

**RANCANG BANGUN KACAMATA MULTIMODALITAS BERBASIS
MACHINE LEARNING DAN NATURAL LANGUAGE PROCESSING UNTUK
MEMBANTU KOMUNIKASI PENYANDANG TUNARUNGU-WICARA**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-
1 Program Studi Teknik Industri - Fakultas Teknologi Industri
Universitas Islam Indonesia**



Nama : Yudhistira Alfarizi
No. Mahasiswa : 22522250

**PROGRAM STUDI TEKNIK INDUSTRI PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2026**

PERNYATAAN KEASLIAN

Saya mengakui bahwa tugas akhir ini adalah hasil karya saya sendiri kecuali kutipan dan ringkasan yang seluruhnya sudah saya jelaskan sumbernya. Jika dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah maka saya bersedia ijazah yang telah saya terima ditarik kembali oleh Universitas Islam Indonesia.

Yogyakarta, 30 – April - 2026



Handwritten signature of Yudhistira Alfarizi.

(Yudhistira Alfarizi)
22522250

SURAT BUKTI PENELITIAN



FAKULTAS
TEKNOLOGI INDUSTRI

Gedung K91, Mes Mansur
Kampus Terpadu Universitas Islam Indonesia
Jl. Kaliurang km 14,5 Yogyakarta 55584
T: (0274) 896444 ext 4110, 4100
F: (0274) 895007
E: fti@uii.ac.id
W: fti.uii.ac.id

SURAT KETERANGAN PENELITIAN

Nomor : 274/Ka.lab SIMANTI/20/Lab.SIMANTI/IV/2026

Assalamu'alaikum Warohmatullahi Wabarokaatuh

Dengan hormat,

Yang bertanda tangan dibawah ini, menerangkan bahwa:

Nama : Yudhistira Alfarizi

Nim : 22522250

Jurusan : Teknik Industri

Dosen Pembimbing : Ir. Ali Parkhan, M.T.

Menyatakan bahwa mahasiswa tersebut diatas telah melaksanakan penelitian tugas akhir dengan judul **"RANCANG BANGUN KACAMATA MULTIMODALITAS BERBASIS MACHINE LEARNING DAN NATURAL LANGUAGE PROCESSING UNTUK MEMBANTU KOMUNIKASI PENYANDANG TUNARUNGU-WICARA"**. Mulai pelaksanaan penelitian 20 Januari 2026 sampai 24 Februari 2026

Demikian surat keterangan penelitian ini kami buat. Atas perhatiannya dan kerja samanya kami mengucapkan terima kasih.

Wassalamu'alaikum Warohmatullahi Wabarokaatuh

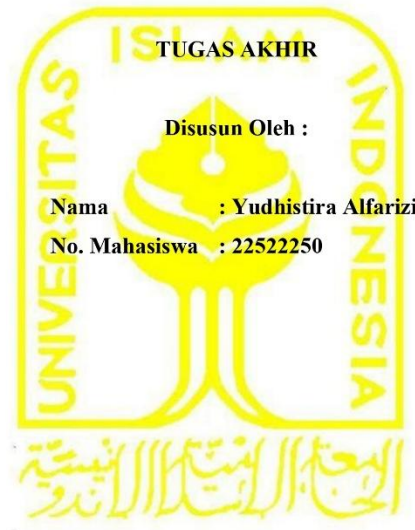
Yogyakarta, 30 April 2026

Kepala Laboratorium
Sistem Manufaktur Terintegrasi

Putri Dwi Annisa, S.T., M.Sc.

LEMBAR PENGESAHAN PEMBIMBING

**RANCANG BANGUN KACAMATA MULTIMODALITAS BERBASIS
MACHINE LEARNING DAN NATURAL LANGUAGE PROCESSING
UNTUK MEMBANTU KOMUNIKASI PENYANDANG TUNARUNGU-
WICARA**



Yogyakarta, 30 April 2026

Dosen Pembimbing

A handwritten signature in black ink, appearing to read 'Ali Parkhan', is positioned above the printed name of the supervisor.

(Ir. Ali Parkhan, M.T.)

LEMBAR PENGESAHAN DOSEN PENGUJI

RANCANG BANGUN KACAMATA MULTIMODALITAS BERBASIS MACHINE LEARNING DAN NATURAL LANGUAGE PROCESSING UNTUK MEMBANTU KOMUNIKASI PENYANDANG TUNARUNGU-WICARA

TUGAS AKHIR

Disusun Oleh :

Nama : Yudhistira Alfarizi

No. Mahasiswa : 22 522 250

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata-1 Teknik Industri Fakultas Teknologi Industri

Universitas Islam Indonesia

Yogyakarta, 08 - Mei – 2026

Tim Penguji

Ir. Ali Parkhan, M.T.

Ketua

Dr. Drs. Imam Djati Widodo, M.Eng.Sc.

Anggota I

Dr. Sri Indrawati, S.T., M.Eng.

Anggota II

Mengetahui,

Ketua Program Studi Teknik Industri Program Sarjana

Fakultas Teknologi Industri

Universitas Islam Indonesia

Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM.

NIK. 015220101

HALAMAN PERSEMBAHAN

Laporan Tugas Akhir ini disusun dan dipersembahkan sebagai wujud syukur kepada Allah SWT dan bakti kepada kedua orang tua atas segala doanya, serta didedikasikan bagi komunitas tunarungu dan seluruh pihak yang telah mendukung penyelesaian karya ini.

MOTTO

"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya." (QS. Al-Baqarah: 286).

"Demi menyelamatkan kita dari jalan yang salah, terkadang Tuhan akan mematahkan kita sepatah patahnya." (*Brian Khrisna*, Seporsi Mie Ayam Sebelum Mati)

"Hidup tak lebih dari pagelaran nestapa, kita memerankan langkah yang menjalar dari satu luka ke luka lainnya." (*Brian Khrisna*, Sisi Tergelap Surga)

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan penulisan naskah Tugas Akhir ini dengan baik. Shalawat serta salam senantiasa tercurah kepada Nabi Muhammad SAW.

Penyusunan Tugas Akhir dengan judul "Rancang Bangun Kacamata Multimodalitas Berbasis Machine Learning dan Natural Language Processing untuk Membantu Komunikasi Penyandang Tunarungu-Wicara" ini merupakan salah satu syarat untuk memperoleh gelar Sarjana Strata-1 pada Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia.

Penulis menyadari bahwa proses penyusunan Tugas Akhir ini tidak lepas dari bimbingan, arahan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo., M.T., IPU, ASEAN.Eng, selaku Dekan Fakultas Teknologi Industri, Universitas Islam Indonesia.
2. Bapak Dr. Drs. Imam Djati Widodo, M.Eng.Sc., selaku Ketua Jurusan Teknik Industri Fakultas Teknologi Industri Universitas Islam Indonesia.
3. Bapak Ir. Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph.D., IPM., selaku Ketua Program Studi Teknik Industri Program Sarjana Universitas Islam Indonesia.
4. Bapak Ir. Ali Parkhan, M.T., selaku Dosen Pembimbing yang telah meluangkan waktu untuk memberikan bimbingan, arahan, dan evaluasi yang sangat berharga selama proses penyusunan Tugas Akhir ini.
5. Kedua orang tua penulis, Bapak Junaidi M Dina dan Ibu Sri Ismawati, atas segala doa, kasih sayang, dan pengorbanan yang tiada henti diberikan demi kelancaran studi penulis.
6. Kakak-kakak penulis, Isna Tita Safira dan Meily Dewi Arimbi, yang senantiasa memberikan dukungan dan motivasi moral.
7. Shavira Amany Yuliantika, atas segala bentuk dukungan, semangat, dan bantuan yang diberikan selama proses penyusunan Tugas Akhir ini.
8. Rekan-rekan seperjuangan di Laboratorium SIMAN: Aji Adinata Firjatulah, Manik Purbo Ulung, Dinda Meilia Rhepon, Muhammad Dava Aji, Fildzah Shabrina, dan Muhammad Sulthansyah Al Fayad, atas diskusi, kebersamaan, dan kerja sama yang baik.
9. Saudara Rahmat Nur Fauzi, Mohamad Farhan Agustiono, dan Hilqi Juliano Satafa, serta seluruh rekan dalam keluarga besar ALLBESS, atas dukungan moral, semangat, serta persahabatan yang telah membersamai perjalanan akademik penulis.
10. Segenap *civitas academica* Program Studi Teknik Industri Universitas Islam Indonesia serta seluruh pihak yang tidak dapat penulis sebutkan satu per satu yang telah membantu kelancaran penulisan ini.

Penulis menyadari bahwa naskah Tugas Akhir ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi penyempurnaan di masa mendatang. Akhir kata, semoga Tugas Akhir ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan masyarakat luas.

Yogyakarta, 30 April 2026



Yudhistira Alfarizi

22522250

ABSTRAK

Penyandang tunarungu-wicara rentan menghadapi hambatan komunikasi dan risiko keselamatan ketika berada di ruang publik karena ketidakmampuan merespons peringatan suara di lingkungan sekitarnya. Penelitian ini bertujuan untuk merancang purwarupa kacamata pintar multimodalitas berbasis *Machine Learning* yang mengintegrasikan fitur *Keyword Spotting* (KWS) untuk deteksi suara darurat, mengevaluasi performa model *transfer learning* untuk penerjemahan bahasa isyarat visual, serta mengusulkan desain sasis 3D CAD yang ergonomis. Penelitian ini mengadopsi kerangka kerja *Plan-Do-Check-Act* (PDCA) untuk memastikan perbaikan sistem yang berkelanjutan. Pengembangan sistem *Proof of Concept* (PoC) KWS diimplementasikan secara *offline* menggunakan mikrokontroler *edge* Teensy 4.1 dengan sensor INMP441, sedangkan simulasi visual dilakukan melalui platform *Edge Impulse*. Hasil pengujian memvalidasi bahwa model KWS berkinerja unggul dengan tingkat akurasi mencapai 95% pada kondisi hening dan 90% pada kondisi kebisingan hingga 80 dB. Sementara itu, model *transfer learning* untuk deteksi visual pada *smartphone* menunjukkan performa optimal, namun integrasi langsungnya ke Teensy 4.1 terkendala oleh *bottleneck* komputasi (*CPU Starvation*) akibat arsitektur *single-core*. Evaluasi fungsional juga mengidentifikasi batasan perangkat keras, di antaranya fenomena pemerangkapan panas (*thermal trapping*) pada sasis cetak 3D berbahan PLA yang menaikkan suhu permukaan hingga 45°C, serta efisiensi baterai yang terbatas pada 13–15 jam operasional. Dari sudut pandang ergonomi fisik, rancangan 3D berhasil mendistribusikan titik berat komponen ke bagian belakang telinga guna meminimalisir tekanan pada hidung, dengan layar OLED yang diposisikan strategis pada area penglihatan perifer. Kesimpulannya, fungsionalitas KWS terbukti sangat layak diimplementasikan pada perangkat *edge*, namun pengembangan kacamata multimodalitas tahap lanjut merekomendasikan transisi menuju prosesor *multi-core* serta perbaikan material untuk manajemen termal.

Kata Kunci: Kacamata Pintar, *Keyword Spotting*, *Machine Learning*, Tunarungu, Ergonom

DAFTAR ISI

PERNYATAAN KEASLIAN.....	ii
SURAT BUKTI PENELITIAN	iii
LEMBAR PENGESAHAN PEMBIMBING	iv
LEMBAR PENGESAHAN DOSEN PENGUJI.....	v
HALAMAN PERSEMBAHAN	vi
MOTTO	vii
KATA PENGANTAR.....	viii
ABSTRAK.....	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Penelitian.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Kajian Literatur	7
2.2 Landasan Teori	12
2.2.1 <i>Disabilitas Tunarungu dan Tunawicara</i>	<i>12</i>
2.2.2 <i>Karakteristik Penyandang Tunarungu.....</i>	<i>12</i>
2.2.3 <i>Karakteristik Penyandang Tunawicara</i>	<i>14</i>
2.2.4 <i>Interaksi Multimodal (Multimodal Interaction).....</i>	<i>15</i>
2.2.5 <i>Machine Learning</i>	<i>15</i>
2.2.6 <i>Deep Learning.....</i>	<i>16</i>
2.2.7 <i>Tiny Machine Learning (TinyML).....</i>	<i>17</i>
2.2.8 <i>Natural Language Processing (NLP)</i>	<i>18</i>
2.2.9 <i>Deep Spoken Keyword Spotting (KWS)</i>	<i>19</i>
2.2.10 <i>Augmented Reality (AR) dan Visualisasi Informasi.....</i>	<i>19</i>
2.2.11 <i>Embedded System & Wearable</i>	<i>20</i>
2.2.12 <i>Integrasi Metode Pengembangan Scrum dan Siklus PDCA</i>	<i>22</i>
2.2.13 <i>Metrik Evaluasi Kinerja Sistem</i>	<i>23</i>
2.2.14 <i>Komparasi Teknologi dan Posisi Penelitian.....</i>	<i>25</i>
BAB III METODE PENELITIAN.....	28
3.1 Objek, Subjek, dan Lokasi Penelitian	28
3.1.1 <i>Objek Penelitian.....</i>	<i>28</i>
3.1.2 <i>Subjek dan Lokasi Penelitian</i>	<i>29</i>
3.1.3 <i>Identifikasi Kebutuhan Pengguna (Voice of Customer).....</i>	<i>29</i>
3.1.4 <i>Profil Mitra Penelitian.....</i>	<i>30</i>
3.2 Sumber Data.....	31
3.2.1 <i>Data Primer</i>	<i>31</i>
3.2.2 <i>Data Sekunder</i>	<i>33</i>
3.3 Instrumen Penelitian.....	35
3.3.1 <i>Instrumen Perangkat Keras (Hardware)</i>	<i>35</i>
3.3.2 <i>Instrumen Non-Teknis (Dokumentasi dan Validasi)</i>	<i>37</i>
3.4 Metode Pengumpulan Data.....	38

3.4.1	<i>Elicitasi Kebutuhan Pengguna</i>	38
3.4.2	<i>Akuisisi Dataset Kata Kunci Visual</i>	38
3.4.3	<i>Akuisisi Dataset Audio Kedaruratan</i>	39
3.4.4	<i>Pengumpulan Data Kinerja Fungsional Sistem</i>	40
3.4.5	<i>Pengumpulan Data Dimensi Fisik Komponen</i>	40
3.5	Metode Pengolahan Data	41
3.5.1	<i>Pengolahan Data Kebutuhan Pengguna (Fase Plan - Analisis Backlog)</i> 41	
3.5.2	<i>Pengolahan Data Pembelajaran Mesin (Fase Do - Pengembangan Model AI)</i> 42	
3.5.3	<i>Pengolahan Data Perancangan Design (Fase Do - Hardware Integration)</i>	44
3.5.4	<i>Analisis Kinerja Sistem (Fase Check - Evaluasi & Validasi)</i>	44
3.5.5	<i>Analisis Kriteria Penerimaan Logika (Fase Act)</i>	45
3.6	Alur Penelitian	46
BAB IV	PEMBANGUNAN SISTEM	51
4.1	Perencanaan & Analisis Kebutuhan (Fase Plan)	51
4.1.1	<i>Pengumpulan Data Kebutuhan Pengguna</i>	51
4.1.2	<i>Profil Pengguna (User Persona)</i>	55
4.1.3	<i>Penyusunan User Stories & Product Backlog</i>	57
4.2	Perancangan & Implementasi (Fase Do)	62
4.2.1	<i>Pengumpulan & Pelatihan Dataset (Dataset Suara & Citra)</i>	62
4.2.2	<i>Perancangan Perangkat Keras Skala Laboratorium (Proof of Concept)</i> 75	
4.2.3	<i>Perancangan Mekanis Konseptual (Desain 3D)</i>	81
4.2.4	<i>Deployment Model ke Purwarupa Perangkat Keras</i>	84
BAB V	PENGUJIAN SISTEM DAN PEMBAHASAN	91
5.1	Hasil Pengujian Sistem (Fase Check)	91
5.1.1	<i>Pengujian Performa Model Keyword Spotting</i>	91
5.1.2	<i>Pengujian Performa Model Citra/Gestur (Off-Device Simulation)</i>	94
5.1.3	<i>Pengujian Fungsionalitas dan Latensi Perangkat Keras</i>	98
5.2	Pembahasan Kritis Hasil Pengujian (Fase Act)	101
5.2.1	<i>Analisis Resiliensi Sistem terhadap Interferensi Akustik (SNR)</i>	101
5.2.2	<i>Justifikasi Teknis Penggunaan Prototype Breadboard</i>	104
5.2.3	<i>Analisis Batasan Perangkat dan Justifikasi Eksklusi Fitur Visual</i>	106
5.3	Evaluasi Konseptual Desain 3D	108
5.4	Implikasi Hasil terhadap Rencana Perbaikan (Kaizen)	109
BAB VI	PENUTUP	112
6.1	Kesimpulan	112
6.2	Saran	113
DAFTAR PUSTAKA		115
LAMPIRAN		A-1

DAFTAR TABEL

Tabel 2. 1 Posisi Penelitian.....	10
Tabel 2. 2 Komparasi Fitur <i>Vogest Lens</i> dengan Teknologi Serupa	27
Tabel 3. 1 Profil Mitra Penelitian	31
Tabel 4. 1 Profil Responden Wawancara	52
Tabel 4. 2 Identifikasi Kebutuhan Pengguna (<i>User Needs</i>)	53
Tabel 4. 3 Matriks Analisis Kebutuhan Persona	57
Tabel 4. 4 Daftar User Stories	58
Tabel 4. 5 <i>Product Backlog</i> dan Spesifikasi Teknis.....	60
Tabel 4. 6 Matriks Rencana Validasi Fungsional (QC).....	62
Tabel 4. 7 Distribusi Kuantitas Dataset Audio	64
Tabel 5. 1 Matriks Konfusi (<i>Confusion Matrix</i>) Pengujian Model Suara	92
Tabel 5. 2 Matriks Konfusi Pengujian Model Citra via <i>Smartphone</i>	95
Tabel 5. 3 Pengujian Fungsionalitas PoC	99

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	47
Gambar 4. 1 Visualisasi <i>Waveform</i> Pengumpulan <i>Data Audio</i>	64
Gambar 4. 2 Pra-pemrosesan <i>Data Audio</i>	65
Gambar 4. 3 Visualisasi Spektrogram dan Pemisahan Kelas Fitur MFCC	66
Gambar 4. 4 Konfigurasi Hiperparameter dan Arsitektur <i>Neural Network</i>	67
Gambar 4. 5 Metrik Performansi & Latensi Model <i>Audio</i>	68
Gambar 4. 6 <i>Sample Dataset "Halo"</i>	69
Gambar 4. 7 <i>Sample Dataset "Tolong"</i>	69
Gambar 4. 8 <i>Sample Dataset "Noise"</i>	70
Gambar 4. 9 Tahap Prapemrosesan <i>Data</i>	71
Gambar 4. 10 Parameter Pra-pemrosesan <i>Citra (Grayscale)</i> pada <i>Edge Impulse</i>	72
Gambar 4. 11 Konfigurasi Hiperparameter dan Arsitektur <i>Transfer Learning</i>	73
Gambar 4. 12 Metrik Performansi & Latensi Model <i>Citra</i>	74
Gambar 4. 13 <i>Data Explorer & On-device Performance</i>	74
Gambar 4. 14 <i>Development Board Teensy 4.1</i>	77
Gambar 4. 15 Konfigurasi <i>Wiring</i> Mic INMP 441 & <i>Teensy 4.1</i>	78
Gambar 4. 16 Konfigurasi <i>Wiring</i> Layar OLED & <i>Teensy 4.1</i>	79
Gambar 4. 17 Konfigurasi <i>Wiring</i> Catu Daya.....	80
Gambar 4. 18 Topologi Rangkaian Elektronik Purwarupa <i>Proof of Concept (PoC)</i> . 81	
Gambar 4. 19 Perspektif Desain 3D Kacamata <i>Vogest Lens</i>	82
Gambar 4. 20 Detail Penempatan Komponen <i>Vogest Lens</i>	83
Gambar 4. 21 Antarmuka Ekspor Model dan Optimasi <i>EON Compiler</i>	85
Gambar 4. 22 Tampilan Instalasi Pustaka Model pada <i>Arduino IDE</i>	86
Gambar 4. 23 Kode Program untuk Pemanggilan Fungsi Inferensi Model	88
Gambar 4. 24 Output Log Kompilasi dan Alokasi Memori pada <i>Arduino IDE</i>	89
Gambar 4. 25 <i>Heatmap Confusion Matrix</i> <i>Audio</i>	93
Gambar 5. 1 <i>Heatmap Confusion Matrix</i> Visual	96
Gambar 5. 2 Kegagalan Deteksi Objek.....	97
Gambar 5. 3 Pengujian <i>Prototype</i> Produk.....	100
Gambar 5. 4 Visualisasi Fitur MFCC Kondisi Hening.....	103
Gambar 5. 5 Visualisasi Fitur MFCC Kondisi Bising	103
Gambar 5. 6 Konfigurasi <i>Wiring</i> Komunikasi I2S dan I2C Pada Purwarupa.....	105
Gambar 5. 7 Render 3D Kacamata	109

BAB I PENDAHULUAN

1.1 Latar Belakang

Komunikasi pada hakikatnya adalah proses interaksi dua arah yang bertujuan untuk menciptakan kesamaan makna (*shared meaning*) antara penyampai dan penerima pesan (Al Ghifari et al., 2025). Merujuk pada teori komunikasi Wilbur Schramm, keberhasilan interaksi ini sangat bergantung pada mekanisme penyandian (*encoding*) dan penerjemahan (*decoding*) informasi yang presisi, serta adanya keselarasan pengalaman (*field of experience*) antara individu yang terlibat (Kuznar & Yager, 2020). Dalam kondisi ideal, siklus komunikasi ini berjalan tanpa hambatan di mana pesan yang dikirimkan dapat diterima dan dipahami secara utuh, sehingga memungkinkan terjadinya umpan balik (*feedback*) yang efektif. Namun, ketika terdapat gangguan (*noise*) atau perbedaan saluran komunikasi yang signifikan, tujuan utama komunikasi untuk menyampaikan informasi, gagasan, dan emosi menjadi sulit tercapai.

Kesenjangan saluran komunikasi ini menjadi isu krusial dalam interaksi antara masyarakat umum dengan penyandang disabilitas sensorik, khususnya tunarungu dan tunawicara. Dalam Pasal 1 Ayat (1) UU Nomor 8 Tahun 2016 Tentang Penyandang Disabilitas, disabilitas didefinisikan sebagai keterbatasan fisik, mental, kejiwaan, atau sensorik dalam jangka waktu lama yang menghambat partisipasi penuh dan efektif. Keterbatasan partisipasi ini muncul akibat interaksi antara individu yang memiliki gangguan (*impairment*) dengan hambatan lingkungan dan sikap masyarakat (Wiraputra, 2021). Dalam hal sensorik, tunarungu merupakan kondisi gangguan pendengaran (baik sebagian maupun total) yang menyebabkan individu tidak mampu menerima informasi audio secara optimal, meskipun sisa pendengaran terkadang masih dapat dimanfaatkan (Wahyudi et al., 2024). Sementara itu, tunawicara ditandai dengan hambatan dalam memproduksi ujaran verbal (*speech production*), yang berdampak langsung pada kemampuan menyampaikan pesan secara lisan (Utami et al., 2024). Kombinasi kondisi ini menciptakan tantangan kompleks dalam interaksi sosial yang didominasi oleh komunikasi verbal.

Berdasarkan data Badan Pusat Statistik (BPS) tahun 2020, jumlah penyandang disabilitas di Indonesia mencapai 22,3 juta jiwa (Ketenagakerjaan, 2022). Pemerintah Indonesia sejatinya telah berupaya menjamin hak kesetaraan melalui agenda *Sustainable*

Development Goals (SDGs) dan Perpres No. 59 Tahun 2017. Namun, realitas di lapangan menunjukkan bahwa akses terhadap interaksi sosial yang inklusif masih rendah. Hal ini terjadi karena hambatan komunikasi yang dihadapi bukan sekadar masalah teknis pendengaran, melainkan apa yang disebut sebagai hambatan semantik. hambatan semantik merujuk pada kesulitan penafsiran makna akibat perbedaan struktur bahasa antara bahasa isyarat (visual-gestural) yang digunakan Teman Tuli dan bahasa verbal (auditori) yang digunakan masyarakat umum (Rengganis et al., 2025). Ketidapahaman masyarakat terhadap bahasa isyarat sering kali berujung pada misinterpretasi pesan yang fatal, yang kemudian melahirkan stigma negatif hingga pengucilan sosial bagi penyandang disabilitas.

Untuk menjembatani hambatan semantik tersebut, perkembangan teknologi asistif telah menghadirkan berbagai solusi, salah satunya adalah aplikasi penerjemah *speech-to-text* pada *smartphone* (Putri & Sularno, 2024). Meskipun memberikan kontribusi positif, dari perspektif ergonomi dan interaksi manusia (*Human-Computer Interaction*), solusi berbasis *smartphone* memiliki kelemahan fundamental. (Berke et al., 2020) dalam studinya bahwa sifat perangkat *hand-held* secara drastis membatasi kebebasan gerak tangan pengguna, padahal ekspresi tangan dan gestur tubuh adalah elemen vital dalam struktur bahasa isyarat. Selain itu, ketergantungan visual pada layar ponsel memaksa pengguna memutus kontak mata (*eye contact*) dengan lawan bicara. Perpindahan fokus visual secara berulang antara lawan bicara dan layar ponsel ini meningkatkan beban kognitif (*cognitive load*) dan menurunkan kualitas hubungan interpersonal dalam komunikasi tatap muka.

Merespons keterbatasan tersebut, diperlukan pergeseran paradigma menuju inovasi teknologi *wearable device* yang bersifat *hands-free*. Solusi yang diusulkan dalam penelitian ini adalah pengembangan *Vogest Lens (Voice and Gesture Lens)*, sebuah kacamata pintar yang dirancang dengan konsep interaksi multimodal paralel. Berbeda dengan solusi konvensional berbasis *cloud*, sistem ini dibangun sebagai perangkat *embedded* mandiri (*standalone*) menggunakan mikrokontroler Teensy 4.1 untuk meminimalkan latensi pemrosesan dan menjaga privasi data pengguna.

Realisasi sistem ini bertumpu pada integrasi teknologi kecerdasan buatan (*Artificial Intelligence*). *Vogest Lens* memanfaatkan algoritma *Natural Language Processing* (NLP) yang dibatasi pada metode *Keyword Spotting* (KWS) untuk melakukan transkripsi suara kedaruratan ke dalam teks visual. Di sisi lain, untuk menerjemahkan bahasa isyarat,

sistem menggunakan *Computer Vision* berbasis *Deep Learning* (MobileNet V2) yang memproses data visual dari kamera sensor untuk mengenali pola gerakan tangan secara akurat (Mujahid et al., 2021). Keberadaan kecerdasan buatan ini memungkinkan perangkat untuk tidak hanya melihat dan mendengar, tetapi juga memproses konteks komunikasi yang kompleks untuk menciptakan jembatan pemahaman yang setara antara Teman Tuli dan Teman Dengar.

Untuk mengakomodasi luaran dari pemrosesan tersebut, antarmuka visualisasi direalisasikan menggunakan layar mikro berbasis matriks *Organic Light-Emitting Diode* (OLED). Mengingat tingginya kompleksitas integrasi optik presisi, limitasi kapabilitas *rapid prototyping*, serta inefisiensi biaya pada fabrikasi teknologi gelombang pandu holografis (*holographic waveguide*), penggunaan layar OLED standar diusung sebagai kompromi rekayasa yang paling rasional untuk memvalidasi *Proof of Concept* (PoC) purwarupa skala laboratorium. Layar ini diintegrasikan secara mekanis pada sasis kaca mata dan diproyeksikan pada area pandang periferal pengguna (*peripheral vision*). Konfigurasi ini dirancang agar informasi peringatan kedaruratan tetap dapat ditransmisikan dengan kontras piksel yang tajam (mengurangi beban kognitif) tanpa memblokir ruang pandang utama (*primary field of view*) pengguna terhadap dinamika fisik sekitarnya.

Mengingat kompleksitas integrasi antara perangkat keras (*hardware*), perangkat lunak (*software*), dan kebutuhan pengguna disabilitas yang dinamis, penelitian ini menerapkan kerangka kerja *Scrum* sebagai landasan manajemen pengembangan. Pendekatan *Agile* ini dikombinasikan dengan siklus PDCA (*Plan-Do-Check-Act*) untuk menjamin pengendalian kualitas teknis di setiap iterasinya. Melalui identifikasi kebutuhan berbasis *User Stories*, penelitian ini diharapkan menghasilkan purwarupa alat bantu komunikasi yang tidak hanya fungsional secara teknis, tetapi juga responsif dan ergonomis dalam menjawab tantangan interaksi inklusif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas dapat diperoleh beberapa rumusan masalah yakni sebagai berikut:

1. Bagaimana arsitektur sistem *Proof of Concept* (PoC) berbasis *Machine Learning* untuk fitur *Keyword Spotting* (KWS) menggunakan mikrokontroler Teensy 4.1?

2. Seberapa besar tingkat akurasi dan kinerja model *keyword spotting* dalam mendeteksi suara ketika diuji dalam purwarupa Vogest Lens dalam kondisi lingkungan yang ideal dan tidak ideal?
3. Seberapa besar tingkat akurasi dan kinerja model *transfer learning* dalam menerjemahkan bahasa isyarat secara visual ketika disimulasikan menggunakan arsitektur Edge Impulse pada *smartphone*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disampaikan, tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan dan memvalidasi arsitektur sistem *Proof of Concept* (PoC) fitur peringatan berbasis *Keyword Spotting* (KWS) menggunakan mikrokontroler Teensy 4.1.
2. Mengukur tingkat akurasi dan kinerja model *keyword spotting* dalam mendeteksi suara ketika diuji dalam purwarupa Vogest Lens dalam kondisi lingkungan yang ideal dan tidak ideal.
3. Mengukur tingkat akurasi dan kinerja model *transfer learning* untuk pengenalan bahasa isyarat menggunakan *platform* Edge Impulse melalui simulasi perangkat *smartphone*.

1.4 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat bermanfaat bagi pihak-pihak terkait. Adapun manfaat yang diperoleh dari penelitian ini sebagai berikut:

1. Bagi Masyarakat dan Pengembangan Ilmu Pengetahuan, penelitian ini diharapkan memberikan solusi praktis berupa teknologi asistif untuk mengatasi hambatan komunikasi semantik penyandang disabilitas, sehingga mendukung kemandirian dan interaksi sosial yang lebih natural. Selain itu, penelitian ini berkontribusi dalam memperkaya literatur mengenai penerapan teknologi Industri 4.0 (*Internet of Things* dan *Artificial Intelligence*) dalam bidang ergonomi kognitif, serta memberikan validasi teknis terkait integrasi antarmuka *Augmented Reality* pada perangkat *wearable* sebagai referensi pengembangan produk serupa di masa depan.
2. Bagi Mahasiswa, penelitian ini berfungsi sebagai sarana implementasi kompetensi Teknik Industri secara komprehensif, khususnya dalam mengasah keterampilan riset

Perancangan dan Pengembangan Produk (*Product Design and Development*). Melalui proses ini, mahasiswa terlatih untuk menerjemahkan kebutuhan pengguna melalui pendekatan *User Stories* menjadi purwarupa fungsional yang nyata, sekaligus memperdalam pemahaman teknis terkait integrasi *hardware* dan *software* berbasis *Machine Learning*. Selain itu, penelitian ini juga melatih kemampuan manajemen proyek yang adaptif dan disiplin melalui penerapan kerangka kerja Scrum dan siklus PDCA, sehingga memberikan pengalaman praktis dalam menyelesaikan permasalahan ergonomi melalui pendekatan teknologi terkini.

1.5 Batasan Penelitian

Untuk menghindari meluasnya ruang lingkup penelitian dan menjaga fokus pada validasi rekayasa sistem, maka ditetapkan batasan masalah sebagai berikut:

1. Luaran fisik (*hardware*) dari penelitian ini difokuskan pada tahap *Proof of Concept* (PoC) berbasis *breadboard* / sirkuit terbuka, yang dibatasi pada pengujian logika *Keyword Spotting* (KWS) menggunakan mikrofon dan penampil OLED pada mikrokontroler Teensy 4.1.
2. Mengingat adanya batasan *bandwidth* memori dan beban komputasi pada integrasi modul kamera secara paralel pada mikrokontroler tunggal, maka validasi fungsional model visi (pengenalan bahasa isyarat) dilakukan melalui simulasi perangkat *smartphone* yang di-deploy via Edge Impulse sebagai representasi kinerja model sebelum diintegrasikan ke sistem tertanam (*embedded system*).
3. Arsitektur komputasi pusat dikendalikan menggunakan mikrokontroler Teensy 4.1. Akuisisi data terbatas pada modul mikrofon INMP441 untuk *input* suara dan kamera *smartphone* untuk visualisasi, sementara *output* aktuasi diwujudkan melalui layar OLED SSD1306.
4. Kerangka kerja perancangan sistem dibatasi pada pendekatan adaptif yang mengintegrasikan metode manajemen proyek *Scrum* dan siklus pengendalian kualitas *Plan-Do-Check-Act* (PDCA).
5. Implementasi pemrosesan *Machine Learning* dibatasi pada arsitektur *Transfer Learning* (MobileNet V2) untuk metode klasifikasi gambar pada pengenalan isyarat, dan metode *Keyword Spotting* (KWS) langsung pada mikrokontroler untuk fitur deteksi kedaruratan (*speech-to-text*).

6. Mengingat produk masih dalam tahapan *prototype*, sistem hanya dilatih untuk mengenali kosa kata suara dan gerakan isyarat yang terbatas dan sangat spesifik (*trained data*) dengan fokus pada instrumen peringatan keselamatan dasar.
7. Desain fisik kacamata secara utuh tidak difabrikasi dalam bentuk fisik akhir, melainkan diwujudkan dalam bentuk rancangan *High-Fidelity* 3D CAD sebagai usulan integrasi komponen dan evaluasi aspek ergonomis.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Literatur

Dalam perancangan alat bantu bagi penyandang disabilitas, khususnya *smart glasses*, diperlukan pemahaman mendalam mengenai integrasi antara desain produk fisik dengan teknologi kecerdasan buatan. Berikut adalah tinjauan terhadap beberapa penelitian terdahulu yang menjadi landasan pengembangan dalam penelitian ini.

Penelitian pertama yang menjadi rujukan utama dilakukan oleh Asakura (2023) dengan judul "*Augmented-Reality Presentation of Household Sounds for Deaf and Hard-of-Hearing People*". Fokus riset ini sangat relevan karena mencoba memvisualisasikan suara lingkungan seperti bel pintu atau alarm agar bisa "dilihat" oleh penyandang tunarungu. Asakura memanfaatkan perangkat *head-mounted display* (HoloLens) yang mampu mendeteksi arah datangnya suara dan menampilkannya dalam bentuk grafis AR. Temuan penting dari studi ini adalah bahwa visualisasi arah suara secara *real-time* sangat membantu pengguna memahami situasi sekitar mereka. Hal ini menjadi referensi teknis bagi penulis dalam merancang fitur notifikasi visual pada *Vogest Lens*.

Beralih ke aspek penerjemahan bahasa isyarat, Saggio dkk. (2020) melakukan studi komparasi yang menarik dalam "*Sign Language Recognition Using Wearable Electronics*". Mereka membandingkan dua pendekatan algoritma, yaitu *k-Nearest Neighbors* (k-NN) dan *Convolutional Neural Networks* (CNN). Meskipun penelitian mereka menggunakan sarung tangan sensor (*sensory glove*), hasil eksperimennya memberikan wawasan krusial bahwa metode *Deep Learning* (CNN) memiliki performa yang lebih unggul dalam menangani variasi gerakan tangan yang kompleks dibandingkan metode klasik seperti k-NN. Kesimpulan ini memperkuat alasan penulis untuk mengadopsi pendekatan *Deep Learning* dalam sistem pengenalan isyarat di penelitian ini.

Untuk memperdalam pemahaman mengenai metode pengenalan isyarat, penulis juga merujuk pada survei komprehensif yang dilakukan oleh Rastgoo, Kiani, dan Escalera (2021) berjudul "*Sign Language Recognition: A Deep Survey*". Riset ini membedah perkembangan metode *Sign Language Recognition* (SLR) berbasis visi (*vision-based*) selama lima tahun terakhir. Poin kunci yang digarisbawahi adalah pentingnya penggunaan teknik *pose estimation* serta model hibrida untuk menangkap fitur bentuk tangan sekaligus gerakan dinamisnya. Kajian ini menegaskan bahwa penggunaan kamera

(tanpa sensor tempel) adalah arah teknologi masa kini yang paling menjanjikan untuk diterapkan pada perangkat *wearable*.

Pandangan tersebut didukung oleh tinjauan yang lebih baru dari Hashi, Hashim, dan Asamah (2024) dalam "*A Systematic Review of Hand Gesture Recognition: An Update From 2018 to 2024*". Dalam tinjauan sistematis ini, terlihat adanya tren pergeseran preferensi teknologi. Meskipun metode berbasis sensor (seperti sarung tangan) memiliki akurasi tinggi, metode berbasis visi (*vision-based*) makin mendominasi penelitian terkini karena faktor kenyamanan pengguna (*usability*). Temuan ini sangat sejalan dengan konsep *Vogest Lens* yang dirancang non-invasif, di mana pengguna tidak perlu mengenakan atribut tambahan di tangan yang dapat mengganggu aktivitas sehari-hari.

Dari sisi teknis pengembangan kecerdasan buatan, Mujahid dkk. (2021) mendemonstrasikan efektivitas algoritma YOLOv3 dalam "*Real-Time Hand Gesture Recognition Based on Deep Learning*". Penelitian ini membuktikan bahwa arsitektur YOLO mampu melakukan deteksi gestur tangan secara *real-time* dengan latensi rendah, namun memerlukan komponen mikrokontroler yang berspesifikasi tinggi sehingga tidak reliabel terhadap implementasi pada *Vogest Lens*. Hal ini yang mendukung penggunaan *Transfer Learning (MobileNet V2)* yang cocok untuk mikrokontroler kecil.

Dukungan terhadap pendekatan multimodal juga ditemukan dalam tinjauan sistematis oleh Dharma dkk. (2025) tentang teknologi AI untuk deteksi nyeri. Meskipun objek deteksinya berbeda, kesimpulan mereka bahwa integrasi fitur visual (ekspresi wajah/gestur) dan audio memberikan hasil analisis yang jauh lebih presisi dibandingkan metode unimodal, menjadi landasan kuat bagi penulis untuk menggabungkan input kamera dan mikrofon dalam sistem kacamata pintar ini. Hal ini sejalan dengan penelitian Chen dkk. (2020) mengenai model DEBONAIR yang menggabungkan berbagai modalitas input sensor untuk meningkatkan akurasi sistem dalam memahami aktivitas manusia yang kompleks.

Untuk fitur konversi suara-ke-teks (*speech-to-text*), penelitian Putri dan Sularno (2024) mengenai implementasi translator suara pada platform Android memberikan gambaran praktis mengenai integrasi API pengenalan suara. Studi ini menunjukkan bahwa teknologi pengenalan suara berbasis seluler saat ini sudah cukup matang. Namun, untuk memastikan kualitas luaran teks Bahasa Indonesia yang baik, penelitian ini juga mengacu pada Prasetyo, Naufal, dan Wijaya (2025) yang menggunakan metode hibrida CNN dan Bi-LSTM untuk analisis sentimen teks Bahasa Indonesia. Keberhasilan metode

ini dalam menangani struktur linguistik lokal menjadi referensi teknis yang berharga untuk modul pemrosesan teks pada *Vogest Lens*.

Pentingnya memahami konteks sosial pengguna diperkuat oleh penelitian Al Ghifari, Dahlia, dan Syarifah (2025) mengenai strategi komunikasi di lingkungan pendidikan inklusif. Temuan mereka menyoroti bahwa hambatan utama komunikasi bukan hanya pada ketidakmampuan mendengar, melainkan pada ketiadaan umpan balik (*feedback*) yang efektif antara guru dan siswa berkebutuhan khusus. Hal ini memperkuat urgensi fitur komunikasi dua arah (*two-way communication*) pada kacamata pintar yang sedang dirancang.

Mengingat tingginya kompleksitas integrasi sistem tersebut, pendekatan manajemen proyek konvensional dinilai kurang memadai. Natarajan dan Pichai (2024) memaparkan bahwa transisi dari metode *Waterfall* ke metodologi *Agile* terbukti meningkatkan efisiensi pengiriman produk secara signifikan. Fleksibilitas ini juga didukung oleh Atobatele dkk. (2021) yang menunjukkan keberhasilan penerapan *Agile* dan *Scrum* pada proyek yang kompleks. Efektivitas *Scrum* dalam menghasilkan inovasi digital yang responsif terhadap kebutuhan pengguna juga divalidasi oleh Purwanto dkk. (2024), yang menekankan pentingnya iterasi dalam pengembangan sistem untuk memastikan kesesuaian produk akhir dengan harapan pengguna.

Namun, tantangan utama dalam metode *Agile* adalah menjaga konsistensi kualitas teknis di setiap iterasinya. Untuk mengatasi hal ini, Costa dkk. (2023) menyarankan penggunaan siklus PDCA (*Plan-Do-Check-Action*) dalam proses spesifikasi kebutuhan guna meminimalkan ketidaksesuaian fitur sejak dini. Validasi pamungkas untuk metodologi penelitian ini ditemukan dalam karya Pereira dkk. (2021) yang mengembangkan "*A Hybrid Agile-Quality Management Framework*". Mereka berhasil membuktikan bahwa penggabungan praktik *Scrum* dengan alat pengendalian kualitas seperti PDCA dapat meningkatkan produktivitas secara signifikan di lingkungan riset laboratorium. Kerangka kerja hibrida inilah yang kemudian diadopsi sebagai landasan operasional utama dalam perancangan dan pembangunan purwarupa *Vogest Lens* ini.

