wajah akan dibelah menjadi dua sesuai dengan titik tengah pada poin 2, dan akan menghasilkan dua bagian wajah, wajah kanan, dan wajah kiri.
7. Selanjutnya dilakukan komparasi luas area wajah pada masing-masing hasil dari poin 6. Bagian wajah yang memiliki luas area yang lebih besar akan dijadikan bahan untuk dilakukan *clonning* dan *flip* sehingga menghasilkan dua buah gambar yang *mirror*.
8. Hasil pada poin 7 dilakukan penggabungan sehingga menghasilkan sebuah wajah utuh yang menghadap kedepan.

Wajah dari hasil metode usulan pada gambar III.4 akan dilakukan proses FER sesuai dengan gambar III.3 dengan memanfaatkan model FER yang sudah ada untuk mendapatkan data emosi wajah dari siswa. Hasil data emosi wajah siswa akan digabungkan dengan data arah pandangan mata dan data arah kepala menghadap untuk disimpulkan *engagement* dari siswa yang melakukan pembelajaran daring dengan materi *asynchronous* video.

III.3 Perancangan dan Pengembangan

Tahap ketiga dalam DSRM adalah mendefinisikan perancangan dan pengembangan berdasarkan rancangan solusi yang telah didefinisikan sebelumnya. Pada tahap ini akan berfokus pada pengembangan metode *pre-processing* untuk mengatasi masalah oklusi wajah oleh tangan pada penerapan FER di LMS. Tabel III.3 merupakan tahapan *input*, *process*, dan *output* pada tahap perancangan dan pengembangan.

Tabel III.3 Tahap perancangan dan pengembangan

<i>Input</i>	Rancangan solusi
Proses	Pengembangan metode <i>pre-processing</i> untuk mengatasi masalah oklusi wajah oleh tangan pada penerapan FER di LMS
<i>Output</i>	Metode <i>pre-processing</i> dapat mengatasi masalah oklusi wajah oleh tangan dalam penerapan FER

Berdasarkan tabel III.3, *input* pada tahap ini adalah berasal dari rancangan solusi pada bagian sebelumnya. Selanjutnya akan dilakukan pengembangan metode *pre-processing* untuk mengatasi masalah oklusi wajah oleh tangan pada penerapan FER di LMS. Pengembangan metode tersebut akan dilakukan dengan metode *Knowledge Discovery in Databases* (KDD), adapun tahapan pembangunannya sesuai dengan tabel III.4.

Tabel III.4 Tahapan pada metode KDD

Tahap	Deskripsi	Hasil
<i>Selection</i>	<ul style="list-style-type: none"> - Menggunakan dataset AffectNet dan/atau FER2013 - Memilih subset gambar yang mengandung wajah dengan oklusi tangan - Dimungkinkan menambahkan oklusi secara sintetik menggunakan augmentasi data 	Kumpulan data wajah teroklusi tangan
<i>Pre-processing</i>	<ul style="list-style-type: none"> - Menerapkan deteksi wajah menggunakan <i>haarcascade opencv</i> - Mendeteksi keberadaan tangan menggunakan <i>MediaPipe</i> 	Data wajah yang sudah diproses dengan anotasi oklusi.
<i>Transformation</i>	<ul style="list-style-type: none"> - Mengambil bagian wajah yang tidak teroklusi - Melakukan <i>flip</i> dan <i>combine</i> 	Dataset yang telah diproses dengan metode baru.
<i>Data Mining</i>	<ul style="list-style-type: none"> - Melatih model CNN atau ResNet50 untuk klasifikasi emosi - Membandingkan performa dengan dan tanpa metode <i>pre-processing</i> usulan 	Model FER yang diuji dengan <i>pre-processing</i> usulan

<i>Interpretation & Evaluation</i>	<ul style="list-style-type: none"> - Menggunakan metrik evaluasi akurasi validasi untuk membandingkan model - Jika metode <i>pre-processing</i> meningkatkan akurasi deteksi emosi, berarti berhasil! - Jika tidak, akan dilakukan iterasi ulang pada tahap <i>pre-processing</i> dan <i>transformation</i>. 	Analisis dampak preprocessing terhadap akurasi model
--	---	--

III.4 Demonstrasi

Tahap keempat dalam metodologi DSRM adalah demonstrasi. Pada tahap ini akan berfokus pada uji coba terhadap model yang telah dikembangkan. Pada Tabel III.5 merupakan *input*, proses, dan *output* pada tahap demonstrasi.

Tabel III.5 Tahapan demonstrasi

<i>Input</i>	Metode <i>pre-processing</i> dapat mengatasi masalah oklusi wajah oleh tangan dalam penerapan FER
Proses	Uji coba metode yang telah dikembangkan
<i>Output</i>	Data emosi wajah siswa

Berdasarkan Tabel III.5, input pada tahapan ini yaitu metode *pre-processing* yang telah dikembangkan pada hasil tahapan perancangan dan pengembangan. Berdasarkan input tersebut, dilakukan uji coba terhadap metode yang telah dibangun. Hasil dari tahapan demonstrasi adalah mendapatkan data *engagement* siswa yang dapat digunakan sebagai bahan evaluasi oleh pengajar ketika melakukan pembelajaran daring menggunakan LMS.

III.5 Evaluasi

Tahap kelima dari DSRM yaitu tahap evaluasi. Pada tahap ini akan berfokus pada pengukuran performa dari model yang telah dikembangkan. Pada Tabel III.6 merupakan tahapan *input*, proses, dan *output* pada tahap evaluasi.

Tabel III.6 Tahap evaluasi

<i>Input</i>	Data emosi wajah siswa
Proses	Evaluasi probabilitas emosi wajah siswa
<i>Output</i>	Persentase pengaruh terhadap probabilitas emosi wajah siswa setelah menerapkan metode usulan

Proses pada tabel III.6 adalah mengevaluasi data emosi wajah siswa hasil proses demonstrasi untuk mengetahui perubahan probabilitas emosi setelah menerapkan metode usulan. Evaluasi akan dilakukan dengan membandingkan probabilitas emosi pada data wajah dengan keadaan *baseline* dan data wajah setelah diterapkan proses *pre-processing* dengan metode usulan.

