

BAB 2

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Pada penelitian ini, terdapat beberapa karya yang relevan terkait dengan perancangan alat bantu komunikasi visual untuk penyandang tuna rungu berbasis kacamata pintar dan teknologi *Speech-to-Text*. Penelitian-penelitian terdahulu membahas berbagai aspek yang mendukung pengembangan sistem, seperti penerapan *Speech-to-Text* untuk membantu komunikasi penyandang tuna rungu, penggunaan perangkat *wearable* sebagai media tampilan informasi, implementasi model *Faster-Whisper* untuk pengenalan ucapan, serta pemanfaatan mikrokontroler dan jaringan WiFi dalam proses transmisi data audio dan teks. Hasil kajian terhadap penelitian terdahulu digunakan sebagai landasan dalam merancang sistem yang mengintegrasikan ESP32-S3, mikrofon digital INMP441, model *Faster-Whisper*, dan layar OLED untuk menghasilkan alat bantu komunikasi visual yang mampu menampilkan hasil transkripsi suara secara *near real-time*. Berikut beberapa penelitian yang relevan dengan penelitian ini.

Penelitian Nguyen (2025) yang berjudul *AI-Powered Smart glasses for Real-time Speech-to-Text Assistance for the Hearing Impaired*, mengembangkan kacamata pintar untuk membantu penyandang gangguan pendengaran memahami percakapan. Sistem menggunakan ESP32-S3, mikrofon digital, layar OLED, dan model *Whisper* untuk mengubah suara menjadi teks. Metode yang digunakan menerapkan arsitektur *client-server*, di mana data audio dikirim melalui jaringan WiFi untuk diproses menjadi teks. Hasil transkripsi kemudian ditampilkan pada layar OLED yang terpasang pada perangkat. Hasil penelitian menunjukkan bahwa sistem mampu menampilkan teks dengan baik dan mendukung komunikasi secara *near real-time*.

Maulana dkk. (2021), dalam penelitiannya *Rancang Bangun Kacamata Pengubah Suara Menjadi Teks*, mengembangkan alat bantu komunikasi berupa kacamata yang menangkap suara dari aplikasi *voice recorder* di *smartphone* dan mengirimkannya ke Arduino Uno via *Bluetooth* HC-05. Hasil pengolahan suara ditampilkan dalam bentuk teks pada layar LCD 2x16. Walau sistem masih

bergantung pada perangkat eksternal seperti ponsel, penelitian ini membuktikan bahwa teknologi *wearable* untuk konversi suara ke teks dapat direalisasikan secara fungsional dan efisien dengan latensi pengiriman data suara ke teks kurang dari 1 detik.

Khoiroh dkk. (2024) dalam judul *Implementasi Speech Recognition Whisper Pada Debat Calon Wakil Presiden Republik Indonesia*. Sistem diuji untuk mengubah suara dari rekaman debat menjadi teks secara otomatis menggunakan model *Whisper*. Pengujian dilakukan untuk menilai akurasi dan performa sistem pada audio dengan karakteristik formal, multi-pembicara, dan potensi kebisingan latar. Hasil pengujian menunjukkan bahwa *Whisper* mampu melakukan transkripsi dengan tingkat akurasi masih sangat tinggi dalam mengenali gagasan setiap cawapres, yaitu 95,19% pada gagasan Muhaimin, 96,71% pada gagasan Gibran, dan 87,16% pada gagasan Mahfud.

Husein (2025) dalam skripsinya yang berjudul *Rancang Bangun Pipeline Notulensi Otomatis dengan Integrasi Model Whisper (Speech-to-Text) dan Model PEGASUS (Abstractive Summarization)* mengembangkan sistem notulensi otomatis yang mengintegrasikan model *Whisper* untuk mengubah suara menjadi teks dan model PEGASUS untuk menghasilkan ringkasan otomatis dari hasil transkripsi. Sistem dirancang untuk membantu proses dokumentasi rapat secara lebih efisien dengan mengurangi kebutuhan pencatatan manual. Hasil penelitian menunjukkan bahwa integrasi kedua model tersebut mampu menghasilkan transkripsi dan ringkasan rapat secara otomatis sehingga meningkatkan efisiensi proses pembuatan notulen serta mempermudah pengelolaan informasi hasil rapat.