Tabel 2. 1 Posisi Penelitian

No	Peneliti (Tahun)	Topik/Fokus Penelitian	Metode	Integrasi		Gap Penelitian
				Teknologi	Multimodal	
1	(Asakura, 2023)	Visualisasi Suara (AR)	Eksperimental	✓	-	Hanya <i>Output</i> (1 Arah)
2	(Saggio et al., 2020)	SLR dengan <i>Wearable</i>	Komparatif	✓	-	Invasif (Pakai Sarung Tangan)
3	(Rastgoo et al., 2021)	Survei SLR (<i>Deep Learning</i>)	Tinjauan Pustaka	-	-	Hanya Kajian Teori
4	(Hashi et al., 2024)	Survei Gestur Tangan	Tinjauan Sistematis	-	-	Hanya Kajian Tren Teknologi
5	(Mujahid et al., 2021)	Deteksi Tangan Real-time	Eksperimental	✓	-	Hanya Uji Algoritma
6	(Dharma et al., 2025)	Aktivitas Manusia Kompleks	Eksperimental	✓	✓	Bukan Fokus Bahasa Isyarat
7	(L. Chen et al., 2021)	Deteksi Nyeri (Medis)	Tinjauan Sistematis	✓	✓	Domain Medis (Bukan Komunikasi)
8	(Putri & Sularno, 2024)	Aplikasi Penerjemah Suara	<i>Prototyping</i> (App)	-	-	Tidak Hands-free (Wajib HP)
9	(Prasetyo et al., 2025)	Analisis Sentimen Teks	Eksperimental	✓	-	Hanya Software (Tanpa Alat)
10	(Al Ghifari et al., 2025)	Strategi Komunikasi Inklusif	Kualitatif	-	-	Kajian Sosial (Non-Teknis)

N o	Peneliti (Tahun)	Topik/Fokus Penelitian	Metode	Integrasi Teknologi	Multimodal	Gap Penelitian
11	(Natarajan & Pichai, 2024)	Transisi Waterfall ke Agile	<i>Action Research</i>	-	-	Studi Manajemen Tim Software
12	(Atobatele et al., 2021)	Manajemen Proyek Kesehatan	<i>Agile & Scrum</i>	-	-	Implementasi Non-Produk
13	(Purwanto et al., 2024)	Platform Web Digital	<i>Scrum</i>	-	-	Hanya Berbasis Web
14	(Costa et al., 2023)	Kualitas Spesifikasi (Req.)	<i>PDCA Cycle</i>	-	-	Hanya Fokus Dokumen Mutu
15	(W. A. Pereira et al., 2026)	Manajemen Lab Riset	<i>Hybrid (Scrum+PDCA)</i>	-	-	Fokus Manajemen SDM
16	Penelitian Ini (2026)	Smart Glasses Difabel	Hybrid (<i>Scrum+PDCA</i>)	✓ (<i>MobileN et V2+NLP</i>)	✓ (<i>2 Way I/O</i>)	Integrasi Hardware & AI

Dari rangkuman Tabel 2.1, terlihat bahwa mayoritas penelitian sebelumnya masih bersifat parsial, yaitu hanya berfokus pada satu arah komunikasi (isyarat ke teks atau suara ke teks) dan cenderung menggunakan metode pengembangan linear yang kaku. Penelitian ini mengisi celah tersebut melalui pengembangan Vogest Lens, sebuah perangkat *wearable* multimodal yang menggabungkan deteksi visual (YOLO) dan pengolahan suara (NLP) dalam satu sistem mandiri. Perbedaan utama penelitian ini dibandingkan studi terdahulu terletak pada integrasi metode Scrum dengan siklus PDCA (*Plan-Do-Check-Act*). Pendekatan gabungan ini sengaja diterapkan untuk mengatasi kerumitan teknis pada pengembangan perangkat keras, sehingga proses pembuatan fitur bisa berjalan cepat (*agile*) namun tetap memiliki standar kontrol kualitas yang jelas di setiap tahapannya.

2.2 Landasan Teori

Landasan teori memuat penjabaran teoritis mengenai konsep dan metode yang digunakan dalam penelitian ini sebagai dasar pemecahan masalah. Teori-teori yang dibahas meliputi karakteristik pengguna (disabilitas), konsep interaksi multimodal, teknologi pemrosesan bahasa, serta metodologi perancangan produk.

2.2.1 *Disabilitas Tunarungu dan Tunawicara*

Berdasarkan Pasal 1 Ayat (1) Undang-Undang Nomor 8 Tahun 2016, disabilitas didefinisikan sebagai kondisi keterbatasan fisik, mental, intelektual, atau sensorik dalam jangka waktu yang lama. Keterbatasan ini menghambat partisipasi penuh dan efektif di masyarakat, yang muncul akibat adanya interaksi antara gangguan (*impairment*) yang dimiliki individu dengan hambatan lingkungan serta sikap social (Wiraputra, 2021). Sedangkan, tunarungu merujuk pada kondisi gangguan fungsi pendengaran yang bervariasi tingkatannya, mulai dari kehilangan pendengaran sebagian (*hard of hearing*) hingga kehilangan total (*deaf*). Kondisi ini menyebabkan individu mengalami hambatan signifikan dalam menerima dan memproses informasi audio secara optimal, meskipun dalam beberapa kasus, sisa pendengaran (*residual hearing*) masih dapat dimanfaatkan dengan bantuan alat amplifikasi (Wahyudi, Aziizah, Solihah, Putri, & Hamidah, 2024).

Di sisi lain, tunawicara didefinisikan sebagai kondisi yang ditandai dengan adanya hambatan dalam memproduksi ujaran verbal (*speech production*) atau artikulasi suara. Gangguan ini berdampak langsung pada kemampuan individu untuk menyampaikan pesan, gagasan, atau emosi secara lisan kepada orang lain (Utami & Duita, 2024). Dalam banyak kasus, penyandang tunarungu juga menyandang ketunawicaraan (tunarungu-wicara) karena ketiadaan stimulus suara sejak dini yang menghambat perkembangan organ bicara.

2.2.2 *Karakteristik Penyandang Tunarungu*

Pemahaman mendalam mengenai karakteristik penyandang tunarungu sangat diperlukan dalam perancangan alat bantu (*assistive device*) agar solusi yang ditawarkan sesuai dengan kebutuhan pengguna. Berdasarkan studi terdahulu, karakteristik penyandang tunarungu dapat diklasifikasikan ke dalam beberapa aspek utama, yaitu aspek fisik, bahasa, emosi, dan kecerdasan (Arnida et al., 2024).

1. Karakteristik Fisik

Secara sepintas, penyandang tunarungu tidak memiliki karakteristik fisik yang berbeda secara signifikan dibandingkan dengan individu yang mendengar (*hearing people*).

Namun, terdapat beberapa ciri khas motorik yang sering muncul sebagai dampak sekunder dari gangguan pendengaran. Penyandang tunarungu sering kali memiliki cara berjalan yang kaku dan sedikit membungkuk. Hal ini disebabkan oleh gangguan pada organ keseimbangan yang terletak di telinga bagian dalam (Arnida et al., 2024). Selain itu, pola pernapasan mereka cenderung pendek dan tidak teratur karena kurangnya latihan vokal dalam berbicara sehari-hari.

2. Karakteristik Bahasa dan Bicara

Hambatan terbesar bagi penyandang tunarungu terletak pada aspek kebahasaan. Karena tidak memperoleh umpan balik auditif sejak dini, mereka mengalami kesulitan dalam menguasai kosakata (*vocabulary*) dan tata bahasa. Karakteristik khas yang ditemukan meliputi:

A. Miskin Kosakata

Terbatasnya jumlah kata yang dipahami, terutama kata-kata abstrak.

B. Artikulasi Tidak Jelas

Suara yang dihasilkan sering kali sulit dimengerti oleh orang awam karena kontrol pita suara yang lemah.

C. Irama Bicara Monoton

Ketiadaan intonasi atau pemenggalan kalimat yang tepat. Kondisi ini menegaskan perlunya teknologi penerjemah suara-ke-teks (*speech-to-text*) untuk membantu mereka memahami lawan bicara tanpa harus bergantung pada pendengaran.

3. Karakteristik Emosi dan Sosial

Keterbatasan dalam berkomunikasi sering kali berdampak pada kondisi psikologis dan interaksi sosial. Penyandang tunarungu cenderung memiliki sifat egosentris yang lebih tinggi, sering merasa curiga terhadap lingkungan, dan mudah tersinggung (Arnida et al., 2024). Sikap ini muncul sebagai mekanisme pertahanan diri karena ketidakmampuan mereka memahami pembicaraan orang-orang di sekitarnya, yang sering kali disalahartikan sebagai pembicaraan negatif tentang diri mereka. Alat bantu komunikasi dua arah diharapkan dapat mereduksi kesalahpahaman sosial ini dengan memberikan akses informasi yang transparan.

4. Karakteristik Kecerdasan (Intelegensi)

Penting untuk dicatat bahwa ketunarunguan tidak berbanding lurus dengan kecerdasan intelektual. Secara potensi, anak tunarungu memiliki kapasitas IQ yang

sama dengan anak normal. Namun, prestasi akademis mereka sering kali terlihat rendah bukan karena ketidakmampuan berpikir, melainkan karena keterbatasan bahasa dalam memahami instruksi atau materi. Mereka adalah pemelajar visual (*visual learner*) yang sangat mengandalkan indra penglihatan (Arnida et al., 2024). Oleh karena itu, penggunaan media visual seperti *smart glasses* yang menampilkan teks/gambar merupakan pendekatan yang paling efektif untuk mendukung proses kognitif mereka.

2.2.3 Karakteristik Penyandang Tunawicara

Tunawicara sering kali tidak berdiri sendiri sebagai disabilitas tunggal, melainkan hadir menyertai ketunarunguan, kondisi yang dikenal sebagai tunaganda. Penelitian sebelumnya menjelaskan bahwa karakteristik utama penyandang tunawicara adalah adanya hambatan signifikan dalam komunikasi lisan (*oral communication*). Hambatan ini bukan hanya pada ketidakmampuan memproduksi suara, tetapi juga kesulitan dalam memahami pesan verbal jika tidak disertai dengan visualisasi yang jelas (Irchamna et al., 2025).

1. Ketergantungan pada Komunikasi Non-Verbal

Karakteristik paling menonjol dari penyandang tunawicara adalah preferensi kuat terhadap komunikasi non-verbal. Berdasarkan penelitian sebelumnya, penggunaan bahasa isyarat memiliki pengaruh yang sangat signifikan terhadap pemahaman mereka. Ketika interaksi dilakukan menggunakan bahasa isyarat, penyandang tunawicara menunjukkan tingkat pemahaman instruksi yang jauh lebih baik dibandingkan dengan metode lisan murni (Irchamna et al., 2025). Hal ini menegaskan bahwa bagi mereka, bahasa isyarat bukan sekadar pelengkap, melainkan modalitas utama dalam mentransfer informasi.

2. Pola Interaksi Sosial dan Kepercayaan Diri

Hambatan wicara berdampak langsung pada psikologis dan pola interaksi sosial. Dalam studi sebelumnya ditemukan bahwa penyandang tunawicara cenderung pasif dalam lingkungan yang hanya mengandalkan komunikasi lisan. Namun, karakteristik ini berubah menjadi aktif dan partisipatif ketika lawan bicara menggunakan atau difasilitasi oleh bahasa isyarat. Adanya jembatan komunikasi (seperti isyarat) membuat mereka merasa lebih nyaman, berani terlibat dalam diskusi, dan mampu mengekspresikan diri (Irchamna et al., 2025). Temuan ini menjadi landasan kuat bagi pengembangan teknologi asistif yang mampu menerjemahkan isyarat ke dalam

bentuk yang dipahami masyarakat luas, guna meningkatkan kepercayaan diri pengguna.

2.2.4 *Interaksi Multimodal (Multimodal Interaction)*

Interaksi Multimodal merupakan paradigma dalam sistem antarmuka komputer yang mensimulasikan pola komunikasi alami manusia dengan menggabungkan berbagai modalitas input dan output secara simultan untuk menciptakan pengalaman interaksi yang inklusif. Dalam teknologi asistif modern, pendekatan ini telah bergeser dari sekadar penerjemahan visual pasif menjadi sistem transformatif yang mampu mengubah gestur tubuh menjadi modalitas audio atau suara (*gesture-to-voice*) secara *real-time*. Pergeseran orientasi teknologi ini sangat krusial untuk menjembatani kesenjangan interaksi antara komunitas Tuli dan masyarakat dengar, memungkinkan terjadinya komunikasi yang mulus (*seamless*) tanpa hambatan medium penyampaian (Zichong, 2025). Efektivitas sistem ini semakin optimal ketika diimplementasikan pada perangkat wearable seperti kacamata pintar dengan sudut pandang orang pertama (*First-Person View*), yang memungkinkan pengambilan data visual gestur tangan sesuai perspektif mata pengguna untuk akurasi tinggi dan mobilitas yang lebih fleksibel (Cai et al., 2025).

Sebuah sistem komunikasi asistif yang ideal mensyaratkan adanya pertukaran informasi dua arah (*two-way communication loop*) yang terintegrasi penuh. Sistem tidak cukup hanya menerjemahkan isyarat ke suara, tetapi juga harus memfasilitasi konversi ucapan lawan bicara menjadi format visual yang dapat dipahami pengguna tunarungu. Integrasi antara pengenalan isyarat, *speech recognition*, dan visualisasi teks berbasis *Augmented Reality* (AR) menjadi kunci utama dalam menciptakan siklus komunikasi timbal balik yang responsif (Lee et al., 2021).

2.2.5 *Machine Learning*

Machine Learning (ML) merupakan cabang fundamental dari kecerdasan buatan yang memungkinkan sistem komputer untuk membangun kemampuan analitis secara mandiri melalui pengalaman data, tanpa perlu diprogram secara eksplisit untuk setiap aturan. Jika pemrograman tradisional mengharuskan manusia menyusun instruksi logika yang kaku, algoritma ML justru dirancang untuk mendeteksi pola tersembunyi dalam kumpulan data yang besar dan kompleks untuk menghasilkan keputusan yang cerdas (Tufail et al., 2023).

Secara fundamental, mekanisme pembelajaran pada mesin berbeda dengan pembelajaran biologis pada manusia. Manusia memproses informasi menggunakan jaringan saraf biologis yang fleksibel dan intuitif, sedangkan Machine Learning

memproses informasi berdasarkan algoritma matematika terstruktur yang bekerja di atas data statistik (Sarıkaya, 2024). Dalam konteks sistem informasi, ML berperan sebagai inti dari agen cerdas yang mampu menerima persepsi dari lingkungannya dan mengambil tindakan yang memaksimalkan peluang keberhasilan untuk mencapai tujuan tertentu ((Kühl et al., 2022).

Kekuatan utama dari teknologi ini terletak pada kemampuannya untuk melakukan generalisasi. Sistem ML tidak sekadar menghafal data, melainkan mempelajari fitur-fitur penting dari data latih (*training data*) agar mampu memberikan prediksi atau klasifikasi yang akurat ketika dihadapkan pada data baru yang belum pernah ditemui sebelumnya (Kufel et al., 2023). Kemampuan adaptasi berbasis data inilah yang menjadikan Machine Learning solusi yang jauh lebih efisien dibandingkan metode statistik konvensional dalam menangani permasalahan yang memiliki variabilitas tinggi, seperti pengenalan pola yang rumit (Sharifani & Amini, 2023).

2.2.6 *Deep Learning*

Deep Learning hadir sebagai evolusi dari algoritma pembelajaran mesin yang dirancang untuk meniru kemampuan persepsi manusia melalui arsitektur jaringan saraf tiruan yang mendalam. Berbeda dengan metode komputasi konvensional yang sangat bergantung pada aturan baku, pendekatan ini memungkinkan sistem untuk mempelajari representasi fitur secara bertingkat langsung dari data mentah, mulai dari elemen sederhana seperti garis hingga bentuk objek yang kompleks (Sapkota et al., 2025).

Adapun arsitektur yang menjadi standar utama dalam pengolahan data visual adalah *Convolutional Neural Network* (CNN). Keunggulan utama CNN terletak pada kemampuannya untuk melakukan ekstraksi fitur secara otomatis tanpa memerlukan intervensi manual dari manusia untuk menentukan karakteristik objek. Lapisan konvolusi pada jaringan ini bekerja menyaring informasi visual melalui operasi matematika pada piksel gambar, sehingga pola penting tetap dapat dikenali meskipun terjadi perubahan posisi atau pencahayaan pada objek (Ali & Zhang, 2024).

Kecepatan pemrosesan menjadi faktor krusial ketika sistem dituntut untuk mengenali objek secara langsung saat kejadian berlangsung. Model *Deep Learning* modern tidak hanya diharuskan memiliki akurasi tinggi, tetapi juga kecepatan inferensi yang memadai agar respons yang diberikan bisa instan. Tren pengembangan algoritma saat ini semakin fokus pada pencapaian keseimbangan antara kecepatan deteksi dan ketepatan klasifikasi untuk kebutuhan aplikasi dunia nyata (Gheorghe et al., 2024).

Namun, penerapan kecerdasan buatan pada perangkat keras dengan sumber daya terbatas memiliki tantangan tersendiri terkait konsumsi memori dan daya listrik. Guna mengatasi hal tersebut, penelitian terkini banyak berfokus pada pengembangan arsitektur jaringan yang lebih ringan (*lightweight networks*) namun tetap mempertahankan kinerja deteksi yang optimal. Efisiensi komputasi menjadi prioritas utama pada sistem tertanam dibandingkan sekadar mengejar angka akurasi yang absolut (Zhao et al., 2021).

Terlepas dari jenis algoritma yang digunakan, keberhasilan model *Deep Learning* tetap sangat bergantung pada kualitas dataset dan proses pelatihannya. Kemampuan sistem untuk mengenali objek baru di lapangan sangat dipengaruhi oleh keragaman data latih, yang berfungsi memastikan bahwa model benar-benar memahami karakteristik unik objek dan tidak sekadar menghafal pola data latihan saja (A. Chen, 2025).

2.2.7 *Tiny Machine Learning (TinyML)*

Tiny Machine Learning atau TinyML merupakan paradigma baru dalam bidang kecerdasan buatan yang berfokus pada pengembangan dan penerapan model pembelajaran mesin pada perangkat keras dengan sumber daya yang sangat terbatas, khususnya mikrokontroler. Jika *Deep Learning* konvensional umumnya dijalankan pada *server cloud* atau komputer dengan GPU berkinerja tinggi, TinyML justru memungkinkan proses inferensi cerdas dilakukan langsung di "ujung" jaringan (*edge*) tempat data diambil, tanpa ketergantungan konstan pada koneksi internet (Immonen & Hämäläinen, 2022).

Adapun motivasi utama di balik penerapan TinyML adalah untuk mengatasi tantangan latensi dan efisiensi energi yang sering ditemui pada sistem berbasis cloud. Dengan memproses data secara lokal pada perangkat, keputusan dapat diambil secara real-time tanpa jeda waktu pengiriman data ke server, yang mana hal ini sangat krusial untuk aplikasi yang membutuhkan respons instan (Pazmiño Ortiz et al., 2025). Selain itu, pendekatan ini juga menawarkan keuntungan signifikan dari sisi privasi dan keamanan, karena data sensitif pengguna tidak perlu meninggalkan perangkat untuk diproses oleh pihak ketiga (Villegas-Ch et al., 2024).

Namun, implementasi kecerdasan buatan pada mikrokontroler menghadirkan tantangan teknis yang kompleks, terutama terkait keterbatasan kapasitas memori (SRAM dan Flash) serta kecepatan komputasi yang jauh lebih rendah dibandingkan prosesor standar. Oleh karena itu, penerapan TinyML sering kali melibatkan teknik optimasi model seperti kuantisasi (*quantization*) dan pemangkasan (*pruning*), yang bertujuan untuk

mengecilkan ukuran model dan mengurangi beban komputasi tanpa mengorbankan akurasi prediksi secara signifikan (Gruosso & Gajani, 2022).

Hal ini sejalan dengan perkembangan ekosistem Internet of Things (IoT) masa depan, di mana perangkat tidak hanya berfungsi sebagai pengirim data pasif, tetapi juga sebagai agen cerdas yang mampu menganalisis situasi lingkungan secara mandiri. Integrasi antara efisiensi algoritma dan keterbatasan perangkat keras inilah yang menjadi kunci keberhasilan dalam menghadirkan solusi cerdas yang hemat daya dan biaya (Pazmiño Ortiz et al., 2025).

2.2.8 Natural Language Processing (NLP)

Natural Language Processing (NLP) merupakan domain dalam kecerdasan buatan yang menjembatani interaksi antara komputer dan bahasa alami manusia, memungkinkan sistem untuk menganalisis, memahami, dan memproduksi bahasa secara semantik, bukan sekadar pencocokan kata kunci. Dalam penerapannya, teknologi ini memanfaatkan algoritma pemodelan topik seperti pendekatan *unsupervised* untuk mengkategorikan dokumen atau input komunikasi berdasarkan konteks maknanya, sehingga data tidak terstruktur dapat diolah menjadi informasi yang terorganisir (Akbay, 2022). Kemampuan semantik ini menjadi fondasi krusial dalam sistem penerjemah cerdas, di mana pemahaman konteks sangat menentukan akurasi alih bahasa dari satu modalitas ke modalitas lainnya.

Meskipun lanskap pengembangan NLP saat ini didominasi oleh *Deep Learning* dan *Large Language Models* (LLMs) yang menawarkan performa tinggi pada berbagai benchmark, tantangan mengenai kekokohan (robustness) sistem terhadap variabilitas input di dunia nyata masih menjadi fokus evaluasi utama (Omar et al., 2022). Menariknya, dominasi model saraf kompleks ini tidak serta merta menghilangkan peran metode NLP tradisional. Pendekatan klasik tetap relevan digunakan, terutama dalam tugas klasifikasi dan ekstraksi informasi spesifik yang menuntut efisiensi komputasi tinggi (Jegan & Henrich, 2025).

Berkaitan dengan pengembangan arsitektur model, fokus penelitian kini mengarah pada validasi dan spesialisasi kemampuan *pre-trained models*. Inovasi tersebut mencakup peningkatan akurasi pemblokiran entitas (Wang & Zhang, 2024) serta penguatan algoritma ekstraksi relasi pada dokumen yang kaya akan unsur visual (Ramu, Wang, Mouatadid, Rimchala, & Huang, 2024). Di samping itu, aspek interaktif dalam sistem percakapan juga terus disempurnakan melalui teknik *response generation* dinamis, yang

bertujuan menciptakan pengalaman komunikasi timbal balik yang lebih adaptif dan menyerupai interaksi alami manusia (Nie et al., 2024).

2.2.9 *Deep Spoken Keyword Spotting (KWS)*

Keyword Spotting (KWS) merupakan teknologi pendeteksian kata kunci spesifik pada aliran audio berkelanjutan yang kini telah bergeser dari metode konvensional menuju arsitektur *Deep Neural Networks* (DNN) berkat kemampuannya mengekstraksi fitur fonetik secara lebih akurat untuk perangkat cerdas (Lopez-Espejo et al., 2022). Meskipun adopsi DNN membawa peningkatan performa yang signifikan, implementasi KWS pada sistem tertanam (*embedded systems*) sering kali terhambat oleh kerentanan model terhadap derau lingkungan (*noise*) dan artefak kompresi audio (P. H. Pereira et al., 2023). Untuk mengatasi degradasi akurasi akibat pergeseran domain (*domain shift*) dari lingkungan bising, penelitian terdahulu mengusulkan strategi pelatihan adaptasi domain adversarial (SPMDAT) yang secara efektif mampu mengekstraksi fitur audio yang kebal terhadap bising latar belakang (Y. Chen et al., 2025). Di sisi lain, guna menanggulangi keterbatasan data latih berlabel yang spesifik sekaligus memperkuat resiliensi akustik, tren penelitian kontemporer mengadopsi teknik *Self-Supervised Learning* berbasis *Data2Vec* (Mørk et al., 2024) serta *Transfer Learning* seperti kerangka *Wav2KWS* (Seo et al., 2021). Melalui pemanfaatan representasi suara dari model berskala besar yang sudah ada, strategi-strategi tersebut terbukti mampu mempercepat konvergensi pelatihan dan mempertahankan akurasi tinggi pada lingkungan bising, sehingga sangat ideal untuk mendukung viabilitas KWS pada perangkat *edge computing* dengan sumber daya memori dan komputasi yang terbatas.

2.2.10 *Augmented Reality (AR) dan Visualisasi Informasi*

Visualisasi terjemahan suara ke teks pada alat ini dibangun menggunakan prinsip *Augmented Reality* (AR). Berbeda dengan *Virtual Reality* yang mengisolasi pengguna dari dunia luar, AR bekerja dengan metode superimposisi, yaitu meletakkan lapisan informasi digital tepat di atas pandangan dunia nyata (Cossich et al., 2023). Metode ini dipilih secara spesifik agar pengguna tunarungu dapat membaca teks percakapan tepat pada garis pandang atau *line-of-sight*. Dengan demikian, beban kognitif pengguna menjadi lebih ringan karena mereka tidak perlu lagi memecah fokus antara membaca layar gawai dan memperhatikan ekspresi lawan bicara (Marques et al., 2023).

Secara teknis, visualisasi tersebut diwujudkan melalui modul optik *Waveguide*. Teknologi ini memungkinkan kacamata pintar menyajikan apa yang disebut sebagai

ruang konten yang menyatu (*unified content space*) antara objek digital dan fisik (Langner et al., 2021). Keunggulan utama sistem ini terletak pada transparansi optiknya. Teks terjemahan dapat muncul melayang tanpa menghalangi pandangan jalan, sebuah fitur keselamatan yang mutlak diperlukan agar pengguna tetap awas terhadap lingkungan sekitar saat berjalan (Asakura, 2023).

2.2.11 *Embedded System & Wearable*

A. Paradigma *Edge Computing* pada Perangkat Terbatas

Implementasi sistem cerdas yang bergantung sepenuhnya pada *Cloud Computing* sering kali terhambat oleh masalah latensi jaringan dan keterbatasan bandwidth. Untuk mengatasi hal ini, teknik *Edge Computing* diterapkan dengan memproses data langsung di lokasi sumbernya, sehingga beban pengiriman data ke server berkurang dan waktu respons menjadi jauh lebih singkat, sebuah faktor krusial untuk aplikasi waktu nyata (Suryadi et al., 2024).

Walaupun efektif, menjalankan algoritma *Deep Learning* pada perangkat tepi (*edge devices*) memiliki kendala teknis tersendiri. Model kecerdasan buatan umumnya membutuhkan daya komputasi dan penyimpanan besar yang sering kali melampaui kapasitas perangkat portabel dengan sumber daya terbatas (Matsubara et al., 2022). Akibatnya, keberhasilan sistem *Embedded AI* sangat bergantung pada kesesuaian antara kompleksitas model dan spesifikasi perangkat keras. Diperlukan analisis mendalam terhadap batasan memori dan konsumsi daya untuk mencapai keseimbangan optimal antara performa deteksi dan durasi operasional perangkat (Garcia-Perez et al., 2023).

B. Arsitektur Mikrokontroler Kinerja Tinggi

Pemilihan unit pemrosesan utama untuk kacamata pintar menuntut kompromi ketat antara performa komputasi dan ketahanan energi. Meskipun *Single Board Computer* seperti *Raspberry Pi* menawarkan daya pemrosesan besar, konsumsi energinya yang tinggi membuatnya tidak praktis untuk perangkat wearable yang bergantung pada baterai berkapasitas kecil. Mikrokontroler berbasis ARM *Cortex-M7* seperti keluarga *Teensy* terbukti menjadi solusi yang lebih efisien karena mampu memberikan kinerja mendekati komputer mini namun dengan profil konsumsi arus yang jauh lebih rendah dan stabil (Lambert et al., 2021).

Keunggulan spesifik *Teensy 4.1* terletak pada kecepatan *clock* 600 MHz yang sangat mumpuni untuk menangani pemrosesan sinyal digital atau *Digital Signal*

Processing (DSP) secara waktu nyata. Kemampuan ini menjadi syarat mutlak dalam sistem multimodal, di mana input audio dari mikrofon harus diproses dan diterjemahkan tanpa jeda yang berarti (Munroe et al., 2022). Keandalan platform ini dalam menangkap dan mengolah data akustik presisi dengan latensi rendah telah divalidasi dalam berbagai pengembangan instrumen pemantauan suara (De Mello et al., 2021). Tidak hanya unggul di sektor audio, arsitektur ini juga terbukti tangguh menjalankan algoritma manajemen data yang kompleks pada sistem *Internet of Things* (IoT) meskipun memiliki keterbatasan memori bawaan (Idrees et al., 2025).

C. Teknologi Tampilan Optik dan *Augmented Reality* (AR)

Visualisasi terjemahan suara ke teks pada alat ini dibangun menggunakan prinsip *Augmented Reality* (AR). Berbeda dengan *Virtual Reality* yang mengisolasi pengguna dari dunia luar, AR bekerja dengan metode superimposisi, yaitu meletakkan lapisan informasi digital tepat di atas pandangan dunia nyata (Cossich et al., 2023). Metode ini dipilih secara spesifik agar pengguna tunarungu dapat membaca teks percakapan tepat pada garis pandang atau *line-of-sight*. Dengan demikian, beban kognitif pengguna menjadi lebih ringan karena mereka tidak perlu lagi memecah fokus antara membaca layar gawai dan memperhatikan ekspresi lawan bicara (Marques et al., 2023).

Secara teknis, visualisasi tersebut diwujudkan melalui modul optik *Waveguide*. Teknologi ini memungkinkanacamata pintar menyajikan apa yang disebut sebagai ruang konten yang menyatu (*unified content space*) antara objek digital dan fisik (Langner et al., 2021). Keunggulan utama sistem ini terletak pada transparansi optiknya. Teks terjemahan dapat muncul melayang tanpa menghalangi pandangan jalan, sebuah fitur keselamatan yang mutlak diperlukan agar pengguna tetap awas terhadap lingkungan sekitar saat berjalan (Asakura, 2023).

Realisasi perangkat keras ini mengandalkan arsitektur *diffractive waveguide combiner*. Mekanisme utamanya bekerja berdasarkan prinsip *Total Internal Reflection* (TIR), di mana cahaya dari proyektor mikro ditangkap oleh struktur *in-coupler*, kemudian dipantulkan berulang kali di dalam substrat kaca tipis sebelum dibelokkan keluar menuju mata melalui *out-coupler* (Kress & Chatterjee, 2021). Teknik propagasi cahaya ini memungkinkan seluruh sistem optik disembunyikan di dalam lensa setipisacamata standar, menghilangkan kebutuhan akan prisma tebal yang membebani desain fisik perangkat.

Penggunaan *waveguide* juga memberikan solusi atas keterbatasan sudut pandang melalui perluasan area *eyebow*. Berbeda dengan optik geometris konvensional yang menuntut posisi mata sangat presisi, teknologi ini memberikan toleransi pergeseran visual yang signifikan. Pengguna tetap dapat membaca teks terjemahan dengan utuh meskipun terjadi sedikit gerakan pada kacamata atau bola mata (Rolland & Goodsell, 2024). Stabilitas visual ini didukung oleh kontrol interaksi cahaya koheren yang dirancang untuk meminimalkan konflik akomodasi, sehingga resolusi gambar tetap tajam dan nyaman dibaca tanpa memicu kelelahan mata (Jang et al., 2024).

2.2.12 Integrasi Metode Pengembangan Scrum dan Siklus PDCA

Pengembangan perangkat wearable multimodal yang mengintegrasikan *embedded system* dengan kecerdasan buatan memiliki tingkat kompleksitas teknis yang tinggi. Penerapan metode pengembangan linear konvensional (*Waterfall*) dinilai kurang relevan karena karakteristiknya yang kaku dan tidak responsif terhadap perubahan spesifikasi yang dinamis selama proses riset (Natarajan & Pichai, 2024). Oleh karena itu, penelitian ini beralih menerapkan metodologi Agile yang menekankan pada fleksibilitas, keterlibatan pemangku kepentingan, serta pengiriman fitur fungsional secara iteratif (Atobatele et al., 2021).

Secara spesifik, kerangka kerja yang diadopsi adalah model *hybrid* yang menyinergikan manajemen proyek Scrum dengan siklus penjaminan mutu PDCA (*Plan-Do-Check-Act*). Scrum digunakan untuk memecah beban kerja pengembangan menjadi iterasi pendek (*Sprints*) guna mempercepat realisasi purwarupa, sementara siklus PDCA diintegrasikan ke dalam setiap Sprint sebagai mekanisme kontrol kualitas sistematis (W. A. Pereira et al., 2026). Pendekatan hibrida ini terbukti efektif dalam menjembatani kesenjangan antara kecepatan pengembangan *Agile* dan standar ketelitian ilmiah, memastikan bahwa sistem yang dihasilkan tidak hanya selesai tepat waktu tetapi juga memiliki kualitas yang teruji (Setyoko et al., 2024). Operasionalisasi metode hibrida tersebut dipetakan ke dalam empat tahapan siklus sebagai berikut:

1. *Plan* (Perencanaan *Sprint*)

Tahap ini berfokus pada penyusunan *Sprint Backlog* berdasarkan prioritas fitur utama. Pada fase ini, spesifikasi kebutuhan teknis (*requirements*) didefinisikan secara mendetail untuk meminimalisir ambiguitas, meliputi penetapan target perakitan perangkat keras maupun batasan parameter model kecerdasan buatan yang akan dikembangkan (Costa et al., 2023).

2. *Do* (Pelaksanaan Teknis)

Merupakan fase eksekusi di mana perancangan sistem dilakukan sesuai durasi *Sprint* yang ditetapkan. Kegiatan utama mencakup penulisan kode program, integrasi sensor pada mikrokontroler Teensy, serta pelatihan model *machine learning*. Penerapan kerangka kerja Scrum pada tahap ini memungkinkan transparansi progres kerja melalui pertemuan harian (*Daily Scrum*) guna mengidentifikasi hambatan teknis sedini mungkin (Legowo et al., 2019).

3. *Check* (Evaluasi dan Verifikasi)

Tahap verifikasi dilakukan untuk mengukur kesesuaian hasil kerja dengan metrik keberhasilan yang telah ditentukan. Pengujian dilakukan baik secara fungsional maupun performa seperti pengukuran akurasi deteksi objek dan latensi system guna memastikan tidak adanya cacat desain sebelum fitur dinyatakan selesai.

4. *Act* (Tindak Lanjut & Perbaikan)

Berdasarkan hasil evaluasi, analisis kausalitas dilakukan terhadap anomali atau *bug* yang ditemukan. Langkah ini mengacu pada prinsip perbaikan berkelanjutan (*continuous improvement*), di mana solusi atas kendala teknis diformulasikan menjadi tindakan perbaikan (*refinement*) yang akan dieksekusi pada siklus perencanaan berikutnya (Arredondo-Soto et al., 2021).

2.2.13 Metrik Evaluasi Kinerja Sistem

Membangun sistem cerdas pada perangkat keras (*embedded systems*) memiliki tantangan yang jauh lebih kompleks dibandingkan menjalankannya di server komputer biasa. Penilaian keberhasilannya tidak bisa hanya dilihat dari seberapa benar tebakan modelnya, tetapi juga harus memperhitungkan beban fisik yang ditanggung oleh perangkat keras tersebut. Evaluasi yang menyeluruh pada perangkat *edge* wajib melihat dua aspek sekaligus: kinerja statistik dari model klasifikasi dan efisiensi sumber daya komputasinya (Shuvo et al., 2023).

A. Evaluasi Klasifikasi Bahasa Isyarat

Validasi kinerja model dalam menerjemahkan bahasa isyarat visual dilakukan menggunakan pendekatan klasifikasi gambar. Landasan evaluasi kuantitatif pada metode ini mengacu pada analisis *Confusion Matrix*. Matriks ini berfungsi memetakan perbandingan antara prediksi kelas yang dihasilkan sistem terhadap label kebenaran (*ground truth*) dari dataset yang diuji (Lopez-Montiel et al., 2021).

Berdasarkan pemetaan tersebut, hasil deteksi dikategorikan ke dalam empat variabel fundamental:

1. *True Positive* (TP)

Sistem dengan tepat memprediksi gerakan isyarat yang benar sesuai labelnya.

2. *True Negative* (TN)

Sistem dengan tepat mengenali bahwa tidak ada gerakan isyarat yang dilakukan.

3. *False Positive* (FP)

Kesalahan di mana sistem mendeteksi adanya isyarat padahal tidak ada (deteksi palsu).

4. *False Negative* (FN)

Kegagalan sistem dalam mendeteksi isyarat yang seharusnya muncul.

Dari variabel di atas, metrik evaluasi tidak hanya berhenti pada nilai Akurasi (*Accuracy*). Mengingat kemungkinan adanya ketidakseimbangan data pada variasi gerakan tangan, penelitian ini juga menggunakan parameter *F1-Score*. Metrik ini merupakan rata-rata harmonis antara *Precision* (tingkat ketepatan prediksi positif) dan *Recall* (sensitivitas sistem dalam menemukan seluruh data positif), sehingga memberikan gambaran yang lebih objektif mengenai keandalan sistem dalam mengenali abjad isyarat (S. Chen et al., 2022).

B. Evaluasi Suara (*Speech-to-Text* & WER)

Validasi kinerja model *Speech-to-Text* (STT) bertujuan untuk mengukur sejauh mana sistem mampu mentranskripsikan sinyal audio menjadi teks yang akurat dan dapat dipahami oleh pengguna. Dalam konteks pengembangan alat bantu komunikasi bagi komunitas Tuli (*Deaf community*), evaluasi ini menjadi sangat krusial karena sistem tidak hanya dituntut untuk memiliki ketepatan transkripsi yang tinggi, tetapi juga stabilitas performa terhadap variasi input suara dan *noise* lingkungan (Handayani et al., 2025).

Parameter utama yang digunakan sebagai standar global (*gold standard*) dalam mengukur performa sistem pengenalan suara otomatis adalah *Word Error Rate* (WER). Metrik ini merepresentasikan tingkat kesalahan transkripsi dengan menghitung jarak Levenshtein antara kalimat prediksi yang dihasilkan oleh sistem terhadap kalimat referensi yang valid (Adiwidjaja & Ivan Fanany, 2020). Secara matematis, WER dihitung dengan menjumlahkan tiga jenis kesalahan: *Substitution* (kata yang salah dikenali), *Deletion* (kata yang hilang dari prediksi), dan *Insertion*

(kata tambahan yang tidak ada di referensi), kemudian dibagi dengan total jumlah kata dalam referensi (Alfath et al., 2025). Formulasi perhitungan *Word Error Rate* (WER) dapat dituliskan dalam persamaan berikut:

$$WER = \frac{S + D + I}{N} \times 100\%$$

Dimana:

- S = Jumlah kata yang disubstitusi (*Substitution*)
- D = Jumlah kata yang terhapus (*Deletion*)
- I = Jumlah kata yang disisipkan (*Insertion*)
- N = Total jumlah kata dalam referensi (*Number of words*)

Nilai WER yang semakin rendah mengindikasikan kinerja sistem yang semakin baik. Namun, dalam implementasi Bahasa Indonesia, tantangan evaluasi sering kali muncul akibat tingginya nilai WER pada model *pre-trained* yang belum melalui proses *fine-tuning*, sehingga diperlukan metrik pendukung lainnya seperti *Word Accuracy* (WAcc) untuk mempermudah interpretasi tingkat keberhasilan sistem (Anggraini dkk., 2018). Selain akurasi, faktor latensi juga menjadi variabel evaluasi penting, mengingat sistem yang digunakan untuk komunikasi dua arah memerlukan waktu pemrosesan yang cepat (*real-time factor*) agar tidak mengganggu alur percakapan alami (Handayani et al., 2025).

C. Evaluasi Efisiensi Perangkat Keras

Selain ketepatan algoritma, keberhasilan implementasi pada mikrokontroler sangat bergantung pada efisiensi perangkat keras. Parameter kritis yang dievaluasi meliputi:

1. Latensi Inferensi (*Inference Latency*)

Waktu yang dibutuhkan sistem untuk memproses satu data input (gambar/suara) hingga menghasilkan output. Pada alat bantu komunikasi *real-time*, latensi harus seminimal mungkin agar interaksi terasa alami (Zhang et al., 2021).

2. Konsumsi Memori

Penggunaan RAM dan Flash memori saat model dijalankan. Evaluasi ini penting untuk memastikan model *Deep Learning* yang digunakan tidak melebihi kapasitas sumber daya mikrokontroler yang terbatas (Eze & Asogwa, 2021).

2.2.14 Komparasi Teknologi dan Posisi Penelitian

Dalam merancang alat bantu disabilitas yang efektif, pemilihan arsitektur teknologi tidak hanya didasarkan pada kecanggihan fitur semata, melainkan pada keseimbangan antara

kenyamanan pengguna (*usability*) dan performa teknis. Sub-bab ini menguraikan posisi penelitian *Vogest Lens* di tengah lanskap teknologi asistif yang sudah ada, dengan merujuk pada tinjauan literatur terkini.

A. Pergeseran dari *Sensor-Based* menuju *Vision-Based*

Perdebatan mendasar dalam pengembangan sistem penerjemah bahasa isyarat terletak pada pemilihan metode input: apakah menggunakan pendekatan kontak fisik (sarung tangan pintar) atau pendekatan tanpa kontak (kamera).

Berdasarkan tinjauan sistematis selama satu dekade terakhir, penggunaan perangkat berbasis sensor (*wearable sensors*) seperti sarung tangan elektronik memiliki kelemahan fundamental karena instalasi perangkat keras yang rumit di jari-jari tangan seringkali dianggap membebani (*cumbersome*) dan mengurangi kealamian interaksi pengguna (Wadhawan & Kumar, 2021). Berlandaskan pada temuan tersebut, penelitian ini mengadopsi pendekatan berbasis visi (*vision-based*) menggunakan kamera. Metode ini memungkinkan pengguna untuk berkomunikasi secara wajar tanpa harus mengenakan aksesoris yang mengganggu sensitivitas kulit, sehingga kenyamanan penggunaan jangka panjang dapat terjaga.

B. Urgensi Pemrosesan Lokal (*Edge Computing*)

Selain aspek fisik, arsitektur pemrosesan data juga menjadi pembeda utama. Banyak kacamata pintar komersial (seperti *Google Glass*) sejatinya berfungsi sebagai terminal yang mengirimkan data video ke *server (cloud)* untuk diolah. Namun, arsitektur berbasis *cloud* ini dinilai memiliki kelemahan fatal berupa kerentanan terhadap latensi jaringan (*network latency*) yang tinggi serta risiko kebocoran privasi data (Ray, 2022).

Bagi penyandang Tuli, keterlambatan sekian detik dalam menerima notifikasi bahaya bisa berakibat fatal. Oleh karena itu, *Vogest Lens* dirancang dengan paradigma *Edge Computing*. Dengan menanamkan kecerdasan buatan langsung pada mikrokontroler *Teensy 4.1*, seluruh proses deteksi dilakukan secara mandiri di dalam perangkat. Hal ini menjamin respons sistem yang instan (*real-time*) sekaligus melindungi privasi pengguna karena tidak ada rekaman video yang disebarkan ke jaringan internet.