III.6 Komunikasi

Tahap terakhir dalam DSRM adalah komunikasi. Pada tahap ini akan berfokus pada dokumentasi hasil yang kemudian dikomunikasikan dalam bentuk tulisan ilmiah. Tabel III.7 merupakan tahapan *input*, proses, dan *output* pada tahap komunikasi.

Tabel III.7 Tahap komunikasi

<i>Input</i>	<ul style="list-style-type: none"> - Rancangan solusi - Metode preprocessing data wajah dengan memanfaatkan bagian setengah wajah - Hasil evaluasi perubahan probabilitas emosi sebelum dan sesudah menerapkan metode usulan
Proses	Penyusunan dokumen ilmiah
<i>Output</i>	<ul style="list-style-type: none"> - Tesis - Paper

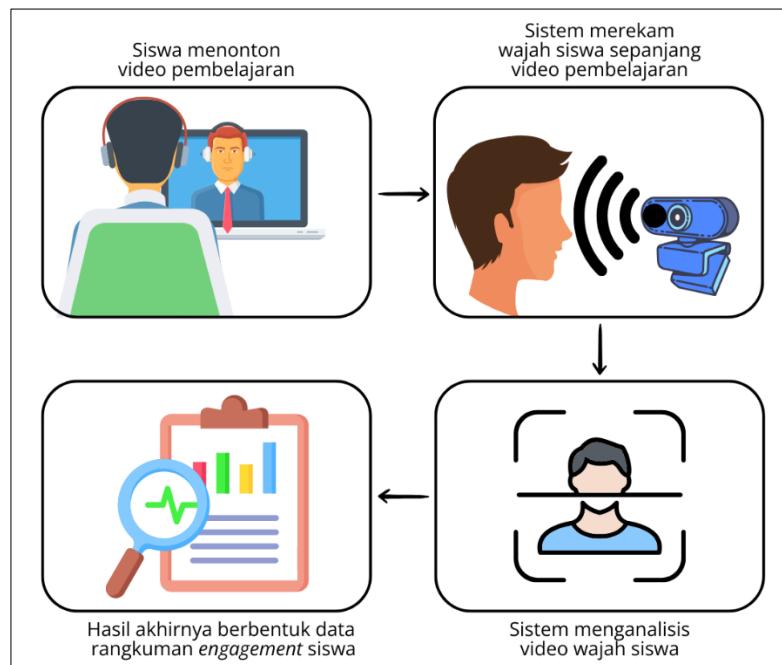
Berdasarkan Tabel III.7, *input* pada tahap evaluasi diperoleh dari seluruh tahapan sebelumnya yang kemudian akan dilakukan penyusunan dokumen ilmiah yang menghasilkan dokumen tesis dan publikasi *paper*.

BAB IV Analisis dan Desain Awal

Bagian ini akan menjelaskan gambaran penelitian yang akan dilakukan. Pada bagian ini juga akan dijelaskan lingkungan yang akan di analisis dan kondisi-kondisi kapan FER akan digunakan untuk menentukan *engagement* dari siswa.

IV.1 Lingkungan Penerapan

Lingkungan penerapan penelitian ini adalah lingkungan pembelajaran daring dengan memanfaatkan *platform learning management system*. Lalu, kondisi penerapan *facial expressions recognition* pada *learning management system* adalah khusus pada kondisi pembelajaran *asynchronous* yang mewajibkan siswa untuk menonton video rekaman pembelajaran yang diberikan oleh pemateri seperti apa yang tertera pada gambar IV.1



Gambar IV.1 Alur penerapan mekanisme umpan balik pada LMS

IV.2 Analisis Wajah

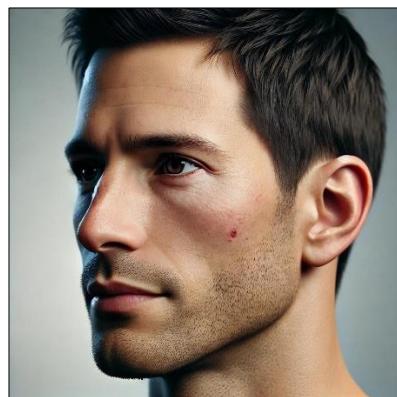
Ada beberapa parameter untuk menentukan *engagement* siswa, pertama adalah arah kepala (*head pose*), lalu arah pandangan mata (*eye gaze*), dan emosi wajah. Sebagai contoh, gambar IV.2 memperlihatkan hasil analisis wajah menggunakan tiga

parameter yang digunakan (mata, kepala, emosi) dari sebuah proses ekstraksi wajah baik yang bisa berasal dari gambar maupun video.