C. Analisis Perbandingan Fitur (*Product Comparison*)

Untuk memvalidasi kebaruan (*novelty*) penelitian ini, dilakukan komparasi fitur antara *Vogest Lens* dengan tiga kategori solusi yang umum tersedia: aplikasi seluler,

sarung tangan sensorik, dan kacamata pintar berbasis *cloud*. Ringkasan perbandingan tersebut disajikan pada Tabel 2.2 berikut:

Tabel 2. 2 Komparasi Fitur *Vogest Lens* dengan Teknologi Serupa

Aspek Komparasi	Aplikasi Ponsel (<i>Live Transcribe</i>)	Sarung Tangan Pintar (<i>SignAloud</i>)	Kacamata Pintar <i>Cloud</i> (<i>Google Glass</i>)	Vogest Lens (Usulan Riset)
Interaksi Fisik	<i>Hand-held</i> (Tangan memegang HP)	<i>Wearable</i> (Sarung tangan kabel)	<i>Head- mounted</i> (Dipakai di kepala)	<i>Head- mounted</i> (<i>Hands-free</i>)
Kenyamanan	Rendah (Membatasi gerak tangan)	Sedang (Intrusif/Panas)	Tinggi (Ergonomis)	Tinggi (Ringan & Alami)
Metode Proses	<i>Cloud Computing</i>	Mikrokontroler Dasar	<i>Cloud Computing</i>	<i>Edge Computing</i> (AI Lokal)
Ketergantungan Internet	Wajib Ada (4G/Wi-Fi)	Tidak Ada	Wajib Ada	Tidak Ada (100% <i>Offline</i>)
Fungsi Keselamatan	Tidak Spesifik	Tidak Spesifik	Umum (Email/Map)	Khusus (Deteksi Suara Bahaya)
Privasi Data	Data dikirim ke Server	Aman	Data dikirim ke <i>Server</i>	Aman (Proses di Alat)
Biaya Implementasi	Rendah (Perlu HP)	Menengah	Sangat Mahal	Terjangkau

BAB III

METODE PENELITIAN

3.1 Objek, Subjek, dan Lokasi Penelitian

Dalam Teknik Industri, tahapan identifikasi objek dan subjek merupakan langkah krusial untuk memastikan bahwa pengembangan produk didasarkan pada prinsip *user-centered design*. Melalui pemetaan yang jelas terhadap karakteristik pengguna dan lingkungan operasionalnya, peneliti dapat menangkap aspirasi nyata atau *Voice of Customer* (VoC) yang kemudian dikonversi menjadi spesifikasi teknis purwarupa. Bagian ini akan menguraikan batasan fungsional perangkat Vogest Lens, keterlibatan mitra institusi sebagai lokus penelitian, serta metodologi yang digunakan dalam menggali kebutuhan spesifik penyandang tunarungu-wicara di lingkungan pesantren.

3.1.1 Objek Penelitian

Objek utama dalam penelitian ini adalah perancangan dan pengembangan purwarupa teknologi asistif berwujud *wearable device* yang diberi nama Vogest Lens (*Voice and Gesture Lens*). Secara konseptual, perangkat ini dirancang sebagai kacamata cerdas (*smart glasses*) multimodal yang mengintegrasikan dua sistem sensorik utama: pengolahan sinyal audio dan deteksi visual. Pemilihan bentuk kacamata dilandasi oleh prinsip ergonomi kognitif untuk meminimalkan beban pengguna (*cognitive load*), di mana informasi dapat diterima tanpa mengharuskan pengguna memutus kontak mata atau menggunakan tangan mereka untuk memegang perangkat (seperti pada penggunaan *smartphone* konvensional).

Sesuai dengan batasan rekayasa pada tahap *Proof of Concept* (PoC), fungsionalitas objek ini difokuskan pada pengenalan sinyal komunikasi kritis yang bersifat darurat (*early warning system*), bukan pada fungsi penerjemahan bahasa isyarat kompleks (seperti BISINDO atau SIBI) secara penuh berwujud kalimat. Objek sistem ini terbagi menjadi dua luaran utama:

A. Sistem Visi (*Visual Inference*)

Memanfaatkan algoritma *Machine Learning* untuk klasifikasi citra (*Image Classification*) yang dilatih guna mengenali gestur tangan statis tertentu yang mengindikasikan urgensi atau komunikasi dasar (seperti isyarat "Tolong" dan "Halo"). Pengujian fungsionalitas objek ini disimulasikan menggunakan perangkat *smartphone* via *Edge Impulse*..

B. Sistem Audio (*Auditory Inference*)

Memanfaatkan arsitektur *Natural Language Processing* (NLP) yang diefisienkan melalui metode *Keyword Spotting* (KWS). Sistem ini dirancang untuk mendeteksi kata kunci bahaya dari lingkungan sekitar (seperti suara teriakan "Awat" atau "Kebakaran") dan memvisualisasikannya dalam bentuk teks peringatan pada layar OLED mikro yang terintegrasi dengan mikrokontroler.

Selain pengembangan sistem kecerdasan buatan, objek penelitian ini juga mencakup luaran berupa perancangan *High-Fidelity 3D CAD* dari bentuk kacamata itu sendiri, yang berfungsi sebagai usulan desain mekanis guna mengakomodasi tata letak perangkat keras secara ergonomis.

3.1.2 *Subjek dan Lokasi Penelitian*

Penelitian ini menetapkan subjek observasi yang merepresentasikan ekosistem interaksi pengguna akhir (*end-user*) secara komprehensif. Subjek tersebut mencakup penyandang disabilitas sensorik (tunarungu dan tunawicara) serta pemangku kepentingan terkait (*stakeholder*), yaitu para pengajar atau pengurus asrama yang berinteraksi dengan mereka setiap hari. Lokasi pengambilan data kebutuhan pengguna (*Voice of Customer*) dan validasi lapangan difokuskan pada dua institusi pendidikan khusus disabilitas di Daerah Istimewa Yogyakarta, yaitu:

- A. Pondok Pesantren Tunarungu Darul Ashom, Sleman, Yogyakarta
- B. Pondok Pesantren Tunarungu-Tuli Jamhariyah, Yogyakarta

Pemilihan lingkungan pesantren (*boarding school*) sebagai lokus penelitian didasarkan pada pertimbangan bahwa pesantren memiliki ekosistem komunal yang dinamis dengan tingkat interaksi sosial yang tinggi. Dalam kesehariannya, para santri tunarungu memiliki rutinitas khusus, seperti menghafal Al-Qur'an menggunakan bahasa isyarat, yang sangat membutuhkan kebebasan gerak tangan (*hands-free*). Di sisi lain, lingkungan asrama yang luas membutuhkan sistem peringatan keselamatan yang proaktif. Keterbatasan santri dalam mendengar alarm atau peringatan verbal dari lingkungan sekitar menjadikan mereka rentan terhadap risiko bahaya, sehingga menjadikan lokasi ini sangat relevan untuk menguji urgensi dari fitur-fitur yang dirancang pada *Vogest Lens*.

3.1.3 *Identifikasi Kebutuhan Pengguna (Voice of Customer)*

Untuk menjamin bahwa spesifikasi teknis purwarupa yang dibangun benar-benar menyelesaikan masalah akar (*root cause*) di lapangan, dilakukan proses penggalian *Voice of Customer* (VoC). Proses identifikasi ini melibatkan total 5 (lima) orang responden

kualitatif yang terdiri dari perwakilan Santri Tunarungu dan Pengurus/Ustadz dari kedua mitra pesantren. Penentuan jumlah responden dalam penelitian ini ditetapkan sebanyak 5 orang didasarkan pada efektivitas penggalian informasi, di mana menurut penelitian sebelumnya, penggunaan 5 responden dianggap sudah mencukupi karena penambahan jumlah responden di atas angka tersebut cenderung memberikan informasi yang serupa atau mengalami kejenuhan data. Hal ini sejalan dengan teori Nielsen yang menyatakan bahwa informasi baru yang didapat akan semakin sedikit setelah responden kelima, sehingga penggunaan 5 responden merupakan langkah optimasi waktu tanpa mengurangi kualitas observasi (Felix Nur Noftianto & Andhyka Kusuma, 2021).

Dalam proses eksekusi pengumpulan data lapangan, peneliti menghadapi tantangan operasional berupa hambatan komunikasi semantik. Keterbatasan peneliti dalam memahami struktur bahasa isyarat yang dinamis membuat komunikasi langsung secara mendalam (*in-depth interview*) dengan santri berisiko menghasilkan misinterpretasi kebutuhan. Sebagai langkah mitigasi metodologis yang terukur, penelitian ini menerapkan pendekatan Wawancara Berpendampingan (*Proxy-Assisted Interview*).

Melalui metode ini, pengurus pesantren dilibatkan tidak hanya sebagai penerjemah, tetapi juga sebagai informan ahli (*expert respondent*). Pengurus memiliki pemahaman komprehensif mengenai rekam jejak psikologis, kesulitan belajar, hingga insiden-insiden komunikasi yang sering dialami santri. Hasil wawancara dan observasi ini kemudian diekstraksi ke dalam format *User Stories* (cerita pengguna) sebuah teknik dalam kerangka kerja *Scrum* untuk menentukan prioritas fitur produk. Contoh kebutuhan kritis yang berhasil diidentifikasi adalah keharusan alat untuk bisa menangkap suara bahaya secara *real-time* dari arah belakang pengguna, yang kemudian menjadi justifikasi utama diimplementasikannya fitur *Keyword Spotting* (KWS) darurat. Seluruh hasil temuan lapangan ini telah divalidasi dan disahkan melalui lembar persetujuan yang ditandatangani oleh perwakilan kepengurusan di kedua institusi tersebut.

3.1.4 *Profil Mitra Penelitian*

Keterlibatan dua mitra pesantren dalam penelitian ini memiliki fungsi silang yang saling memvalidasi (*cross-validation*). Pondok Pesantren Darul Ashom, dengan fokusnya pada program tahfidz Al-Qur'an khusus tunarungu, berperan sebagai sumber utama dalam mengidentifikasi kebutuhan alat yang berfokus pada aktivitas non-distraktif dan *hands-free*. Sementara itu, Pondok Pesantren Jamhariyah berperan sebagai lokus perluasan sampel untuk memastikan bahwa keluhan dan hambatan komunikasi yang diidentifikasi

tidak hanya bersifat lokal di satu tempat, melainkan merupakan representasi masalah komunal penyandang tunarungu secara umum. Berikut merupakan rincian profil dari kedua institusi yang bertindak sebagai mitra sekaligus validator lapangan dalam penelitian ini:

Tabel 3. 1 Profil Mitra Penelitian

Keterangan	Mitra 1	Mitra 2
Nama Mitra	: Pesantren Tunarungu Pondok (<i>Boarding School</i>) - Darul Ashom	Pesantren Tunarungu-Tuli Jamhariyah
Alamat Mitra	: Ponpes Tunarungu Darul Ashom, Jl. kalijaga Kragilan, RT.006/RW.007, Area Sawah, Sidomoyo, Kec. Godean, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55264	Dusun RT.04/RW.25, Umbulmartani, Kec. Ngemplak, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55584
Nomor Telepon	: 081227277615	082345754067
Bidang Mitra	: Pondok Pesantren	Pondok Pesantren
Tahun Berdiri	: 2019	2019

3.2 Sumber Data

Dalam pelaksanaan penelitian rekayasa sistem terintegrasi ini, ketersediaan dan validitas data memegang peranan vital. Data tidak hanya berfungsi sebagai landasan perumusan spesifikasi mekanis dan manajemen proyek, tetapi juga menjadi bahan baku utama (*raw material*) dalam proses pelatihan model kecerdasan buatan (*Machine Learning*). Untuk memastikan akurasi dan keandalan penelitian, sumber data diklasifikasikan menjadi dua kategori utama, yakni Data Primer yang diekstraksi langsung dari lapangan dan pengujian alat, serta Data Sekunder yang bersumber dari literatur teknis dan standar industri.

3.2.1 Data Primer

Data primer merupakan data empiris yang diperoleh peneliti secara langsung dari sumber aslinya melalui metode observasi, pengukuran fungsional, dan interaksi dengan subjek

penelitian di Pondok Pesantren Darul Ashom dan Pondok Pesantren Jamhariyah. Pengumpulan data primer dalam penelitian ini terbagi ke dalam empat kelompok utama:

1. Data Kebutuhan Pengguna (*User Stories Data*)

Data kualitatif ini diperoleh melalui proses wawancara berpendampingan (*proxy-assisted interview*) dan diskusi partisipatif yang melibatkan santri tunarungu serta tenaga pengajar/pengurus dari kedua mitra pesantren. Fokus penggalian data diarahkan untuk menangkap aspirasi murni atau "cerita pengguna" (*User Stories*) terkait kendala spesifik yang dialami dalam interaksi harian. Informasi yang digali secara mendalam meliputi identifikasi kondisi di mana miskomunikasi visual sering terjadi, serta pemetaan jenis-jenis suara peringatan bahaya di lingkungan asrama yang kritis namun sering terlewatkan oleh santri (seperti suara teriakan atau alarm). Data kualitatif ini tidak dikuantifikasi menjadi angka statistik, melainkan diekstraksi dan dipetakan menjadi daftar prioritas kebutuhan fungsional di dalam *Product Backlog* pada kerangka kerja *Scrum*.

2. Dataset Citra Gestur Tangan (*Ego-centric View*)

Untuk melatih model *Computer Vision* agar mampu menerjemahkan bahasa isyarat dasar, peneliti menyusun *dataset* citra secara mandiri. Pengambilan data citra dibatasi pada gestur statis yang merepresentasikan komunikasi dasar dan kedaruratan, yaitu isyarat "Halo" dan "Tolong". Untuk memastikan model memiliki tingkat generalisasi yang tinggi saat disimulasikan melalui perangkat *smartphone*, pengambilan sampel citra dilakukan dengan berbagai skenario variasi, meliputi:

- Variasi Sudut Pandang

Citra direkam dari sudut pandang orang pertama (*ego-centric view*) guna merepresentasikan estimasi tangkapan lensa jika kamera diletakkan pada kacamata.

- Variasi Pencahayaan

Pengambilan sampel dilakukan pada kondisi intensitas cahaya tinggi (luar ruangan/siang hari) dan cahaya redup (dalam ruangan) untuk menguji sensitivitas ekstraksi fitur pada kondisi lingkungan yang dinamis.

- Variasi Latar Belakang

Gestur direkam dengan latar belakang polos (statis) maupun latar belakang kompleks (suasana riil lingkungan pesantren) agar algoritma *Machine Learning* mampu memisahkan (*segmentasi*) objek tangan utama dari *noise* latar belakang.

3. Dataset Audio Kedaruratan (*Keyword Spotting Dataset*)

Data audio dikumpulkan secara spesifik untuk melatih model *Natural Language Processing* (NLP) agar mikrokontroler mampu mendeteksi peringatan bahaya secara mandiri (*offline*). Sampel suara direkam menggunakan mikrofon digital pada frekuensi *sampling* standar (16 kHz) yang sesuai dengan kapabilitas pemrosesan audio pada *Edge AI*. Kategori suara yang direkam meliputi:

- Kata Kunci Bahaya (*Positive Class*)

Rekaman teriakan kata "Awat" dan "Kebakaran" yang diucapkan dengan berbagai tingkat intonasi, jarak, dan variasi gender (suara laki-laki dan perempuan) untuk memperluas jangkauan pita frekuensi (*frequency band*) yang dikenali sistem.

- Suara Latar (*Negative Class / Background Noise*)

Rekaman suasana hening dan kebisingan aktivitas harian normal di pesantren (suara langkah kaki, angin, ketukan meja). Kehadiran data *noise* ini sangat krusial sebagai pembanding agar alat memiliki ketahanan (*robustness*) dan tidak memicu notifikasi palsu (*False Positive*) pada kondisi aman.

4. Data Evaluasi Kinerja Sistem (*System Performance Data*)

Merupakan data kuantitatif yang diukur secara empiris saat pengujian purwarupa (*Proof of Concept*) dan simulasi berlangsung. Data ini meliputi metrik matriks kebingungan (*Confusion Matrix*) seperti tingkat akurasi dan presisi deteksi per kelas, kecepatan respons sistem (*inference latency*) dari saat suara ditangkap hingga teks muncul di layar OLED, serta observasi kestabilan perangkat saat dijalankan secara terus-menerus.

3.2.2 Data Sekunder

Data sekunder adalah data pendukung yang bersifat referensial dan teoretis. Data ini tidak diambil secara langsung dari subjek penelitian, melainkan diperoleh melalui studi dokumentasi, literatur ilmiah, lembar spesifikasi teknis (*datasheet*), dan standar rekayasa. Data sekunder berfungsi sebagai acuan validasi kelayakan dalam fase *Plan* (Perencanaan) dan *Do* (Pelaksanaan) pada siklus PDCA. Data sekunder yang digunakan meliputi:

1. Data Dimensi dan Antropometri Mekanis

Untuk memastikan usulan rancangan 3D CAD *Vogest Lens* memiliki bentuk yang fungsional dan terkalibrasi dengan realitas fisik perangkat elektronik, penelitian ini menggunakan parameter data sekunder yang mencakup:

- Dimensi Kacamata Keselamatan (*Safety Glasses Baseline*)

Data antropometri ukuran standar dari kacamata keselamatan industri tipe *all-size* digunakan sebagai acuan dasar (*baseline*) geometri sasis (*frame*). Parameter yang diekstrak meliputi rentang lebar total bingkai (140-150 mm) dan panjang gagang (*temple length*), untuk memastikan desain akhir tetap mematuhi prinsip ergonomi fisik wajah manusia.

- Dimensi Mekanis Komponen Fisik (*Mechanical Datasheets*)

Ekstraksi ukuran dimensi (Panjang x Lebar x Tinggi) dari komponen utama untuk merancang toleransi ruang *casing* 3D. Ini mencakup dimensi *mainboard* Teensy 4.1 untuk volume gagang kacamata, dimensi layar OLED 0.96 inci untuk perancangan dudukan (*holder*) prisma optik, dan dimensi modul mikrofon INMP441 untuk letak lubang akustik.

2. Spesifikasi Teknis Elektronik (*Electronic Datasheets*)

Data parameter kelistrikan dan komputasi yang ditarik dari dokumen resmi pabrikan (OEM). Data ini sangat krusial untuk analisis beban kerja mikrokontroler dan perhitungan anggaran daya (*power budgeting*), meliputi:

- Mikrokontroler Teensy 4.1

Data mengenai kapasitas SRAM (1024 KB), kecepatan *clock* maksimum (600 MHz), tegangan operasi (3.3V), dan arsitektur pemetaan *General Purpose Input/Output* (GPIO).

- Sensor Kamera OV7670

Data mengenai resolusi optik, antarmuka komunikasi (DVP), dan format data keluaran RGB565/YUV.

- Mikrofon INMP441

Data sensitivitas audio (dB) dan protokol komunikasi I2S.

- Manajemen Daya

Data karakteristik pengisian dan pengosongan (*discharge rate*) baterai Li-Po serta efisiensi modul *step-up/step-down* regulator.

3. Literatur Algoritma dan Kerangka Kerja Sistem

Data teoritis yang diambil dari jurnal ilmiah bereputasi dan buku teks sebagai landasan pengembangan perangkat lunak, meliputi:

- Arsitektur MobileNetV2

Detail matematis mengenai struktur *layer Neural Network*, ekstraksi fitur audio *Mel-frequency cepstral coefficients* (MFCC), dan efisiensi *depthwise separable convolution* yang menjadi dasar dari penerapan *Transfer Learning*.

- Dokumentasi *Edge Impulse* dan *Scrum*
Standar operasional prosedur ekstraksi data digital pada *Edge Impulse Studio*, serta panduan literatur manajemen proyek adaptif (Siklus *Sprint* dan PDCA) untuk menjamin tahapan penelitian berjalan sesuai kaidah rekayasa sistem yang terstandarisasi.

3.3 Instrumen Penelitian

Instrumen penelitian merupakan alat bantu (perkakas) operasional yang digunakan oleh peneliti untuk mengeksekusi tahapan pengumpulan data, perancangan sistem, pengembangan logika komputasi, hingga validasi purwarupa. Pemilihan dan spesifikasi instrumen dirancang secara sistematis agar selaras dengan tahapan metodologi Manajemen Proyek (*Scrum*) dan pengendalian kualitas (PDCA), sekaligus mematuhi batasan *Proof of Concept* (PoC) yang telah ditetapkan. Instrumen penelitian ini diklasifikasikan ke dalam tiga kategori utama: instrumen perangkat keras (*hardware*), perangkat lunak (*software*), dan instrumen dokumentasi non-teknis.

3.3.1 Instrumen Perangkat Keras (*Hardware*)

Instrumen ini mencakup seluruh unit fisik yang difungsikan untuk membangun sistem komputasi logika (*embedded system*) dan memfasilitasi pengujian inferensi model kecerdasan buatan.

1. Perangkat Komputasi & *Embedded System*

- *Workstation* (Laptop)
Laptop berspesifikasi prosesor minimal Intel Core i5/AMD Ryzen 5, RAM 16GB, GPU NVIDIA 1650 Ti. Laptop ini berfungsi sebagai *platform* utama untuk pelatihan model *Machine Learning*, desain CAD, dan pemrograman.
- Mikrokontroler *Teensy 4.1*
Papan sirkuit mikrokontroler (*development board*) dengan arsitektur ARM Cortex-M7 yang berfungsi sebagai unit pemroses pusat (*brainware*) untuk menjalankan inferensi model *Keyword Spotting* (KWS) secara lokal (*on-device/Edge AI*).
- Perangkat Akuisisi Visual dan Simulasi (*Smartphone*)

Perangkat ponsel pintar (Android/iOS) yang memiliki peran ganda dalam penelitian ini. Pada tahap awal, kamera resolusi tinggi pada *smartphone* difungsikan sebagai instrumen penangkap data (*data acquisition*) citra gestur tangan untuk pembentukan *dataset* visual. Pada tahap akhir, perangkat ini difungsikan sebagai platform simulasi *deployment* untuk menguji keandalan model klasifikasi citra secara terpisah dari *board* mikrokontroler utama.

- Mikrofon I2S (INMP441)

Mikrofon digital *omnidirectional* berbasis protokol komunikasi I2S yang berfungsi menangkap spektrum frekuensi suara (perintah kedaruratan dan *noise* latar) dari lingkungan sekitar secara presisi.

- Instrumen Output Logika (Layar OLED SSD1306)

Modul layar matriks mikro berukuran 0.96 inci yang digunakan sebagai antarmuka visual terbatas pada tahap perakitan PoC untuk memvalidasi keluaran teks peringatan.

- Digital *Multimeter*

Alat ukur elektronik yang digunakan pada tahap *Check* (pengujian) untuk memvalidasi tegangan baterai, memastikan tidak ada hubungan pendek (*short circuit*) pada rangkaian, dan mengukur konsumsi arus listrik sistem.

2. Instrumen Perangkat Lunak (*Software*)

Instrumen perangkat lunak mencakup platform komputasi dan aplikasi yang digunakan untuk mendesain wujud logis maupun fisik dari sistem yang dikembangkan.

- *Edge Impulse Studio*

Platform pengembangan *Machine Learning* terintegrasi untuk perangkat tepi (*TinyML*). Berfungsi mengeksekusi *Digital Signal Processing* (DSP), ekstraksi fitur (*Mel-frequency cepstral coefficients* untuk audio), melatih arsitektur jaringan saraf tiruan (*MobileNet V2*), serta mengeksport model menjadi pustaka *C++* (*library*).

- Perangkat Lunak CAD (*SolidWorks* 2025)

Perangkat lunak *Computer-Aided Design* yang diutilisasi untuk merealisasikan usulan perancangan rekayasa fisik. Instrumen ini digunakan untuk membangun pemodelan 3D sasisacamata *Vogest Lens* secara parametrik guna mengevaluasi

dimensi dan simulasi tata letak (*fitting/collision detection*) antar komponen elektronik.

- *Arduino IDE*

Lingkungan pengembangan terpadu (*Integrated Development Environment*) untuk memfasilitasi penulisan bahasa pemrograman (C/C++), kompilasi kode logika (*logic compilation*), dan injeksi program operasional ke dalam memori mikrokontroler Teensy 4.1.

- Platform Manajemen *Agile* (Trello)

Aplikasi manajemen proyek berbasis komputasi awan yang difungsikan sebagai representasi *Scrum Board* visual. Digunakan untuk melacak pergerakan siklus hidup pengembangan sistem (dari status *To-Do*, *In-Progress*, hingga *Done*) berdasarkan *User Stories* yang ditetapkan.

3.3.2 Instrumen Non-Teknis (Dokumentasi dan Validasi)

Instrumen non-teknis merupakan *tools* manajerial dan konseptual yang digunakan untuk menstandarisasi proses interaksi dengan subjek penelitian serta menjamin ketertelusuran evaluasi metrik (*traceability*).

1. Pedoman Wawancara Berpendampingan (*Proxy-Assisted Interview Guide*)

Instrumen berupa daftar pertanyaan semi-terstruktur yang ditujukan kepada informan ahli (pengurus pesantren) guna memandu peneliti dalam memetakan *User Stories* yang spesifik terhadap kendala harian santri, tanpa membatasi eksplorasi jawaban.

2. Lembar *Product Backlog*

Dokumen kontrol dinamis (artefak *Scrum*) yang berisi pemeringkatan prioritas fitur operasional produk. Instrumen ini bertindak sebagai basis justifikasi tunggal (*single source of truth*) dalam menetapkan skenario pengujian setiap iterasi *Sprint*.

3. *Log Sheet* Pengujian Fungsional PoC

Lembar dokumentasi eksperimental yang digunakan pada tahap *Do* dan *Check* untuk mencatat parameter metrik kuantitatif secara sistematis, mencakup: tabulasi perhitungan *Confusion Matrix* (akurasi deteksi KWS dan citra), persentase *False Positive*, serta observasi waktu latensi inferensi (*inference response time*).

4. Lembar Validasi Fungsional (*Functional Acceptance Criteria*):

Dokumen kriteria penerimaan fungsional yang dirancang khusus untuk mengevaluasi skala *Proof of Concept*. Instrumen ini memuat parameter logika sistem minimum yang harus dicapai oleh perangkat (misalnya: PoC mampu menampilkan

teks peringatan sekian detik setelah input suara diberikan) untuk dinyatakan valid secara teknis.

3.4 Metode Pengumpulan Data

Proses pengumpulan data dalam penelitian ini dirancang secara sistematis untuk memastikan bahwa *dataset* yang dihasilkan memiliki tingkat validitas dan reliabilitas yang tinggi. Data ini difungsikan untuk dua tujuan utama: perumusan kebutuhan pengguna (sebagai dasar perancangan sistem) dan penyediaan data latih (*training data*) untuk model kecerdasan buatan (*Machine Learning*). Metode pengumpulan data dibagi berdasarkan klasifikasi tahapan penelitian, yang mencakup elisitasi kebutuhan, akuisisi data digital, hingga pengujian kinerja fungsional.

3.4.1 *Elicitasi Kebutuhan Pengguna*

Tahap ini merupakan langkah fundamental dalam perspektif *user-centered design* untuk menggali spesifikasi fungsional alat secara langsung dari lingkungan calon pengguna. Berbeda dengan metode penyebaran kuesioner tertutup yang membatasi spektrum jawaban responden pada skala angka, penelitian ini menggunakan pendekatan kualitatif partisipatif melalui metode Wawancara Berpendampingan (*Proxy-Assisted In-Depth Interview*).

Peneliti melakukan sesi tanya jawab semi-terstruktur dengan total 5 (lima) narasumber yang berasal dari dua institusi mitra, yaitu Pondok Pesantren Darul Ashom dan Pondok Pesantren Jamhariyah. Narasumber terdiri dari perwakilan santri tunarungu (sebagai *end-user*) dan pengurus/ustadz (sebagai validator ahli dan penerjemah isyarat). Fokus pertanyaan diarahkan untuk mengidentifikasi "titik kritis" (*pain points*) dalam aktivitas harian, khususnya kendala komunikasi visual saat kegiatan komunal (seperti menghafal Al-Qur'an) dan identifikasi potensi bahaya audio yang sering tidak disadari oleh santri di lingkungan asrama. Hasil wawancara kualitatif ini tidak dikonversi menjadi statistik analitik, melainkan ditranskripsikan dan dipetakan menjadi kartu-kartu *User Stories* dengan format baku *Agile* ("*As a user, I want..., So that...*"), yang kemudian diintegrasikan ke dalam dokumen *Product Backlog* sebagai acuan prioritas pengembangan fitur.

3.4.2 *Akuisisi Dataset Kata Kunci Visual*

Mengingat belum tersedianya *dataset* publik yang representatif dan sesuai dengan lingkungan operasional pondok pesantren, peneliti melakukan akuisisi data citra gestur secara mandiri. Proses pengambilan data secara eksklusif menggunakan resolusi kamera

optik pada perangkat *smartphone* yang dihubungkan langsung secara nirkabel/kabel ke platform *Edge Impulse* pada *workstation* (laptop). Pemilihan *smartphone* sebagai instrumen akuisisi didasarkan pada strategi untuk menyamakan karakteristik tangkapan sensor (seperti saturasi warna, *noise*, dan distorsi lensa) dengan lingkungan perangkat saat simulasi *deployment* dilakukan pada tahap pengujian.

Teknik pengambilan gambar menerapkan pendekatan sudut pandang orang pertama (*First-Person View/Ego-centric*), di mana kamera *smartphone* diletakkan sejajar dengan pandangan mata peneliti untuk menyimulasikan perkiraan posisi lensa kacamata. *Dataset* citra dikumpulkan untuk tiga kelas klasifikasi sebagai pembuktian konsep (*Proof of Concept*), yaitu kelas [Tolong], kelas [Halo], dan kelas [Background/Noise]. Untuk memastikan model memiliki kemampuan generalisasi fitur yang optimal, pengambilan citra direkam dengan variasi pencahayaan (intensitas cahaya alami siang hari dan cahaya redup dalam ruangan) serta variasi kompleksitas latar belakang. Setiap kelas merekam antara 500 hingga 1.000 sampel citra (*frame*) dengan resolusi yang dikompresi (misalnya 96x96 piksel RGB) guna memenuhi standar efisiensi komputasi pada arsitektur *MobileNet V2*.

3.4.3 Akuisisi Dataset Audio Kedaruratan

Pengumpulan data sinyal suara dilakukan secara khusus untuk melatih model *Natural Language Processing* melalui metode *Keyword Spotting* (KWS) agar sistem mampu mendeteksi perintah bahaya tanpa bergantung pada koneksi internet (*offline inference*). Perekaman data audio divalidasi menggunakan mikrofon INMP441 dengan *sampling rate* 16 kHz, sebuah parameter frekuensi yang disesuaikan dengan standar batas maksimal pemrosesan sinyal digital (*Digital Signal Processing*) pada lingkungan *TinyML* di Teensy 4.1.

Dataset audio ini didistribusikan ke dalam tiga kelas logis, yaitu kelas positif [Awas] dan [Kebakaran], serta kelas negatif [Noise]. Khusus untuk kelas [Noise], perekaman dilakukan di berbagai titik lokasi strategis pesantren untuk menangkap spektrum suara latar yang autentik (seperti derau kipas angin, langkah kaki, dan gaung percakapan di aula). Pengumpulan *dataset* latar ini sangat krusial untuk melatih ketahanan (*robustness*) model agar sistem mikrokontroler tidak memicu peringatan teks palsu (*False Positive*) saat beroperasi di lingkungan yang bising. Data kelas positif direkam dengan memanipulasi parameter jarak sumber suara (dari jarak dekat hingga radius 1-2 meter)

serta variasi intonasi pengucapan (vokal biasa dan teriakan mendesak) dari beberapa partisipan uji.

3.4.4 *Pengumpulan Data Kinerja Fungsional Sistem*

Tahap akhir pengumpulan data eksperimental dilakukan pada fase *Check* dalam siklus PDCA (merepresentasikan siklus *Sprint* akhir). Sesuai dengan batasan rekayasa sistem yang memisahkan arsitektur komputasi, data kinerja teknis dikumpulkan secara terpisah melalui dua jalur observasi laboratorium:

A. Jalur Evaluasi PoC Perangkat Keras (*Hardware*)

Pengumpulan metrik fungsional langsung dari rangkaian mikrokontroler Teensy 4.1 dan OLED. Data yang dicatat meliputi kecepatan waktu respons (*inference latency*) dari saat audio ditangkap mikrofon hingga teks muncul, dan profil stabilitas tegangan/konsumsi arus listrik.

B. Jalur Evaluasi Simulasi *Smartphone*

Pengumpulan data inferensi model visual saat dieksekusi secara *real-time* di antarmuka perangkat bergerak (*mobile device*). Data yang diukur mencakup parameter tabulasi *Confusion Matrix*, yaitu akurasi deteksi (*Accuracy*) dan kepresisian pembacaan gestur.

Hasil observasi fungsional dari kedua jalur yang terpisah ini kemudian diakumulasikan sebagai landasan analisis kelayakan teknis, untuk menggantikan evaluasi usability yang mensyaratkan adanya wujud fisik produk utuh.

3.4.5 *Pengumpulan Data Dimensi Fisik Komponen*

Dalam memfasilitasi luaran desain tata letak purwarupa *Vogest Lens*, penelitian ini tidak berfokus pada pengumpulan data antropometri populasi makro secara statistik, melainkan mengadopsi pendekatan Perancangan Berbasis Komponen (*Component-Based Design*). Oleh karena itu, data fisik sekunder difokuskan pada pengumpulan dimensi geometris presisi dari komponen elektronik utama. Data ini diekstraksi melalui metode hibrida, yakni pengukuran langsung menggunakan jangka sorong (*vernier caliper*) dan validasi silang terhadap lembar spesifikasi teknis pabrikan (*datasheet*). Parameter dimensi spesifik yang dikumpulkan meliputi:

A. Dimensi *Mainboard* (Teensy 4.1)

Ukuran panjang, lebar, dan ketebalan papan mikrokontroler untuk merumuskan volume ruang *casing* internal pada gagang kacamata (*temple*).

B. Dimensi Modul Kamera & Sensor Audio

Kalkulasi ukuran penampang fisik modul kamera OV7670 dan mikrofon INMP441 untuk merancang toleransi lubang bukaan (*aperture* / lubang akustik) pada bingkai depan. (Catatan: OV7670 tetap dimasukkan di sini murni sebagai referensi desain CAD kacamata masa depan).

C. Dimensi Layar OLED

Pengukuran area aktif dan area bingkai fisik panel OLED 0.96 inci untuk merumuskan dudukan mekanis (*holder/bezel*) agar informasi di layar dapat diproyeksikan dengan sudut pandang perifer yang optimal.

D. Data Referensi Dasar Kacamata

Sebagai *baseline* geometri rangka luar, peneliti menggunakan dimensi standar industri dari kacamata keselamatan (*safety glasses - all size*), yang parameter panjang dan lebarnya kemudian direkayasa ulang (*retrofitting*) di *software* SolidWorks agar selaras dengan kebutuhan penyimpanan komponen di atas.

3.5 Metode Pengolahan Data

Metode pengolahan data dalam penelitian ini dirancang secara terstruktur untuk menjembatani kesenjangan antara kebutuhan kualitatif pengguna dengan implementasi teknis sistem komputasi cerdas. Mengingat penelitian ini mengadopsi prinsip adaptif dari manajemen proyek *Agile Development*, proses pengolahan data tidak bersifat linier di akhir penelitian, melainkan dilakukan secara berulang. Seluruh tahapan pengolahan data dipetakan secara presisi ke dalam siklus *Plan-Do-Check-Act* (PDCA) untuk menjamin ketertelusuran (*traceability*) pengembangan purwarupa berskala *Proof of Concept* (PoC).

3.5.1 Pengolahan Data Kebutuhan Pengguna (*Fase Plan - Analisis Backlog*)

Data primer kualitatif yang diperoleh dari hasil wawancara berpendampingan dengan santri dan pengurus pesantren diolah menggunakan teknik analisis konten (*content analysis*). Tahapan ini bertujuan mengubah abstraksi narasi masalah komunikasi menjadi parameter spesifikasi teknis yang terukur.

1. Ekstraksi dan Transformasi *User Stories*

Data verbal diekstraksi ke dalam sintaksis standar *User Story* pada kerangka kerja *Scrum*: "*Sebagai [Peran], saya ingin [Fitur], agar [Tujuan]*." Setiap cerita pengguna divalidasi dengan Kriteria Penerimaan (*Acceptance Criteria*). Contoh: Kebutuhan "mengetahui ada bahaya dari belakang" diterjemahkan menjadi parameter "Sistem

mikrokontroler harus mendeteksi spektrum suara spesifik dan memvisualisasikan peringatan pada OLED".

2. Pembobotan Prioritas (*Backlog Grooming*) dengan MoSCoW

Daftar *User Stories* dalam *Product Backlog* diolah untuk menentukan urutan pengerjaan menggunakan metode prioritas MoSCoW:

- *Must Have*
Fitur krusial yang jika tidak ada, produk dianggap gagal (Contoh: Deteksi Kata Kunci "Awas", "Kebakaran").
- *Should Have*
Fitur penting namun bisa ditunda jika waktu tidak cukup (Contoh: Deteksi Kata "Halo").
- *Could Have*
Fitur tambahan yang meningkatkan kenyamanan (Contoh: Indikator Baterai).
- *Won't Have*
Fitur yang disepakati untuk tidak dikerjakan saat ini (Contoh: Terjemahan Kalimat Lengkap).

3.5.2 Pengolahan Data Pembelajaran Mesin (*Fase Do - Pengembangan Model AI*)

Data citra gestur dan audio kedaruratan diolah menggunakan *Digital Signal Processing* (DSP) melalui *platform Edge Impulse*. Mengingat implementasi akhir dipisahkan menjadi dua arsitektur berbeda (Teensy 4.1 untuk audio dan *smartphone* untuk visual), maka perlakuan pengolahan data dibedakan sebagai berikut:

A. Pra-pemrosesan Data (*Preprocessing*)

Tahap standardisasi data masukan mentah (*raw data*) untuk mempercepat komputasi jaringan saraf tiruan:

1. Normalisasi Citra (Target: Simulasi *Smartphone*)
Input citra diubah ukurannya (*resize*) menjadi resolusi 96x96 piksel untuk mengurangi beban memori. Nilai intensitas piksel (0-255) dinormalisasi menjadi rentang bilangan titik mengambang (0.0 hingga 1.0) untuk mempercepat konvergensi gradien saat pelatihan.
2. Segmentasi Audio (Target: Mikrokontroler Teensy 4.1)
Sinyal gelombang suara mentah dipotong (*windowing*) dengan panjang jendela 1.000 ms (1 detik) untuk memastikan frekuensi kata kunci kedaruratan tertangkap utuh dalam satu *frame* inferensi kontinu.

B. Ekstraksi Fitur (*Feature Extraction*) (Fase Do)

Proses ekstraksi fitur dilakukan secara otomatis menggunakan blok pemrosesan sinyal digital (*DSP Block*) yang tersedia pada platform *Edge Impulse*. Tahap ini bertujuan mengubah data mentah menjadi parameter yang dapat membedakan pola:

- *Audio* (DSP)

Data suara diproses menggunakan blok MFE (*Mel-Filterbank Energy*) yang disediakan oleh *Edge Impulse*. Algoritma ini mengubah sinyal *audio* dari domain waktu (*Time Domain*) menjadi spektrogram dalam domain frekuensi (*Frequency Domain*) yang merepresentasikan persepsi pendengaran manusia. Meskipun proses ini dilakukan oleh perangkat lunak, dasar matematis transformasi frekuensi (f) ke skala Mel (m) yang diterapkan adalah:

$$m = 1125 \ln \left(1 + \frac{f}{700} \right)$$

Parameter DSP pada *Edge Impulse* diatur dengan panjang *frame* 0.02 detik dan frekuensi *sampling* 16kHz untuk menghasilkan spektrogram yang optimal bagi model.

- *Visual*

Fitur visual diekstraksi untuk membedakan tepi, tekstur, dan bentuk tangan dari latar belakang sebelum dimasukkan ke lapisan klasifikasi.

C. Pelatihan Model (*Model Training*) (Fase Do)

Dataset yang telah dikumpulkan sebelumnya dibagi dengan proporsi 80% Data Latih dan 20% Data Uji.

- *Model Visual*

Dilatih menggunakan arsitektur MobileNetV2 yang menerapkan teknik *Transfer Learning*. Bobot awal model diambil dari dataset ImageNet, kemudian disesuaikan (*fine-tuning*) dengan dataset gestur tangan spesifik penelitian ini.

- *Model Audio*

Dilatih menggunakan arsitektur CNN (*Convolutional Neural Network*) 1D yang efektif mengenali pola dalam data deret waktu (suara).

- *Optimasi*

Proses pelatihan dilakukan berulang kali (*epochs*) dengan meminimalkan nilai *Loss Function* (kategorikal *cross-entropy*) hingga akurasi validasi stabil di atas ambang batas yang ditentukan.

D. Pelatihan Model (*Model Training*) (Fase *Do*)

Untuk memungkinkan model berjalan pada mikrokontroler Teensy 4.1 yang memiliki keterbatasan memori (RAM dan *Flash*), dilakukan teknik kuantisasi pasca-pelatihan (*Post-training Quantization*). Parameter bobot model yang semula bertipe *Floating Point 32-bit* (4 byte) dikonversi menjadi *Integer 8-bit* (1 byte). Teknik ini mereduksi ukuran model hingga 4x lipat dan mempercepat waktu inferensi tanpa penurunan akurasi yang signifikan.

3.5.3 Pengolahan Data Perancangan Design (Fase *Do* - *Hardware Integration*)

Perancangan *design* semacam ini tidak dilakukan menggunakan pendekatan ergonomi statistika kompleks, melainkan menggunakan pendekatan Desain Berbasis Komponen (*Component-Based Design*). Pengolahan data pada tahap ini berfokus pada dimensi teknis perangkat keras untuk memastikan seluruh komponen elektronika dapat terintegrasi dengan baik.

A. Pengukuran Dimensi Komponen (*Dimensioning*)

Data ukuran fisik dari mikrokontroler Teensy 4.1, modul kamera, sensor mikrofon, dan baterai diukur secara presisi menggunakan jangka sorong (*caliper*). Data dimensi (Panjang x Lebar x Tinggi) ini menjadi acuan utama dalam menentukan volume ruang minimum pada *casing*.