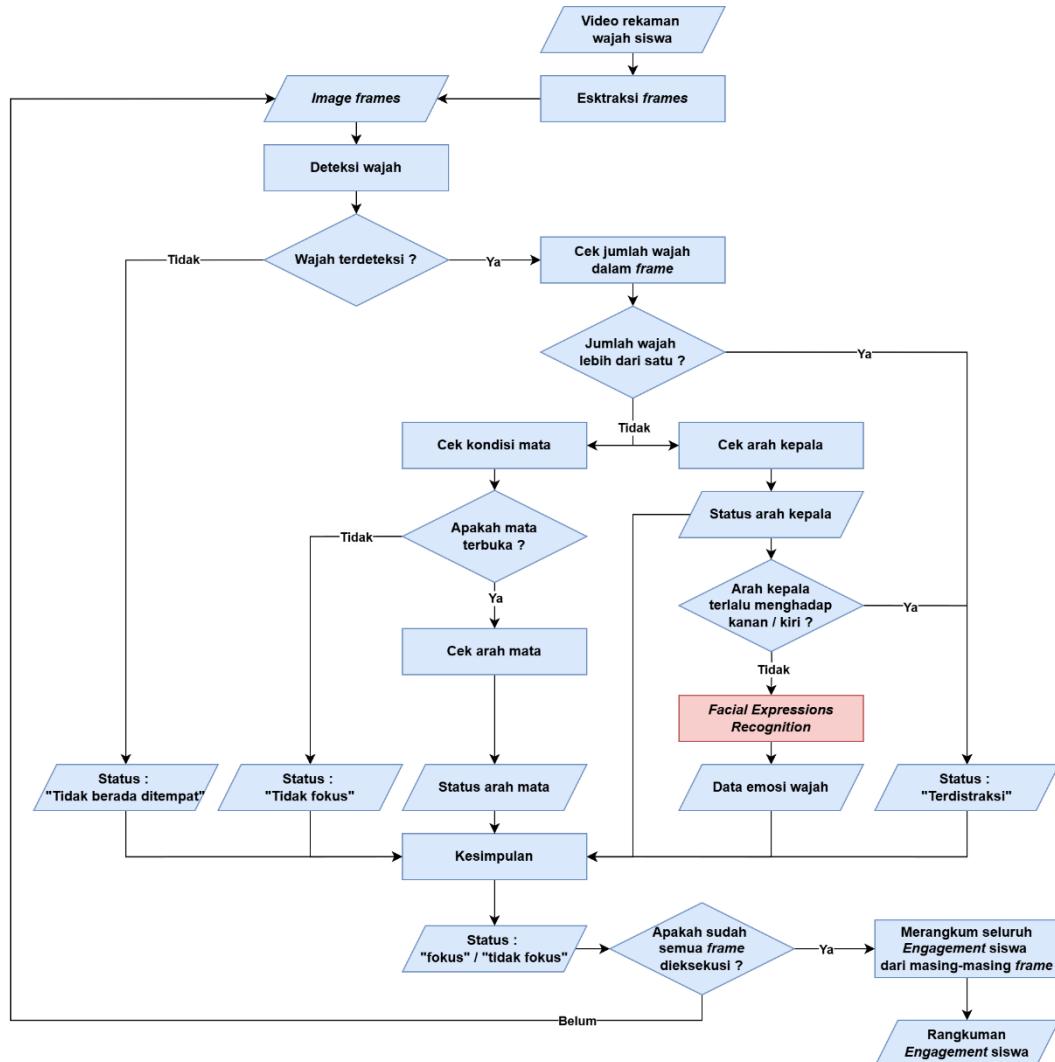
Gambar IV.2 Contoh hasil analisis wajah dengan parameter arah pandangan mata, arah kepala, dan emosi wajah

Adapun penggunaan pendekatan FER tidak selalu pada setiap gambar yang di eksekusi, contoh: jika kondisi kepala dan mata sedang menatap ke samping (tidak melihat layar monitor) seperti pada gambar IV.3, maka sistem akan langsung menyimpulkan bahwa siswa tersebut dalam keadaan tidak fokus tanpa harus dilakukan proses pengenalan emosi wajah.



Gambar IV.3 Contoh kondisi wajah yang arah kepala dan mata tidak menghadap monitor atau menghadap kesamping (gambar dibuat dengan AI)

Pengenalan emosi wajah hanya diterapkan ketika mata dan kepala masih terdeteksi dalam keadaan melihat monitor seperti gambar IV.2. Lebih lengkapnya algoritma untuk menentukan *engagement* siswa dapat dilihat pada gambar IV.4.



Gambar IV.4 Diagram alir algoritma kondisi penggunaan FER untuk menentukan *engagement* siswa

Penjelasan diagram alir pada gambar IV.4 adalah sebagai berikut:

1. Video rekaman wajah siswa diekstrak menjadi kumpulan *frame* gambar sepanjang video rekaman per 20 *frame*.
2. Kumpulan frame tersebut diproses satu per satu untuk dianalisis
3. Satu frame tersebut dieksekusi pertama dengan melakukan deteksi wajah
4. Jika tidak ada wajah yang terdeteksi, maka sistem akan otomatis menyimpulkan bahwa pada *frame* tersebut siswa tidak berada di tempat,

namun jika ada wajah yang terdeteksi, maka sistem akan melanjutkan ke proses memeriksa jumlah wajah yang terdeteksi.

5. Jika jumlah wajah yang terdeteksi lebih dari satu wajah, maka sistem akan menyimpulkan bahwa pada *frame* tersebut, siswa terdistraksi oleh orang lain
6. Jika jumlah wajah pada poin 5 hanya terdeteksi satu wajah, maka proses selanjutnya adalah deteksi keadaan mata dan deteksi arah kepala
7. Pada saat deteksi keadaan mata didapati mata dalam keadaan tertutup, maka sistem akan langsung menyimpulkan bahwa pada *frame* tersebut, siswa tidak sedang fokus, namun ketika mata terdeteksi dalam kondisi terbuka, maka prosesnya akan dilanjutkan dengan deteksi arah pandangan mata yang kemudian menghasilkan kondisi arah mata ke kanan atau ke kiri
8. Dalam deteksi arah kepala, jika kepala terlalu ke kanan atau ke kiri (melebihi ambang batas), maka sistem akan langsung menyimpulkan bahwa pada *frame* tersebut siswa dalam keadaan terdistraksi, namun jika arah kepala tidak melebihi ambang batas yang ditentukan, maka akan didapatkan data arah kepala, ke kanan atau ke kiri dan akan dilanjutkan dengan proses pengenalan emosi wajah
9. Proses pengenalan emosi wajah akan dilakukan dengan menerapkan *pre-processing* dengan metode usulan untuk mendapatkan data emosi wajah siswa
10. Hasil deteksi arah mata (poin 7), arah kepala (poin 8), dan data emosi wajah (poin 9) akan dianalisis untuk menyimpulkan *engagement* siswa pada satu *frame*.
11. Proses akan diiterasi sebanyak frame yang tersedia hasil dari proses pada poin 1
12. Jika semua frame sudah diekseksusi, maka proses akhir adalah membuat rangkuman dari kumpulan *engagement* per *frame* sehingga dapat disimpulkan bahwa siswa tersebut dominan untuk fokus atau tidak fokus dalam pembelajaran daring, selain itu dalam keadaan fokus, juga akan ditampilkan emosi wajah dominan dari siswa yang bersangkutan.