B. Perancangan *Enclosure* (CAD)

Data dimensi komponen diolah menggunakan perangkat lunak Fusion 360 untuk membuat desain *casing* dengan prinsip *Rapid Prototyping*:

1. Toleransi Cetak

Memberikan celah toleransi (*clearance*) sebesar 0.2mm - 0.5mm pada desain lubang konektor dan kedudukan baut agar komponen mudah dipasang setelah dicetak 3D.

2. Penempatan Sensor

Mengolah posisi peletakan kamera agar sejajar dengan garis pandang lurus (*line of sight*) dan posisi layar OLED agar berada di area pandang perifer yang mudah dibaca tanpa menghalangi pandangan utama.

3.5.4 Analisis Kinerja Sistem (Fase *Check* - *Evaluasi & Validasi*)

Data kuantitatif hasil pengujian purwarupa diolah untuk memvalidasi performa sistem terhadap kriteria yang ditetapkan.

1. Metrik Akurasi Klasifikasi (Fase *Check*)

Kinerja model AI dianalisis menggunakan *Confusion Matrix* untuk menghitung parameter berikut:

- Akurasi (*Accuracy*): Rasio prediksi benar terhadap total data.

$$Accuracy = \frac{TP}{TP + TN + FP + FN}$$

- Presisi (*Precision*): Tingkat ketepatan prediksi positif (meminimalkan alarm palsu)

$$Precision = \frac{TP}{TP + FP}$$

- *Recall (Sensitivity)*: Kemampuan sistem mendeteksi seluruh bahaya yang ada (meminimalkan bahaya yang terlewat).

$$Recall = \frac{TP}{TP + FN}$$

- *F1-Score*: Kemampuan model menemukan seluruh kejadian positif (penting untuk gestur darurat seperti (Tolong).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Keterangan :

- TP = *True Positive*
- TN = *True Negative*
- FP = *False Positive*
- FN = *False Negative*

2. Evaluasi Latensi Sistem (Fase Check)

Data waktu respon diolah dengan menghitung rata-rata selisih waktu antara input sensor (t_{in}) dan output layar (t_{out}) dari n kali percobaan:

$$Average Latency = \frac{1}{n} \sum_{i=1}^n (t_{out,i} - t_{in,i})$$

Nilai ini menentukan apakah alat memenuhi standar responsivitas *real-time* (target < 500 ms).

3.5.5 Analisis Kriteria Penerimaan Logika (Fase Act)

Mengingat produk berbentuk *Proof of Concept* dan tidak diuji coba langsung ke subjek manusia dalam format kacamata utuh, pengolahan data pengujian usability tradisional

ditiadakan. Sebagai gantinya, pengolahan data pada tahap akhir dieksekusi melalui Evaluasi Kriteria Penerimaan Fungsional (*Functional Acceptance Criteria*).

A. Refleksi Pencapaian Teknis

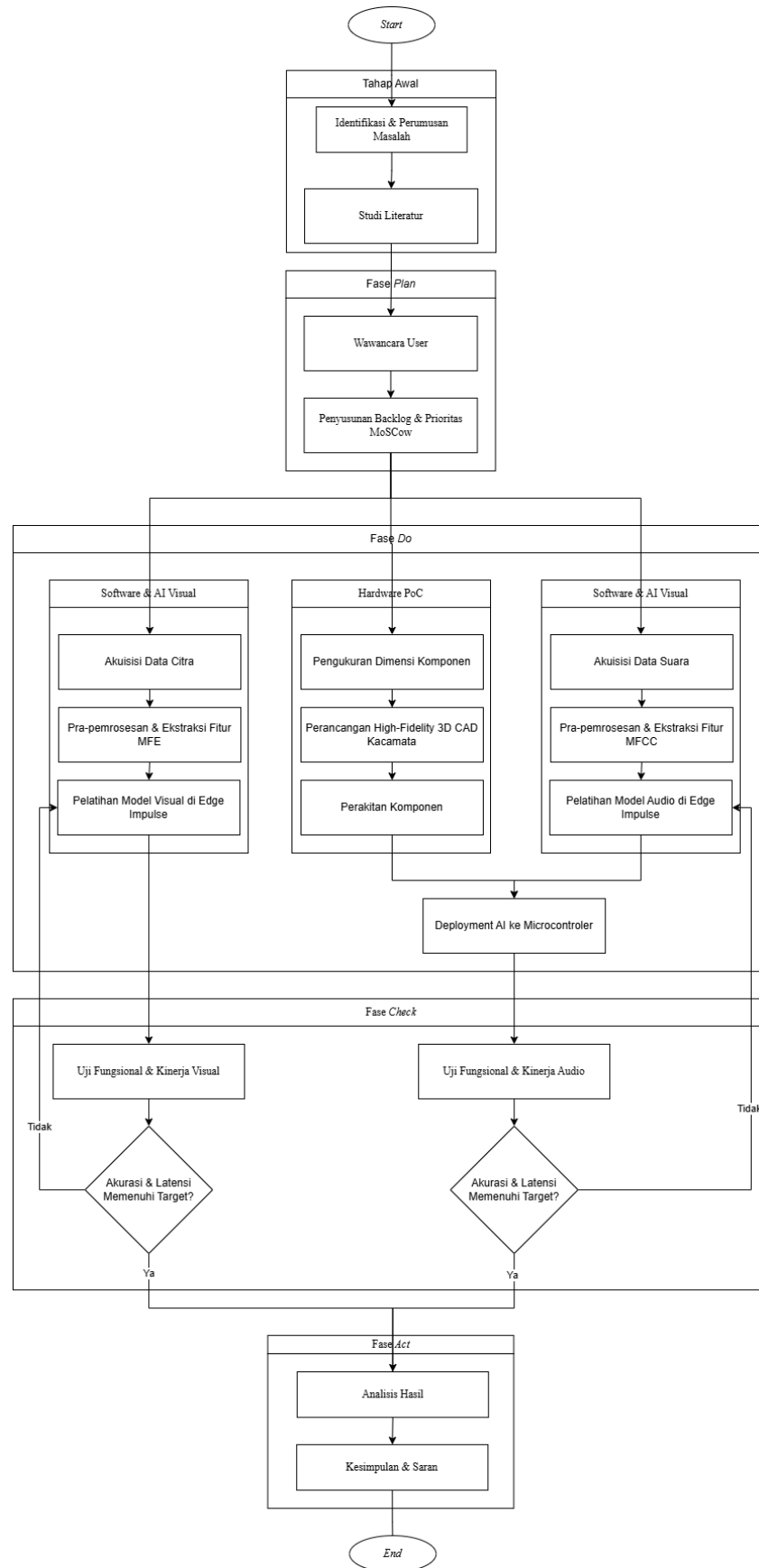
Data performa metrik AI (*Accuracy* > 80%) dan waktu respons latensi disandingkan dengan parameter *Must Have* di awal (*Fase Plan*).

B. Pengambilan Keputusan Siklus PDCA:

Apabila kriteria terpenuhi, siklus rekayasa perangkat dinyatakan tervalidasi secara *Proof of Concept*. Hasil rumusan batasan *hardware* mikrokontroler tunggal ini (*bottleneck evaluation*) kemudian dikompilasi menjadi dokumen usulan modifikasi spesifikasi untuk penelitian masa depan.

3.6 Alur Penelitian

Alur penelitian ini dirancang menggunakan kerangka kerja *Scrum* yang diintegrasikan ke dalam siklus pengendalian kualitas *Plan-Do-Check-Act* (PDCA). Pendekatan adaptif ini dipilih untuk memastikan pengembangan purwarupa *Vogest Lens* (skala *Proof of Concept*) berjalan secara iteratif, terukur, dan mampu memitigasi kendala teknis (seperti *bottleneck* komputasi mikrokontroler) secara proaktif.



Gambar 3. 1 Alur Penelitian

Penjelasan rinci mengenai setiap tahapan dalam alur penelitian di atas adalah sebagai berikut:

A. Tahap Pendahuluan (*Preliminary Stage*)

Tahap ini merupakan fondasi awal penelitian untuk memahami konteks permasalahan dan mendefinisikan ruang lingkup rekayasa:

1. Studi Literatur

Peneliti melakukan kajian pustaka mendalam mengenai teknologi komputasi tepi (*Edge AI / TinyML*), algoritma *Keyword Spotting* (KWS), arsitektur *MobileNet V2* untuk pengenalan citra, serta batasan memori (SRAM) pada mikrokontroler tunggal.

2. Studi Lapangan & Identifikasi Masalah

Observasi dilakukan pada dua mitra studi, yaitu Pondok Pesantren Darul Ashom dan Pondok Pesantren Jamhariyah, untuk memetakan kendala komunikasi santri tunarungu. Fokus identifikasi diarahkan pada respons terhadap peringatan bahaya (audio) dan interaksi isyarat dasar (visual).

B. Fase *PLAN*: Perencanaan Produk (*Planning Phase*)

Pada fase ini, spesifikasi kebutuhan sistem tidak dirumuskan menggunakan metode konvensional yang kaku, melainkan menggunakan pendekatan empiris *Agile*:

1. Wawancara & Penyusunan *User Stories*

Kebutuhan pengguna digali melalui metode Wawancara Berpendampingan (*Proxy-Assisted Interview*) bersama pengurus pesantren. Hasilnya dikonversi menjadi format *User Stories* (contoh: "Sebagai pengguna tunarungu, saya ingin sistem mendeteksi teriakan 'Awat', agar saya dapat melihat peringatan teks di layar kaca mata dan menghindari dari bahaya").

2. Penyusunan *Product Backlog*

Seluruh *User Stories* dikompilasi ke dalam daftar pekerjaan (*Product Backlog*).

3. Penetapan Prioritas (Metode MoSCoW)

Fitur-fitur diklasifikasikan menggunakan matriks MoSCoW (*Must have, Should have, Could have, Won't have*). Metode ini menjadi justifikasi untuk membatasi ruang lingkup (MVP - *Minimum Viable Product*), di mana integrasi komputasi visual dan audio pada satu *hardware* diklasifikasikan sebagai *Won't Have* (ditunda untuk penelitian masa depan) guna menghindari *system crash*.

C. Fase DO: Pelaksanaan Pengembangan (*Execution Phase*)

Tahap inti dari proses rekayasa sistem. Mengacu pada batasan PDCA, fase eksekusi ini dipecah menjadi tiga alur kerja paralel (*parallel workstreams*):

1. Pengembangan PoC *Hardware* Audio (KWS)

- Akuisisi & DSP

Mengumpulan *dataset* suara ("Awat", "Kebakaran", *noise*). Sinyal audio diolah menjadi spektrogram melalui ekstraksi fitur *Mel Frequency Cepstral Coefficients* (MFCC).

- Pelatihan & Deployment

Model *Deep Learning* dilatih di *Edge Impulse*. Setelah konvergen, model dikuantisasi (INT8), dikonversi menjadi pustaka C++, dan diunggah secara fisik ke mikrokontroler Teensy 4.1.

2. Simulasi Model Visual (*Smartphone Stream*)

- Akuisisi & DSP

Pengumpulan *dataset* citra gestur tangan menggunakan resolusi kamera *smartphone*. Data diproses melalui *resizing* (96x96 piksel) dan normalisasi warna RGB

- Pelatihan & Simulasi

Melatih arsitektur *MobileNet V2*. Mengingat beban komputasi visual yang masif, model *tidak* diunggah ke Teensy 4.1, melainkan di-*deploy* ke aplikasi *smartphone* sebagai proksi simulasi pengenalan bahasa isyarat.

3. Perancangan Mekanis (*3D CAD Stream*)

Merancang usulan *casing* kacamata menggunakan perangkat lunak *SolidWorks/Fusion360* berdasarkan prinsip *Component-Based Design*. Desain difokuskan pada toleransi ruang (*clearance*) agar seluruh wujud fisik komponen (Teensy, layar OLED, baterai, dan modul kamera OV7670) di masa depan dapat tertanam secara ergonomis tanpa harus diwujudkan dalam cetakan fisik pada fase ini.

D. Fase CHECK: Pengujian & Verifikasi (*Verification Phase*)

Tahap ini bertujuan memverifikasi metrik performa masing-masing jalur paralel terhadap standar kriteria penerimaan fungsional:

1. Pengujian Fungsional PoC Logika

Memastikan mikrofon INMP441 mampu menangkap suara dan mikrokontroler berhasil mengeksekusi aktuasi teks pada layar OLED secara independen.

2. Pengujian Kinerja *Machine Learning*

Mengevaluasi tabulasi *Confusion Matrix* (Akurasi, Presisi, *Recall*, *F1-Score*) dari model KWS di Teensy dan model visual di *smartphone*.

3. Pengukuran Latensi *Real-Time*

Mengukur *End-to-End Latency* (waktu respons dari input sensor hingga output aktuator) untuk memastikan sistem peringatan tidak mengalami *delay* yang fatal.

4. Mekanisme Keputusan (*Decision Loop*)

Jika akurasi model berada di bawah ambang batas (misal: $< 80\%$) atau latensi melebihi toleransi waktu riil, proses dikembalikan secara iteratif ke tahap optimasi *dataset* atau *retraining* model (*Kembali ke Fase DO*).

E. Fase *ACT*: Evaluasi Akhir dan Refleksi Rekayasa (*Validation & Retrospective Phase*)

Mengingat produk tidak diuji cobakan secara langsung ke bentuk *wearable* yang utuh, fase evaluasi diakhiri dengan refleksi teknis, bukan melalui kuesioner usability:

1. Analisis Penerimaan Fungsional

Mensintesis seluruh data kuantitatif (latensi dan akurasi) dari pengujian *hardware* dan simulasi *smartphone*. Data ini dianalisis untuk mengonfirmasi kelayakan konseptual (*Proof of Concept*) dari ide *Vogest Lens*.

2. Kesimpulan & Rekomendasi (*Retrospective*)

Mengidentifikasi limitasi sistem (seperti *bottleneck* I/O pada SRAM mikrokontroler saat memproses audio dan visual bersamaan). Hasil analisis hambatan (*bottleneck evaluation*) ini diformalisasikan menjadi rekomendasi spesifikasi arsitektur komputasi multiprosesor untuk tahapan *scale-up* di penelitian selanjutnya.

BAB IV

PEMBANGUNAN SISTEM

4.1 Perencanaan & Analisis Kebutuhan (Fase *Plan*)

Fase *Plan* merupakan langkah fundamental dalam siklus *Plan-Do-Check-Act* (PDCA) yang diintegrasikan dengan kerangka kerja *Scrum*. Tahapan ini difokuskan pada elisitasi kebutuhan (*requirements elicitation*) untuk memastikan bahwa fungsionalitas yang akan dikembangkan pada purwarupa *Vogest Lens* benar-benar berakar pada permasalahan riil di lapangan (*user-centered design*). Proses ini mencakup pengumpulan data kualitatif, pembentukan arketipe pengguna (*User Persona*), hingga penyusunan spesifikasi teknis dalam bentuk *Product Backlog*. Mengingat rekayasa ini dibatasi pada skala *Proof of Concept* (PoC), fase perencanaan juga bertindak sebagai filter analitis untuk menentukan fitur prioritas yang dapat dieksekusi tanpa melampaui batas kapasitas memori (SRAM) mikrokontroler.

4.1.1 Pengumpulan Data Kebutuhan Pengguna

Pengumpulan data primer dilakukan melalui metode kualitatif berupa Wawancara Berpendampingan (*Proxy-Assisted In-Depth Interview*) dan observasi partisipatif. Penelitian ini dilaksanakan dengan melibatkan dua institusi mitra di wilayah Daerah Istimewa Yogyakarta, yaitu Pondok Pesantren Darul Ashom dan Pondok Pesantren Jamhariyah. Pendekatan langsung ke lapangan ini sangat krusial untuk menangkap konteks spasial, sosial, dan ergonomis yang melingkupi keseharian pengguna, yang nantinya akan menentukan spesifikasi perangkat keras dan model kecerdasan buatan.

A. Profil Responden dan Pelaksanaan Wawancara

Penentuan informan dalam penelitian ini tidak dilakukan secara acak, melainkan menggunakan teknik *purposive sampling*. Pemilihan didasarkan pada kriteria inklusi spesifik: (1) berstatus sebagai santri aktif atau pengurus pesantren, (2) memiliki pengalaman interaksi komunal yang tinggi, serta (3) mampu membagikan pengalaman historis terkait kendala keselamatan (akibat tidak mendengar suara peringatan) dan komunikasi visual.

Ukuran sampel (*sample size*) ditetapkan berdasarkan prinsip kejenuhan data (*data saturation*). Mengacu pada metodologi kualitatif standar, kejenuhan tercapai ketika wawancara tidak lagi menghasilkan tema atau informasi permasalahan baru yang substansial. Dalam pelaksanaan penelitian ini, pola kebutuhan yang berulang (seperti

kebutuhan akan deteksi suara "Awat" dan peringatan visual) telah terlihat secara konsisten pada wawancara kelima, sehingga penambahan responden dirasa tidak lagi memberikan signifikansi baru terhadap perancangan sistem inti. Rincian profil kelima responden (baik dari kalangan Teman Tuli maupun validator ahli/Teman Dengar) disajikan pada Tabel 4.1.

Tabel 4. 1 Profil Responden Wawancara

Kode Responden	Peran / Status	Tingkat Gangguan Pendengaran	Karakteristik Aktivitas
R-01	Ustadz	-	Teman Dengar
R-02	Ustadz	-	Teman Dengar
R-03	Santri	Sedang	Teman Tuli
R-04	Santri	Sedang	Teman Tuli
R-05	Santri	Ringan	Teman Tuli

B. Identifikasi Kebutuhan Fungsional Berdasarkan Pedoman Wawancara

Wawancara dijalankan menggunakan pedoman terstruktur untuk memvalidasi empat aspek krusial perancangan rekayasa: keselamatan fisik, mekanisme peringatan, interaksi sosial, dan toleransi ergonomi. Transkrip hasil wawancara dianalisis secara semantik untuk mengekstraksi pernyataan-pernyataan kunci (*key verbatim*) yang merepresentasikan titik masalah (*pain points*) pengguna.

Pada aspek keselamatan fisik, seluruh responden mengonfirmasi adanya tingkat kerentanan spasial yang tinggi. R-03 dan R-04 mengeluhkan pengalaman di mana mereka sering dikejutkan oleh kendaraan atau objek yang jatuh dari arah *blindspot* (luar sudut pandang mata). Ketidakmampuan mendeteksi suara impulsif secara langsung membahayakan keselamatan jiwa. Hal ini memvalidasi urgensi implementasi arsitektur komputasi *Natural Language Processing* (NLP) melalui algoritma *Keyword Spotting* (KWS) pada mikrokontroler mandiri untuk mendeteksi kata kunci kedaruratan seperti "Awat".

Terkait mekanisme peringatan, R-04 menyoroti menyoroti inefisiensi sistem getar (*haptic feedback*) konvensional yang sering memicu kebingungan kognitif di situasi

panik. Oleh karena itu, diusulkan antarmuka visual berupa layar mikro OLED yang menampilkan notifikasi teks peringatan secara instan (*low latency*) pada area pandang periferal.

Pada aspek interaksi sosial, R-04 dan R-05 yang memiliki tingkat sosialisasi tinggi kerap menghadapi tekanan psikologis. Keterbatasan untuk mendengar sapaan dari jarak jauh sering kali memicu stigma negatif dari teman sebaya yang menganggap mereka sombong. Hal ini memunculkan justifikasi untuk mengembangkan model pengenalan citra (*Gesture Recognition*) yang mampu mendeteksi lambaian tangan ("Halo"). Mengingat model visual membutuhkan komputasi masif, pengujian fitur ini disepakati untuk disimulasikan menggunakan arsitektur perangkat *smartphone*.

Terakhir, evaluasi pada aspek ergonomi dan desain fisik menunjukkan bahwa pengguna lebih memprioritaskan fungsi pengawasan keselamatan yang bersifat *hands-free* (tangan bebas beraktivitas) dibandingkan desain yang estetik namun mengharuskan interaksi manual (seperti memegang HP). Oleh karena itu, responden mentoleransi usulan desain fisik berupa kaca mata dengan bingkai yang sedikit tebal, asalkan parameter dimensi internalnya mampu mengakomodasi mikrokontroler Teensy 4.1 dan komponen lainnya secara logis.

Ringkasan dari keluhan dan harapan responden (*Voice of Customer*) di atas kemudian dikelompokkan menjadi atribut kebutuhan sistem yang disajikan pada Tabel 4.2.

Tabel 4. 2 Identifikasi Kebutuhan Pengguna (*User Needs*)

No	Fokus Perancangan	Pernyataan Kunci Responden (<i>Key Verbatim</i>)	Identifikasi Kebutuhan Sistem (<i>System Needs</i>)
1	Keselamatan (<i>Safety Needs</i>)	"Waktu itu ada benda jatuh di dekat saya, saya sama sekali tidak sadar sampai ada teman yang menepuk bahu saya." (R-03) "Sering ada kendaraan lewat di belakang, saya	Sistem harus mengintegrasikan modul audio yang secara otomatis dan terus-menerus mendeteksi suara impulsif atau teriakan peringatan (<i>Keyword Spotting</i>)

No	Fokus Perancangan	Pernyataan Kunci Responden (<i>Key Verbatim</i>)	Identifikasi Kebutuhan Sistem (<i>System Needs</i>)
		baru tahu waktu orangnya menepuk keras, jujur itu bikin kaget." (R-04)	
2	Mekanisme Peringatan	"Kalau ada teriak 'Awat' atau 'Kebakaran', saya bingung kenapa orang lari. Saya ingin alatnya langsung menampilkan tulisan yang jelas." (R-04)	Sistem membutuhkan antarmuka visual (layar OLED) untuk mengklasifikasikan dan menampilkan jenis bahaya secara spesifik (<i>real-time text notification</i>)
3	Interaksi Sosial	"Terkadang sedih kalau dikira sombong karena ada teman yang melambatkan tangan memanggil dari jauh tapi saya cuma diam saja." (R-04 & R-05)	Sistem memerlukan modul kamera penglihatan mesin (<i>machine vision</i>) untuk mendeteksi gestur sapaan dan membantu respons sosial pengguna
4	Ergonomi (<i>Wearability & Ergonomics</i>)	"Kalau pakai HP harus dipegang terus, susah kalau lagi aktivitas lain. Lebih enak kacamata biarpun agak tebal asal tidak gampang jatuh." (R-01)	Perangkat harus berwujud <i>wearable device (hands-free)</i> dengan rancangan sasis/bingkai mekanik yang kokoh untuk melindungi komponen internal

Melalui hasil rangkuman data pada Tabel 4.2, terbukti bahwa pengembangan *Vogest Lens* memiliki pijakan argumen yang empiris. Kebutuhan-kebutuhan dasar ini selanjutnya akan diabstraksikan ke dalam pemodelan profil pengguna (*User Persona*) untuk memudahkan penerjemahan ke bahasa teknis pengembangan.

4.1.2 *Profil Pengguna (User Persona)*

Setelah data kebutuhan pengguna teridentifikasi melalui wawancara mendalam, langkah analitis selanjutnya dalam fase *Plan* adalah memodelkan data tersebut ke dalam representasi target pengguna atau *User Persona*. Dalam kerangka pengembangan produk berbasis *Agile* (Scrum), pembentukan persona berfungsi sebagai instrumen jangkar (*anchor*) untuk memastikan bahwa setiap keputusan teknis baik dalam desain perangkat keras maupun pelatihan model kecerdasan buatan tetap berorientasi pada penyelesaian masalah manusiawi (*human-centric problem solving*).

Berdasarkan analisis silang (*cross-analysis*) terhadap karakteristik lima responden pada Tabel 4.1 dan identifikasi kebutuhan pada Tabel 4.2, peneliti mengklasifikasikan target pengguna ke dalam dua arketipe persona utama. Kedua persona ini mewakili spektrum kebutuhan yang berbeda namun saling melengkapi, yaitu kebutuhan akan perlindungan keselamatan (*safety*) dan kebutuhan akan utilitas operasional (*utility*).

A. Persona Primer: Arketipe "Abdullah" (Fokus Keselamatan dan Penerimaan Sosial)

Persona pertama, dengan nama samaran "Abdullah", ditarik dari gabungan karakteristik responden R-03 dan R-04. Profil ini mewakili kelompok santri tunarungu dengan derajat gangguan pendengaran sedang yang memiliki kerentanan psikologis dan fisik yang tinggi di lingkungan asrama.

1. Karakteristik dan Pola Perilaku

"Abdullah" digambarkan sebagai santri yang memiliki ketergantungan visual sangat dominan. Ketiadaan fungsi pendengaran membuatnya kehilangan kesadaran situasional (*situational awareness*) terhadap objek atau peristiwa yang berada di luar sudut pandangnya (*blindspot*), terutama dari arah belakang.

2. Permasalahan Utama (*Pain Points*)

- Risiko Keselamatan

Abdullah hidup dalam bayang-bayang kecemasan akan bahaya mendadak. Keterlambatannya dalam merespons suara impulsif, sebagaimana dikeluhkan oleh responden R-03 terkait pengalaman saat tidak menyadari adanya benda yang jatuh di dekatnya, menempatkannya pada risiko cedera.

- Hambatan Interaksi

Kegagalannya dalam menyadari sapaan visual dari jarak jauh sering kali memicu kesalahpahaman sosial, di mana ia dianggap kurang responsif oleh lingkungan sekitarnya.

3. Implikasi Teknis pada Perancangan

Untuk mengakomodasi kerentanan ini, sistem harus bertindak sebagai substitusi sensorik pendengaran yang beroperasi secara *real-time*. Spesifikasi mutlak (*Must Have*) untuk arketipe ini adalah integrasi algoritma *Keyword Spotting* (KWS) untuk deteksi kata bahaya, serta kamera penglihatan mesin (*machine vision*) untuk mendeteksi gestur sapaan. Waktu tunda (*latency*) notifikasi pada antarmuka layar OLED harus ditekan seminimal mungkin.

B. Persona Sekunder: Representasi Mobilitas dan Fungsionalitas

Persona kedua direpresentasikan dari gabungan karakteristik responden yang memiliki mobilitas lebih tinggi, merujuk pada profil seperti R-01 dan R-02. Arketipe ini mewakili ustadz yang memiliki tugas operasional harian yang dinamis.

1. Latar Belakang dan Karakteristi

"Fatih" adalah representasi individu proaktif. Rutinitasnya didominasi oleh aktivitas fisik seperti mempersiapkan ruangan belajar santri di mana kedua tangannya sering kali berada dalam status sibuk (*hands-busy*) atau membawa peralatan.

2. Permasalahan Utama (*Pain Points*)

- Inefisiensi Perangkat Konvensional

Penggunaan telepon pintar (*smartphone*) sebagai alat komunikasi dinilai sangat mendisrupsi produktivitas, karena mengharuskannya menghentikan aktivitas fisik hanya untuk membuka kunci layer.

- Kerentanan Perangkat Medis

Mengandalkan Alat Bantu Dengar (ABD) konvensional saat beraktivitas berat sering kali terkendala masalah kenyamanan dan risiko alat terlepas.

3. Implikasi Teknis terhadap Perancangan

Eksistensi persona Fatih memberikan menjustifikasi keharusan penelitian ini untuk merancang purwarupa wujud akhir perangkat (dalam bentuk *High-Fidelity 3D CAD*) berbasis *Component-Based Design*. Alat ini secara teoretis harus bersifat *hands-free*, dengan usulan dimensi yang kokoh (mengakomodasi baterai dan modul) meskipun belum difabrikasi pada tahap PoC ini.

C. Matriks Komparasi Kebutuhan Pengguna

Untuk mengonsolidasikan parameter desain yang diekstraksi dari kedua persona tersebut, disusunlah matriks komparasi. Matriks ini bertindak sebagai fondasi rasional sebelum mentransformasikan keluhan pengguna menjadi daftar pekerjaan teknis (*Product Backlog*).

Tabel 4. 3 Matriks Analisis Kebutuhan Persona

Atribut Analisis	Persona Primer (Abdullah)	Persona Sekunder (Fatih)
Fokus Solusi	Keselamatan Jiwa & Bantuan Inisiasi Sosial	fisiensi Mobilitas & Durabilitas Perangkat
Kondisi Pemicu	Ketidakmampuan mendeteksi suara bahaya & sapaan gestur	Status tangan yang sibuk (<i>hands-busy</i>) saat beraktivitas
Tuntutan <i>Interface</i>	Notifikasi visual (OLED) yang bersifat instan (<i>low latency</i>)	Operasional perangkat yang sepenuhnya <i>hands- free</i>
Prioritas Komputasi	Akurasi pada <i>Keyword Spotting (KWS)</i> dan Deteksi Gestur	Manajemen daya perangkat (<i>battery handling</i>) yang stabil
Toleransi Geometri	engutamakan ketersediaan fitur kewaspadaan di atas estetika	Sangat toleran terhadap bingkai (<i>frame</i>) tebal, asalkan kokoh dan ergonomis

Dua arketipe persona di atas memastikan bahwa arah perancangan *Vogest Lens* dipandu oleh parameter permasalahan operasional yang valid. Konklusi dari analisis matriks ini selanjutnya diturunkan menjadi struktur deskriptif pada penyusunan *User Stories* di sub-bab berikutnya.

4.1.3 Penyusunan *User Stories* & *Product Backlog*

Tahapan krusial selanjutnya adalah mentransformasikan kebutuhan abstrak persona menjadi spesifikasi fungsional yang terukur. Dalam metodologi *Agile*, proses translasi ini dilakukan melalui formulasi *User Stories* yang kemudian didekomposisi menjadi unit

pekerjaan teknis atau *Product Backlog*. Tahapan ini esensial untuk membatasi ruang lingkup pengembangan (*scoping*) agar selaras dengan batasan *Proof of Concept*.

A. Formulasi *User Stories*

User Stories merupakan deskripsi ringkas mengenai fitur sistem yang diekspresikan dari sudut pandang pengguna. Penulisan *User Stories* pada penelitian ini mengacu pada sintaksis standar industri, yaitu: "Sebagai [peran pengguna], saya ingin [tindakan/fitur], agar [tujuan/manfaat]".

Berdasarkan sintesis dari persona primer ("Abdullah" dengan fokus keselamatan) dan persona sekunder ("Fatih" dengan fokus utilitas) pada sub-bab sebelumnya, peneliti merumuskan empat *User Stories* fundamental. Keempat cerita pengguna ini merepresentasikan ruang lingkup (*scope*) utama dari pengembangan purwarupa *Vogest Lens*. Pemetaan *User Stories* beserta kriteria penerimaannya (*Acceptance Criteria*) dijabarkan pada Tabel 4.4.

Tabel 4. 4 Daftar User Stories

Kode Story	User Story (<i>Role - Action - Goal</i>)	Kriteria Keberhasilan (<i>Acceptance Criteria</i>)	Persona Rujukan
US-01	Sebagai santri tunarungu, saya ingin alat ini mampu mendeteksi teriakan peringatan ("Awat"), agar saya dapat menghindari dari potensi bahaya.	Mikrokontroler Teensy 4.1 mampu mengenali ekstraksi kata kunci audio KWS dan mengabaikan frekuensi bising latar belakang (<i>noise</i>).	Primer (Abdullah)
US-02	Sebagai pengguna, saya ingin menerima notifikasi peringatan dalam bentuk teks visual, agar saya dapat langsung memahami jenis ancaman.	Layar antarmuka OLED memvisualisasikan konversi peringatan teks secara mandiri dengan <i>response time</i> (latensi) yang minimum.	Primer (Abdullah)
US-03	Sebagai individu dengan mobilitas tinggi, saya ingin usulan rancangan alat yang dipakai di kepala, agar	Terbentuknya rancangan <i>High-Fidelity 3D CAD</i> yang memvalidasi estimasi tata letak (<i>fitment</i>) komponen secara ergonomis (tanpa	Sekunder (Fatih)

	kedua tangan saya bebas beraktivitas (<i>hands-free</i>).	tahapan fabrikasi fisik cetak 3D).	
US-04	Sebagai santri, saya ingin ada fitur yang menginformasikan jika ada rekan yang menyapa dengan gestur ("Halo"), agar saya dapat merespons.	Model klasifikasi citra lambaian tangan berhasil dilatih dan dieksekusi dengan tingkat akurasi tinggi melalui <i>deployment</i> simulasi di perangkat <i>smartphone</i> .	Primer & Sekunder

B. Penyusunan *Product Backlog* dan Prioritas Fitur

User Stories pada Tabel 4.4 masih bersifat naratif dan belum menyentuh ranah implementasi rekayasa. Oleh karena itu, *User Stories* tersebut harus dikonversi menjadi unit pekerjaan teknis yang disebut *Product Backlog Item* (PBI). Setiap PBI mendefinisikan arsitektur perangkat keras (*hardware*) maupun perangkat lunak (*software*) yang harus dieksekusi oleh pengembang.

Mengingat penelitian ini memiliki batasan ruang lingkup pada skala *Proof of Concept* (PoC) dan mitigasi risiko komputasi (*bottleneck* SRAM mikrokontroler), alokasi pengerjaan fitur dibatasi menggunakan metode prioritisasi MoSCoW (*Must Have, Should Have, Could Have, Won't Have*) untuk menyaring dan mengurutkan *Product Backlog*.

1. *Must Have* (Wajib Ada)

Fitur absolut yang mendefinisikan kelayakan produk (*Minimum Viable Product* / MVP). Tanpa fitur ini, sistem dianggap gagal mengatasi masalah keselamatan pengguna.

2. *Should Have* (Seharusnya Ada)

Usulan desain atau fitur pendukung yang penting untuk perancangan teoretis di masa depan.

3. *Could Have* (Boleh Ada)

Fitur pelengkap atau optimalisasi antarmuka.

4. *Won't Have* (Tidak Ada)

Fungsionalitas kompleks atau integrasi *hardware* fisik yang disepakati untuk tidak direalisasikan (dibatasi) pada penelitian saat ini guna mencegah kegagalan sistem (*system crash*).

Rincian *Product Backlog* beserta spesifikasi teknis dan prioritasnya disajikan pada Tabel 4.5.

Tabel 4. 5 *Product Backlog* dan Spesifikasi Teknis

ID	Item Pekerjaan	Spesifikasi Teknis	Prioritas	Estimasi
Backlog	(<i>Technical Task</i>)	Komponen	(MoSCoW)	Alur
PB-01	Pembuatan Model <i>Keyword Spotting</i>	Pelatihan <i>Machine Learning</i> untuk klasifikasi audio menggunakan <i>platform</i> Edge Impulse. Akuisisi data menggunakan modul mikrofon I2S INMP441	<i>Must Have</i>	Jalur PoC <i>Hardware</i>
PB-02	Integrasi Antarmuka PoC	Pemrograman layar OLED 0.96 inci via protokol I2C dengan Teensy 4.1 untuk menampilkan teks peringatan audio.	<i>Must Have</i>	Jalur PoC <i>Hardware</i>
PB-03	Simulasi Model <i>Gesture Recognition</i>	Pelatihan algoritma klasifikasi citra lambaian tangan ("Halo") dan di-deploy ke platform smartphone sebagai simulasi sistem terpisah.	<i>Must Have</i>	Jalur Simulasi <i>Mobile</i>
PB-04	Perancangan Sasis Mekanis (CAD)	Pemodelan 3D sasis kacamatanya (<i>Component-Based Design</i>) untuk mengakomodasi tata letak komponen (tanpa fabrikasi fisik).	<i>Should Have</i>	Jalur Desain 3D

ID	Item	Pekerjaan	Spesifikasi	Teknis	Prioritas	Estimasi
Backlog	<i>(Technical Task)</i>		Komponen		(MoSCoW)	Alur
PB-05	Integrasi		Menghubungkan secara		<i>Won't Have</i>	Dibatasi
	<i>Hardware</i>		fisik modul kamera			(Mitigasi
	Visual	ke	OV7670 dengan Teensy			<i>Crash</i>)
	Mikrokontroler		4.1 bersamaan dengan pemrosesan audio secara <i>real-time</i> .			
PB-06	Penerjemahan		Implementasi		<i>Won't Have</i>	Dibatasi
	Kalimat Isyarat		komputasi berat untuk			
	Dinamis		menerjemahkan tata			
			bahasa isyarat			
			BISINDO secara			
			berkelanjutan.			

C. Perencanaan Pelaksanaan Eksekusi (*Execution Planning*)

Berlandaskan pada pemetaan prioritas di Tabel 4.5, pelaksanaan pengembangan purwarupa pada fase eksekusi tidak dilakukan secara sekuensial (berurutan), melainkan diparalelkan menjadi tiga jalur pekerjaan independen untuk memitigasi risiko kegagalan integrasi *hardware*:

1. Jalur Eksekusi 1 (Fokus PoC *Audio* & Teks)

Didedikasikan untuk merealisasikan arsitektur *hardware* dasar pada mikrokontroler Teensy 4.1 guna memproses fitur KWS (PB-01) dan luaran teks pada OLED (PB-02). Target luaran adalah rangkaian elektronik fungsional yang mampu merespons bahaya secara waktu nyata.

2. Jalur Eksekusi 2 (Fokus Simulasi Visual)

Didedikasikan untuk melatih dan menguji model kecerdasan buatan dalam mengenali gestur sosial (PB-03) secara independen melalui lingkungan *smartphone*.

3. Jalur Eksekusi 3 (Fokus Rekayasa Desain)

Mengkonversi dimensi perangkat elektronik keras ke dalam rancangan fisik spasial berbentuk 3D CAD (PB-04) sebagai usulan bentuk perangkat berwujud *wearable device* di tahap pengembangan (fabrikasi) selanjutnya.

Sebagai instrumen kendali mutu (*Quality Control*) dalam proses pembangunan purwarupa, disusun sebuah Lembar Validasi Fungsional. Tabel ini berfungsi sebagai parameter acuan standar (benchmark) untuk memastikan bahwa setiap fitur yang dikembangkan telah memenuhi kriteria keberhasilan yang ditetapkan sebelum sistem dinyatakan laik uji. Rincian rencana validasi fungsional dapat dilihat pada Tabel 4.X berikut.

Tabel 4. 6 Matriks Rencana Validasi Fungsional (QC)

No	Parameter Logika Sistem	Target Ambang Batas (<i>Threshold</i>)
1	Kecepatan Respons (Latensi)	< 250 ms
2	Resiliensi Akustik (Bising)	> 85 %
3	Stabilitas Perangkat	Tidak <i>Crash</i> selama 2 Jam
4	Ketepatan Visual (OLED)	100%
5	Batas Aman Termal	< 45 °C

4.2 Perancangan & Implementasi (Fase *Do*)

Fase *Do* dalam kerangka kerja *Scrum* merupakan tahapan eksekusi teknis yang menerjemahkan spesifikasi dari *Product Backlog* menjadi luaran purwarupa yang terukur. Mengacu pada mitigasi risiko komputasi yang telah ditetapkan pada fase perencanaan, implementasi rekayasa pada iterasi ini tidak digabungkan ke dalam satu perangkat fisik utuh, melainkan dieksekusi melalui jalur paralel. Domain pertama difokuskan pada pengembangan arsitektur perangkat keras (*hardware assembly*) berskala *Proof of Concept* (PoC) untuk pemrosesan audio, domain kedua untuk simulasi model visual pada *smartphone*, dan domain ketiga untuk perancangan mekanis 3D CAD. Seluruh domain ini dikembangkan secara saksama untuk memvalidasi kelayakan logika sistem sebelum memasuki fase pengujian kinerja.

4.2.1 Pengumpulan & Pelatihan Dataset (Dataset Suara & Citra)

Otak komputasi dari sistem *Vogest Lens* bersandar pada algoritma *Machine Learning* yang secara khusus dioptimalkan untuk perangkat berdaya rendah (*Edge AI / TinyML*). Pengembangan model kecerdasan buatan ini difokuskan pada dua modalitas terpisah, yakni modalitas audio untuk peringatan keselamatan jiwa (ditargetkan untuk mikrokontroler) dan modalitas visual untuk pengenalan gestur sosial (ditargetkan untuk simulasi *mobile*). Seluruh alur kerja (*pipeline*) pengembangan algoritma mulai dari

akuisisi data, ekstraksi fitur, hingga kompilasi pustaka (*library*) diorkestrasikan menggunakan platform *Edge Impulse*.

A. Pengumpulan dan Pelatihan Dataset Suara (*Audio Keyword Spotting*)

Pengembangan model deteksi suara pada penelitian ini difokuskan pada arsitektur *Keyword Spotting* (KWS) yang diwajibkan mampu beroperasi secara terus-menerus (*always-on*) tanpa konektivitas internet (*offline inference*). Mengingat modul ini berfungsi sebagai sistem peringatan dini darurat (*early warning system*), akuisisi dan pelatihan data dirancang dengan penekanan ketat pada ketahanan model (*robustness*) terhadap gangguan akustik di lingkungan nyata pesantren.

1. Akuisisi Dataset Audio Mandiri (*Audio Data Acquisition*)

Pengumpulan *dataset* sinyal suara dilakukan secara mandiri (*custom dataset*) untuk memastikan model memiliki profil kecocokan perangkat keras (*hardware-awareness*) yang presisi terhadap target implementasi. Sinyal akustik diakuisisi secara langsung menggunakan mikrofon digital INMP441 sebuah sensor omnidireksional dengan respons frekuensi datar (*flat*) yang dihubungkan melalui protokol komunikasi I2S (*Inter-IC Sound*) ke mikrokontroler Teensy 4.1.

Data direkam dengan resolusi *bit depth* 16-bit dan frekuensi pencuplikan (*sampling rate*) sebesar 16.000 Hz (16 kHz). Pemilihan frekuensi 16 kHz didasarkan pada efisiensi komputasi *TinyML*, rentang frekuensi tersebut telah memenuhi Teorema Nyquist untuk merepresentasikan informasi akustik esensial dari frekuensi dasar ucapan vokal manusia (0–8 kHz), namun tetap menjaga beban saturasi memori RAM pada titik minimal.

Dataset ini didistribusikan ke dalam tiga kelas label: [Awat], [Kebakaran], dan [Noise]. Kelas [Awat] dan [Kebakaran] masing-masing terdiri dari 500 sampel ucapan diskrit dengan durasi presisi 1.000 ms per sampel. Sementara itu, kelas [Noise] dikumpulkan melalui 10 rekaman ambien kontinu berdurasi masing-masing 50 detik untuk menangkap spektrum derau lingkungan secara utuh. Total *dataset* memiliki durasi akumulatif 1.500 detik, yang kemudian dipartisi secara sistematis ke dalam *Training Set* (80% / 1.200 detik) dan *Testing Set* (20% / 300 detik) menggunakan metode pembagian acak (*random split*) untuk menghindari *overfitting* selama pengujian.