DAFTAR PUSTAKA

- Afzal, S., Khan, H. A., Piran, M. J., & Lee, J. W. (2024). A Comprehensive Survey on Affective Computing: Challenges, Trends, Applications, and Future Directions. *IEEE Access*, 12, 96150–96168. <https://doi.org/10.1109/ACCESS.2024.3422480>
- Alotaibi, N. S. (2024). The Impact of AI and LMS Integration on the Future of Higher Education: Opportunities, Challenges, and Strategies for Transformation. *Sustainability (Switzerland)*, 16(23). <https://doi.org/10.3390/su162310357>
- Aly, M., Ghallab, A., & Fathi, I. S. (2023). Enhancing Facial Expression Recognition System in Online Learning Context Using Efficient Deep Learning Model. *IEEE Access*, 11(November), 121419–121433. <https://doi.org/10.1109/ACCESS.2023.3325407>
- Awais, M., Raza, M., Singh, N., Bashir, K., Manzoor, U., Islam, S. U., & Rodrigues, J. J. P. C. (2021). LSTM-Based Emotion Detection Using Physiological Signals: IoT Framework for Healthcare and Distance Learning in COVID-19. *IEEE Internet of Things Journal*, 8(23), 16863–16871. <https://doi.org/10.1109/JIOT.2020.3044031>
- Bao, J., Tao, X., & Zhou, Y. (2024a). An Emotion Recognition Method Based on Eye Movement and Audiovisual Features in MOOC Learning Environment. *IEEE Transactions on Computational Social Systems*, 11(1), 171–183. <https://doi.org/10.1109/TCSS.2022.3221128>
- Bao, J., Tao, X., & Zhou, Y. (2024b). An Emotion Recognition Method Based on Eye Movement and Audiovisual Features in MOOC Learning Environment. *IEEE Transactions on Computational Social Systems*, 11(1), 171–183. <https://doi.org/10.1109/TCSS.2022.3221128>
- Beekman, R., Tijssen, C. C., Visser, L. H., & Schellens, R. L. L. A. (2002). Dropped head as the presenting symptom of primary hyperparathyroidism [3]. *Journal of Neurology*, 249(12), 1738–1739. <https://doi.org/10.1007/s00415-002-0898-7>
- Bekmanova, G., Yergesh, B., Sharipbay, A., & Mukanova, A. (2022). Emotional Speech Recognition Method Based on Word Transcription. *Sensors*, 22(5). <https://doi.org/10.3390/s22051937>
- Benabbes, K., Housni, K., Hmedna, B., Zellou, A., & Mezouary, A. El. (2023). A New Hybrid Approach to Detect and Track Learner's Engagement in e-Learning. *IEEE Access*, 11(June), 70912–70929. <https://doi.org/10.1109/ACCESS.2023.3293827>
- Bhardwaj, P., Gupta, P. K., Panwar, H., Siddiqui, M. K., Morales-Menendez, R., & Bhaik, A. (2021). Application of Deep Learning on Student Engagement in e-learning environments. *Computers and Electrical Engineering*, 93(August 2020), 107277. <https://doi.org/10.1016/j.compeleceng.2021.107277>
- Cao, T., Liu, C., & Chen, J. (2021). Nonfrontal Expression Recognition in the Wild Based on PRNet Frontalization and Muscle Feature Strengthening. *Mathematical Problems in Engineering*, 2021. <https://doi.org/10.1155/2021/6620752>

- Costley, J., Southam, A., Bailey, D., & Haji, S. A. (2022). How use of learning management system mediates the relationships between learner interactions and learner outcomes. *Interactive Technology and Smart Education*, 19(2), 184–201. <https://doi.org/10.1108/ITSE-12-2020-0236>
- Cui, Y., Wang, S., & Zhao, R. (2021). Machine Learning-Based Student Emotion Recognition for Business English Class. *International Journal of Emerging Technologies in Learning*, 16(12), 94–107. <https://doi.org/10.3991/ijet.v16i12.23313>
- Das, A., Sarma, M. Sen, Hoque, M. M., Siddique, N., & Dewan, M. A. A. (2024). AVaTER: Fusing Audio, Visual, and Textual Modalities Using Cross-Modal Attention for Emotion Recognition. *Sensors*, 24(18), 5862. <https://doi.org/10.3390/s24185862>
- Fernandez, J., Martínez, R., Innocenti, B., & López, B. (2024). Contribution of EEG Signals for Students' Stress Detection. *IEEE Transactions on Affective Computing*, 1–12. <https://doi.org/10.1109/TAFFC.2024.3503995>
- Freire-obregón, J. S. J. L. D., & Castrillón-santana, M. (2024). Multimodal emotion recognition based on a fusion of audiovisual information with temporal dynamics. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-20227-6>
- Gaikwad, K. S. (2017). Knowledge Discovery in Database Systems. *International Journal of Computing, Communication and Instrumentation Engineering*, 4(2), 532–535. <https://doi.org/10.15242/ijccie.ae0417103>
- Gamage, G., De Silva, D., Mills, N., Alahakoon, D., & Manic, M. (2024). Emotion AWARE: an artificial intelligence framework for adaptable, robust, explainable, and multi-granular emotion analysis. *Journal of Big Data*, 11(1). <https://doi.org/10.1186/s40537-024-00953-2>
- Guan, Y. (2024). Characteristics of teaching English and American literature based on multimodality under language acquisition theory. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–16. <https://doi.org/10.2478/amns.2023.2.00978>
- Gupta, S., Kumar, P., & Tekchandani, R. (2023a). A multimodal facial cues based engagement detection system in e-learning context using deep learning approach. *Multimedia Tools and Applications*, 82(18), 28589–28615. <https://doi.org/10.1007/s11042-023-14392-3>
- Gupta, S., Kumar, P., & Tekchandani, R. (2023b). An optimized deep convolutional neural network for adaptive learning using feature fusion in multimodal data. *Decision Analytics Journal*, 8(May), 100277. <https://doi.org/10.1016/j.dajour.2023.100277>
- Gupta, S., Kumar, P., & Tekchandani, R. K. (2023c). 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. <https://doi.org/10.1007/s11042-022-13558-9>
- Harb, A., Gad, A., Yaghi, M., Alhalabi, M., Zia, H., Yousaf, J., Khelifi, A., Ghoudi, K., & Ghazal, M. (2023). Diverse distant-students deep emotion recognition and visualization. *Computers and Electrical Engineering*, 111(PB), 108963. <https://doi.org/10.1016/j.compeleceng.2023.108963>
- Herrero-Alvarez, R., Callejas-Castro, E., Miranda, G., & Leon, C. (2024). Analysis of Sentiment Toward Computer Science in Pre-University