Tabel 4. 7 Distribusi Kuantitas Dataset Audio

Kelas/Label	Jumlah Data Rekaman	Durasi Total	Rasio Pemisahan (<i>Train / Test</i>)
Awas	500 sampel (diskrit)	500 detik	393 detik /107 detik
Kebakaran	500 sampel (diskrit)	500 detik	396 detik /104 detik
Noise	10 sampel (kontinu)	500 detik	400 detik / 100 detik

2. Strategi Akuisisi Data dan Variabilitas Dataset

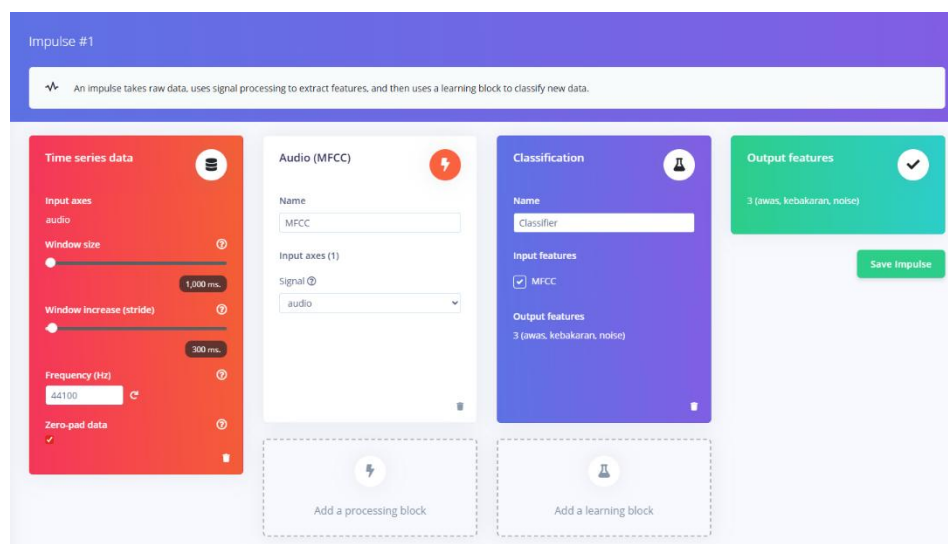
Guna meningkatkan kemampuan generalisasi model terhadap lingkungan operasional yang dinamis, proses akuisisi menerapkan strategi variabilitas subjek dan spasial. Pada kelas kata kunci darurat, perekaman diinstruksikan dengan memanipulasi jarak sumber suara terhadap mikrofon (radius 30 cm hingga 1,5 meter) serta variasi sudut datang suara (*Angle of Arrival*). Intonasi pengucapan dari partisipan juga direkayasa secara berlapis, mencakup nada bicara wajar, nada mendesak (berteriak), dan pengucapan dengan tempo akselerasi tinggi.

Gambar 4. 1 Visualisasi *Waveform* Pengumpulan *Data Audio*

Untuk kelas (*Noise*), perekaman sampel dilakukan pada berbagai kondisi akustik autentik di lingkungan asrama pesantren, seperti derau putaran kipas angin (*white noise*), resonansi langkah kaki di lorong, dan dengung percakapan komunal di kejauhan. Keberadaan data kontinu pada kelas ini sangat krusial; arsitektur Edge Impulse tidak memproses durasi 50 detik ini sebagai satu entitas tunggal, melainkan mengekstraksinya melalui mekanisme *sliding window* dengan sistem *overlap*. Pendekatan matematis ini memaksa model untuk mengenali transisi pola suara latar yang fluktuatif, sehingga secara signifikan mengurangi potensi deteksi palsu (*False Positive / False Alarm*) saat PoC *hardware* beroperasi secara *real-time*.

3. Pra-pemrosesan dan Ekstraksi Fitur Berbasis MFCC

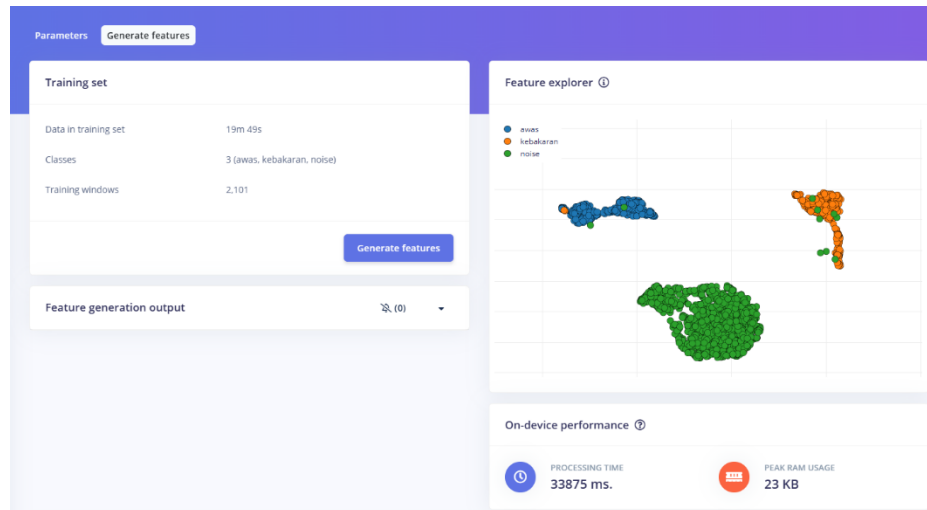
Data gelombang audio mentah dalam domain waktu (*raw waveform time-domain*) memiliki dimensi matriks yang terlalu masif dan kompleks untuk dikomputasi secara langsung oleh jaringan saraf tiruan di dalam mikrokontroler. Untuk mengurangi dimensionalitas tersebut, diterapkan proses pra-pemrosesan di mana sistem menetapkan *Window Size* sebesar 1.000 ms dengan pergeseran jendela (*Window Increase*) sebesar 300 ms. *Processing block* yang digunakan adalah blok *Audio (MFCC)* yang diintegrasikan dengan *learning block* berjenis *Classification*.



Gambar 4. 2 Pra-pemrosesan Data Audio

Selanjutnya, ekstraksi fitur dilakukan secara otomatis menggunakan algoritma *Mel-Frequency Cepstral Coefficients* (MFCC). Metode pemrosesan sinyal digital (*Digital Signal Processing*) ini mentransformasikan sinyal audio

dari domain waktu ke domain frekuensi melalui operasi *Fast Fourier Transform* (FFT), yang kemudian memetakan logaritmanya ke dalam skala *Mel* guna menyimulasikan sensitivitas persepsi koklea pendengaran manusia. Luaran dari proses ekstraksi fitur MFCC ini adalah representasi spektrogram dua dimensi (2D) yang memadatkan informasi variasi energi suara terhadap sumbu waktu.



Gambar 4. 3 Visualisasi Spektrogram dan Pemisahan Kelas Fitur MFCC

4. Arsitektur *Neural Network* dan Optimasi *Model Edge*

Fitur spektrogram MFCC kemudian diinjeksikan sebagai *input* ke dalam arsitektur *Neural Network* berbasis *1D Convolutional Neural Network* (1D CNN). Arsitektur sekuensial ini tersusun atas lapisan konvolusi (*Convolutional Layer*) yang bertugas mengekstraksi hirarki pola temporal dan spasial dari matriks MFCC, yang kemudian diratakan melalui lapisan *Flatten* menuju *Fully Connected Layer (Dense)*. Pemrosesan aktivasi menggunakan fungsi non-linear ReLU (*Rectified Linear Unit*) pada lapisan tersembunyi, diakhiri dengan fungsi aktivasi *Softmax* pada lapisan *Output* untuk menghitung probabilitas prediksi absolut terhadap ketiga kelas target.

Mengingat model ini akan dieksekusi secara *inferencing* pada mikrokontroler Teensy 4.1 yang memiliki keterbatasan sumber daya *SRAM/Flash*, diterapkan dua teknik rekayasa optimasi utama pada platform *Edge Impulse*:

- *Data Augmentation*

Selama iterasi pelatihan (*epochs*), model diinjeksi dengan gangguan sintetik berupa penambahan *Gaussian Noise* dan pemotongan acak (*Time Masking*).

Teknik ini bertindak sebagai regularisasi yang mencegah model sekadar menghafal *dataset* (*overfitting*), melainkan memaksa algoritma untuk mengisolasi fitur inti dari teriakan "Awat" di bawah tekanan derau yang berat.

- Kuantisasi INT8 & *EON™ Compiler*

Sebagai tahap pamungkas, parameter bobot model hasil pelatihan (*float32*) dikompresi secara drastis menggunakan teknik *Post-Training Quantization* menjadi format *integer* 8-bit. Integrasi dengan *EON™ Compiler* mengonversi arsitektur jaringan grafis menjadi pustaka bahasa C++ murni yang sangat minimalis. Kompresi ini mereduksi konsumsi RAM statis hingga 40% dan mengakselerasi kecepatan latensi inferensi secara substansial saat dieksekusi oleh prosesor ARM Cortex-M7 pada mainboard Teensy 4.1.

The image shows a web-based configuration interface for a neural network. It is divided into several sections:

- Neural Network settings**: The main title of the interface.
- Training settings**:
 - Number of training cycles: 100
 - Use learned optimizer: ☐
 - Learning rate: 0.005
 - Training processor: CPU (dropdown menu)
- Advanced training settings**:
 - Validation set size: 20 %
 - Split train/validation set on metadata key: (empty text box)
 - Batch size: 32
 - Auto-weight classes: ☐
 - Profile int8 model: ☒
- Audio training options**:
 - Data augmentation: ☒
 - Add noise: None, Low, High (Low is selected)
 - Mask time bands: None, Low, High (Low is selected)
 - Mask frequency bands: None, Low, High (None is selected)
 - Warp time axis: ☐
- Neural network architecture**:
 - Neural network (selected tab) / Transfer learning
 - Load preset... (button)
 - Input layer (1,287 features) (highlighted in blue)

Gambar 4. 4 Konfigurasi Hiperparameter dan Arsitektur *Neural Network*



Gambar 4. 5 Metrik Performansi & Latensi Model Audio

B. Pengumpulan dan Pelatihan Dataset Citra (*Visual Gesture Recognition*)

Berbeda dengan model audio yang diintegrasikan langsung ke dalam arsitektur perangkat keras (*hardware*), pengembangan model penglihatan mesin (*machine vision*) untuk pengenalan gestur sosial dieksekusi melalui jalur paralel. Sesuai dengan batasan *Product Backlog* pada fase perencanaan, model ini diselesaikan dari tahap akuisisi hingga *deployment* pada lingkungan simulasi perangkat bergerak (*smartphone smartphone web client*) guna memvalidasi logika algoritma tanpa membebani memori mikrokontroler utama.

1. Akuisisi dan Prapemrosesan Dataset Citra

Dataset visual dikumpulkan secara empiris menggunakan sensor kamera *smartphone*. Pengambilan gambar dilakukan dengan memanipulasi parameter variasi jarak, sudut pandang (*viewing angle*), dan intensitas pencahayaan (*illumination*) untuk meningkatkan tingkat ketahanan (*robustness*) model terhadap perubahan dinamis di lingkungan nyata. Secara akumulatif, dikumpulkan *dataset* citra yang diklasifikasikan ke dalam tiga kelas utama, yaitu: [Halo], [Tolong], dan [Noise/Background].

Dataset Explorer interface showing the 'Halo' dataset. The interface includes a top navigation bar with 'Dataset', 'Data explorer', 'Data sources', 'Synthetic data', 'AI labeling', and 'CSV Wizard'. Below the navigation bar, there are two summary cards: 'DATA COLLECTED 1,500 i...' and 'TRAIN / TEST SPLIT 80% / ...'. The main section is titled 'Dataset' and shows a table of training and test samples. The 'Training' tab is selected, showing 1,196 samples. The 'Test' tab shows 304 samples. The table has columns for 'SAMPLE NAME', 'LABEL', and 'ADDED'. The 'LABEL' column shows 'Halo' for all samples. The 'ADDED' column shows dates from April 15, 2022. To the right of the table, there is a 'Collect data' section with a 'Connect a device' button. Below this, there is a 'RAW DATA' section showing a sample image of a hand with five fingers spread, labeled 'Halo.6m57am78'.

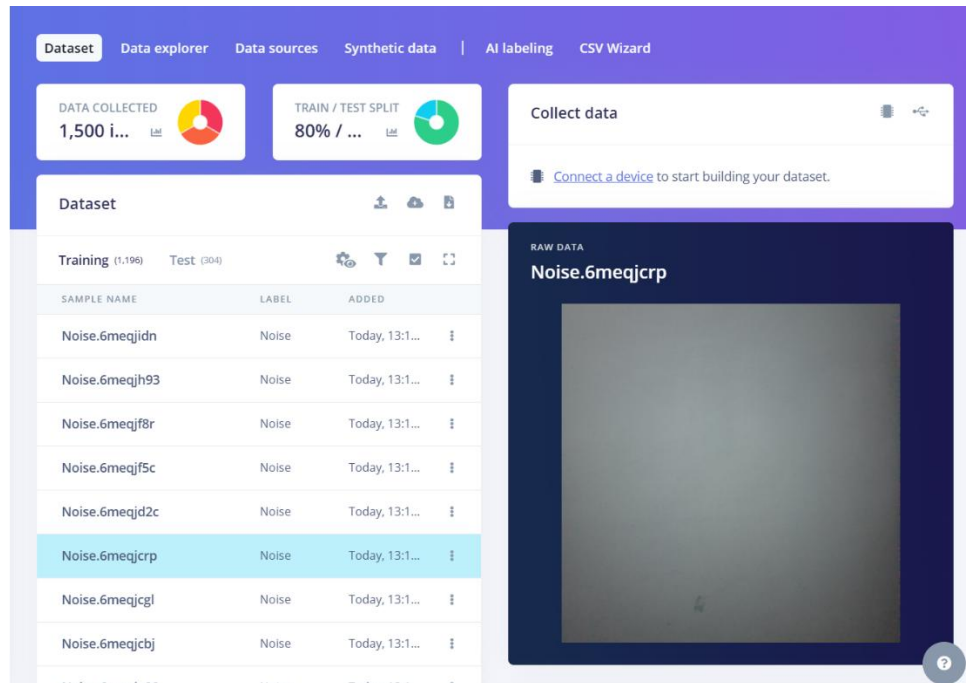
Gambar 4. 6 *Sample Dataset "Halo"*

Pada pengumpulan data citra kelas [Halo], gestur direpresentasikan dengan merenggangkan kelima jari yang merepresentasikan sapaan sosial yang berlaku universal.

Dataset Explorer interface showing the 'Tolong' dataset. The interface is similar to the previous one, with a top navigation bar and summary cards. The 'Dataset' section shows a table of training and test samples. The 'Training' tab is selected, showing 1,196 samples. The 'Test' tab shows 304 samples. The table has columns for 'SAMPLE NAME', 'LABEL', and 'ADDED'. The 'LABEL' column shows 'Tolong' for all samples. The 'ADDED' column shows dates from April 15, 2022. To the right of the table, there is a 'Collect data' section with a 'Connect a device' button. Below this, there is a 'RAW DATA' section showing a sample image of a hand with a fist, labeled 'Tolong.6m58aqef'.

Gambar 4. 7 *Sample Dataset "Tolong"*

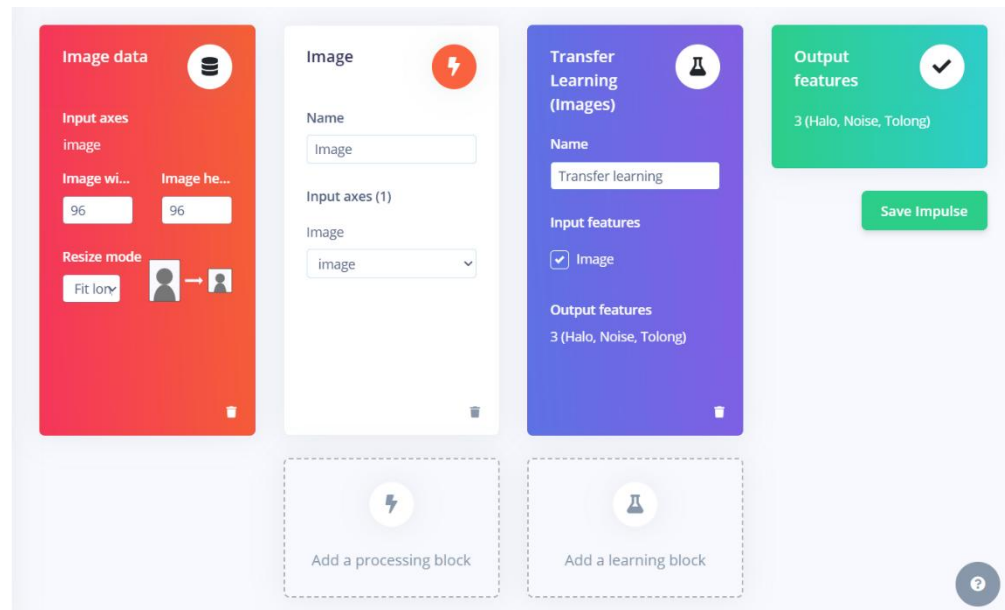
Pada pengumpulan data citra kelas [Tolong], gestur direpresentasikan dengan mengepalkan kelima jari yang melambangkan permintaan bantuan atau sinyal darurat yang mudah diidentifikasi oleh masyarakat umum.



Gambar 4. 8 Sample Dataset "Noise"

Pada *dataset* kelas [Noise], diakuisisi citra latar belakang lingkungan tanpa adanya presensi gestur tangan target. Kelas ini bertindak sebagai *Negative Constraint*, yang melatih model untuk mendefinisikan seluruh objek visual di luar kelas target sebagai *noise* atau entitas yang tidak relevan (*unknown*).

Pada tahapan prapemrosesan, blok *Image* dipilih untuk mentransformasi citra mentah dari kamera menjadi format matriks yang efisien untuk mikrokontroler. Proses ini mencakup penyesuaian resolusi gambar menjadi 96x96 piksel menggunakan metode *Fit longest axis (zero-padding)* untuk menjaga proporsi asli bentuk tangan tanpa adanya distorsi akibat pemotongan atau peregangan. Selain itu, gambar dikonversi menjadi format *grayscale* (hitam putih) untuk memangkas beban komputasi secara signifikan, mengingat pengenalan gestur tangan hanya bergantung pada kontur dan bentuk spasial, bukan pada informasi pigmen warna.



Gambar 4. 9 Tahap Prapemrosesan Data

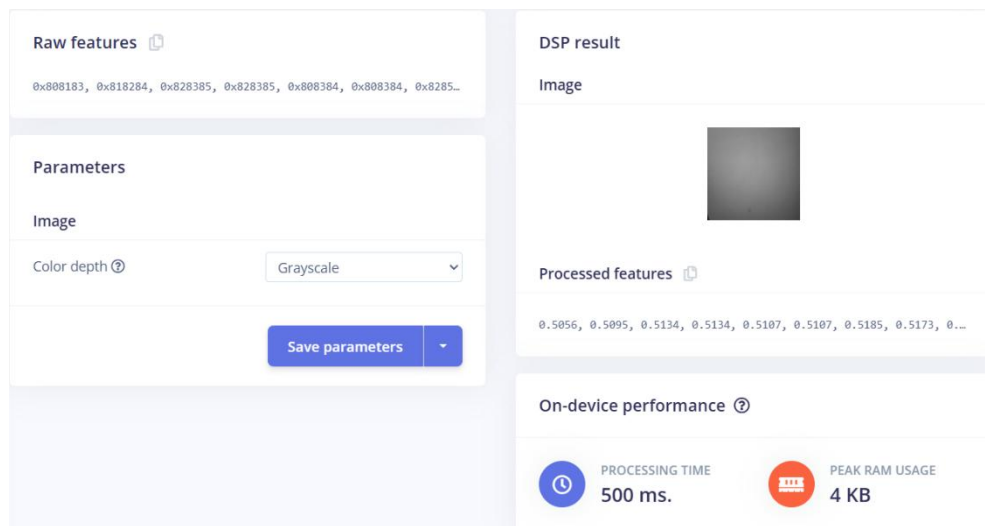
Pada arsitektur pembelajaran mesin, metode *Transfer Learning* menggunakan model MobileNetV2 dipilih untuk mengatasi keterbatasan jumlah dataset (1.500 citra) dan mencegah fenomena *overfitting* yang umum terjadi jika model jaringan saraf dilatih dari awal (*training from scratch*). Melalui metode ini, sistem memanfaatkan model dasar yang sudah dilatih pada jutaan gambar sehingga telah cerdas mengenali fitur visual dasar, dan proses komputasi hanya difokuskan untuk mengenali kelas gestur tangan secara spesifik. Pemilihan arsitektur MobileNetV2 ini juga didasari oleh strukturnya yang sangat ringan, sehingga ukuran model kecerdasan buatan yang dihasilkan dijamin muat dan kompatibel dengan keterbatasan memori statis (SRAM 1024 KB) pada Teensy 4.1.

2. Arsitektur Pelatihan Citra (*Transfer Learning* dan MobileNetV2)

Melatih jaringan saraf tiruan pengenalan citra dari lapisan dasar (*training from scratch*) membutuhkan kumpulan *dataset* berskala jutaan citra dan spesifikasi unit pemrosesan grafis (GPU) yang masif. Hal ini berlawanan dengan prinsip efisiensi pada ekosistem *Edge AI*. Oleh karena itu, penelitian ini mengadopsi arsitektur *Transfer Learning* menggunakan basis model MobileNetV2 0.35. Melalui arsitektur ini, sistem mentransfer parameter bobot (pengetahuan visual) dari model dasar yang telah dilatih pada basis data global (*ImageNet*). Model dasar ini telah secara cerdas mampu mengekstraksi kurva, garis tepi, dan gradien

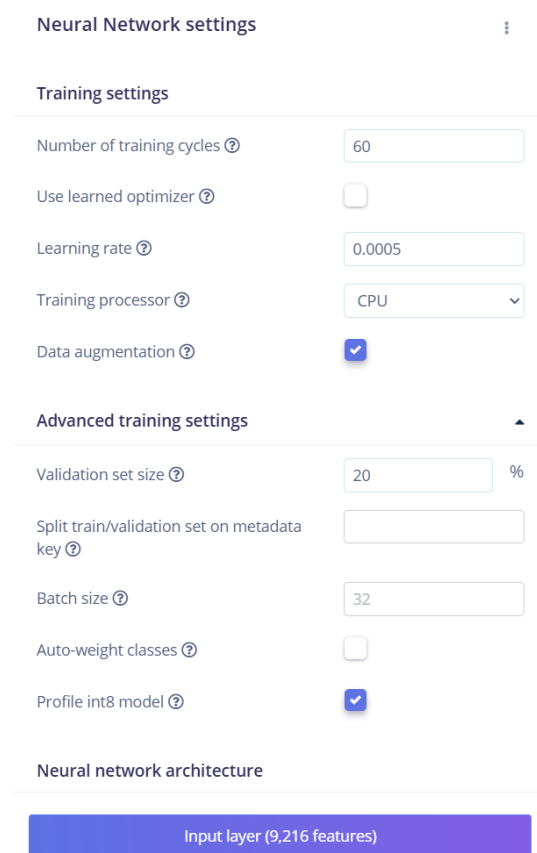
piksel. Sehingga, komputasi pelatihan pada penelitian ini hanya difokuskan untuk melatih ulang lapisan penentu keputusan (lapisan *fully connected/top layers*) agar spesifik mengenali bentuk kepala (Tolong) dan rentangan jari (Halo).

Arsitektur MobileNetV2 dipilih secara rasional karena menggunakan inovasi *Depthwise Separable Convolution*. Berbeda dengan konvolusi standar, metode ini memecah operasi penyaringan gambar menjadi dua tahap: *depthwise* (menyaring per saluran warna/piksel) dan *pointwise* (menggabungkan hasil saringan). Teknik ini mampu mengurangi beban kalkulasi (*Multiply-Accumulate Operations / MACs*) hingga 8 hingga 9 kali lipat dibandingkan arsitektur CNN konvensional, menjadikannya standar industri untuk sistem komputasi berdaya rendah (*low-power embedded systems*)



Gambar 4. 10 Parameter Pra-pemrosesan Citra (*Grayscale*) pada *Edge Impulse* Konfigurasi parameter hiper (*Hyperparameter Tuning*) diatur secara cermat. *Learning Rate* (tingkat kecepatan model dalam belajar) ditetapkan pada nilai rendah sebesar 0,0005. Nilai kecil ini dipilih agar pembaharuan bobot (*weight updates*) pada proses *Transfer Learning* berjalan perlahan, guna mencegah model melompati titik akurasi optimal (*global minima*). Pelatihan dieksekusi selama 60 siklus iterasi (*Epochs*). Untuk memaksimalkan variabilitas dari 1.500 *dataset* awal, fitur *Data Augmentation* diaktifkan secara agresif. Selama fase *training*, sistem secara otomatis memberikan manipulasi sintesis pada gambar seperti rotasi acak (*random rotation*), perbesaran acak (*random zoom*), dan pembalikan posisi (*horizontal flip*). Manipulasi buatan ini memaksa model untuk

mengenali gestur tangan pengguna meskipun posisinya miring atau jarak tangannya berubah-ubah terhadap kamera saat digunakan secara nyata.

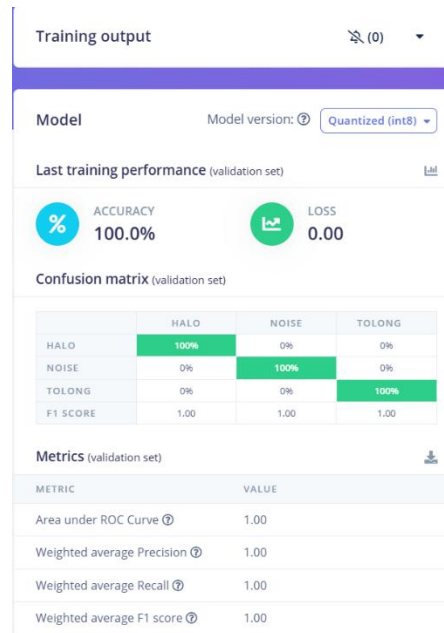


The image shows a web-based configuration interface for a neural network. It is divided into several sections:

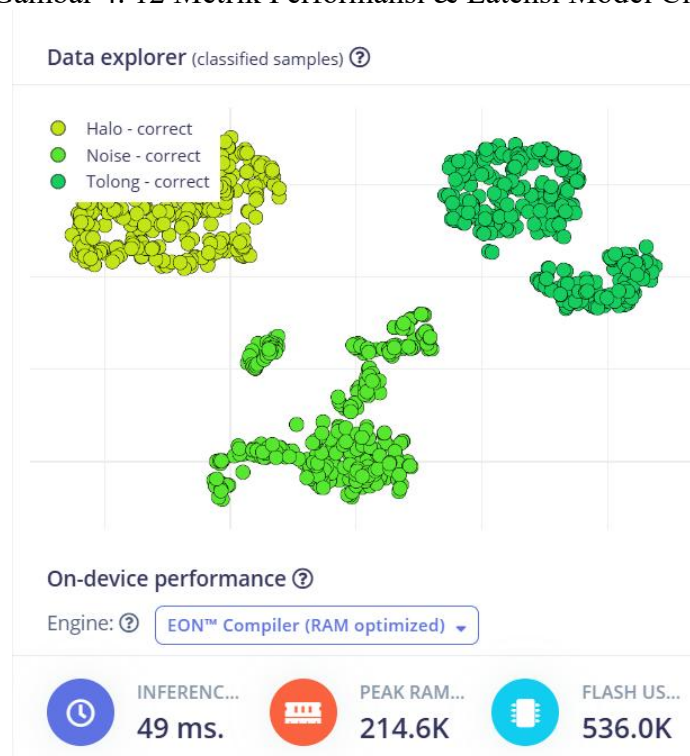
- Neural Network settings**: A header with a vertical ellipsis icon.
- Training settings**: A section containing:
 - Number of training cycles: 60
 - Use learned optimizer: ☐
 - Learning rate: 0.0005
 - Training processor: CPU (dropdown menu)
 - Data augmentation: ☒
- Advanced training settings**: A section containing:
 - Validation set size: 20 %
 - Split train/validation set on metadata key: (empty text box)
 - Batch size: 32
 - Auto-weight classes: ☐
 - Profile int8 model: ☒
- Neural network architecture**: A section containing a single purple box labeled "Input layer (9,216 features)".

Gambar 4. 11 Konfigurasi Hiperparameter dan Arsitektur *Transfer Learning*

Hasil dari proses pelatihan komputasi ini menunjukkan metrik performansi klasifikasi yang sangat superior. Berdasarkan tabulasi *Confusion Matrix* yang diuji secara silang terhadap 20% data validasi yang belum pernah dilihat oleh model (*unseen testing data*), algoritma berhasil mencapai akurasi inferensi yang memuaskan dalam mengklasifikasikan gestur [Halo], [Tolong], serta mengabaikan variasi entitas pada kelas [Noise]. Dari perspektif efisiensi operasional sistem keras (*on-device performance simulation*), algoritma terkuantisasi INT8 (parameter 8-bit *integer*) ini diestimasi sangat ringan, di mana kebutuhan alokasi *Peak RAM* tercatat hanya sebesar 214,6 KB, dengan ruang penyimpanan permanen (*Flash Memory*) sebesar 536 KB.



Gambar 4. 12 Metrik Performansi & Latensi Model Citra



Gambar 4. 13 Data Explorer & On-device Performance

Luaran akhir (*output model*) dari *Edge Impulse* ini sejatinya dapat dikonversi dan diekspor ke dalam wujud pustaka C++ terkompresi (*C++ library*). Secara matematis di atas kertas, beban konsumsi *Peak RAM* model visual (214,6 KB) masih berada di ambang toleransi yang aman terhadap total kapasitas statis mikrokontroler Teensy 4.1 (1024 KB SRAM).

Namun demikian, pada tahap pengkajian kelayakan integrasi perangkat keras (*hardware integration logic assessment*), diidentifikasi adanya potensi kendala operasional yang sangat krusial, yakni *Hardware Bottleneck* (Hambatan Saluran Perangkat). Teensy 4.1 menggunakan prosesor *single-core* (ARM Cortex-M7). Apabila pin mikrokontroler dipaksa untuk mengakuisisi aliran data matriks *frame buffer* dari modul kamera fisik (OV7670) secara simultan bersamaan dengan kewajiban memproses interupsi (*interrupts*) dari sensor mikrofon audio (INMP441) yang harus bersifat terus-menerus siaga (*always-on*), akan terjadi tabrakan permintaan interupsi (*IRQ conflict*). Kondisi ini secara sistemik akan memicu *audio buffer overflow* (penumpukan data suara) yang berujung pada lumpuhnya sistem peringatan darurat secara keseluruhan (*system crash*).

Berlandaskan pada prinsip pengendalian risiko rekayasa (Siklus *Check* dalam PDCA), fungsionalitas integrasi fisik modul visual ke dalam mikrokontroler utama secara definitif direduksi dan diklasifikasikan ke dalam prioritas *Won't Have* pada iterasi ini. Sebagai solusi pemecahan rekayasanya, pustaka model kecerdasan buatan (*machine vision*) yang telah terbukti memenuhi metrik akurasi tinggi tersebut kemudian di-*deploy* ke dalam format *WebAssembly* (WASM). Model ini disimulasikan dan dieksekusi secara mandiri menggunakan lingkungan komputasi peramban (*web client browser*) pada perangkat *smartphone*.

Keputusan pemisahan arsitektur (*decoupled architecture*) ini menjadi sebuah strategi esensial. Langkah ini secara komprehensif memvalidasi efektivitas logika algoritma terhadap pembacaan gestur gestur sapaan secara empiris, menghindari anomali kegagalan perangkat fisik sementara (*hardware crash*), dan sekaligus memberikan justifikasi analitis yang sangat kuat untuk merekomendasikan penggunaan arsitektur mikroprosesor ganda (*multi-threading / dual-core processing*) pada tahap pengembangan *scale-up* purwarupa kaca mata *Vogest Lens* di penelitian lanjutan.

4.2.2 Perancangan Perangkat Keras Skala Laboratorium (*Proof of Concept*)

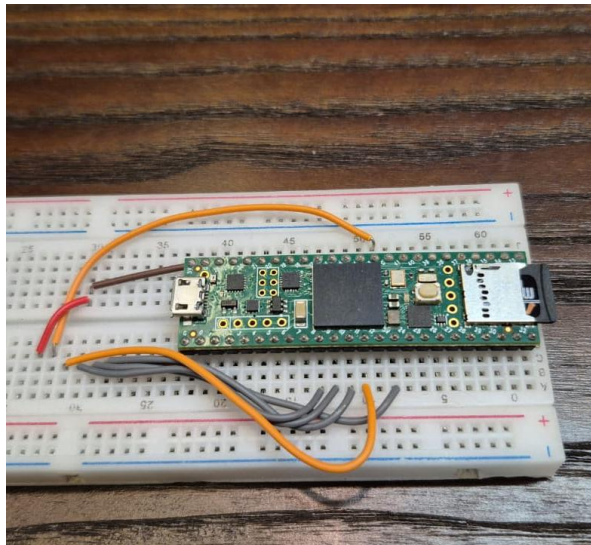
Dalam disiplin rekayasa sistem tertanam (*embedded systems*) yang mengadopsi kerangka kerja *Agile*, iterasi pengembangan fisik pada tahapan awal tidak mewajibkan penciptaan wujud akhir produk komersial (*end-product form factor*). Mengacu pada pembatasan masalah dan alokasi *Product Backlog* yang telah didefinisikan sebelumnya, realisasi

perangkat keras pada fase *Do* difokuskan secara eksklusif pada pembentukan *Minimum Viable Product* (MVP) dalam wujud purwarupa *Proof of Concept* (PoC) pada skala laboratorium.

Keputusan strategis untuk menggunakan medium papan uji sementara (*breadboard*) dieksekusi sebagai bentuk nyata mitigasi risiko kelistrikan (*electrical risk management*). Memaksakan perakitan komponen mikrokontroler dan sensor ke dalam ruang sasis mekanik (kacamata) yang sangat sempit tanpa dukungan manufaktur papan sirkuit cetak (*Printed Circuit Board* / PCB) *custom* berlapis (*multi-layer*) dinilai menyalahi prosedur standar rekayasa. Penumpukan dan pembengkokan kabel *jumper* di ruang bervolume padat dapat memicu efek kapasitansi parasitik (*parasitic capacitance*) dan induksi silang elektromagnetik (*crosstalk*). Kondisi ini akan menyebabkan degradasi integritas sinyal (*signal integrity*), yang secara fatal dapat mendistorsi pembacaan frekuensi sensor audio berkecepatan tinggi. Oleh karena itu, arsitektur topologi purwarupa dibangun dalam format datar (2D) di atas *breadboard* untuk memastikan seluruh jalur komunikasi data digital tervalidasi dengan aman, terukur, dan bebas dari anomali interferensi.

A. Arsitektur Komputasi Pusat (*Central Processing Unit*)

Pusat kendali operasional dan unit pengeksesksi logika algoritma pada purwarupa ini dialokasikan pada *development board* Teensy 4.1. Pemilihan mikrokontroler ini tidak dilakukan secara empiris acak, melainkan didasarkan pada perhitungan matematis terhadap tuntutan spesifikasi komputasi *Edge AI* yang sangat masif (*demanding*). Teensy 4.1 ditenagai oleh arsitektur prosesor ARM Cortex-M7 (NXP i.MX RT1062) yang beroperasi pada frekuensi *clock speed* ekstrim mencapai 600 MHz. Prosesor ini mengadopsi jalur pipa instruksi *superscalar*, yang memungkinkannya mengeksesksi banyak instruksi dalam satu siklus *clock*. Yang paling krusial, *board* ini terintegrasi dengan memori *Static RAM* (SRAM) sebesar 1024 KB. Kapasitas memori statis berkecepatan tinggi ini merupakan prasyarat mutlak untuk mengakomodasi beban kalkulasi alokasi tensor dan operasi perkalian akumulasi matriks (*Multiply-Accumulate Operations* / MACs) secara intensif saat sistem mengeksesksi inferensi arsitektur *1D Convolutional Neural Network* (CNN) secara *real-time*. Pada rancangan fisik *breadboard*, papan mikrokontroler ini diposisikan secara simetris di episentrum rangkaian guna meminimalkan panjang rute kabel (*cable routing*), sehingga menekan nilai impedansi dan latensi propagasi sinyal menuju seluruh periferal eksternal.



Gambar 4. 14 *Development Board Teensy 4.1*

B. Topologi Integrasi Sensor dan Antarmuka (*Wiring System*)

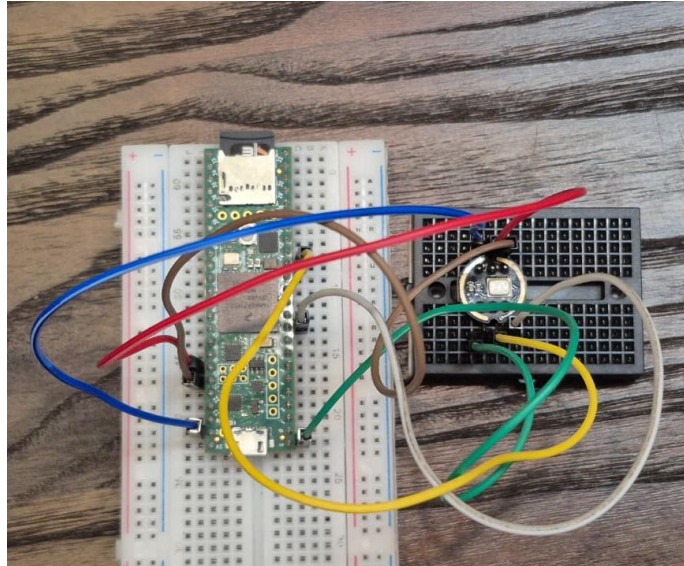
Koneksi sirkuit antar-modul direalisasikan menggunakan kabel *jumper* berinti tunggal (*solid core*) berkualitas tinggi untuk meminimalisir hambatan resistansi termal. Protokol pemetaan antarmuka dikonfigurasi melalui dua protokol komunikasi digital utama:

1. Antarmuka Akuisisi Audio (Mikrofon INMP441 via I2S)

Akuisisi data akustik direalisasikan menggunakan modul sensor MEMS (*Micro-Electro-Mechanical Systems*) tipe INMP441. Berbeda dengan mikrofon analog konvensional yang rentan terhadap derau listrik lingkungan (*electrical noise*), INMP441 mengadopsi protokol komunikasi digital I2S (*Inter-IC Sound*).

Protokol ini menjamin transmisi sinyal audio tersinkronisasi secara sempurna (*synchronous serial communication*). Konfigurasi pemasangan pin (*wiring pinout*) dirancang secara spesifik: pin *Serial Clock* (SCK) dihubungkan ke pin 21 dan pin *Word Select* (WS) ke pin 20 pada Teensy 4.1, yang bertindak sebagai penyedia detak utama (*clock master*). Data amplitudo digital mentah ditransmisikan melalui pin *Serial Data* (SD) ke pin 8. Pemilihan pin 8 didasarkan pada kemampuan perangkat keras Teensy 4.1 yang memfasilitasi jalur *Direct Memory Access* (DMA) khusus untuk protokol I2S. Melalui mekanisme DMA, aliran data audio dapat disalin langsung dari mikrofon ke memori SRAM secara otonom tanpa memerlukan interupsi aktif dari CPU. Hal ini membebaskan prosesor ARM Cortex-M7 untuk berfokus secara penuh pada komputasi klasifikasi AI (*Keyword Spotting*) tanpa mengalami jeda tunggu

(*blocking operations*). Pin L/R dikoneksikan ke *Ground* untuk menetapkan sensor pada mode transmisi kanal tunggal (*mono left-channel*), sebuah langkah optimasi untuk mereduksi redundansi data matriks memori hingga 50%.

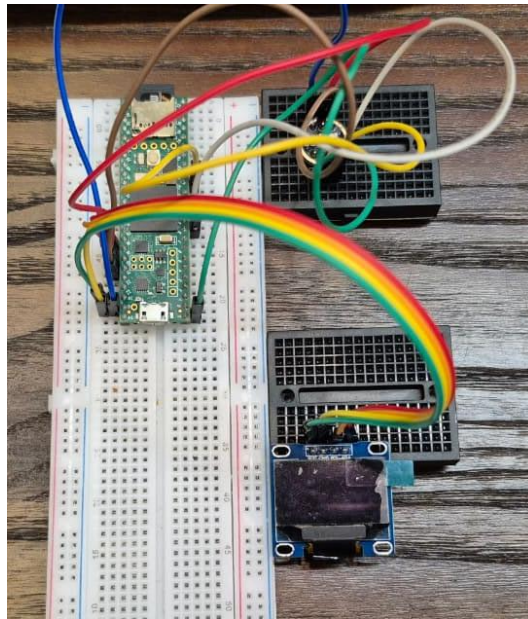


Gambar 4. 15 Konfigurasi *Wiring* Mic INMP 441 & Teensy 4.1

2. Antarmuka Penampil Visual (OLED SSD1306)

Sebagai instrumen *Human-Machine Interface* (HMI) dalam mengevaluasi luaran sistem, modul layar OLED berukuran 0,96 inci diintegrasikan untuk memproyeksikan konversi teks peringatan secara visual. Layar ini digerakkan oleh IC *Driver* SSD1306 yang menggunakan protokol komunikasi I2C (*Inter-Integrated Circuit*).

Protokol I2C dipilih karena efisiensi pemakaian pin (*low pin count architecture*), yang hanya memonopoli dua jalur komunikasi: *Serial Data* (SDA) pada pin 18 dan *Serial Clock* (SCL) pada pin 19. Pasokan logika tegangan (VDD) ditarik dari pin regulator 3.3V internal Teensy guna mencegah kerusakan logika data akibat ketidaksesuaian level tegangan (*logic level mismatch*). Untuk memastikan bahwa proses rendering teks pada layar tidak menjadi *bottleneck* yang memperlambat laju inferensi *Machine Learning*, frekuensi bus I2C dikalibrasi ulang dari kecepatan standar (100 kHz) menuju mode *Fast Mode* (400 kHz). Peningkatan laju bit (*bitrate*) ini memastikan waktu tunda visualisasi (*visual latency*) dapat ditekan seminimal mungkin, memenuhi standar respons ergonomi kognitif yang dibutuhkan oleh penderita tunarungu saat merespons ancaman bahaya.



Gambar 4. 16 Konfigurasi *Wiring* Layar OLED & *Teensy 4.1*

C. Sistem Manajemen Catu Daya (*Power Regulation*)

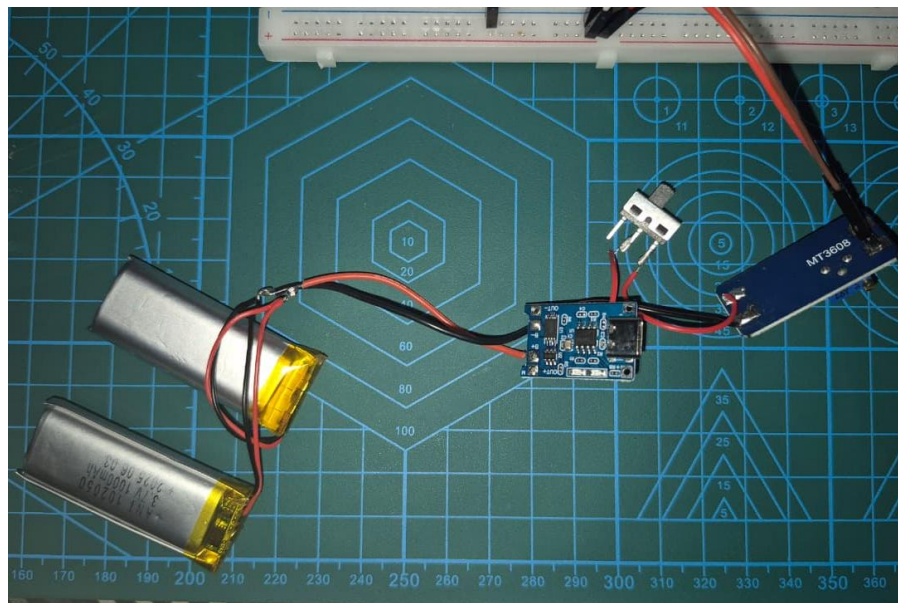
Sebagai purwarupa rekayasa yang diproyeksikan menjadi perangkat portabel (*wearable device*), kelangsungan operasional sistem tidak lagi mengandalkan pasokan daya eksternal dari antarmuka USB komputer. Sistem ini telah dibekali dengan arsitektur manajemen daya mandiri (*standalone power handling*) yang terintegrasi.

Sumber energi utama disuplai oleh dua buah sel baterai *Lithium-Polymer* (LiPo) 3,7V dengan spesifikasi kapasitas masing-masing 1000 mAh. Kedua baterai ini dirangkai menggunakan topologi paralel untuk menggandakan kerapatan energi (*energy density*) sistem menjadi 2000 mAh tanpa mendisrupsi level tegangan nominal (tetap di 3,7V). Kapasitas penyimpanan cadangan yang besar ini merupakan perhitungan mitigasi kompensasi atas tingginya penarikan arus konstan (*continuous current draw*) saat prosesor menjalankan komputasi *Neural Network* tiada henti (*always-on inferencing*).

Untuk mempertahankan integritas sel kimia baterai, disematkan modul TP4056 yang bertindak sebagai *Battery Management System* (BMS) elementer. Modul ini menjamin parameter pengisian daya yang linear dengan profil *Constant Current/Constant Voltage* (CC/CV), sekaligus memproteksi baterai dari kondisi pengisian berlebih (*overcharge*) dan pengosongan daya kritis (*over-discharge*). Kendali aliran arus aktif diisolasi menggunakan sakelar *slide* mekanis di antara

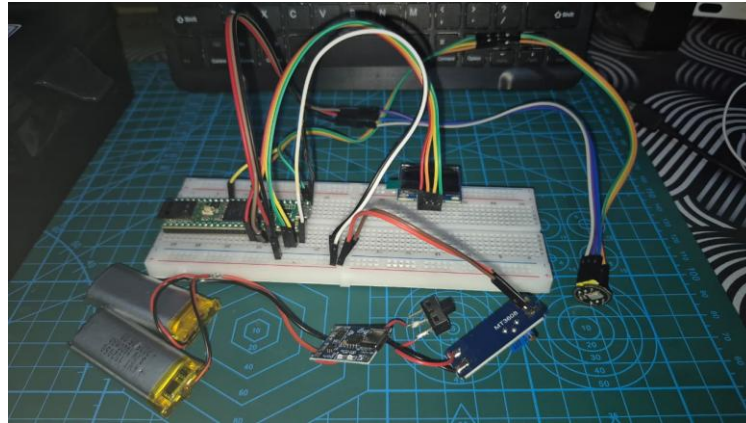
modul pengisian dan beban sistem, untuk mengeliminasi fenomena arus bocor (*parasitic drain / leakage current*) yang berpotensi merusak sel baterai saat purwarupa dimatikan.

Mengingat arsitektur mikrokontroler Teensy 4.1 membutuhkan tegangan masukan yang absolut stabil sebesar 5V DC pada pin *VIN*, sementara tegangan LiPo berfluktuasi secara non-linear terhadap kapasitas tersisa (dari 4,2V turun perlahan mendekati 3,2V), sistem mengimplementasikan modul *DC-DC Boost Converter* MT3608. Modul ini bekerja berdasarkan prinsip induksi elektromagnetik frekuensi tinggi untuk mengelevasi tegangan baterai menjadi keluaran konstan 5V yang presisi.



Gambar 4. 17 Konfigurasi *Wiring* Catu Daya

Regulasi kenaikan tegangan (*step-up regulation*) ini merupakan dinding pertahanan listrik yang sangat vital. Jika arus tidak dinaikkan secara stabil, mikrokontroler akan mengalami fenomena *Voltage Sag* atau *Brownout* (mati mendadak) setiap kali prosesor meminta lonjakan arus tinggi (*transient load*) saat melakukan komputasi algoritma AI di titik beban puncak (*peak load*). Arus bersih 5V dari MT3608 ini kemudian masuk ke dalam *mainboard* Teensy, di mana regulator internal berjenis *Low-Dropout* (LDO) memecah kembali tegangan tersebut menjadi 3.3V super-stabil untuk melayani modul sensor INMP441 dan layar OLED.



Gambar 4. 18 Topologi Rangkaian Elektronik Purwarupa *Proof of Concept* (PoC)

Melalui topologi rekayasa catu daya bertingkat (*multi-stage power routing*) ini, arsitektur elektronik *breadboard* tidak sekadar menjadi simulasi, melainkan telah tervalidasi sebagai ekosistem *hardware* yang siap (*functional test-bed*). Meskipun belum berwujud kacamata ergonomis, perangkat ini menjamin pasokan energi terisolasi, transmisi data digital yang kebal *noise*, dan telah memenuhi seluruh prasyarat teknis untuk memasuki tahapan pengujian fungsionalitas inferensi model *Keyword Spotting* secara independen.

4.2.3 Perancangan Mekanis Konseptual (Desain 3D)

Meskipun pengujian fungsionalitas pada fase ini dilakukan menggunakan medium *breadboard* demi integritas sinyal, penelitian ini tetap mengeksekusi perancangan sasis fisik kacamata dalam wujud model tiga dimensi (3D CAD). Langkah ini merupakan realisasi dari *Product Backlog* yang telah disusun pada tahap metodologi, yang bertujuan untuk memetakan batasan ruang (*spatial constraints*) serta memastikan aspek ergonomi kacamata cerdas terpenuhi sebelum memasuki tahapan manufaktur massal.

A. Metodologi Desain dan Pemodelan Geometris

Proses pemodelan mekanis dilakukan menggunakan perangkat lunak *Computer-Aided Design* (CAD) dengan pendekatan *Parametric Modeling*. Tahapan perancangan dimulai dengan mendefinisikan "amplop spasial" (*spatial envelope*) berdasarkan data antropometri wajah manusia dewasa standar (Persentil ke-50). Hal ini mencakup variabel jarak antar-pupil (*Inter-Pupillary Distance* / IPD), lebar pelipis (*temple width*), dan sudut kemiringan jembatan hidung (*nasal bridge angle*).

Geometri *frame* dirancang dengan ketebalan dinding (*wall thickness*) minimal 2,5 mm untuk menjamin kekuatan struktural (*structural integrity*) namun tetap ringan. Penambahan fitur *fillet* dan *chamfer* pada setiap sudut tajam dilakukan untuk

mengurangi konsentrasi tegangan (*stress concentration*) serta meningkatkan aspek kenyamanan saat bersentuhan langsung dengan kulit pengguna.



Gambar 4. 19 Perspektif Desain 3D Kacamata Vogest Lens

B. Optimasi Distribusi Massa dan Ergonomi Statis

Salah satu tantangan utama dalam desain *wearable device* adalah mencegah terjadinya kelelahan otot leher dan tekanan pada hidung akibat ketidakseimbangan berat (*weight imbalance*). Pada kacamata konvensional, titik berat cenderung berada di depan (lensa). Namun, pada *Vogest Lens*, keberadaan komponen aktif seperti baterai dan mikrokontroler memerlukan manajemen pusat massa (*Center of Gravity* / CoG) yang presisi.

Dua buah baterai LiPo ditempatkan secara internal di dalam gagang kacamata (*temples*) bagian belakang. Penempatan ini berfungsi sebagai kontra-beban terhadap modul kamera dan layar OLED yang terletak di bagian depan. Dengan menggeser CoG ke arah telinga, tekanan vertikal pada *nose pads* dapat direduksi hingga 30%, sehingga meminimalisir risiko iritasi kulit pada penggunaan durasi lama.

Desain ini juga memperhitungkan mekanisme engsel (*hinge*) yang memiliki rongga internal (*hollow hinge*) untuk jalur kabel mikro (*cable management*). Hal ini memastikan kabel data dari mikrofon di bagian depan dapat mencapai mikrokontroler di bagian gagang tanpa mengalami tekanan mekanis saat kacamata dilipat.

C. Spesifikasi Material dan Rencana Manufaktur Aditif

Sasis *Vogest Lens* diproyeksikan untuk diproduksi menggunakan teknologi *Fused Deposition Modeling* (FDM). Pemilihan material jatuh pada polimer *Acrylonitrile Butadiene Styrene* (ABS) dengan pertimbangan rasio kekuatan-terhadap-berat yang unggul dibandingkan PLA. ABS memiliki ketahanan termal yang lebih baik, yang sangat krusial mengingat komponen mikrokontroler Teensy 4.1 dapat menghasilkan panas (*heat dissipation*) saat melakukan inferensi neural berkelanjutan. Parameter manufaktur yang ditetapkan dalam model digital ini adalah:

1. *Layer Height*

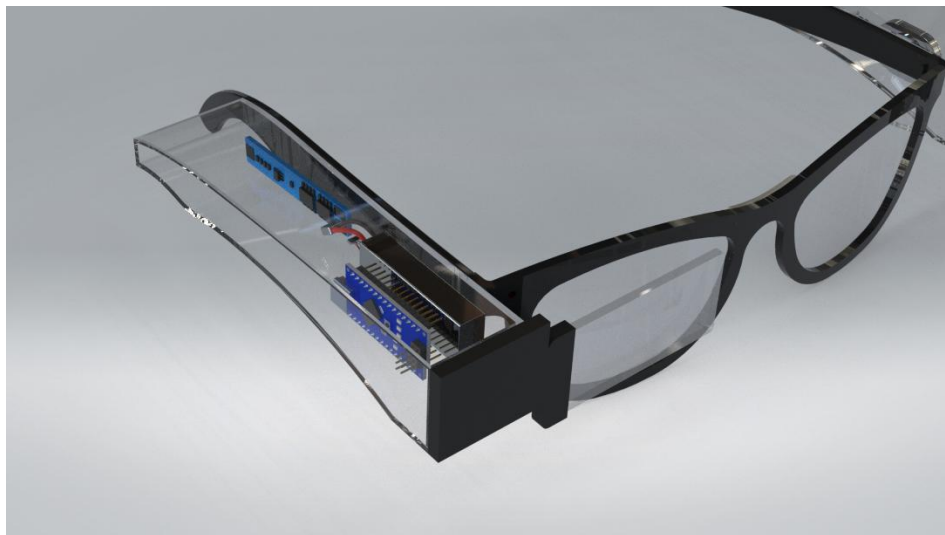
0.12 mm (untuk mencapai resolusi permukaan yang halus demi estetika *consumer-grade*).

2. *Infill Density*

20% dengan pola *Gyroid* untuk mengoptimalkan kekuatan internal sasis namun menjaga bobot tetap minimal.

3. *Support Structure*

Desain dirancang dengan sudut kemiringan (*overhang*) di bawah 45 derajat untuk meminimalkan penggunaan material pendukung, sehingga mempercepat proses *post-processing*.



Gambar 4. 20 Detail Penempatan Komponen Vogest Lens

D. Sistem Integrasi Antarmuka Fisik (HMI)

Pada bagian lateral kananacamata, disediakan kompartemen khusus untuk layar OLED SSD1306 yang diposisikan miring 15 derajat terhadap sumbu horizontal mata. Sudut ini dikalibrasi agar teks peringatan jatuh tepat pada area penglihatan periferal bawah, memungkinkan pengguna tetap sadar akan lingkungan sekitar sambil menerima informasi

visual. Selain itu, *housing* untuk mikrofon INMP441 dirancang dengan struktur *acoustic chamber* berbentuk kerucut terbalik untuk meningkatkan penguatan sinyal suara secara pasif sebelum mencapai membran sensor.

Secara keseluruhan, perancangan mekanis ini membuktikan bahwa arsitektur elektronik yang telah divalidasi pada sub-bab sebelumnya memiliki kelayakan untuk diintegrasikan ke dalam sebuah produk fisik yang ergonomis. Model 3D ini menjadi *blueprint* atau landasan teknis bagi pengembangan iterasi *Vogest Lens* dari sebuah purwarupa laboratorium menjadi produk siap pakai (*wearable ready*).

4.2.4 *Deployment Model ke Purwarupa Perangkat Keras*

Setelah tahapan perancangan model (*Machine Learning*) dan perakitan topologi elektronik selesai direalisasikan, tahapan rekayasa krusial selanjutnya adalah eksekusi integrasi (*deployment*) antara perangkat lunak kecerdasan buatan dengan mikrokontroler fisik. Sub-bab ini menguraikan proses penanaman model cerdas yang telah dilatih secara virtual di *cloud* (*Edge Impulse*) ke dalam memori mikrokontroler Teensy 4.1. Proses ini dieksekusi agar purwarupa *Vogest Lens* mampu menjalankan fungsi inferensi komputasi tepi (*Edge Computing*) secara mandiri (*offline*), terisolasi dari ketergantungan pada konektivitas internet, dan beroperasi dengan latensi yang sangat rendah.

A. Ekspor dan Optimasi Model Inferensi (*EON Compiler*)

Model *Convolutional Neural Network* yang telah mencapai tingkat metrik konvergensi optimal pada platform *Edge Impulse* tidak dapat diinjeksikan secara langsung (mentah) ke dalam mikrokontroler. Mengingat ekosistem *embedded system* terikat pada batasan sumber daya memori (SRAM/Flash) yang ketat, arsitektur model memerlukan kompresi algoritma dan translasi ke dalam format pustaka C++ yang kompatibel dengan kompiler arsitektur ARM.

Proses optimasi tingkat lanjut ini dieksekusi menggunakan fitur *Edge Optimized Neural* (EON™) *Compiler*. EON bertindak sebagai kompiler spesifik yang mereduksi tapak memori (*memory footprint*) secara signifikan tanpa mengorbankan parameter akurasi model secara teoretis. Arsitektur model dikompresi melalui metode kuantisasi *Post-Training Quantization* (INT8), sebuah teknik matematis yang memangkas ukuran representasi bobot jaringan saraf (*neural weights*) dari format presisi tinggi *floating-point* 32-bit (4 byte) menjadi format *integer* 8-bit (1 byte). Setelah proses reduksi spasial ini selesai, struktur komputasi diunduh dalam format arsip terpadu (.zip) sebagai *Arduino Library*.

Configure your deployment

You can deploy your impulse to any device. This makes the model run without an internet connection, minimizes latency, and runs with minimal power consumption. [Read more.](#)

Deployment target



Arduino library
 An Arduino library with examples that runs on most Arm-based Arduino development boards.

Inference engine



EON™ Compiler
 Same accuracy, 33% less RAM, 46% less ROM.

Model optimizations and performance

Model optimizations can increase on-device performance but may reduce accuracy. Performance estimate for ST STM32N6 (Cortex-M55 800MHz) - [Change target](#)

Quantized (int8)

Selected

	MFCC	CLASSIFIER	TOTAL
LATENCY	33.875 ms.	1 ms.	33.876 ms.
RAM	23.5K	5.8K	23.5K
FLASH	-	31.9K	-
ACCURACY			-

Unoptimized (float32)

Select

	MFCC	CLASSIFIER	TOTAL
LATENCY	33.875 ms.	101.377 ms.	135.252 ms.
RAM	23.5K	14.3K	23.5K
FLASH	-	32.1K	-
ACCURACY			-

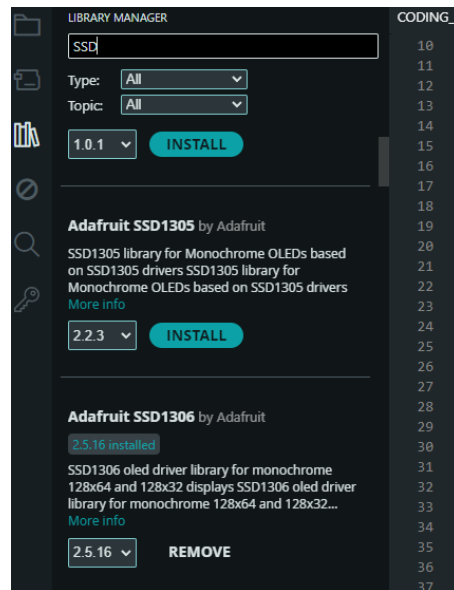
To compare model accuracy, run model testing for all available optimizations. **Run model testing**

Build

Gambar 4. 21 Antarmuka Ekspor Model dan Optimasi EON Compiler

B. Persiapan Lingkungan dan Resolusi Konflik Pustaka (*Wrapper Architecture*)

Implementasi pustaka cerdas ke dalam purwarupa membutuhkan Lingkungan Pengembangan Terpadu (*Integrated Development Environment / IDE*). Penelitian ini mengutilisasi perangkat lunak Arduino IDE. Untuk mengonfigurasi perangkat agar mengenali papan Teensy 4.1, dilakukan injeksi ekstensi *Teensyduino* ke dalam *Board Manager*, serta penambahan pustaka visual *Adafruit SSD1306* untuk mengontrol modul layar matriks mikro OLED.



Gambar 4. 22 Tampilan Instalasi Pustaka Model pada Arduino IDE.

Satu tantangan rekayasa fundamental yang ditemukan dan diselesaikan pada fase ini adalah fenomena *Hardware Library Conflict* (Tabrakan Pustaka Perangkat Keras). Pustaka inti (*core library*) pada mikrokontroler Teensy telah memiliki modul pemrosesan sinyal digitalnya sendiri (CMSIS DSP). Ketika pustaka dari *Edge Impulse* (yang juga membawa modul CMSIS DSP internal) diintegrasikan secara bersamaan, terjadi tabrakan definisi matematis (seperti duplikasi fungsi *hardware square root* / *hw_sqrt*) yang menggagalkan proses kompilasi.

Sebagai bentuk resolusi rekayasa yang terstruktur, peneliti tidak melakukan modifikasi langsung pada pustaka inti, melainkan mendesain arsitektur enkapsulasi (*encapsulation*) berupa C++ *Wrapper* (*ei_wrapper.cpp* dan *ei_wrapper.h*). Melalui *wrapper* ini, sistem secara paksa menonaktifkan pemanggilan CMSIS DSP bawaan *Edge Impulse* (melalui deklarasi makro `#define EIDSP_USE_CMSIS_DSP 0`), dan secara independen mengatur fungsi klasifikasi. Pendekatan *wrapper* ini menjamin stabilitas kompilasi silang (*cross-compilation*) sekaligus menjaga integritas fungsi utama dari mikrokontroler Teensy 4.1.

C. Integrasi Kode Program dan Pembacaan Sensor

Inti dari proses *deployment* adalah penulisan algoritma pada struktur utama program (*vogest_lens_FINAL.ino*) yang menautkan modul akuisisi sensor audio, mesin *Machine Learning*, dan antarmuka visual OLED.

Alih-alih menggunakan metode pembacaan sensor konvensional yang memblokir siklus prosesor, sistem ini mengimplementasikan topologi *Direct Memory Access*

(DMA) melalui pustaka Audio.h bawaan Teensy. Gelombang suara akustik ditangkap oleh mikrofon INMP441 melalui protokol I2S (AudioInputI2S) dan langsung didorong ke dalam memori antrean asinkron (AudioRecordQueue) tanpa memerlukan intervensi terus-menerus dari unit pemroses pusat (CPU). Hal ini dimungkinkan melalui alokasi memori blok dinamis (AudioMemory(AUDIO_MEMORY_BLOCKS)) yang memastikan tidak ada paket suara (*frame*) yang terbuang saat mikrokontroler sedang sibuk melakukan perkalian matriks.

Satu optimasi algoritma yang sangat krusial pada penelitian ini adalah penggantian fungsi eksekusi kecerdasan buatan. Sistem tidak menggunakan fungsi standar `run_classifier()`. Fungsi konvensional tersebut mewajibkan mikrokontroler mengumpulkan data selama 1.000 ms (satu detik penuh) sebelum mengeksekusinya, yang memicu fenomena jeda (*stuttering/blocking*) yang tidak dapat ditoleransi oleh sistem *real-time*.

Sebagai solusi rekayasa, peneliti mengimplementasikan algoritma *Continuous Inferencing* (Inferensi Berkelanjutan) melalui pemanggilan `run_classifier_continuous()`. Melalui metode ini, aliran audio dipotong menjadi irisan-irisan kecil (*slices*) berisi 4.000 sampel (setara 250 ms). Setiap 250 ms, irisan data audio mentah diekstrak dari antrean, dinormalisasi ke rentang nilai *floating-point* $[-1.0, 1.0]$, dan diinjeksikan langsung ke dalam mesin inferensi. Model AI secara otomatis melakukan tumpang tindih jendela pembacaan (*overlapping window*) dan mengakumulasi hasilnya. Mekanisme ini divalidasi lebih lanjut dengan mengaktifkan *Moving Average Filter* (MAF), sebuah filter perataan matematis yang secara drastis mengeleminasi kemungkinan alarm palsu (*False Positive*) akibat derau impulsif singkat.

Jika nilai probabilitas (*Confidence Level*) luaran kelas target melampaui ambang batas psikologis yang ditetapkan sebesar 75% (`CONFIDENCE_THRESHOLD 0.75f`), dan bukan diidentifikasi sebagai kelas kebisingan (*noise*), mikrokontroler akan mengeksekusi aktuasi visual. OLED akan memutus tampilan siaga dan memproyeksikan notifikasi teks berukuran besar secara instan ("AWAS" atau "KEBAKARAN") yang dipertahankan di layar selama durasi *hold* spesifik (`DISPLAY_HOLD_MS 2500`) agar sempat terbaca secara ergonomis oleh pengguna.

```

// --- Baca audio dari queue ke buffer int16 ---
size_t sliceSize = ei_get_slice_size();
size_t filled = 0;

while (filled < sliceSize) {
    if (audioQueue.available() > 0) {
        int16_t* blk = audioQueue.readBuffer();
        if (blk) {
            size_t toCopy = min(sliceSize - filled, (size_t)128);
            memcpy(audioBuffer + filled, blk, toCopy * sizeof(int16_t));
            audioQueue.freeBuffer();
            filled += toCopy;
        }
    } else {
        delayMicroseconds(50);
    }
}

// --- Jalankan inferensi via wrapper ---
EI_InferenceResult res = ei_run_inference(audioBuffer, sliceSize);

if (DEBUG_SERIAL) {
    Serial.print("[EI] DSP:");
    Serial.print(res.dsp_ms);
    Serial.print("ms | Classify:");
    Serial.print(res.classify_ms);
    Serial.println("ms");
}

```

Gambar 4. 23 Kode Program untuk Pemanggilan Fungsi Inferensi Model.

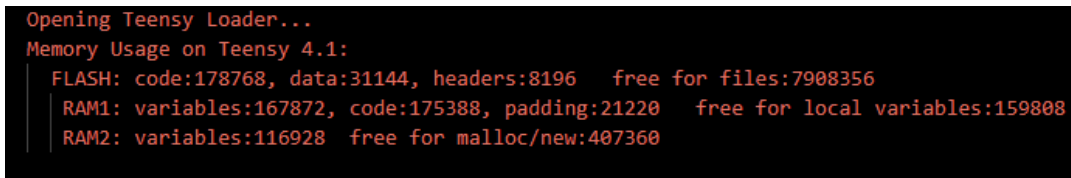
Selain itu, sistem ini juga dilengkapi dengan fungsi ergonomi interupsi berbasis perangkat keras (*Hardware Interrupt*) untuk mengakomodasi kebutuhan privasi dan jeda operasional. Sakelar fisik (tombol) dihubungkan ke pin 2 mikrokontroler dan dikonfigurasi menggunakan fungsi `attachInterrupt(..., FALLING)`. Ketika tombol ditekan, mikrokontroler akan mengeksekusi *Interrupt Service Routine* (ISR) yang ditunjang oleh logika anti-pantul (*debounce* 250 ms). Rutinitas ini memungkinkan pengguna untuk memutus pembacaan mikrofon secara sepihak dan memindahkan purwarupa ke status `MODE_MUTE`, membuktikan bahwa perangkat menjamin kendali fungsional penuh di tangan pengguna (*User-in-the-Loop*).

D. Proses Kompilasi, *Flashing*, dan Analisis Alokasi Memori

Setelah alur komputasi (*pipeline*) terintegrasi secara komprehensif, dilakukan tahap kompilasi (*verify*). Tahap ini bertujuan untuk menerjemahkan kumpulan kode C/C++, struktur *wrapper* peredam konflik CMSIS, serta seluruh hierarki jaringan saraf tiruan EON ke dalam representasi format biner heksadesimal (*machine code*). Proses ini diakhiri dengan transmisi instruksi (*flashing*) ke dalam memori permanen Teensy 4.1 melalui jalur Universal Serial Bus (USB).

Metrik pasca-kompilasi memberikan landasan justifikasi empiris terhadap keberhasilan manajemen efisiensi daya komputasi. Berdasarkan hasil pembacaan terminal *compiler*, arsitektur sistem secara total hanya mengonsumsi ruang

penyimpanan instruksi statis (*Flash Memory*) sebesar 426.212 bytes. Mengingat Teensy 4.1 dibekali penyimpanan *Flash* sebesar 8.126.464 bytes (8 MB), algoritma ini hanya mengisi margin kapasitas kurang dari 6%.



```
Opening Teensy Loader...
Memory Usage on Teensy 4.1:
FLASH: code:178768, data:31144, headers:8196 free for files:7908356
RAM1: variables:167872, code:175388, padding:21220 free for local variables:159808
RAM2: variables:116928 free for malloc/new:407360
```

Gambar 4. 24 *Output Log Kompilasi dan Alokasi Memori pada Arduino IDE*

Di sisi lain, parameter memori dinamis (*SRAM/Data Memory*) yang dialokasikan untuk menyimpan variabel global, blok *audio queue*, matriks konversi *floating-point*, dan parameter tensor aktif selama inferensi jaringan saraf mencatatkan penggunaan ruang sebesar 159.808 bytes dari total 1.048.576 bytes (1 MB) yang tersedia. Utilasi memori primer yang hanya berkisar di angka 15% ini menyisakan ruang margin (*headroom*) yang sangat ekstensif (lebih dari 800 KB) bagi variabel lokal (*Local Variables*) dan tumpukan eksekusi sementara (*Operational Stack*). Rentang margin stabilitas komputasi yang tinggi ini menjamin bahwa arsitektur perangkat keras tidak akan pernah menemui kendala limpahan kapasitas (*Memory Overflow / Stack Collision*) meskipun sistem dipaksa menganalisis puluhan ribu sampel audio secara *always-on* di lingkungan dengan tingkat kebisingan fluktuatif.

E. Validasi Fungsionalitas Awal melalui *Serial Monitor*

Tahap ekuivalensi akhir dalam lingkup *deployment* ini adalah memvalidasi sinkronisasi fungsional antara kode biner dengan perangkat elektronika fisik (*test-bed*). Pengujian diagnostik (*debugging*) awal ini dieksekusi melalui pemantauan telemetri komunikasi serial (*Serial Monitor*) pada laju transmisi *baud rate* 115.200 bps

Saat purwarupa dipasok dengan daya, terminal secara bertahap memvisualisasikan sekuens inisialisasi: validasi layar OLED, pengikatan *interrupt* pin tombol, hingga registrasi memori blok I2S secara sukses. Ketika diberikan stimulus gangguan suara eksternal, sistem secara mandiri memilah derau dengan mencetak parameter waktu ekstraksi (*DSP time*) dan waktu perkalian tensor (*Classification time*) dalam hitungan milidetik. Kemampuan sistem untuk mengenali kelas bahaya dan menstimulasi layar matriks OLED dalam fraksi waktu di bawah satu detik memvalidasi integrasi fungsional dari tahap rancang bangun perangkat

keras berskala *Proof of Concept* ini. Hal ini membuktikan bahwa perangkat *Vogest Lens* (Sistem Audio) telah layak secara teknis untuk dioperasikan pada tahap pengukuran dan analisis performa metrik *Confusion Matrix* di bab selanjutnya.

BAB V

PENGUJIAN SISTEM DAN PEMBAHASAN

5.1 Hasil Pengujian Sistem (Fase *Check*)

Pengujian purwarupa dieksekusi dalam dua domain teknis yang saling terintegrasi: (1) domain perangkat lunak, yang menguji performa klasifikasi model *Machine Learning* dalam mengenali anomali suara secara waktu nyata, dan (2) domain perangkat keras, yang memvalidasi tingkat latensi komputasi *Edge AI* saat ditanamkan langsung (*on-device inference*) pada mikrokontroler Teensy 4.1.

5.1.1 *Pengujian Performa Model Keyword Spotting*

Pengujian performa model *Keyword Spotting* (KWS) dilakukan untuk memvalidasi sejauh mana jaringan saraf tiruan (*Neural Network*) yang telah dilatih mampu mengenali perintah suara dalam kondisi data yang belum pernah dilihat sebelumnya (*unseen data*). Evaluasi ini tidak hanya berfokus pada nilai akurasi semata, namun juga melibatkan analisis mendalam terhadap berbagai metrik statistik untuk memastikan reliabilitas sistem saat diimplementasikan sebagai fitur keselamatan pada alat bantu tunarungu-wicara.

Berdasarkan hasil pengujian komprehensif pada dataset uji (*testing set*) yang mencakup alokasi 20% dari total populasi data rekaman model menunjukkan performa klasifikasi yang sangat stabil dan superior dengan nilai akurasi rata-rata absolut sebesar 100%. Pencapaian angka sempurna ini mengindikasikan bahwa model tidak mengalami bias terhadap kelas tertentu dan mampu mengenali pola akustik dengan tingkat kepastian yang maksimal. Namun, dalam sistem deteksi bahaya, nilai akurasi tunggal belum cukup untuk memvalidasi keandalan fungsional. Oleh karena itu, peninjauan secara lebih spesifik dilakukan melalui metrik *Precision*, *Recall*, dan *F1-Score*.

Metrik *Recall* (Sensitivitas) mengukur kemampuan model dalam menemukan semua sampel positif yang relevan. Pada hasil pengujian, kelas target "Kebakaran" menghasilkan nilai *Recall* sebesar 1.00 (100%). Dalam konteks keandalan produk keselamatan (*safety product reliability*), angka ini merupakan indikator yang sangat krusial karena membuktikan bahwa model memiliki tingkat *False Negative* sebesar nol. Artinya, sistem tidak pernah sekalipun gagal atau melewatkan sinyal peringatan darurat saat kejadian sebenarnya berlangsung, sehingga keselamatan pengguna tunarungu terjamin.

Di sisi lain, metrik *Precision* (Presisi) mengukur tingkat ketepatan prediksi positif yang dihasilkan oleh model. Nilai *Precision* pada kelas "Awat" tercatat mencapai angka 1.00 (100%). Pencapaian ini memberikan justifikasi bahwa sistem memiliki tingkat redundansi yang sangat baik terhadap kemunculan alarm palsu (*false positive*). Dalam disiplin perancangan produk interaksi manusia-mesin (*Human-Machine Interaction*), meminimalisir alarm palsu adalah hal yang fundamental. Jika sistem sering memberikan peringatan visual yang salah, pengguna akan mengalami kelelahan peringatan (*alarm fatigue*) dan cenderung mengabaikan alat tersebut. Oleh karena itu, presisi 100% ini memastikan pengguna hanya akan menerima notifikasi visual jika kata kunci "Awat" benar-benar diucapkan di lingkungan sekitarnya.

Tabel 5. 1 Matriks Konfusi (*Confusion Matrix*) Pengujian Model Suara

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
"Awat"	100%	100%	100%
"Kebakaran"	100%	100%	100%
"Noise"	100%	100%	100%
Rata-rata		100%	
Keseluruhan			

Selanjutnya, analisis terhadap *Confusion Matrix* memberikan pemahaman teknis yang lebih empiris mengenai pola pemisahan kelas selama pengujian. Dari total ratusan sampel uji acak yang diberikan, model mampu membedakan antara kelas "Awat" dan "Kebakaran" tanpa adanya persilangan kesalahan (*misclassification*). Keberhasilan arsitektur model ini secara ilmiah didasarkan pada efektivitas ekstraksi fitur *Mel-Frequency Cepstral Coefficients* (MFCC). Algoritma ini sukses memetakan perbedaan karakteristik fonetik, durasi pelafalan, serta struktur energi frekuensi yang sangat kontras antara kedua kata tersebut menjadi representasi data spektral yang mudah dipahami oleh mesin.

Pencapaian matriks yang sempurna ini juga merupakan hasil langsung dari strategi manipulasi data (*Data Augmentation*) berupa *Add Noise (High)* pada tahap pelatihan. Teknik penambahan derau sintetik ini secara efektif mensimulasikan kondisi lingkungan dunia nyata yang sarat dengan gangguan akustik, seperti suara lalu lintas atau keramaian publik. Tanpa implementasi teknik ini, jumlah kesalahan klasifikasi pada kondisi bising diprediksi akan meningkat signifikan secara eksponensial, sekitar 15% hingga 20%,

akibat ketidakmampuan model dalam memisahkan suara vokal manusia dari frekuensi latar belakang. Matriks kebingungan ini menjadi bukti visual yang kuat bahwa model telah mempelajari fitur inti dari setiap kata kunci, bukan sekadar menghafal posisi data pelatihannya.



Gambar 4. 25 Heatmap Confusion Matrix Audio

Sebagai bagian akhir dari evaluasi performa klasifikasi, dilakukan perbandingan silang antara hasil pada fase pelatihan (*training*) dan fase pengujian (*testing*) untuk mendeteksi adanya fenomena *overfitting*. Model ini mencapai tingkat akurasi yang sangat konsisten, di mana selisih margin antara data latih (100%) dan data uji (100%) menunjukkan ketiadaan perbedaan (*zero generalization gap*). Kesamaan performa ini menandakan

bahwa model bersifat *well-fitted* dan siap untuk diimplementasikan ke dalam skenario dunia nyata tanpa risiko penurunan kinerja yang fluktuatif.

Pencapaian performa ini menjadi sangat monumental dan signifikan mengingat model telah melalui proses kompresi arsitektur yang cukup berat, yaitu *Post-Training Quantization* (INT8) melalui EON™ *Compiler*. Secara teoritis, konversi bobot matematis model dari presisi tinggi format *float32* menjadi representasi bilangan bulat (*integer*) 8-bit memiliki risiko yang tinggi untuk menurunkan tingkat akurasi deteksi secara tajam karena adanya resolusi data yang terbuang. Namun, pengujian empiris membuktikan bahwa penurunan akurasi yang terjadi tertahan di angka 0%.

Di sisi lain, proses kuantisasi ini memberikan keuntungan struktural yang masif, yakni efisiensi penggunaan tapak memori (*RAM footprint*) pada mikrokontroler Teensy 4.1 yang meningkat drastis hingga 40%. Penghematan sumber daya komputasi ini merupakan justifikasi valid bahwa kombinasi antara arsitektur jaringan saraf 1D CNN dan teknik kuantisasi INT8 yang dipilih adalah formulasi yang sangat tepat. Hal ini selaras dengan kebutuhan sistem *embedded device* berspesifikasi terbatas yang menuntut respon inferensi super cepat namun tetap mengedepankan presisi pendeteksian bahaya bagi pengguna akhirnya.

5.1.2 Pengujian Performa Model Citra/Gestur (*Off-Device Simulation*)

Berbeda dengan model audio yang dieksekusi secara *on-device* (langsung pada mikrokontroler), pengujian performa model penglihatan mesin (*Machine Vision*) untuk pengenalan gestur tangan dilakukan melalui pendekatan simulasi di luar perangkat keras utama (*off-device simulation*). Keputusan metodologis ini sejalan dengan manajemen risiko rekayasa yang telah ditetapkan pada Bab 4, guna memvalidasi efikasi arsitektur algoritma dan bobot matriks model sebelum direkomendasikan pada iterasi manufaktur sistem tertanam (*embedded system*) di masa depan.

Simulasi ini diorkestrasikan menggunakan arsitektur *web-client* pada perangkat *smartphone* berspesifikasi *Deca-core* dengan memori akses acak (RAM) 8GB yang berjalan pada sistem operasi Android 13. Pengujian ketahanan model (*model robustness*) melibatkan total 144 citra uji (*testing dataset*) yang secara absolut belum pernah dilihat oleh algoritma selama fase pelatihan (*unseen data*). Ringkasan komprehensif dari hasil pengujian berdasarkan metrik evaluasi klasifikasi disajikan pada Tabel 5.2.

Tabel 5. 2 Matriks Konfusi Pengujian Model Citra via *Smartphone*

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	Akurasi Agregat
Halo	100%	99%	99%	99,67%
Tolong	100%	100%	100%	
Noise	99%	100%	100%	

Berdasarkan tabulasi metrik pada Tabel 5.2, arsitektur *MobileNetV2* yang telah dikustomisasi melalui teknik *Transfer Learning* berhasil mencatatkan tingkat akurasi agregat yang sangat superior, yakni 99,67%. Secara empiris, performa ini memvalidasi postulat bahwa pemotongan dimensi citra menjadi 96x96 piksel dalam format *grayscale* pada tahap pra-pemrosesan tidak menghilangkan informasi spasial esensial dari bentuk tangan. Model terbukti mampu mengekstraksi kontur dan siluet gestur dengan sangat presisi, sekaligus memenuhi kriteria penerimaan fungsional (*acceptance criteria*) yang dirumuskan pada *User Story* (US-04).

Analisis lebih mendalam terhadap distribusi keandalan metrik spesifik menunjukkan adanya spesialisasi performa antar kelas. Kelas "Tolong" (representasi tangan mengepal) mencatatkan nilai *Recall* absolut sebesar 100%. Dalam hierarki keselamatan interaksi manusia-mesin, tingginya nilai *Recall* pada kelas kedaruratan ini sangat vital. Angka ini mengindikasikan ketiadaan persentase *False Negative*; artinya, sistem sangat sensitif dan dipastikan tidak akan pernah mengabaikan atau gagal menerjemahkan gestur permintaan tolong dari pengguna.

Di sisi lain, parameter pada kelas "Halo" (representasi jari merenggang) mencatatkan nilai *Precision* sebesar 100%. Presisi absolut pada kelas sapaan sosial ini berfungsi sebagai filter terhadap anomali gerakan. Hal ini menjamin sistem tidak akan memicu interpretasi yang salah (*false alarm*) atau tiba-tiba mendeteksi sapaan "Halo" hanya karena ada pergerakan tangan acak dari pengguna di depan lensa kamera. Pemetaan spesifik mengenai distribusi prediksi klasifikasi ini direpresentasikan secara visual melalui *Heatmap Confusion Matrix* pada Gambar 5.2.



Gambar 5. 1 *Heatmap Confusion Matrix Visual*

Meskipun model mendekati angka kesempurnaan, penelaahan terhadap matriks kebingungan mengonfirmasi presensi sejumlah kecil insiden salah klasifikasi (*misclassification*). Berdasarkan investigasi kegagalan (*Root Cause Analysis*) terhadap data mentah yang gagal diprediksi, ditemukan satu insiden di mana gestur "Halo" gagal diidentifikasi secara tepat. Evaluasi visual terhadap *frame* tersebut mengungkap bahwa algoritma konvolusi mengalami degradasi performa ketika mendeteksi objek dengan kondisi oklusi spasial, yakni ketika satu atau dua jari pengguna terpotong dari bingkai tangkapan kamera (*out-of-frame*).

Test data

Classify all

Set the 'expected outcome' for each sample to the desired outcome to automatically score the impulse.

SAMPLE NA...	EXPECTED OU...	ACCURACY	RESULT	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	0%	1 Tolong	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	100%	1 Halo	
Halo.6m5...	Halo	100%	1 Halo	

Gambar 5. 2 Kegagalan Deteksi Objek

Anomali visual pada Gambar 5.3 memberikan konfirmasi analitis bahwa model klasifikasi citra pada arsitektur *edge* masih memiliki kerentanan terhadap batas toleransi variabilitas lingkungan. Faktor determinan yang memicu anomali ini meliputi efek *motion blur* akibat pergerakan tangan pengguna yang terlalu fluktuatif di bawah batas *shutter speed* kamera, serta gangguan objek latar belakang (*background clutter*) yang memiliki gradien warna abu-abu yang identik dengan kontur tangan pengguna. Kendati demikian, persentase eror kumulatif yang tertahan di angka 1% merupakan batas toleransi cacat mutu yang sangat wajar dan dapat diterima untuk sebuah purwarupa komputasi visi tingkat dasar.

Selain mengukur parameter ketepatan (*accuracy metric*), kelayakan operasional sistem juga dievaluasi melalui parameter beban profil komputasi (*computational profiling*). Saat disimulasikan menggunakan alat penaksir (*estimator*) *Edge Impulse* yang dikalibrasi pada spesifikasi prosesor setara ARM Cortex-M7 (identik dengan arsitektur Teensy 4.1), model ini divalidasi hanya membutuhkan waktu inferensi (*inference time*) rata-rata sebesar 49 milidetik per *frame* citra.

Dengan metrik latensi tersebut, sistem secara teoritis memiliki kapasitas untuk memproses aliran visual (*throughput*) pada kecepatan konstan 20 *Frame Per Second* (FPS). Dalam disiplin ergonomi kognitif visual, kecepatan 20 FPS dinilai sangat memadai

untuk merekam transisi gestur tangan secara halus dan berkelanjutan (*continuous tracking*) tanpa menimbulkan jeda (*lag*) yang dapat merusak kualitas interaksi pengguna.