- Education. *IEEE Access*, 12(April), 71205–71218. <https://doi.org/10.1109/ACCESS.2024.3402991>
- Hunukumbure, A. D., Horner, P. J., Fox, J., & Thakerar, V. (2021). An online discussion between students and teachers: a way forward for meaningful teacher feedback? *BMC Medical Education*, 21(1), 1–9. <https://doi.org/10.1186/s12909-021-02730-8>
- Jain, V., & Learned-Miller, E. (2010). Fddb: A benchmark for face detection in unconstrained settings. *UMass Amherst Technical Report, January*. http://works.bepress.com/erik_learned_miller/55/
- Kanwal, S., & Asghar, S. (2021). Speech Emotion Recognition Using Clustering Based GA-Optimized Feature Set. *IEEE Access*, 9, 125830–125842. <https://doi.org/10.1109/ACCESS.2021.3111659>
- Khan, U. A., Xu, Q., Liu, Y., Lagstedt, A., Alamäki, A., & Kauttonen, J. (2024). Exploring contactless techniques in multimodal emotion recognition: insights into diverse applications, challenges, solutions, and prospects. In *Multimedia Systems* (Vol. 30, Nomor 3). Springer Berlin Heidelberg. <https://doi.org/10.1007/s00530-024-01302-2>
- Khoshnam, F., & Baraani-Dastjerdi, A. (2022). A dual framework for implicit and explicit emotion recognition: An ensemble of language models and computational linguistics. *Expert Systems with Applications*, 198(October 2021), 116686. <https://doi.org/10.1016/j.eswa.2022.116686>
- Kim, J., & Lee, D. (2023). Facial Expression Recognition Robust to Occlusion and to Intra-Similarity Problem Using Relevant Subsampling. *Sensors*, 23(5). <https://doi.org/10.3390/s23052619>
- Le, H. D., Lee, G. S., Kim, S. H., Kim, S., & Yang, H. J. (2023). Multi-Label Multimodal Emotion Recognition With Transformer-Based Fusion and Emotion-Level Representation Learning. *IEEE Access*, 11(February), 14742–14751. <https://doi.org/10.1109/ACCESS.2023.3244390>
- Li, S. (2025). Application of entertainment e-learning mode based on genetic algorithm and facial emotion recognition in environmental art and design courses. *Entertainment Computing*, 52(May 2024), 100798. <https://doi.org/10.1016/j.entcom.2024.100798>
- Li, S., & Deng, W. (2022a). A Deeper Look at Facial Expression Dataset Bias. *IEEE Transactions on Affective Computing*, 13(2), 881–893. <https://doi.org/10.1109/TAFFC.2020.2973158>
- Li, S., & Deng, W. (2022b). Deep Facial Expression Recognition: A Survey. *IEEE Transactions on Affective Computing*, 13(3), 1195–1215. <https://doi.org/10.1109/TAFFC.2020.2981446>
- Lin, H. C. K., Liao, Y. C., & Wang, H. T. (2022). Eye Movement Analysis and Usability Assessment on Affective Computing Combined with Intelligent Tutoring System. *Sustainability (Switzerland)*, 14(24). <https://doi.org/10.3390/su142416680>
- Liu, T., Wang, M., Yang, B., Liu, H., & Yi, S. (2025). ESERNet: Learning spectrogram structure relationship for effective speech emotion recognition with swin transformer in classroom discourse analysis. *Neurocomputing*, 612(July 2024), 128711. <https://doi.org/10.1016/j.neucom.2024.128711>
- Liu, Z. T., Rehman, A., Wu, M., Cao, W. H., & Hao, M. (2021). Speech emotion recognition based on formant characteristics feature extraction and

- phoneme type convergence. *Information Sciences*, 563, 309–325. <https://doi.org/10.1016/j.ins.2021.02.016>
- Llurba, C., Fretes, G., & Palau, R. (2024). Classroom Emotion Monitoring Based on Image Processing. *Sustainability (Switzerland)*, 16(2). <https://doi.org/10.3390/su16020916>
- Mahmud, M. M., Freeman, B., & Abu Bakar, M. S. (2022). Technology in education: efficacies and outcomes of different delivery methods. *Interactive Technology and Smart Education*, 19(1), 20–38. <https://doi.org/10.1108/ITSE-01-2021-0021>
- Malakar, S., Chiracharit, W., & Chamnongthai, K. (2024). Masked Face Recognition with Generated Occluded Part using Image Augmentation and CNN Maintaining Face Identity. *IEEE Access*, 12(August), 126356–126375. <https://doi.org/10.1109/ACCESS.2024.3446652>
- Mamieva, D., Abdusalomov, A. B., Kutlimuratov, A., Muminov, B., & Whangbo, T. K. (2023). Multimodal Emotion Detection via Attention-Based Fusion of Extracted Facial and Speech Features. *Sensors*, 23(12). <https://doi.org/10.3390/s23125475>
- Mao, J., Qian, Z., & Lucas, T. (2023). Sentiment Analysis of Animated Online Education Texts Using Long Short-Term Memory Networks in the Context of the Internet of Things. *IEEE Access*, 11(August), 109121–109130. <https://doi.org/10.1109/ACCESS.2023.3321303>
- Mattioli, M., & Cabitza, F. (2024). Not in My Face: Challenges and Ethical Considerations in Automatic Face Emotion Recognition Technology. *Machine Learning and Knowledge Extraction*, 6(4), 2201–2231. <https://doi.org/10.3390/make6040109>
- Mehta, N. K., Prasad, S. S., Saurav, S., Saini, R., & Singh, S. (2022). Three-dimensional DenseNet self-attention neural network for automatic detection of student's engagement. *Applied Intelligence*, 52(12), 13803–13823. <https://doi.org/10.1007/s10489-022-03200-4>
- Mensah, J. A., Nortey, E. N. N., Ocran, E., Iddi, S., & Asiedu, L. (2024). De-occlusion and recognition of frontal face images: a comparative study of multiple imputation methods. *Journal of Big Data*, 11(1). <https://doi.org/10.1186/s40537-024-00925-6>
- Meriem, B., Benlahmar, H., Naji, M. A., Sanaa, E., & Wijdane, K. (2022). Determine the Level of Concentration of Students in Real Time from their Facial Expressions. *International Journal of Advanced Computer Science and Applications*, 13(1), 159–166. <https://doi.org/10.14569/IJACSA.2022.0130119>
- Moise, G., Dragomir, E. G., Șchiopu, D., & Iancu, L. A. (2024). Towards Integrating Automatic Emotion Recognition in Education: A Deep Learning Model Based on 5 EEG Channels. *International Journal of Computational Intelligence Systems*, 17(1). <https://doi.org/10.1007/s44196-024-00638-x>
- Morar, A., Moldoveanu, F., Moldoveanu, A., & Oana Balan, V. A. (2017). GPU accelerated 2D and 3D image processing. *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017*, 11, 653–656. <https://doi.org/10.15439/2017F265>

- Ngo, D., Nguyen, A., Dang, B., & Ngo, H. (2024). Facial Expression Recognition for Examining Emotional Regulation in Synchronous Online Collaborative Learning. *International Journal of Artificial Intelligence in Education*, 34(3), 650–669. <https://doi.org/10.1007/s40593-023-00378-7>
- Orosoo, M., Rajkumari, Y., Ramesh, K., Fatma, G., Nagabhaskar, M., Gopi, A., & Rengarajan, M. (2024). Enhancing English Learning Environments Through Real-Time Emotion Detection and Sentiment Analysis. *International Journal of Advanced Computer Science and Applications*, 15(7), 875–889. <https://doi.org/10.14569/IJACSA.2024.0150787>
- Pan, T., Ye, Y., Zhang, Y., Xiao, K., & Cai, H. (2024). Online multi-hypergraph fusion learning for cross-subject emotion recognition. *Information Fusion*, 108(February), 102338. <https://doi.org/10.1016/j.inffus.2024.102338>
- Pascual, A. M., Valverde, E. C., Kim, J. I., Jeong, J. W., Jung, Y., Kim, S. H., & Lim, W. (2022). Light-FER: A Lightweight Facial Emotion Recognition System on Edge Devices. *Sensors*, 22(23), 1–9. <https://doi.org/10.3390/s22239524>
- Peterson, A. T. (2023). Asynchrony and promotive interaction in online cooperative learning. *International Journal of Educational Research Open*, 5(September), 100300. <https://doi.org/10.1016/j.ijedro.2023.100300>
- Pordoy, J., Farman, H., Dicheva, N., Anwar, A., Nasralla, M. M., Khilji, N., & Rehman, I. U. (2024). Multi-Frame Transfer Learning Framework for Facial Emotion Recognition in e-Learning Contexts. *IEEE Access*, 12(October), 151360–151381. <https://doi.org/10.1109/ACCESS.2024.3478072>
- Poux, D., Allaert, B., Ihaddadene, N., Bilasco, I. M., Djeraba, C., & Bennamoun, M. (2022). Dynamic Facial Expression Recognition under Partial Occlusion with Optical Flow Reconstruction. *IEEE Transactions on Image Processing*, 31, 446–457. <https://doi.org/10.1109/TIP.2021.3129120>
- Rathod, M., Dalvi, C., Kaur, K., Patil, S., Gite, S., Kamat, P., Kotecha, K., Abraham, A., & Gabralla, L. A. (2022). Kids' Emotion Recognition Using Various Deep-Learning Models with Explainable AI. *Sensors*, 22(20). <https://doi.org/10.3390/s22208066>
- Roy, S., Gaur, V., Raza, H., & Jameel, S. (2023). CLEFT: Contextualised Unified Learning of User Engagement in Video Lectures With Feedback. *IEEE Access*, 11(January), 17707–17720. <https://doi.org/10.1109/ACCESS.2023.3245982>
- Sassi, A., Jaafar, W., Cherif, S., Abderrazak, J. Ben, & Yanikomeroglu, H. (2023). Video Traffic Analysis for Real-Time Emotion Recognition and Visualization in Online Learning. *IEEE Access*, 11(July), 99376–99386. <https://doi.org/10.1109/ACCESS.2023.3313973>
- Shi, L. (2024). The Integration of Advanced AI-Enabled Emotion Detection and Adaptive Learning Systems for Improved Emotional Regulation. *Journal of Educational Computing Research*, 63(1), 173–201. <https://doi.org/10.1177/07356331241296890>
- Shobana, B. T., & Kumar, G. A. S. (2021). I-Quiz: An Intelligent Assessment Tool for Non-Verbal Behaviour Detection. *Computer Systems Science and Engineering*, 40(3), 1007–1021. <https://doi.org/10.32604/CSSE.2022.019523>