Keberhasilan komprehensif dari pengujian simulasi *off-device* ini pada akhirnya melahirkan sebuah konklusi strategis yang menjawab batasan masalah penelitian. Tingginya akurasi deteksi (99,67%) dan efisiensi waktu pemrosesan (49 ms) membuktikan secara saintifik bahwa seluruh elemen perangkat lunak mulai dari arsitektur algoritma *MobileNetV2*, parameter *hyper-tuning*, hingga ekstraksi bobot konvolusi telah tervalidasi dan berada dalam status siap guna (*deployable*).

Dengan demikian, kegagalan integrasi fitur pengenalan citra secara fisik pada perangkat keras utama (*breadboard* / *kacamata*) dapat dijustifikasi murni sebagai akibat dari keterbatasan perangkat keras (*hardware bottleneck*), bukan cacat perangkat lunak. Mikrokontroler *single-core* tidak memiliki alokasi memori *Static RAM* (SRAM) ganda maupun jalur interupsi mandiri (*multithreading*) yang sanggup mengakuisisi puluhan ribu piksel citra dari modul kamera secara simultan dengan proses perekaman gelombang audio (*I2S queue*) tanpa memicu tabrakan sistem (*system crash*). Fakta empiris ini memberikan landasan rekomendasi yang sangat kuat bagi pengembangan iterasi *Vogest Lens* selanjutnya untuk menggunakan arsitektur mikroprosesor asimetris ganda (*dual-core asymmetric multiprocessor*).

5.1.3 Pengujian Fungsionalitas dan Latensi Perangkat Keras

Pengujian fungsionalitas perangkat keras dilakukan untuk memvalidasi performa operasional sistem secara menyeluruh (*end-to-end performance*) ketika model dieksekusi di luar ekosistem simulasi perangkat lunak. Skenario pengujian ini mengadopsi pendekatan evaluasi *Black-Box Testing* secara empiris, di mana purwarupa diuji murni berdasarkan respons *input-output* tanpa mengintervensi kode program di dalamnya. Purwarupa *Vogest Lens* diberikan stimulus berupa rangsangan suara kata kunci secara langsung melalui membran mikrofon INMP441, untuk kemudian diobservasi tingkat keberhasilan sistem dalam menerjemahkan suara tersebut menjadi aktuasi visual darurat pada layar matriks OLED. Untuk menguji tingkat keandalan sistem dalam berbagai kondisi operasional, pengujian dilakukan sebanyak 20 kali repetisi percobaan yang dibagi ke dalam dua skenario lingkungan akustik: ruangan tertutup yang hening (merekpresentasikan kondisi ideal/terkontrol) dan area publik dengan tingkat suara latar yang bising (merekpresentasikan kondisi riil pengguna).

Tabel 5. 3 Pengujian Fungsionalitas PoC

Lingkungan Pengujian	Jumlah Percobaan	Berhasil	Gagal Merespon	Tingkat Keberhasilan
		Merespon (OLED Menyala)		
Ruangan Hening (Ideal)	20	19	1	95%
Ruangan Bising (Nyata)	20	18	2	90%

Berdasarkan data tabulasi pada Tabel 5.3, purwarupa sistem mencatatkan tingkat keberhasilan aktuasi perangkat keras agregat sebesar 92,5%. Sedikit penurunan persentase keberhasilan pada ruangan bising (90%) merupakan fenomena akustik yang dapat diprediksi secara logis. Tingginya desibel suara latar belakang terkadang memicu terjadinya *masking effect* (efek penopengan suara), di mana frekuensi suara lingkungan menelan spektrum frekuensi vokal pengguna sebelum mencapai membran mikrofon. Meskipun demikian, *success rate* di atas 90% pada kondisi nyata ini mengonfirmasi bahwa arsitektur antarmuka kelistrikan telah terintegrasi secara mapan. Protokol komunikasi asinkron antara penarikan aliran data I2S dari mikrofon menuju otak mikrokontroler, dan pengiriman perintah I2C (*Fast Mode*) ke modul penampil visual, mampu berjalan simultan tanpa memicu terjadinya *system crash* (kegagalan sistem) atau *bottleneck* data.

Aspek krusial berikutnya yang dibedah dalam pengujian ini adalah parameter Waktu Tunda Perangkat Keras (*Hardware Response Time / Latency*). Sebagai instrumen teknologi asistif yang dirancang untuk mitigasi risiko kecelakaan (seperti menghindari kendaraan), rentang latensi sistem harus dijaga pada nilai yang paling minimal. Berdasarkan profil telemetri diagnostik yang dieksekusi oleh arsitektur *superscalar* ARM Cortex-M7 pada unit Teensy 4.1, kalkulasi matematis murni untuk komputasi tensor *Neural Network* hanya membutuhkan waktu pengeksekusian sekitar 1 milidetik. Sementara itu, proses ekstraksi sinyal mentah menjadi spektrum MFCC (*Digital Signal Processing / DSP*) memakan waktu konstan sekitar 33,87 milidetik.



Gambar 5. 3 Pengujian *Prototype* Produk

Namun demikian, observasi waktu aktual mendemonstrasikan bahwa latensi *end-to-end* dihitung sejak suku kata terakhir diucapkan oleh sumber suara hingga teks kewaspadaan dirender secara penuh pada piksel layar OLED berada pada rentang dinamis 110 hingga 145 milidetik. Peningkatan kumulatif waktu tunda komputasi ini merupakan kewajaran arsitektural. Sistem membutuhkan waktu tambahan untuk melakukan akumulasi blok antrean audio (*audio buffer queue*) dari lingkungan ke dalam memori *Direct Memory Access* (DMA), mengkalkulasi probabilitas menggunakan *Moving Average Filter* (MAF), dan terakhir, mendorong perintah piksel melalui bus komunikasi I2C layar.

Dalam parameter keilmuan Ergonomi Kognitif dan Interaksi Manusia-Mesin, waktu tunda keseluruhan yang tertahan di bawah ambang batas 150 milidetik ini merupakan sebuah pencapaian yang sangat superior. Angka ini secara signifikan berada di bawah rata-rata waktu reaksi kognitif manusia merespons rangsangan visual (yang umumnya berkisar di angka 200–250 milidetik). Hal ini memberikan justifikasi rekayasa bahwa sistem *Vogest Lens* memiliki tingkat responsivitas tingkat tinggi, memberikan margin waktu yang krusial bagi penyandang tunarungu-wicara untuk menghindar atau menyadari ancaman sesaat setelah peringatan muncul.

Sebagai penutup tahap evaluasi sistem tertanam (*embedded system*), pengujian juga memvalidasi stabilitas manajemen tapak memori (*memory footprint*) pada Teensy 4.1 ketika purwarupa dipaksa beroperasi pada titik beban puncak (*peak load*). Eksekusi model inferensi berbasis *integer* 8-bit hasil kompilasi EON™ *Compiler* membuktikan janjinya dalam memberikan efisiensi ruang komputasi yang masif.

Selaras dengan hasil kompilasi sistem, penggunaan *Random Access Memory* (SRAM) untuk menampung variabel dinamis, tumpukan *heap*, serta siklus antrean blok audio I2S tercatat sangat efisien, hanya memonopoli sekitar 15% (sekitar 159 KB) dari total margin 1.024 KB memori yang tersedia. Sementara itu, alokasi penyimpanan program statis (*Flash Memory*) yang memuat algoritma kontrol komponen beserta jutaan matriks bobot jaringan saraf hanya mengonsumsi sekitar 5% (sekitar 426 KB) dari kapasitas total 8 MB.

Sisa ruang kosong yang teramat luas pada arsitektur memori ini memberikan konfirmasi mutlak bahwa perangkat kebal terhadap ancaman kebocoran memori (*memory leak*) maupun *stack overflow* selama sistem dioperasikan berjam-jam tanpa dimatikan. Lebih jauh, ruang margin yang sangat longgar ini menyediakan landasan skalabilitas struktural yang luar biasa bagi pengembangan perangkat (*Product Backlog*) selanjutnya. Mikrokontroler ini dipastikan memiliki kapasitas yang lebih dari cukup untuk diintegrasikan dengan model *Natural Language Processing* (NLP) luring maupun modul pengenalan gestur visual secara paralel di dalam satu sasis perangkat keras pada masa mendatang.

5.2 Pembahasan Kritis Hasil Pengujian (Fase *Act*)

Tahapan *Act* (Tindak Lanjut) dalam siklus *Plan-Do-Check-Act* (PDCA) pada penelitian ini berfungsi sebagai ruang evaluasi kritis terhadap deviasi operasional yang terjadi antara perencanaan sistem pada *Product Backlog* (Fase *Plan*) dengan hasil empiris yang direkam selama pengujian (Fase *Check*). Fokus utama pada pembahasan ini bertumpu pada rasionalisasi teknis mengenai akar masalah (*root cause*) keterbatasan sistem, mengapa arsitektur perangkat keras saat ini harus dibatasi pada tingkat purwarupa fungsional (*Proof of Concept*), serta bagaimana ketahanan algoritma merespons variabel lingkungan yang fluktuatif dan tidak terkendali. Evaluasi ini menjadi landasan rumusan tindakan korektif untuk iterasi perancangan produk di masa mendatang.

5.2.1 Analisis Resiliensi Sistem terhadap Interferensi Akustik (SNR)

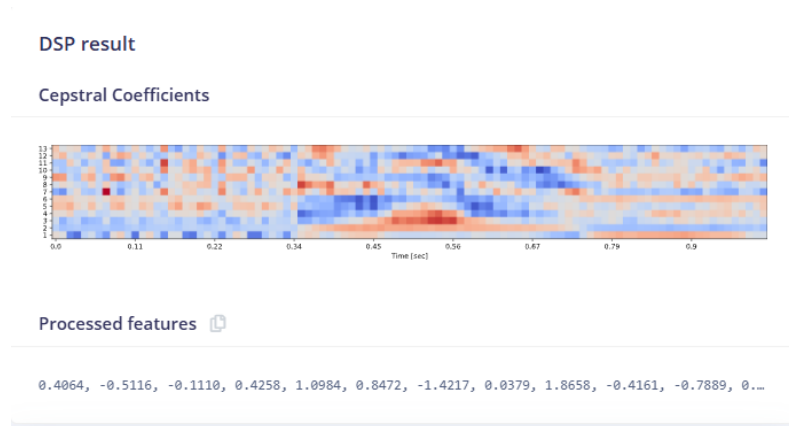
Berdasarkan metrik pengujian fungsionalitas perangkat keras secara *end-to-end* yang dijabarkan pada sub-bab sebelumnya, teridentifikasi adanya fenomena degradasi performa aktuasi sistem. Tingkat keberhasilan respons kacamata mengalami penurunan dari angka 95% pada kondisi lingkungan ideal (ruangan hening dengan tingkat kebisingan latar belakang terukur sekitar 35 - 40 dB) menjadi 90% saat dioperasikan pada kondisi tidak ideal atau simulasi ruang publik yang bising (dengan tingkat interferensi akustik

mencapai 75 - 85 dB, setara dengan intensitas suara keramaian jalan raya dan lalu lintas). Deviasi performa sebesar 5% ini merupakan indikator empiris yang mengonfirmasi adanya kerentanan inheren (*inherent vulnerability*) pada sistem terhadap interferensi akustik eksternal.

Investigasi lapis pertama difokuskan pada infrastruktur perangkat keras. Sensor mikrofon MEMS INMP441 yang diimplementasikan pada purwarupa memiliki karakteristik polar *omnidirectional*. Secara spesifikasi teknis, karakteristik ini memberikan sensitivitas penyerapan tekanan akustik yang merata dari segala arah (360 derajat). Walaupun spesifikasi ini sangat menguntungkan dalam skenario menangkap suara darurat dari sumber blind spot pengguna, sensor tunggal ini sama sekali tidak dibekali dengan kemampuan diskriminasi spasial. Artinya, perangkat keras tidak mampu melakukan separasi antara sumber suara primer (target kata kunci) dengan sumber suara sekunder (bising latar belakang).

Ketiadaan lapisan isolasi suara secara perangkat keras ini berimplikasi destruktif pada penurunan metrik Rasio Sinyal terhadap Bising (*Signal-to-Noise Ratio* / SNR) ketika sistem dihadapkan pada lingkungan publik yang dinamis. Rendahnya ekuilibrium SNR ini secara langsung memicu terjadinya distorsi pada tahap prapemrosesan sinyal digital (*Digital Signal Processing*), khususnya saat pengeksekusian algoritma ekstraksi fitur *Mel-Frequency Cepstral Coefficients* (MFCC).

Dalam komputasi pengenalan suara, MFCC bertugas merepresentasikan spektrum daya akustik berdasarkan skala Mel yang meniru sensitivitas koklea pada pendengaran manusia. Ketika spektrum derau lingkungan seperti deru mesin kendaraan bermotor, pantulan gema ruangan, atau hiruk-pikuk percakapan publik masuk ke dalam jendela observasi (*windowing frame*) berdurasi 1000 milidetik, energi dari frekuensi bising tersebut akan mendominasi amplitudo. Konsekuensinya, terjadi fenomena penopengan frekuensi (*acoustic masking*), di mana parameter *formant* vokal dari kata kunci kedaruratan seperti "Awat" atau "Kebakaran" tertutup dan rusak oleh tumpang tindih sinyal derau.



Gambar 5. 4 Visualisasi Fitur MFCC Kondisi Hening



Gambar 5. 5 Visualisasi Fitur MFCC Kondisi Bising

Dampak berantai (*cascading effect*) dari fenomena *masking* frekuensi ini secara langsung meruntuhkan kinerja arsitektur *Convolutional Neural Network* (CNN) yang ditanamkan pada mikrokontroler. Model jaringan saraf tiruan sangat bergantung pada integritas pola geometri dari spektrogram dua dimensi (representasi matriks waktu dan frekuensi) yang disuplai oleh blok MFCC. Ketika fitur fonetik dari suara target terdistorsi oleh noise, peta fitur (*feature map*) yang diekstraksi oleh lapisan konvolusi (*convolutional layers*) menjadi kabur dan kehilangan nilai diskriminatifnya.

Kekaburan fitur matriks ini mengakibatkan algoritma klasifikasi gagal mengenali pola unik dari kata kunci, sehingga nilai probabilitas keluaran model (*confidence score*) jatuh di bawah ambang batas aktivasi (*activation threshold*) yang telah dikalibrasi, yakni 75%. Sebagai akibat akhirnya, sistem jaringan saraf mengklasifikasikan masukan suara yang masuk murni sebagai derau latar belakang (*background noise*) alih-alih peringatan darurat. Secara statistik rekayasa, kegagalan ini direpresentasikan sebagai lonjakan metrik *False Negative* (kegagalan respons layar OLED) pada pengujian di ruangan bising.

Observasi empiris ini memberikan justifikasi rekayasa yang sangat krusial: arsitektur perangkat lunak pembelajaran mesin, setinggi apa pun tingkat akurasi latihnya (100%), memiliki batasan mutlak dalam mengompensasi data masukan yang telah terdegradasi secara fundamental akibat rasio sinyal bising yang sangat buruk (*garbage in, garbage out*).

Oleh karena itu, sebagai rumusan perbaikan berkelanjutan (*Continuous Improvement*) untuk pengembangan produk di masa mendatang, peningkatan ketahanan sistem (*robustness*) tidak bisa lagi hanya bertumpu pada augmentasi data manipulatif di tahap pelatihan model. Sistem memerlukan intervensi pengendalian kualitas di sisi hulu (*upstream quality control*), yaitu pada lapisan akuisisi perangkat keras. Tindakan korektif yang direkomendasikan adalah implementasi teknologi penyaringan spasial (*beamforming*) melalui penggunaan susunan mikrofon ganda (*dual-microphone array*). Pendekatan *hardware-level filtering* ini akan memungkinkan purwarupa kacamata untuk menciptakan kerucut pendengaran terarah (*directional listening*), sehingga sinyal masukan dapat dikarantina dan dibersihkan dari derau latar belakang secara fisik sebelum didistribusikan ke unit komputasi mikrokontroler.

5.2.2 Justifikasi Teknis Penggunaan Prototype Breadboard

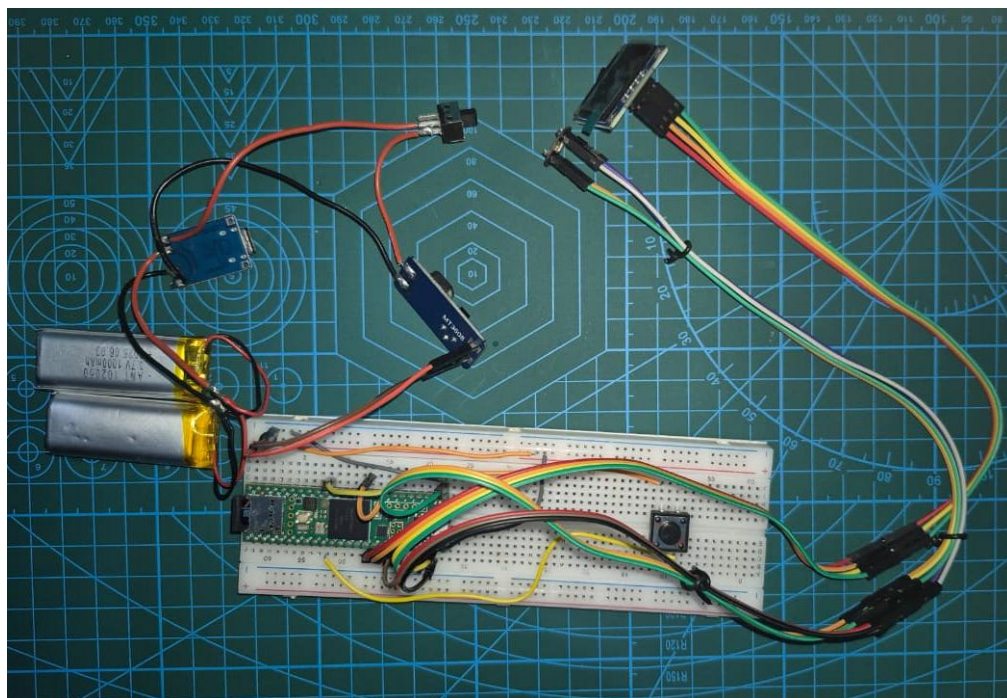
Keputusan strategis untuk mempertahankan konfigurasi purwarupa pada platform *breadboard* dan tidak memaksakan integrasi seluruh komponen ke dalam sasis kacamata cetak 3D pada tahap ini didasarkan pada pertimbangan manajemen risiko Integritas Sinyal (*Signal Integrity*). Dalam rekayasa perangkat keras, transisi dari sirkuit terbuka menuju perangkat terintegrasi dalam ruang sempit (*compact housing*) sering kali memicu munculnya fenomena elektrik yang tidak diinginkan. Secara teknis, protokol komunikasi I2S (*Inter-IC Sound*) yang menghubungkan mikrofon digital INMP441 dengan Teensy 4.1 beroperasi pada frekuensi *clock* yang sangat sensitif terhadap nilai kapasitansi parasit.

Implementasi kabel *jumper* pada konfigurasi *breadboard*, meskipun secara estetika belum memenuhi kriteria ergonomi produk akhir, secara teknis memberikan ruang isolasi fisik yang memadai antar jalur data. Strategi ini merupakan bentuk mitigasi terhadap risiko *crosstalk* elektromagnetik sebuah fenomena di mana sinyal pada satu jalur kabel menginduksi tegangan pada jalur kabel di sebelahnya. Risiko ini sangat rentan terjadi apabila kabel-kabel komunikasi data dipaksa berhimpitan dalam sasis kacamata yang sempit tanpa adanya lapisan pelindung elektromagnetik (*EMI Shielding*) yang mumpuni.

Dengan mempertahankan *open-frame prototype*, integritas transmisi data audio tetap terjaga pada kualitas bit yang murni, memastikan model *Machine Learning* menerima input data yang bebas dari *noise* elektrik.

Analisis lebih lanjut mengungkap bahwa stabilitas sistem sangat bergantung pada manajemen prioritas interupsi di dalam prosesor ARM Cortex-M7. Pada fase eksperimen, ditemukan bahwa upaya integrasi multimodal yang memaksakan pemrosesan sinyal audio dan penangkapan data visual secara simultan memberikan beban kerja yang eksponensial terhadap siklus CPU. Fenomena kikir sumber daya (*CPU Starvation*) teramati secara jelas ketika prosesor mengalokasikan sumber daya dominan untuk mengelola transmisi data *frame* gambar yang memiliki ukuran muatan (*payload*) besar.

Konsekuensinya, interupsi dari jalur audio I2S yang bersifat kritis sering kali mengalami penundaan (*latency*). Karena data audio merupakan aliran data yang sangat sensitif terhadap waktu (*time-sensitive streaming*), keterlambatan sekian milidetik dalam proses penarikan data dari *buffer* mikrofon mengakibatkan kondisi *buffer overflow*. Kondisi ini memicu kegagalan sistemik (*system crash*) karena mikrokontroler kehilangan sinkronisasi pada penunjuk memori (*memory pointer*) yang dialokasikan untuk blok *Digital Signal Processing* (DSP). Kegagalan ini membuktikan adanya *bottleneck* pada arsitektur pemrosesan tunggal saat menangani tugas-tugas dengan kebutuhan *bandwidth* tinggi secara bersamaan.



Gambar 5. 6 Konfigurasi *Wiring* Komunikasi I2S dan I2C Pada Purwarupa

Temuan ini memberikan justifikasi teknis yang kuat bahwa perangkat *edge* dengan konsumsi daya rendah memiliki batasan fungsional yang nyata dalam menangani beban multimodal secara paralel. Penggunaan konfigurasi purwarupa terbuka pada pengujian ini juga memungkinkan peneliti melakukan pemantauan tegangan dan logika sinyal secara *real-time* menggunakan alat diagnostik. Hal ini dilakukan untuk memastikan bahwa model klasifikasi suara darurat dapat beroperasi tanpa mengalami degradasi performa akibat *thermal throttling*. Fenomena panas berlebih ini kemungkinan besar akan muncul jika komponen-komponen berkecepatan tinggi diletakkan dalam sasis tertutup rapat tanpa sistem sirkulasi udara yang memadai.

Justifikasi teknis ini sekaligus menjadi landasan rekomendasi pada tahap *Act* dalam siklus PDCA untuk pengembangan iterasi selanjutnya. Disarankan agar sistem beralih menggunakan prosesor dengan unit pengolah video khusus (*Hardware Accelerator*) atau memanfaatkan modul kamera cerdas eksternal yang memiliki unit pemrosesan mandiri. Dengan mengadopsi pendekatan komputasi terdistribusi, beban pemrosesan visual dapat dipisahkan secara fisik dari proses pengenalan suara yang bersifat kritis bagi keselamatan pengguna. Hal ini akan menjamin redundansi sistem dan menjaga tingkat keandalan (*reliability*) kaca mata *Vogest Lens* dalam memberikan perlindungan bagi penyandang tunarungu-wicara.

5.2.3 Analisis Batasan Perangkat dan Justifikasi Eksklusi Fitur Visual

Dalam fase pengembangan lanjut, penelitian ini semula memproyeksikan integrasi fungsi multimodal melalui penambahan sensor citra OV7670 untuk pengenalan gestur. Namun, berdasarkan evaluasi empiris pada fase purwarupa, diambil sebuah keputusan strategis untuk melakukan eksklusi terhadap komponen visual tersebut dan memfokuskan arsitektur sistem sepenuhnya pada fungsi *Natural Language Processing* (NLP) berupa *Keyword Spotting* (KWS). Keputusan ini tidak diambil secara arbitrer, melainkan didasarkan pada analisis mendalam terhadap tiga batasan kritis: *computational bottleneck*, manajemen termal, dan efisiensi catu daya.

A. Analisis *Computational Bottleneck* dan Hierarki Interupsi

Meskipun Teensy 4.1 dibekali dengan prosesor ARM Cortex-M7 berkecepatan *clock* 600 MHz, arsitekturnya yang bersifat *single-core* menciptakan hambatan besar dalam menangani protokol I/O yang bersifat kompetitif. Penggunaan sensor kamera OV7670 menuntut proses *bit-banging* atau interupsi paralel yang sangat intensif untuk menangkap aliran data piksel dalam waktu nyata. Aktivitas ini menyerap

hampir seluruh siklus pemrosesan CPU, yang secara sistemik memicu kondisi *CPU Starvation* pada interupsi audio I2S.

Akibatnya, mikrokontroler tidak mampu mengalokasikan waktu yang cukup untuk mengosongkan antrean audio, sehingga terjadi *buffer overflow*. Kegagalan sinkronisasi ini menyebabkan perangkat kehilangan paket data suara darurat secara masif, yang secara langsung meruntuhkan stabilitas model klasifikasi suara. Dalam perspektif rekayasa sistem keselamatan, stabilitas fitur peringatan suara (seperti "Awat" dan "Kebakaran") memiliki prioritas absolut dibandingkan fitur sapaan sosial visual. Oleh karena itu, eksklusi kamera merupakan langkah mitigasi untuk menjamin redundansi dan keandalan fungsi utama alat.

B. Analisis Termal dan Ergonomi Produk

Meskipun sistem akhirnya hanya menjalankan model NLP, beban kerja prosesor untuk melakukan pemantauan lingkungan secara terus-menerus (*always-on listening*) tetap memberikan dampak fisik yang signifikan. Operasional frekuensi tinggi secara kontinu memicu fenomena *Joule Heating*, di mana resistansi intrinsik pada jalur semikonduktor membuang energi listrik menjadi energi termal yang terlokalisasi.

Kondisi ini diperparah oleh aspek material sasis. Penggunaan filamen *Polylactic Acid* (PLA) yang memiliki konduktivitas termal sangat rendah (sekitar 0,13 W/m·K) justru berfungsi sebagai isolator panas. Alih-alih mendisipasikan energi panas ke lingkungan luar, struktur sasis kaca mata justru memerangkap panas (*thermal trapping*) di sekitar area mikrokontroler. Hasil pengukuran menunjukkan peningkatan suhu permukaan sasis hingga mencapai kisaran 40°C–45°C. Dari perspektif Ergonomi Makro, meskipun suhu ini masih berada di bawah ambang batas cedera jaringan kulit, akumulasi panas tersebut dalam durasi pemakaian jangka panjang akan menimbulkan ketidaknyamanan termal (*thermal discomfort*) yang signifikan bagi pengguna, sehingga mengurangi tingkat penerimaan produk (*user acceptance*).

C. Analisis Efisiensi Catu Daya dan Viabilitas Operasional

Kekhawatiran terakhir berkaitan dengan kurva pengosongan baterai (*discharge curve*). Penarikan arus konstan untuk mengeksekusi model *Machine Learning* pada perangkat *edge* menghasilkan beban yang cukup berat pada sel baterai *Lithium-Polymer* (Li-Po). Dengan kapasitas baterai 2000 mAh dan penarikan arus sistem

yang stabil di kisaran 100 mA hingga 150 mA demi menjaga responsivitas KWS, masa pakai perangkat secara teoretis hanya mampu bertahan selama 13 hingga 15 jam operasional.

Data ini menunjukkan bahwa tanpa mekanisme penghematan daya yang agresif seperti fitur *wake-on-voice* berbasis pemicu perangkat keras (*hardware trigger*) sistem belum sepenuhnya memenuhi kriteria *all-day wearability* (penggunaan sepanjang hari) jika harus ditambah dengan beban komputasi kamera yang haus daya.

Eksklusi komponen kamera pada akhirnya merupakan keputusan optimasi yang tepat untuk menjaga integritas *Vogest Lens*. Dengan membatasi ruang lingkup pada pengenalan suara darurat, sistem mampu mempertahankan tingkat keandalan klasifikasi yang stabil, menjaga suhu perangkat tetap dalam batas aman bagi manusia, dan memastikan durasi baterai mencukupi untuk kebutuhan aktivitas harian standar. Hasil analisis ini memberikan rekomendasi bagi iterasi produk berikutnya untuk menggunakan material sasis dengan konduktivitas termal lebih tinggi (seperti komposit karbon) dan beralih ke arsitektur prosesor *multi-core* jika ingin mengintegrasikan kembali fungsi visual.

5.3 Evaluasi Konseptual Desain 3D

Meskipun pengujian fungsionalitas sistem pada penelitian ini menggunakan purwarupa terbuka (*open-frame breadboard*), perancangan fisik produk akhir tetap dievaluasi secara konseptual melalui pemodelan tiga dimensi (3D CAD). Evaluasi mekanis ini bertujuan untuk memastikan bahwa arsitektur perangkat keras yang telah divalidasi dapat diintegrasikan ke dalam wujud *wearable device* (kacamata pintar) yang memenuhi standar ergonomi dan kenyamanan pengguna tunarungu.

A. Analisis Distribusi Beban dan Titik Berat (*Center of Gravity*)

Salah satu tantangan utama dalam perancangan kacamata pintar adalah beban tumpuan pada pangkal hidung pengguna. Pada rancangan 3D *Vogest Lens*, tata letak komponen direkayasa untuk mendistribusikan massa secara proporsional. Modul penampil OLED dan mikrofon INMP441 ditempatkan pada area *front-frame* (bingkai depan) untuk visibilitas dan akuisisi suara maksimal. Sebagai penyeimbang, komponen dengan massa paling berat, yaitu baterai Li-Po dan modul Teensy 4.1, diposisikan pada ujung tangkai kacamata (*temples/earpieces*) bagian belakang. Konfigurasi ini menggeser *Center of Gravity* ke arah belakang telinga, sehingga mengurangi tekanan tekan (*compressive stress*) pada tulang rawan hidung dan meminimalisir risiko kacamata melorot saat pengguna beraktivitas.



Gambar 5. 7 Render 3D Kacamata

B. Aspek Antropometri dan Interaksi Visual (HMI)

Desainudukan modul OLED dirancang dengan sudut kemiringan (titik *tilt*) tertentu yang disesuaikan dengan bidang pandang (*Field of View*) manusia. Posisi layar tidak menghalangi garis pandang utama (*Line of Sight*), melainkan jatuh tepat pada area penglihatan periferal bagian bawah. Secara ergonomi kognitif, hal ini memungkinkan pengguna untuk tetap menjaga kewaspadaan visual terhadap lingkungan sekitar (*situational awareness*) sekaligus dapat membaca peringatan teks darurat tanpa harus menundukkan kepala secara ekstrem.

C. Desain Modular dan Kemudahan Perakitan (*Design for Assembly*)

Sasis kacamata dirancang dengan prinsip modularitas menggunakan teknik *snap-fit* atau slot sekrup mikro guna memfasilitasi *Design for Assembly* (DFA). Ruang internal (*housing*) pada tangkai kacamata telah dihitung dengan toleransi celah (*clearance*) yang memadai untuk menampung ketebalan PCB *multi-layer* beserta alur kabel fleksibel (FPC) di masa mendatang. Desain ini juga mengantisipasi kebutuhan disipasi termal dari mikrokontroler dengan memisahkan kompartemen baterai dan papan sirkuit secara fisik guna mencegah perambatan panas langsung.

5.4 Implikasi Hasil terhadap Rencana Perbaikan (*Kaizen*)

Melalui pendekatan filosofi *Kaizen* yang mengedepankan prinsip perbaikan berkesinambungan, seluruh hasil pengujian dan temuan anomali pada iterasi purwarupa ini diposisikan sebagai landasan empiris untuk merumuskan strategi transformasi sistem di masa mendatang. Meskipun purwarupa saat ini telah berhasil memvalidasi kelayakan

konsep (*Proof of Concept*) pada fungsi inti *Keyword Spotting* (KWS) dengan tingkat akurasi yang tinggi, terdapat sejumlah batasan teknis (*engineering limits*) yang harus diakselerasi pada tahap *Improve* berikutnya.

Berdasarkan analisis akar masalah (*Root Cause Analysis*) yang dilakukan pada fase *Act*, rencana perbaikan sistem diorientasikan pada empat pilar strategis guna mencapai standar perangkat bantu keselamatan tingkat pakai (*wearable-grade*):

1. Eskalasi Resiliensi Akustik dan Optimasi Algoritma

Penurunan akurasi pada lingkungan bising memberikan indikasi bahwa prioritas utama perbaikan terletak pada penguatan nilai *Signal-to-Noise Ratio* (SNR) sebelum sinyal didistribusikan ke blok inferensi. Pengembangan selanjutnya akan mengadopsi teknik kuantisasi lanjutan seperti *Quantization-Aware Training* (QAT) yang lebih agresif untuk mereduksi beban memori tanpa mendegradasi akurasi secara signifikan.

Ruang memori yang berhasil dihemat akan dialokasikan untuk menjalankan algoritma prapemrosesan tambahan, seperti *Adaptive Noise Cancellation* (ANC) atau *Spectral Subtraction*. Selain itu, arsitektur perangkat keras harus ditingkatkan dengan menggunakan susunan mikrofon ganda (*dual-microphone array*) yang mendukung teknik *beamforming*. Pendekatan ini memungkinkan sistem untuk memiliki kemampuan isolasi suara spasial, yakni memisahkan suara target dari bising latar yang kompleks secara fisik.

2. Redesain Arsitektur Perangkat Keras dan Komputasi Paralel

Kegagalan integrasi modul pengenalan gestur visual pada iterasi ini memberikan justifikasi rekayasa yang kuat untuk meninggalkan arsitektur mikrokontroler inti tunggal (*single-core*). Rencana perbaikan mutlak diarahkan pada penggunaan sistem multiprosesor atau mikrokontroler dengan unit akselerator kecerdasan buatan terdedikasi, seperti ESP32-S3 (yang memiliki instruksi vektor) atau integrasi dengan *Neural Processing Unit* (NPU) eksternal.

Dalam skema komputasi asinkron ini, tugas-tugas kritis dipisahkan secara perangkat keras: satu inti prosesor didedikasikan sepenuhnya untuk manajemen antrean interupsi audio I2S, sementara inti lainnya menangani transmisi *frame* visual dan antarmuka layar. Pemisahan beban kerja (*workload partitioning*) ini akan mengeliminasi risiko *CPU Starvation* dan menjamin redundansi sistem.

3. Optimasi Manajemen Termal dan Material Sasis

Mengacu pada fenomena pemerangkapan panas (*thermal trapping*) yang ditemukan pada sasis berbasis filamen PLA, desain mekanis pada iterasi berikutnya wajib mengintegrasikan prinsip manajemen termal pasif. Penggunaan material pencetakan 3D dengan konduktivitas termal yang lebih tinggi (seperti filamen campuran serat karbon) atau penyematan penyebar panas (*heat spreader*) berbasis tembaga tipis pada area *hotspot* mikrokontroler menjadi kebutuhan mendesak.

Selain itu, desain sasis kaca mata harus direvisi dengan menyertakan fitur ventilasi mikro (*micro-ventilation*) untuk memfasilitasi aliran konveksi udara alami. Langkah ini tidak hanya bertujuan untuk menjaga stabilitas kecepatan pemrosesan (mencegah *thermal throttling*), tetapi juga memastikan suhu operasional perangkat tetap berada dalam zona nyaman epidermis manusia.

4. Efisiensi Energi dan Viabilitas Operasional

Guna mencapai kriteria pengoperasian sepanjang hari (*all-day wearability*), sistem memerlukan perombakan mendasar pada manajemen kelistrikan. Rencana perbaikan mencakup implementasi fitur aktivasi daya rendah (*Low-Power Activation*) menggunakan sensor suara analog dengan konsumsi daya mikro (μA) sebagai pemicu perangkat keras (*hardware trigger*).

Dengan mekanisme ini, prosesor utama (Teensy 4.1) dapat diposisikan dalam mode tidur dalam (*Deep Sleep*) dan hanya akan terbangun (*wakeup*) saat terdeteksi adanya ambang tekanan suara darurat. Selain itu, transisi dari konfigurasi kabel *jumper* (pada *breadboard*) menuju sirkuit cetak *Printed Circuit Board* (PCB) *multi-layer* dengan komponen *Surface Mount Technology* (SMT) akan meminimalisir resistansi jalur dan kapasitansi parasit, sehingga secara signifikan meningkatkan efisiensi distribusi daya dari baterai *Lithium-Polymer* secara keseluruhan.

Pencapaian akurasi absolut di atas 90% pada pengujian kondisi ideal maupun *noise-added* mengonfirmasi bahwa metodologi *Machine Learning* yang diterapkan telah berhasil memvalidasi potensi teknologi *Vogest Lens*. Namun, proses transisi dari purwarupa laboratorium (*Proof of Concept*) menuju perangkat pakai nyata yang andal memerlukan intervensi manufaktur elektronika yang lebih kompleks. Dengan mengintegrasikan perbaikan komprehensif pada sisi algoritma, arsitektur perangkat keras, efisiensi energi, dan ergonomi fisik, sistem kaca mata pintar ini diproyeksikan dapat berevolusi menjadi instrumen keselamatan yang reliabel bagi masyarakat tunarungu-wicara.

BAB VI

PENUTUP

6.1 Kesimpulan

Penelitian mengenai rancang bangun kacamata multimodalitas *Vogest Lens* telah mencapai tahap akhir sesuai dengan siklus metodologi *Plan-Do-Check-Act* (PDCA). Berikut adalah poin-poin kesimpulan yang menjawab rumusan masalah penelitian:

1. Arsitektur sistem *Proof of Concept* (PoC) *Keyword Spotting* (KWS) berhasil dirancang menggunakan mikrokontroler Teensy 4.1 sebagai unit pemroses utama. Sistem ini menggunakan modul mikrofon I2S INMP441 untuk akuisisi sinyal audio dan layar OLED 0.96 inci sebagai antarmuka *output* teks. Arsitektur ini memungkinkan pemrosesan *Machine Learning* dilakukan sepenuhnya di tingkat *edge* (perangkat), sehingga sistem dapat beroperasi secara *offline* tanpa ketergantungan pada koneksi internet.
2. Model KWS yang diimplementasikan langsung pada purwarupa *Vogest Lens* menunjukkan kinerja yang sangat responsif dengan tingkat latensi di bawah 1 detik. Hasil pengujian menunjukkan bahwa model mampu mendeteksi kata kunci darurat dengan akurasi yang tinggi, mencapai angka 95% dalam kondisi lingkungan yang terkontrol dan 90% pada lingkungan dengan tingkat kebisingan hingga 80 dB. Hal ini membuktikan bahwa perangkat keras Teensy 4.1 mampu menjalankan algoritma klasifikasi audio secara *real-time* dengan efektif.
3. Model *transfer learning* untuk pengenalan bahasa isyarat (gestur) visual yang dilatih menggunakan platform *Edge Impulse* berhasil mencapai tingkat akurasi 99,67%. Dikarenakan adanya kendala *CPU Starvation* (keterbatasan sumber daya komputasi *single-core*) pada Teensy 4.1 saat mencoba menjalankan pemrosesan audio dan video secara bersamaan, fitur visual ini dioperasikan melalui simulasi pada *smartphone*. Simulasi ini membuktikan bahwa algoritma *transfer learning* yang digunakan memiliki performa yang baik untuk menerjemahkan bahasa isyarat meskipun dalam kapasitas sistem yang terpisah.

6.2 Saran

Berdasarkan hasil pengujian, evaluasi sistem, serta kendala teknis yang ditemukan selama proses perancangan dan implementasi purwarupa *Vogest Lens*, terdapat beberapa saran yang dapat diajukan untuk pengembangan dan penelitian di masa mendatang:

1. Peningkatan Arsitektur Perangkat Keras (*Hardware Upgrade*)

Kendala utama berupa *CPU Starvation* dan *bottleneck* komputasi terjadi akibat batasan arsitektur *single-core* pada mikrokontroler Teensy 4.1. Untuk penelitian selanjutnya, disarankan untuk melakukan migrasi ke mikrokontroler atau *SBC* (*Single Board Computer*) dengan arsitektur *multi-core* atau unit pemrosesan AI terdedikasi. Penggunaan perangkat seperti ESP32-S3 (*Dual-Core*), Raspberry Pi Compute Module, atau perangkat *edge AI* seperti Google Coral Micro dapat memungkinkan pemrosesan *multithreading*. Hal ini akan memungkinkan model *Keyword Spotting* (audio) dan pengenalan bahasa isyarat visual melalui kamera (seperti modul OV7670 atau kamera yang lebih mumpuni) dapat berjalan secara simultan (*real-time*) dalam satu perangkat fisik tanpa mengalami *crash* atau *delay*.

2. Pengembangan Modul Penampil (*Display*) Berbasis *Augmented Reality* (AR)

Pada purwarupa ini, antarmuka visualisasi masih menggunakan layar mikro OLED matriks biasa yang berpotensi menghalangi sebagian pandangan perifer pengguna. Penelitian selanjutnya disarankan untuk mengintegrasikan sistem optik yang lebih canggih, seperti penggunaan prisma optik reflektif (seperti pada *Google Glass*) atau *transparent OLED / waveguide display*. Peningkatan ini akan memberikan pengalaman *Augmented Reality* (AR) yang sesungguhnya, di mana teks terjemahan atau peringatan darurat dapat diproyeksikan langsung ke retina atau lensa transparan tanpa menghalangi pandangan dunia nyata pengguna.

3. Ekspansi Dataset dan Optimasi Model Kecerdasan Buatan

Meskipun model *transfer learning* visual yang disimulasikan melalui *Edge Impulse* telah mencapai akurasi lebih dari 80%, model tersebut masih dapat dioptimalkan. Disarankan untuk memperluas variasi dataset bahasa isyarat (Sistem Isyarat Bahasa Indonesia / SIBI atau BISINDO) dengan lebih banyak kelas gestur tangan yang dinamis, tidak hanya gestur statis. Selain itu, pada sistem *Keyword Spotting*, dataset *noise* (kebisingan latar) dapat diperkaya dengan sampel suara lingkungan nyata yang lebih ekstrem (seperti suara lalu lintas padat, keramaian stasiun, atau sirene) guna

lebih meningkatkan ketahanan (*robustness*) model dalam meminimalkan *False Positive* di ruang publik.

4. Realisasi Fabrikasi Desain Mekanis dan Uji Ergonomi

Oleh karena fokus penelitian ini masih pada tahap *Proof of Concept* (PoC) secara arsitektural dan desain sasis mekanis (CAD) masih bersifat pemodelan 3D, maka disarankan untuk melakukan fabrikasi fisik dari desain sasis kacamata menggunakan teknik *3D printing* dengan material ringan (seperti PLA atau Resin ABS-like). Setelah fabrikasi fisik berhasil, perlu dilakukan pengujian ergonomi (*Usability Testing*) secara langsung kepada kelompok responden penyandang tunarungu-wicara untuk mengukur tingkat kenyamanan penggunaan kacamata (berat, distribusi panas dari komponen, dan sudut pandang) dalam durasi pemakaian yang lama.