- Sowjanya, U. L., & Krithiga, R. (2024). Decoding Student Emotions: An Advanced CNN Approach for Behavior Analysis Application Using Uniform Local Binary Pattern. *IEEE Access*, 12(August), 106273–106284. <https://doi.org/10.1109/ACCESS.2024.3436531>
- Sukumaran, A., & Manoharan, A. (2024). Multimodal Engagement Recognition From Image Traits Using Deep Learning Techniques. *IEEE Access*, 12(December 2023), 25228–25244. <https://doi.org/10.1109/ACCESS.2024.3353053>
- Tanko, D., Dogan, S., Burak Demir, F., Baygin, M., Engin Sahin, S., & Tuncer, T. (2022). Shoelace pattern-based speech emotion recognition of the lecturers in distance education: ShoePat23. *Applied Acoustics*, 190, 108637. <https://doi.org/10.1016/j.apacoust.2022.108637>
- Tao, X., Shannon-Honson, A., Delaney, P., Li, L., Dann, C., Li, Y., & Xie, H. (2022). Data Analytics on Online Student Engagement Data for Academic Performance Modeling. *IEEE Access*, 10(October), 103176–103186. <https://doi.org/10.1109/ACCESS.2022.3208953>
- Tian, X., Nunes, B. P., Liu, Y., & Manrique, R. (2024). Predicting Student Engagement Using Sequential Ensemble Model. *IEEE Transactions on Learning Technologies*, 17, 939–950. <https://doi.org/10.1109/TLT.2023.3342860>
- Trabelsi, Z., Alnajjar, F., Parambil, M. M. A., Gochoo, M., & Ali, L. (2023). Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student's Behavior Recognition. *Big Data and Cognitive Computing*, 7(1), 1–17. <https://doi.org/10.3390/bdcc7010048>
- Trigka, M., & Dritsas, E. (2025). A Comprehensive Survey of Deep Learning Approaches in Image Processing. *Sensors*, 25(2). <https://doi.org/10.3390/s25020531>
- Ülgen Sönmez, Yesim Varol, A. (2024). In-depth investigation of speech emotion recognition studies from past to present –The importance of emotion recognition from speech signal for AI-. *Intelligent Systems with Applications*, 22(March). <https://doi.org/10.1016/j.iswa.2024.200351>
- Villegas-Ch, W., Garcia-Ortiz, J., & Sanchez-Viteri, S. (2024). Personalization of Learning: Machine Learning Models for Adapting Educational Content to Individual Learning Styles. *IEEE Access*, 12(July), 121114–121130. <https://doi.org/10.1109/ACCESS.2024.3452592>
- Wang, S. (2021). Online learning behavior analysis based on image emotion recognition. *Traitement du Signal*, 38(3), 865–873. <https://doi.org/10.18280/ts.380333>
- Wickline, V. B., Hall, A. S., Lavrisa, R., McCook, K., Woodcock, M., Bani, M., Strepparava, M. G., Russo, S., & Nowicki, S. (2025). Facial occlusion with medical masks: Impacts on emotion recognition rates for emotion types and intensities. *Quarterly Journal of Experimental Psychology*. <https://doi.org/10.1177/17470218241308569>
- Xu, Z. (2023). Application of Convolution Neural Network Algorithm in Online Education Emotion Recognition. *International Journal of Web-Based*

- Learning and Teaching Technologies*, 18(2), 1–13.
<https://doi.org/10.4018/IJWLTT.331077>
- Yu, S., Androsov, A., Yan, H., & Chen, Y. (2024). Bridging computer and education sciences: A systematic review of automated emotion recognition in online learning environments. *Computers and Education*, 220(3663), 105111. <https://doi.org/10.1016/j.compedu.2024.105111>
- Zaareer, Z., Alhatemi, S. A. A., Alotaibi, A. A., Soliman, A. F., Abdulghani, S. R., Hassan, S., Abouelnaga, H. M., Azaizeh, A., & Jaradat, A. (2024). Investigating the Impact of LMS Quality , Technical Support and Perceived Usefulness on Student Satisfaction in Saudi Universities. *Evolutionary Studies in Imaginative Culture*, 8.
- Zeng, H., Shu, X., Wang, Y., Wang, Y., Zhang, L., Pong, T. C., & Qu, H. (2021). EmotionCues: Emotion-Oriented Visual Summarization of Classroom Videos. *IEEE Transactions on Visualization and Computer Graphics*, 27(7), 3168–3181. <https://doi.org/10.1109/TVCG.2019.2963659>
- Zhang, N., Liu, N., Han, J., Wan, K., & Shao, L. (2023). Face De-Occlusion With Deep Cascade Guidance Learning. *IEEE Transactions on Multimedia*, 25, 3217–3229. <https://doi.org/10.1109/TMM.2022.3157036>
- Zhang, S., Yang, Y., Chen, C., Zhang, X., Leng, Q., & Zhao, X. (2024). Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: A systematic review of recent advancements and future prospects. *Expert Systems with Applications*, 237(PC), 121692. <https://doi.org/10.1016/j.eswa.2023.121692>
- Zhang, X., & Qin, X. (2022). Research on Sentiment Analysis Algorithm for Comments on Online Ideological and Political Courses. *International Journal of Advanced Computer Science and Applications*, 13(11), 174–179. <https://doi.org/10.14569/IJACSA.2022.0131119>
- Zhang, X., Zhang, F., & Xu, C. (2022). Joint Expression Synthesis and Representation Learning for Facial Expression Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(3), 1681–1695. <https://doi.org/10.1109/TCSVT.2021.3056098>
- Zhu-Zhou, F., Gil-Pita, R., García-Gómez, J., & Rosa-Zurera, M. (2022). Robust Multi-Scenario Speech-Based Emotion Recognition System. *Sensors*, 22(6). <https://doi.org/10.3390/s22062343>
- Zhu, C., Wan, X., Xie, S., Li, X., & Gu, Y. (2022). Occlusion-robust Face Alignment using A Viewpoint-invariant Hierarchical Network Architecture. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2022-June, 11102–11111. <https://doi.org/10.1109/CVPR52688.2022.01083>
- Zhu, H., Xu, H., Ma, X., & Bian, M. (2022). Facial Expression Recognition Using Dual Path Feature Fusion and Stacked Attention. *Future Internet*, 14(9). <https://doi.org/10.3390/fi14090258>
- Zhu, K., He, X., Lv, Z., Zhang, X., Hao, R., He, X., Wang, J., He, J., Zhang, L., & Mu, Z. (2023). A 3D Occlusion Facial Recognition Network Based on a Multi-Feature Combination Threshold. *Applied Sciences (Switzerland)*, 13(10). <https://doi.org/10.3390/app13105950>

LAMPIRAN

Lampiran A Potongan Paparan Konsep Perubahan RENIP IPDN 2020-2045

Konsep Perubahan

RENIP IPDN
Tahun 2020-2045

1

Aspek	Fenomena	Peluang	Tantangan dan Ancaman
	<ul style="list-style-type: none"> Penggunaan Artificial Intelligent di berbagai aspek kehidupan, termasuk di bidang pendidikan. Perkembangan Industri 4.0 memperkuat peran teknologi di kehidupan manusia, melalui integrasi pengembangan teknologi dengan keahlian manusia. Terdapat 29 ribu serangan siber selama tahun 2023 yang berhasil di blokir. 	<ul style="list-style-type: none"> Pengembangan pembelajaran berbasis <i>Smart Class</i>. Personalisasi proses pembelajaran menggunakan <i>Artificial Intelligent</i> yang berdampak pada hasil pembelajaran yang lebih optimal. Akselerasi pekerjaan administrasi di bidang pendidikan. otomasi dan efisiensi kegiatan belajar mengajar di lingkungan IPDN. 	<p>Perkembangan kejahatan siber memberikan tantangan bagi keamanan data Nasional, sehingga menjadi tantangan tersendiri bagi instansi pemerintah untuk meningkatkan keamanan siber di dalam negeri.</p>
	<ul style="list-style-type: none"> Peningkatan suhu udara dimuka bumi menjadi salah satu penyebab perubahan iklim secara global. Indonesia terutama mengalami peningkatan suhu udara dimuka bumi rata-rata di angka 2,7°C pada bulan juni 2024. Indonesia merupakan negara kepulauan yang dikelilingi beberapa Samudra, diketahui oleh 4 lempeng Samudra, dan tersebutnya banyak gunung aktif di Indonesia. 	<p>Perubahan iklim yang cukup drastis membuka peluang munculnya penelitian-penelitian kebijakan pemerintah berbasis lingkungan dalam menanggulangi perubahan iklim di wilayah indonesia</p>	<ul style="list-style-type: none"> Perubahan iklim secara ekstrem yang dapat mempengaruhi ketersediaan sumberdaya yang dimiliki Indonesia dan pola hidup masyarakat. Letak geografis Indonesia yang beragam memiliki potensi kejadian bencana alam dalam jumlah yang banyak.
	<p>Pemerintah saat ini tengah mencanangkan perubahan Undang-Undang No.20 Tahun 2003 terkait dengan Sistem Pendidikan Nasional dengan mengintegrasikan UU Sisdiknas, UU Guru dan Dosen, dan UU Dikti.</p>	<p>Terbukanya IPDN untuk melakukan pengelolaan SDM berbasis pada keahlian dan kompetensi sehingga dapat menghasilkan lulusan yang sesuai dengan kebutuhan pasar.</p>	<p>Pengusulan perubahan sistem pendidikan nasional membuat IPDN selaku institusi pendidikan perlu melakukan penyusuaian baik dari segi kurikulum, sistem pembelajaran, sumberdaya manusia, hingga aspek organisasi.</p>

13

Peta Jalan IPDN 2025 -2029 – Membangun Keunggulan

Sasaran	Tantangan dan Ancaman Tahun 2025-2029
<ul style="list-style-type: none"> Mendirikan program studi (prodi) baru. Terbentuknya lembaga penelitian sebagai <i>revenue center</i>. Meningkatnya kemampuan riset di bidang ilmu pemerintahan. Mengembangkan kurikulum di bidang ilmu pemerintahan sesuai dengan analisis kebutuhan dan permasalahan pengelolaan birokrasi pemerintahan. Modernisasi pembelajaran dengan memperkuat <i>E-learning</i> dan <i>hybrid</i>. Meningkatkan jumlah dosen yang memiliki kualifikasi doktor, lektor kapela dan guru besar. Tersedianya fasilitas pendukung proses pendidikan dengan berfokus pada "Smart Class". Penguatan kelembagaan dan tata kelola organisasi. Terjalinnya Kerjasama dengan stakeholder terkait dalam mengimplementasikan Tridharma Perguruan Tinggi dalam dan luar negeri. 	<ul style="list-style-type: none"> Revolusi Industri 5.0 yaitu pengintegrasian teknologi canggih saat ini dengan keahlian manusia. Pembelajaran berbasis digital Kesenjangan referensi dan gaya hidup antara siswa dan pengajar/pendidik Karakteristik generasi Z yang lebih individualis dan egois. transformasi sosial: penuntasan pemenuhan pelayanan dasar kesehatan, pendidikan, dan perlindungan sosial, serta peningkatan kualitas SDM untuk membentuk manusia produktif Transformasi tata kelola: perbaikan kelembagaan yang tepat fungsi, penyempurnaan fondasi penataan regulasi, peningkatan kualitas ASN berbasis merit sistem, peningkatan pelayanan publik berbasis teknologi informasi, serta penguatan kapasitas masyarakat sipil