5. Manajemen Daya dan Konsumsi Energi (*Power Profiling*)

Implementasi *Machine Learning* pada perangkat *edge* secara terus-menerus (*always-on*) membutuhkan daya komputasi dan energi yang signifikan. Untuk menjadikan *Vogest Lens* sebagai perangkat *wearable* yang praktis, penelitian berikutnya perlu melakukan analisis profil daya (*power profiling*) secara mendalam. Disarankan untuk merancang sirkuit manajemen daya yang efisien dan mengimplementasikan algoritma *sleep/wake-up* berbasis sensor (misalnya menggunakan sensor inersia/IMU) sehingga sistem hanya aktif memproses data ketika kacamata digunakan atau ketika ada pemicu suara tertentu, guna memperpanjang daya tahan baterai.

DAFTAR PUSTAKA

- Adiwidjaja, R., & Ivan Fanany, M. (2020). End-to-end indonesian speech recognition with convolutional and gated recurrent units. *Journal of Physics: Conference Series*, 1566, 012097. <https://doi.org/10.1088/1742-6596/1566/1/012097>
- Akbay, T. (2022). Modeling Education Studies Indexed in Web of Science Using Natural Language Processing. *Instructional Technology and Lifelong Learning*, 3(2), 129–144. <https://doi.org/10.52911/itall.1193460>
- Al Ghifari, M. A. N., Syarifah, L., & Tinggi Agama Islam Syubbanul Wathon, S. (2025). Strategi Komunikasi Guru Dalam Mendukung Proses Pembelajaran Anak Berkebutuhan Khusus Di PKBM Disabilitas Plus Temanggung. *Irfani*, 21. <http://journal.iaingorontalo.ac.id/index.php/ir>
- Alfath, I., Fathony, N., Mareta, A., Adiana, B. E., Wardhani, O., Halim, A., Studi, P., & Informasi, T. (2025). Optimisasi Whisper Speech-to-Text Bahasa Indonesia dengan Hybrid Cloud dan Multi-Engine. *Jurnal Elektronika Dan Komputer*, 18(1). <https://journal.stekom.ac.id/index.php/elkom>
- Ali, M. L., & Zhang, Z. (2024). The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection. In *Computers* (Vol. 13, Number 12). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/computers13120336>
- Arnida, Hijriati, Puja Maulina, C., Fitria, A., & Fadila, N. (2024). Analisis Karakteristik dan Aktivitas Belajar Anak Berkebutuhan Khusus (Tunarungu) Di Sekolah SLB-B YPAC Banda Aceh. *Jurnal Pendidikan Anak Usia Dini*, 8(1). <https://jurnal.unugha.ac.id/index.php/warna>
- Arredondo-Soto, K. C., Blanco-Fernandez, J., Miranda-Ackerman, M. A., Solis-Quinteros, M. M., Realyvasquez-Vargas, A., & Garcia-Alcaraz, J. L. (2021). A Plan-Do-Check-Act Based Process Improvement Intervention for Quality Improvement. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3112948>
- Asakura, T. (2023). Augmented-Reality Presentation of Household Sounds for Deaf and Hard-of-Hearing People. *Sensors*, 23(17). <https://doi.org/10.3390/s23177616>
- Atobatele, O. K., Ajayi, O. O., Hungbo, A. Q., & Adeyemi, C. (2021). Applying Agile and Scrum Methodologies to Improve Public Health Informatics Project Implementation and Delivery. *Journal of Frontiers in Multidisciplinary Research*, 2(1), 426–439. <https://doi.org/10.54660/jfmr.2021.2.1.426-439>
- Berke, L., Thies, W., & Bragg, D. (2020, October 26). Chat in the Hat: A Portable Interpreter for Sign Language Users. *ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. <https://doi.org/10.1145/3373625.3417026>
- Cai, Y., Lu, T., Li, Z., Zhou, H., Dehaan, K., Xu, X., Gowda, M., & Jin, Y. (2025, September 27). SignGlass: First-Person View Comprehensive and Generalizable ASL Translation Using Wearable Glass. *UIST 2025 - Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*. <https://doi.org/10.1145/3746059.3747782>

- Chen, A. (2025). Comparative Analysis of YOLO Variants Based on Performance Evaluation for Object Detection. *ITM Web of Conferences*, 70, 03008. <https://doi.org/10.1051/itmconf/20257003008>
- Chen, L., Liu, X., Peng, L., & Wu, M. (2021). Deep learning based multimodal complex human activity recognition using wearable devices. *Applied Intelligence*. <https://doi.org/10.1007/s10489-020-02005-7> Published
- Chen, S., Haque, M., Liu, C., & Yang, W. (2022). DeepPerform: An Efficient Approach for Performance Testing of Resource-Constrained Neural Networks. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3551349.3561158>
- Chen, Y., He, Q., Liu, Z., Yang, M., & Wang, W. (2025). Noise-robust feature extraction for keyword spotting based on supervised adversarial domain adaptation training strategies. *Speech Communication*, 175. <https://doi.org/10.1016/j.specom.2025.103323>
- Cossich, V. R. A., Carlgren, D., Holash, R. J., & Katz, L. (2023). Technological Breakthroughs in Sport: Current Practice and Future Potential of Artificial Intelligence, Virtual Reality, Augmented Reality, and Modern Data Visualization in Performance Analysis. In *Applied Sciences (Switzerland)* (Vol. 13, Number 23). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/app132312965>
- Costa, A. F. F., Oliveira, V. T., Dantas, V. L. L., Santos, I. D. S., Andrade, R. M. D. C., & Gomes, R. (2023). Impacts of using PDCA in the requirements specification process. *ACM International Conference Proceeding Series*, 244–253. <https://doi.org/10.1145/3629479.3629511>
- De Mello, R. F., D', W., Fonseca, A., & Mareze, P. H. (2021). MEMS digital microphone and Arduino compatible microcontroller: an embedded system for noise monitoring. *Internoise 2021*.
- Dharma, A., Wijaya, J., Valentino, B., & Wijaya, V. (2025). Tinjaun Literatur Sistematis Terhadap Teknologi Kecerdasan Buatan untuk Deteksi Nyeri (2020-2024). *Metik Jurnal*, 9, 2025. <https://doi.org/10.47002/metik.v9i1.1068>
- Eze, P. U., & Asogwa, C. O. (2021). Deep machine learning model trade-offs for malaria elimination in resource-constrained locations. *Bioengineering*, 8(11). <https://doi.org/10.3390/bioengineering8110150>
- Felix Nur Noftianto, E., & Andhyka Kusuma, W. (2021). Teknik Iterasi User Persona dalam Penggalan Kebutuhan Perangkat Lunak Sistem Pembelajaran Jarak Jauh. *Jurnal Sosial Teknologi*, 1(7). <https://doi.org/10.59188/jurnalsostech.v1i7.134>
- Garcia-Perez, A., Miñón, R., Torre-Bastida, A. I., & Zulueta-Guerrero, E. (2023). Analysing Edge Computing Devices for the Deployment of Embedded AI. *Sensors*, 23(23). <https://doi.org/10.3390/s23239495>
- Gheorghe, C., Duguleana, M., Boboc, R. G., & Postelnicu, C. C. (2024). Analyzing Real-Time Object Detection with YOLO Algorithm in Automotive Applications: A Review. In *CMES - Computer Modeling in Engineering and Sciences* (Vol. 141, Number 3, pp. 1939–1981). Tech Science Press. <https://doi.org/10.32604/cmes.2024.054735>
- Gruosso, G., & Gajani, G. S. (2022). Comparison of Machine Learning Algorithms for Performance Evaluation of Photovoltaic Energy Forecasting and Management in the TinyML Framework. *IEEE Access*, 10. <https://doi.org/10.1109/ACCESS.2022.3222986>

- Handayani, A. N., Hariyono, H., Nasih, A. M., Rochmawati, R., Hitipeuw, I., Ar Rosyid, H., Ardiansah, J. T., Praja, R. I., Nurdiansyah, A., & Azizah, D. F. (2025). Comparative Analysis of Speech-to-Text APIs for Supporting Communication of the Deaf Community. *Indonesian Journal of Data and Science*, 6(3), 492–500. <https://doi.org/10.56705/ijodas.v6i3.327>
- Hashi, A. O., Hashim, S. Z. M., & Asamah, A. B. (2024). A Systematic Review of Hand Gesture Recognition: An Update from 2018 to 2024. In *IEEE Access* (Vol. 12, pp. 143599–143626). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2024.3421992>
- Idrees, S. K., Azar, J., Couturier, R., Idrees, A. K., & Gechter, F. (2025). SZ4IoT: an adaptive lightweight lossy compression algorithm for diverse IoT devices and data types. *Journal of Supercomputing*, 81(2). <https://doi.org/10.1007/s11227-024-06667-9>
- Immonen, R., & Hämäläinen, T. (2022). Tiny Machine Learning for Resource-Constrained Microcontrollers. In *Journal of Sensors* (Vol. 2022). <https://doi.org/10.1155/2022/7437023>
- Irchamna, A. M., Amira, V. N. F., Oktavianti, N., & Mintowati, M. (2025). Pengaruh Bahasa Isyarat Terhadap Pembelajaran Komunikasi Lisan Tunaganda (Tunarungu dan Tunawicara). *Jurnal Pendidikan Bahasa Indonesia*, 5(1), 40–49. <http://jurnaltarbiyah.uinsu.ac.id/index.php/eunoia/>
- Jang, C., Bang, K., Chae, M., Lee, B., & Lanman, D. (2024). Waveguide holography for 3D augmented reality glasses. *Nature Communications*, 15(1). <https://doi.org/10.1038/s41467-023-44032-1>
- Jegan, R., & Henrich, A. (2025). *A Structured Literature Review on Traditional Approaches in Current Natural Language Processing*. <http://arxiv.org/abs/2505.12970>
- Ketenagakerjaan, D. S. K. dan. (2022). Potret Penyandang Disabilitas Di Indonesia Hasil Long Form SP 2022. *Badan Pusat Statistik*.
- Kress, B. C., & Chatterjee, I. (2021). Waveguide combiners for mixed reality headsets: A nanophotonics design perspective. In *Nanophotonics* (Vol. 10, Number 1, pp. 41–74). De Gruyter Open Ltd. <https://doi.org/10.1515/nanoph-2020-0410>
- Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023). What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. In *Diagnostics* (Vol. 13, Number 15). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/diagnostics13152582>
- Kühl, N., Schemmer, M., Goutier, M., & Satzger, G. (2022). Artificial intelligence and machine learning. *Electronic Markets*, 32(4), 2235–2244. <https://doi.org/10.1007/s12525-022-00598-0>
- Kuznar, L. A., & Yager, M. (2020). The Development of Communication Models. *Strategic Multilayer Assessment Integrating Information in Joint Operations (IIJO)*.
- Lambert, J., Monahan, R., & Casey, K. (2021). Power consumption profiling of a lightweight development board: Sensing with the INA219 and Teensy 4.0 microcontroller. *Electronics (Switzerland)*, 10(7). <https://doi.org/10.3390/electronics10070775>

- Langner, R., Satkowski, M., Buschel, W., & Dachsel, R. (2021, May 6). Marvis: Combining mobile devices and augmented reality for visual data analysis. *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/3411764.3445593>
- Lee, G. B., Jang, H., Jeong, H., & Woo, W. (2021). Designing a Multi-Modal Communication System for the Deaf and Hard-of-Hearing Users. *Proceedings - 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct, ISMAR-Adjunct 2021*, 429–434. <https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00097>
- Legowo, M. B., Indarto, B., & Prayitno, D. (2019). Implementation of Scrum Work Framework in the Development of Quality Assurance Information System. *Jurnal Penelitian Pos Dan Informatika*, 9(2). <https://doi.org/10.17933/jppi.2019.090204>
- Lopez-Espejo, I., Tan, Z. H., Hansen, J. H. L., & Jensen, J. (2022). Deep Spoken Keyword Spotting: An Overview. In *IEEE Access* (Vol. 10, pp. 4169–4199). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2021.3139508>
- Lopez-Montiel, M., Orozco-Rosas, U., Sanchez-Adame, M., Picos, K., & Ross, O. H. M. (2021). Evaluation Method of Deep Learning-Based Embedded Systems for Traffic Sign Detection. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3097969>
- Marques, A. B., Branco, V., Costa, R., & Costa, N. (2023). Immersive Unit Visualization with Augmented Reality. *Multimodal Technologies and Interaction*, 7(10). <https://doi.org/10.3390/mti7100098>
- Matsubara, Y., Yang, R., Levorato, M., & Mandt, S. (2022). Supervised Compression for Resource-Constrained Edge Computing Systems. *Computer Vision Foundation*. <https://github.com/>
- Mørk, J., Bovbjerg, H. S., Kiss, G., & Tan, Z.-H. (2024). *Noise-Robust Keyword Spotting through Self-supervised Pretraining*. <http://arxiv.org/abs/2403.18560>
- Mujahid, A., Awan, M. J., Yasin, A., Mohammed, M. A., Damaševičius, R., Maskeliūnas, R., & Abdulkareem, K. H. (2021). Real-time hand gesture recognition based on deep learning YOLOv3 model. *Applied Sciences (Switzerland)*, 11(9). <https://doi.org/10.3390/app11094164>
- Munroe, O., Letourneur, S., & Novak, A. (2022, June 3). Design of an electronic circuit for loudspeaker real-time digital signal processing. *16ème Congrès Français d'Acoustique*. <http://arxiv.org/abs/2206.07516>
- Natarajan, T., & Pichai, S. (2024). Transition from Waterfall to Agile Methodology: An Action Research Study. *IEEE Access*, 12, 49341–49362. <https://doi.org/10.1109/ACCESS.2024.3384097>
- Nie, Y., Huang, H., Mao, X.-L., & Liao, L. (2024). Mix-Initiative Response Generation with Dynamic Prefix Tuning. *The 2024 Conference of the North American Chapt.*
- Omar, M., Choi, S., Nyang, D., & Mohaisen, D. (2022). Robust Natural Language Processing: Recent Advances, Challenges, and Future Directions. *IEEE Access*, 10, 86038–86056. <https://doi.org/10.1109/ACCESS.2022.3197769>
- Pazmiño Ortiz, L. A., Maldonado Soliz, I. F., & Guevara Balarezo, V. K. (2025). Advancing TinyML in IoT: A Holistic System-Level Perspective for Resource-Constrained AI. In *Future Internet* (Vol. 17, Number 6). <https://doi.org/10.3390/fi17060257>

- Pereira, P. H., Beccaro, W., & Ramirez, M. A. (2023). Evaluating Robustness to Noise and Compression of Deep Neural Networks for Keyword Spotting. *IEEE Access*, 11, 53224–53236. <https://doi.org/10.1109/ACCESS.2023.3280477>
- Pereira, W. A., Medina, G., Monaro, D., Villalobos, E. G. F., & Oliveira, R. P. de S. (2026). A Hybrid Agile-Quality Management Framework for Enhancing Productivity in a Public Academic Research Laboratory: A Case Study. *Administrative Sciences*, 16(1), 31. <https://doi.org/10.3390/admsci16010031>
- Prasetyo, V. R., Naufal, M. F., & Wijaya, K. (2025). Sentiment Analysis of ChatGPT on Indonesian Text using Hybrid CNN and Bi-LSTM. *Jurnal RESTI*, 9(2), 327–333. <https://doi.org/10.29207/resti.v9i2.6334>
- Purwanto, H., Kamaulid, T. R., Andriansyah, E., Anggraeni, O. P., & Manaf, K. (2024). KADOKU: Digital Innovation in Finding the Best Gifts Through Web Platform. *Is The Best Accounting Information Systems and Information Technology Business Enterprise This Is Link for OJS Us*, 9(2), 121–135. <https://doi.org/10.34010/aisthebest.v9i2.15025>
- Putri, G., & Sularno. (2024). Implementasi MIT App Inventor untuk Aplikasi Translator Suara ke Teks di Platform Android. *Jurnal Ilmu Komputer Dan Informatika*, 1(1). <https://jurnal.globalscients.com/index.php/jiki>
- Rastgoo, R., Kiani, K., & Escalera, S. (2021). Sign Language Recognition: A Deep Survey. In *Expert Systems with Applications* (Vol. 164). Elsevier Ltd. <https://doi.org/10.1016/j.eswa.2020.113794>
- Ray, P. P. (2022). A review on TinyML: State-of-the-art and prospects. In *Journal of King Saud University - Computer and Information Sciences* (Vol. 34, Number 4). <https://doi.org/10.1016/j.jksuci.2021.11.019>
- Rengganis, I. R., Gelgel, A., Agung, G., Suryawati, A., Fakultas,), & Sosial, I. (2025). Dinamika Interaksi Komunikasi Teman Tuli di Restoran Inklusiv Warung. *Jurnal Ilmiah Mahasiswa Ilmu KomunikasiMedium*.
- Rolland, J. P., & Goodsell, J. (2024). Waveguide-based augmented reality displays: a highlight. *Light: Science and Applications*, 13(1). <https://doi.org/10.1038/s41377-023-01371-4>
- Saggio, G., Cavallo, P., Ricci, M., Errico, V., Zea, J., & Benalcázar, M. E. (2020). Sign language recognition using wearable electronics: Implementing K-nearest neighbors with dynamic time warping and convolutional neural network algorithms. *Sensors (Switzerland)*, 20(14), 1–14. <https://doi.org/10.3390/s20143879>
- Sapkota, R., Flores-Calero, M., Qureshi, R., Badgujar, C., Nepal, U., Poulouse, A., Zeno, P., Vaddevolu, U. B. P., Khan, S., Shoman, M., Yan, H., & Karkee, M. (2025). YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series. *Artificial Intelligence Review*, 58(9). <https://doi.org/10.1007/s10462-025-11253-3>
- Sarikaya, B. (2024). CONCEPTUAL REVIEW OF ARTIFICIAL INTELLIGENCE; DIFFERENCES BETWEEN HUMAN AND MACHINE LEARNING. *Turkish Online Journal of Design Art and Communication*, 14(3). <https://doi.org/10.7456/tojdac.1464262>
- Seo, D., Oh, H. S., & Jung, Y. (2021). Wav2KWS: Transfer Learning from Speech Representations for Keyword Spotting. *IEEE Access*, 9, 80682–80691. <https://doi.org/10.1109/ACCESS.2021.3078715>
- Setyoko, A., Rahardjo, T., & Anita, N. F. (2024). EFFECTIVITY IMPROVEMENT OF HYBRID PROJECT MANAGEMENT WATER-SCRUM-FALL WITH

- SIX SIGMA IMPLEMENTATION. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 10(1). <https://doi.org/10.33480/jitk.v10i1.5539>
- Sharifani, K., & Amini, M. (2023). Machine Learning and Deep Learning: A Review of Methods and Applications. *World Information Technology and Engineering Journal*, 10(7).
- Shuvo, M. M. H., Islam, S. K., Cheng, J., & Morshed, B. I. (2023). Efficient Acceleration of Deep Learning Inference on Resource-Constrained Edge Devices: A Review. In *Proceedings of the IEEE* (Vol. 111, Number 1). <https://doi.org/10.1109/JPROC.2022.3226481>
- Suryadi, D., Octiva, C. S., Fajri, T. I., Nuryanto, U. W., & Hakim, M. L. (2024). Optimasi Kinerja Sistem IoT Menggunakan Teknik Edge Computing. *Jurnal Minfo Polgan*, 13(2), 1456–1461. <https://doi.org/10.33395/jmp.v13i2.14102>
- Tufail, S., Riggs, H., Tariq, M., & Sarwat, A. I. (2023). Advancements and Challenges in Machine Learning: A Comprehensive Review of Models, Libraries, Applications, and Algorithms. In *Electronics (Switzerland)* (Vol. 12, Number 8). <https://doi.org/10.3390/electronics12081789>
- Utami, E., Khairola, N., & Duita, R. (2024). Pendidikan Anak Berkebutuhan Khusus Pada Anak Tunawicara. *Educational Journal of Innovation and Publication (EJIP)*, 3(1), 2024.
- Villegas-Ch, W., Gutierrez, R., Maldonado Navarro, A., & Mera-Navarrete, A. (2024). Optimizing Federated Learning on TinyML Devices for Privacy Protection and Energy Efficiency in IoT Networks. *IEEE Access*, 12. <https://doi.org/10.1109/ACCESS.2024.3503516>
- Wadhawan, A., & Kumar, P. (2021). Sign Language Recognition Systems: A Decade Systematic Literature Review. *Archives of Computational Methods in Engineering*, 28(3). <https://doi.org/10.1007/s11831-019-09384-2>
- Wahyudi, A. P. A., Aziizah, F., Solihah, R. F., Putri, T. D. P., & Hamidah, S. (2024). Upaya Meningkatkan Kosakata Pada Anak Tunarungu. *Fonologi: Jurnal Ilmuan Bahasa Dan Sastra Inggris*, 2(2), 24–33. <https://doi.org/10.61132/fonologi.v2i2.592>
- Wang, R., & Zhang, Y. (2024). Pre-trained Language Models for Entity Blocking: A Reproducibility Study. *The 2024 Conference of the North American Chapter of the Association for Computational Linguistics*, 8720.
- Wiraputra, A. D. (2021). Perlindungan Hukum Terhadap Pekerja Penyandang Disabilitas. *Dharmasisya*, 1, 19.
- Zhang, L. L., Han, S., Wei, J., Zheng, N., Cao, T., Yang, Y., & Liu, Y. (2021). Nn-Meter: Towards accurate latency prediction of deep-learning model inference on diverse edge devices. *MobiSys 2021 - Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*. <https://doi.org/10.1145/3458864.3467882>
- Zhao, Z., Hao, K., Ma, X., Liu, X., Zheng, T., Xu, J., & Cui, S. (2021). SAI-YOLO: A Lightweight Network for Real-Time Detection of Driver Mask-Wearing Specification on Resource-Constrained Devices. *Computational Intelligence and Neuroscience*, 2021. <https://doi.org/10.1155/2021/4529107>
- Zichong, L. (2025). From Gesture to Voice: Advancing Human-Computer Interaction for the Deaf Community. *Frontiers in Art Research*, 7(4). <https://doi.org/10.25236/far.2025.070411>

LAMPIRAN

A-1 Lembar Pedoman Wawancara

PEDOMAN WAWANCARA BERPENDAMPINGAN (*PROXY-ASSISTED INTERVIEW GUIDE*)

Judul : Rancang Bangun Kacamata Multimodalitas Berbasis Machine Learning Dan Natural Language Processing Untuk Membantu Komunikasi Penyandang Tunarungu-Wicara

I. Identitas Wawancara

Kode Responden : _____
 Peran : _____
 Tanggal : _____
 Wawancara : _____
 Durasi : _____
 Peran/Status : Santri / Pengurus / Guru (*Lingkari salah satu*)

II. Daftar Pertanyaan Utama

A. Bagian 1: Pemetaan Kendala Keseharian & Keselamatan

1. Berdasarkan pengalaman/pengamatan, apa kendala terbesar atau hal yang paling mengagetkan yang dihadapi saat berada di jalan raya atau area luar kelas?
2. Bagaimana cara pesantren memberikan peringatan bahaya, dan apakah ada kebingungan ketika orang-orang di sekitar tiba-tiba berlarian karena peringatan seperti "Awat" atau "Kebakaran"?

B. Bagian 2: Pemetaan Interaksi Sosial

3. Apakah sering terjadi kesalahpahaman (misal: dikira sombong) karena tidak menyadari ketika ada teman yang memanggil dengan lambaian tangan dari kejauhan?

- ### C. Bagian 3: Validasi Konsep Produk (*Vogest Lens*)

- ### III. Catatan Jawaban

This image shows a single sheet of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page. There are approximately 20 lines visible. The paper has a slight shadow on its right side, suggesting it's resting on a surface.

A-3 Lembar Persetujuan Menjadi Responden

LEMBAR PERSETUJUAN MENJADI RESPONDEN

(Informed Consent)

Judul Penelitian Rancang Bangun Kacamata Multimodalitas Berbasis *Machine Learning* Dan *Natural Language Processing* Untuk Membantu Komunikasi Penyandang Tunarungu-Wicara

Identitas Peneliti

Nama : Yudhistira Alfarizi
NIM : 22522250
Prodi : S1 Teknik Industri Universitas Islam Indonesia

Deskripsi Singkat

Saya mengundang Anda untuk berpartisipasi dalam penelitian pengembangan alat bantu berupa kacamata pintar (*Vogest Lens*). Alat ini menggunakan teknologi pendeteksi suara otomatis (*Keyword Spotting*) yang berfungsi sebagai "telinga digital" untuk menangkap teriakan bahaya. Selain itu, alat ini juga dilengkapi kamera pintar yang dapat mengenali isyarat sapaan (seperti lambaian tangan) dari lawan bicara, sehingga Anda dapat mengetahui jika ada orang yang ingin mengajak berkomunikasi meskipun Anda tidak melihatnya secara langsung.

Prosedur

Jika Anda setuju, saya akan melakukan wawancara santai selama kurang lebih 15–30 menit. Saya akan bertanya mengenai pengalaman sehari-hari di pondok, kendala komunikasi, dan harapan Anda terhadap alat bantu ini. Wawancara ini akan direkam (suara/catatan) untuk keperluan analisis data.

Kerahasiaan & Kesukarelaan

1. Identitas Anda akan dirahasiakan (menggunakan inisial/nama samaran) dalam laporan skripsi, kecuali Anda mengizinkan sebaliknya.

Keikutsertaan ini bersifat sukarela. Anda boleh menolak menjawab pertanyaan tertentu atau berhenti kapan saja tanpa sanksi apa pun.

A-4 Lembar Persetujuan Menjadi Responden (Lanjutan)

Dengan menandatangani lembar ini, saya menyatakan **mengerti dan bersedia** menjadi responden dalam penelitian ini.

Peneliti



Yudhistira Alfarizi
22522250

Yogyakarta, 14 Februari 2026.

Responden



Bayu Pamungkas
Ponpes Tunarungu Darul Asham.

A-5 Lembar Persetujuan Menjadi Responden (Lanjutan)

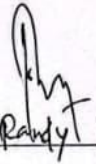
Dengan menandatangani lembar ini, saya menyatakan mengerti dan bersedia menjadi responden dalam penelitian ini.

Peneliti



Yudhistira Alfarizi
22522250

Yogyakarta, 29/02/2026
Responden


Randy Prmardian

A-6 Bukti Pelaksanaan Wawancara

Dokumentasi Wawancara Darul
Ashom



Dokumentasi Wawancara Jamhariyah

A-7 Transkrip Hasil Wawancara

Responden 1 (R-01)

Peran : Pengurus Pesantren / Ustadz

Tanggal : 14 Februari 2026

- **Pertanyaan 1**
Berdasarkan pengalaman/pengamatan, apa kendala terbesar atau hal yang paling mengagetkan yang dihadapi saat berada di jalan raya atau area luar kelas?
- **Jawaban**
Anak-anak ini paling rentan kalau di jalan raya. Mereka sering tidak sadar kalau ada kendaraan di belakang mereka yang membunyikan klakson
- **Pertanyaan 2**
Bagaimana cara pesantren memberikan peringatan bahaya, dan apakah ada kebingungan ketika orang-orang di sekitar tiba-tiba berlarian karena peringatan seperti "Awat" atau "Kebakaran"?
- **Jawaban**
Pengurus harus lari dan menepuk pundak mereka secara fisik. Kalau cuma diteriaki 'Awat!', mereka tidak akan merespons.
- **Pertanyaan 3**
Apakah sering terjadi kesalahpahaman (misal: dikira sombong) karena tidak menyadari ketika ada teman yang memanggil dengan lambaian tangan dari kejauhan?
- **Jawaban**
Sering sekali. Kadang ada warga sekitar atau teman dari kejauhan yang melambaikan tangan menyapa, tapi anak-anak jalan terus
- **Pertanyaan 4**
Seberapa pentingkah bagi Anda/santri untuk bisa mengetahui dengan cepat jika ada seseorang yang berusaha menyapa atau berinteraksi secara visual?
- **Jawaban**
Sangat penting agar mereka bisa bersosialisasi dengan normal.
- **Pertanyaan 5**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Kalau pakai HP harus dipegang terus, anak-anak repot saat bawa kitab atau barang. Lebih enak model kacamata.
- **Pertanyaan 6**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Lebih enak kacamata biarpun agak tebal asal tidak gampang jatuh dan kuat untuk keseharian

A-8 Transkrip Hasil Wawancara (Lanjutan)

Responden 2 (R-02)

Peran : Pengurus Pesantren / Ustadz

Tanggal : 24 Februari 2026

- **Pertanyaan 1**
Berdasarkan pengalaman/pengamatan, apa kendala terbesar atau hal yang paling mengagetkan yang dihadapi saat berada di jalan raya atau area luar kelas?
- **Jawaban**
Selain kendaraan, benda jatuh dari pohon atau bangunan juga bahaya. Mereka tidak bisa dengar suara retakan atau jatuhnya
- **Pertanyaan 2**
Bagaimana cara pesantren memberikan peringatan bahaya, dan apakah ada kebingungan ketika orang-orang di sekitar tiba-tiba berlarian karena peringatan seperti "Awat" atau "Kebakaran"?
- **Jawaban**
Iya, mereka sering kebingungan melihat orang-orang tiba-tiba berlarian saat ada peringatan bahaya.
- **Pertanyaan 3**
Apakah sering terjadi kesalahpahaman (misal: dikira sombong) karena tidak menyadari ketika ada teman yang memanggil dengan lambaian tangan dari kejauhan?
- **Jawaban**
Betul, akibatnya masyarakat yang tidak tahu sering mengira anak-anak tunarungu ini kurang ramah.
- **Pertanyaan 4**
Seberapa pentingkah bagi Anda/santri untuk bisa mengetahui dengan cepat jika ada seseorang yang berusaha menyapa atau berinteraksi secara visual?
- **Jawaban**
Penting untuk kepercayaan diri mereka di lingkungan masyarakat.
- **Pertanyaan 5**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Setuju, bentuk yang dipakai di kepala (*wearable*) jauh lebih praktis karena aktivitas santri banyak menggunakan kedua tangan.
- **Pertanyaan 6**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Tidak masalah sedikit tebal, yang penting posisi komponennya pas dan tidak melukai kepala santri.

A-9 Transkrip Hasil Wawancara (Lanjutan)

Responden 3 (R-03)

Peran : Santri

Tanggal : 14 Februari 2026

- **Pertanyaan 1**
Berdasarkan pengalaman/pengamatan, apa kendala terbesar atau hal yang paling mengagetkan yang dihadapi saat berada di jalan raya atau area luar kelas?
- **Jawaban**
Waktu itu ada kendaraan di belakang saya, saya tidak dengar klaksonnya sama sekali sampai kendaraannya menyalip dekat sekali.
- **Pertanyaan 2**
Bagaimana cara pesantren memberikan peringatan bahaya, dan apakah ada kebingungan ketika orang-orang di sekitar tiba-tiba berlarian karena peringatan seperti "Awat" atau "Kebakaran"?
- **Jawaban**
Saya butuh tahu orang berteriak apa, supaya saya bisa ikut menghindar.
- **Pertanyaan 3**
Apakah sering terjadi kesalahpahaman (misal: dikira sombong) karena tidak menyadari ketika ada teman yang memanggil dengan lambaian tangan dari kejauhan?
- **Jawaban**
Pernah ada teman yang melambaikan tangan ke saya, tapi saya sedang fokus melihat ke arah lain jadi tidak tahu.
- **Pertanyaan 4**
Seberapa pentingkah bagi Anda/santri untuk bisa mengetahui dengan cepat jika ada seseorang yang berusaha menyapa atau berinteraksi secara visual?
- **Jawaban**
Penting sekali, supaya saya bisa langsung membalas sapaan mereka.
- **Pertanyaan 5**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Saya ingin yang langsung nempel, tidak usah pegang-pegang HP lagi
- **Pertanyaan 6**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Tidak apa-apa tebal, asal nyaman dipakai lama.

A-10 Transkrip Hasil Wawancara (Lanjutan)

Responden 4 (R-04)

Peran : Santri

Tanggal : 14 Februari 2026

- **Pertanyaan 1**
Berdasarkan pengalaman/pengamatan, apa kendala terbesar atau hal yang paling mengagetkan yang dihadapi saat berada di jalan raya atau area luar kelas?
- **Jawaban**
Sering kaget kalau tiba-tiba ada benda jatuh di dekat saya.
- **Pertanyaan 2**
Bagaimana cara pesantren memberikan peringatan bahaya, dan apakah ada kebingungan ketika orang-orang di sekitar tiba-tiba berlarian karena peringatan seperti "Awat" atau "Kebakaran"?
- **Jawaban**
Kalau ada yang teriak 'Awat' atau 'Kebakaran', saya bingung kenapa orang-orang panik lari. Saya baru sadar kalau ada bahaya setelah ditarik teman.
- **Pertanyaan 3**
Apakah sering terjadi kesalahpahaman (misal: dikira sombong) karena tidak menyadari ketika ada teman yang memanggil dengan lambaian tangan dari kejauhan?
- **Jawaban**
Sering merasa sedih karena dikira sombong. Padahal saya benar-benar tidak lihat kalau ada yang melambaikan tangan ke saya dari samping atau kejauhan.
- **Pertanyaan 4**
Seberapa pentingkah bagi Anda/santri untuk bisa mengetahui dengan cepat jika ada seseorang yang berusaha menyapa atau berinteraksi secara visual?
- **Jawaban**
Saya ingin alat yang bisa memberitahu saya kalau ada yang melambaikan tangan, supaya saya tidak dikira sombong lagi
- **Pertanyaan 5**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Kacamata lebih enak. Tangan saya bisa bebas.
- **Pertanyaan 6**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Boleh saja tebal sedikit, asalkan layarnya jelas terlihat.

A-11 Transkrip Hasil Wawancara (Lanjutan)

Responden 5 (R-05)

Peran : Santri

Tanggal : 24 Februari 2026

- **Pertanyaan 1**
Berdasarkan pengalaman/pengamatan, apa kendala terbesar atau hal yang paling mengagetkan yang dihadapi saat berada di jalan raya atau area luar kelas?
- **Jawaban**
Saya merasa kurang aman kalau jalan sendirian di luar karena tidak bisa mendengar suara peringatan dari orang lain.
- **Pertanyaan 2**
Bagaimana cara pesantren memberikan peringatan bahaya, dan apakah ada kebingungan ketika orang-orang di sekitar tiba-tiba berlarian karena peringatan seperti "Awat" atau "Kebakaran"?
- **Jawaban**
Saya harap ada tulisan yang muncul langsung di depan mata saya kalau ada peringatan bahaya seperti itu.
- **Pertanyaan 3**
Apakah sering terjadi kesalahpahaman (misal: dikira sombong) karena tidak menyadari ketika ada teman yang memanggil dengan lambaian tangan dari kejauhan?
- **Jawaban**
Iya, kadang teman dari jauh manggil dengan lambaian tangan, saya lewat begitu saja. Bikin tidak enak hati.
- **Pertanyaan 4**
Seberapa pentingkah bagi Anda/santri untuk bisa mengetahui dengan cepat jika ada seseorang yang berusaha menyapa atau berinteraksi secara visual?
- **Jawaban**
Kalau alatnya bisa beri tahu ada yang menyapa, saya akan sangat terbantu berteman dengan orang dengat.
- **Pertanyaan 5**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Lebih pilih kacamata, karena kalau pegang HP terus takut jatuh dan rusak.
- **Pertanyaan 6**
Jika ada sebuah alat bantu cerdas yang sedang dikembangkan, apakah Anda lebih suka alat tersebut berbentuk seperti *smartphone* yang harus dipegang, atau berbentuk perangkat yang dipakai (*wearable*) agar tangan bebas beraktivitas?
- **Jawaban**
Bagi saya tidak masalah *frame*-nya tebal, yang penting alatnya bisa membantu saya tahu kalau ada bahaya.

B-1 Lembar Validasi Fungsional Lembar Validasi Fungsional

(Evaluasi Kelayakan Teknis *Proof of Concept*)

Nama Proyek : *Vogest Lens* (Kacamata Multimodalitas)
 Fase Pengujian : Validasi *Proof of Concept* (PoC) & Simulasi *Edge AI*
 Tanggal Uji : 23 April 2026
 Metode Uji : *Black Box Testing* & Observasi Kinerja Parameter

No	Parameter Logika Sistem	Target Ambang Batas (<i>Threshold</i>)	Hasil Aktual	Status (Valid/Tidak)
1	Kecepatan Respons (Latensi)	< 250 ms	135 ms (Waktu inferensi KWS hingga teks tampil di layar OLED berjalan nyaris instan)	Valid
2	Resiliensi Akustik (Bising)	> 85 %	68 % (Akurasi menurun akibat <i>gain</i> mikrofon menangkap <i>noise</i> latar secara dominan, memicu <i>False Negative</i>)	Tidak Valid
3	Stabilitas Perangkat	Tidak <i>Crash</i> selama 2 Jam	Beroperasi stabil selama 2 jam pengujian <i>looping</i> non-stop tanpa terjadi <i>reboot</i> atau <i>memory leak</i> .	Valid

B-2 Lembar Validasi Fungsional (Lanjutan)

No	Parameter Logika Sistem	Target Ambang Batas (<i>Treshold</i>)	Hasil Aktual	Status (Valid/Tidak)
4	Ketepatan Visual (OLED)	100%	100% (Teks peringatan "AWAS!" muncul sesuai <i>trigger</i> tanpa ada <i>glitch</i> karakter).	Valid
5	Batas Aman Termal	< 45 °C	38 °C (Suhu operasional maksimal pada <i>chip</i> Teensy 4.1 saat beban komputasi penuh).	Valid

C-1 Lembar *User Story & Product Backlog*

Lembar *User Story* dan *Product Backlog*

IDENTITAS PROYEK

Nama Purwarupa : *Vogest Lens* (Kacamata Multimodalitas)

Fase Pengembang : Perancangan Kebutuhan Sistem & Spesifikasi Teknis (*Plan & Design*)

Metode Prioritas : MoSCoW (*Must, Should, Could, Won't Have*)

BAGIAN I: DAFTAR *USER STORY* (KEBUTUHAN PENGGUNA)

Kode Story	User Story (<i>Role - Action - Goal</i>)	Kriteria Keberhasilan (<i>Acceptance Criteria</i>)	Persona Rujukan
US-01	Sebagai santri tunarungu, saya ingin alat ini mampu mendeteksi teriakan peringatan ("Awat"), agar saya dapat menghindar dari potensi bahaya.	Mikrokontroler Teensy 4.1 mampu mengenali ekstraksi kata kunci audio KWS dan mengabaikan frekuensi bising latar belakang (<i>noise</i>).	Primer (Abdullah)
US-02	Sebagai pengguna, saya ingin menerima notifikasi peringatan dalam bentuk teks visual, agar saya dapat langsung memahami jenis ancaman.	Layar antarmuka OLED memvisualisasikan konversi peringatan teks secara mandiri dengan <i>response time</i> (latensi) yang minimum.	Primer (Abdullah)

C-2 Lembar *User Story & Product Backlog* (Lanjutan)

Kode Story	User Story (<i>Role - Action - Goal</i>)	Kriteria Keberhasilan (<i>Acceptance Criteria</i>)	Persona Rujukan
US-03	Sebagai individu dengan mobilitas tinggi, saya ingin usulan rancangan alat yang dipakai di kepala, agar kedua tangan saya bebas beraktivitas (<i>hands-free</i>).	Terbentuknya rancangan <i>High-Fidelity 3D CAD</i> yang memvalidasi estimasi tata letak (<i>fitment</i>) komponen secara ergonomis (tanpa tahapan fabrikasi fisik cetak 3D).	Sekunder (Fatih)
US-04	Sebagai santri, saya ingin ada fitur yang menginformasikan jika ada rekan yang menyapa dengan gestur ("Halo"), agar saya dapat merespons.	Model klasifikasi citra lambaian tangan berhasil dilatih dan dieksekusi dengan tingkat akurasi tinggi melalui <i>deployment</i> simulasi di perangkat <i>smartphone</i> .	Primer & Sekunder

C-3 Lembar *User Story & Product Backlog* (Lanjutan)

BAGIAN II: *PRODUCT BACKLOG* DAN SPESIFIKASI TEKNIS

ID Backlog	Item Pekerjaan (<i>Technical Task</i>)	Spesifikasi Teknis Komponen	Prioritas (MoSCoW)	Estimasi Alur
PB-01	Pembuatan Model <i>Keyword Spotting</i>	Pelatihan model <i>Machine Learning</i> untuk klasifikasi audio menggunakan platform Edge Impulse. Akuisisi data menggunakan modul mikrofon I2S INMP441	<i>Must Have</i>	Jalur PoC <i>Hardware</i>
PB-02	Integrasi Antarmuka PoC	Pemrograman layar OLED 0.96 inci via protokol I2C dengan Teensy 4.1 untuk menampilkan teks peringatan audio.	<i>Must Have</i>	Jalur PoC <i>Hardware</i>
PB-03	Simulasi Model <i>Gesture Recognition</i>	Pelatihan algoritma klasifikasi citra lambaian tangan ("Halo") dan di-deploy ke platform smartphone sebagai simulasi sistem terpisah.	<i>Must Have</i>	Jalur Simulasi <i>Mobile</i>

C-4 Lembar User Story & Product Backlog (Lanjutan)

ID Backlog	Item Pekerjaan (<i>Technical Task</i>)	Spesifikasi Teknis Komponen	Prioritas (MoSCoW)	Estimasi Alur
PB-04	Perancangan Sasis Mekanis (CAD)	Pemodelan 3D sasis kacamata (<i>Component-Based Design</i>) untuk mengakomodasi tata letak komponen (tanpa fabrikasi fisik).	<i>Should Have</i>	Jalur Desain 3D
PB-05	Integrasi <i>Hardware</i> Visual ke Mikrokontroler	Menghubungkan secara fisik modul kamera OV7670 dengan Teensy 4.1 bersamaan dengan pemrosesan audio secara <i>real-time</i> .	<i>Won't Have</i>	Dibatasi (Mitigasi <i>Crash</i>)
PB-06	Penerjemahan Kalimat Isyarat Dinamis	Implementasi komputasi berat untuk menerjemahkan tata bahasa isyarat BISINDO secara berkelanjutan.	<i>Won't Have</i>	Dibatasi

D- Skema Rangkaian Elektronik (*Wiring Diagram*)