Terakhir, penelitian oleh Fathony dkk. (2025) berjudul *Optimisasi Whisper Speech-to-Text Bahasa Indonesia dengan Hybrid Cloud dan Multi-Engine*. Pendekatan ini memanfaatkan mesin *Faster-Whisper large-v3* serta teknik *audio preprocessing*, dengan struktur *client-server* yang terdiri dari server GPU (Google Colab Tesla T4) dan *client* lokal (Ubuntu 22.04, 8-core, 32 GB RAM). Uji coba yang dilakukan terhadap lima jenis audio berbeda menunjukkan keberhasilan pemrosesan sebesar 80%, dengan WER mulai dari 27,69% untuk audio formal hingga 645,16% untuk kondisi informal yang sangat bising. Sistem juga mengonsumsi 21,3% GPU dan 35,4% RAM, dan meski waktu pemrosesan stabil

untuk *file* berukuran biasa, sistem mengalami kegagalan (*timeout*) ketika menangani *file* besar (> 50 MB).

Tabel 2. 1 Penelitian Terdahulu

Penulis	Judul	Metode	Hasil
(Nguyen, 2025)	<i>AI-Powered Smart glasses for Real-time Speech-to-Text Assistance for the Hearing Impaired</i>	Menggunakan ESP32-S3, mikrofon digital, OLED, dan model <i>Whisper</i> dengan arsitektur <i>client-server</i> melalui jaringan WiFi untuk mengubah suara menjadi teks.	Sistem mampu melakukan transkripsi dan menampilkan hasilnya pada OLED secara <i>near real-time</i> sehingga dapat membantu komunikasi penyandang gangguan pendengaran.
(Maulana dkk., 2021)	<i>Rancang Bangun Kacamata Pengubah Suara Menjadi Teks</i>	Perancangan sistem berbasis Arduino Uno yang menerima <i>input</i> suara dari <i>smartphone</i> via <i>Bluetooth</i>	Sistem berhasil menampilkan teks dengan latensi <1 detik, namun masih bergantung pada <i>smartphone</i> untuk <i>input</i> suara.
(Khoiroh dkk., 2024)	<i>Implementasi Speech Recognition Whisper Pada Debat Calon Wakil Presiden Republik Indonesia</i>	Menggunakan <i>Whisper Speech Recognition</i> (mode 1 berbasis transformer) untuk menganalisis transkripsi debat politik dengan fokus pada akurasi pengenalan suara dalam konteks percakapan multi-pembicara.	<i>Whisper</i> mampu melakukan transkripsi dengan tingkat akurasi masih sangat tinggi dalam mengenali gagasan setiap cawapres, yaitu 95,19% pada gagasan Muhaimin, 96,71% pada gagasan Gibran, dan 87,16% pada gagasan Mahfud.

Penulis	Judul	Metode	Hasil
(Husein, 2025)	<i>Rancang Bangun Pipeline Notulensi Otomatis dengan Integrasi Model Whisper dan PEGASUS</i>	Mengintegrasikan model <i>Whisper</i> untuk mengubah suara menjadi teks dan model PEGASUS untuk menghasilkan ringkasan otomatis dari hasil transkripsi rapat.	Sistem mampu menghasilkan transkripsi dan ringkasan rapat secara otomatis sehingga meningkatkan efisiensi proses pembuatan notulen.
(Fathony dkk., 2025)	<i>Optimisasi Whisper Speech-to-Text Bahasa Indonesia dengan Hybrid Cloud dan Multi-Engine</i>	Menggunakan pendekatan <i>Hybrid Cloud</i> (gabungan cloud publik dan privat) serta <i>Multi-Engine Whisper</i> API untuk meningkatkan akurasi transkripsi <i>real-time</i> dalam Bahasa Indonesia.	Hasil terhadap lima jenis audio berbeda menunjukkan keberhasilan pemrosesan sebesar 80%, dengan WER mulai dari 27,69% untuk audio formal hingga 645,16% untuk kondisi informal yang sangat bising.

Berdasarkan kajian terhadap penelitian-penelitian terdahulu pada Tabel 2.1, dapat diketahui bahwa berbagai penelitian telah menerapkan teknologi *Speech-to-Text* menggunakan model *Whisper* maupun perangkat *wearable* untuk mendukung komunikasi dan pengolahan informasi berbasis suara. Namun, sebagian besar penelitian masih berfokus pada kebutuhan notulensi, transkripsi rekaman, atau sistem yang bergantung pada perangkat eksternal dan belum secara khusus dirancang sebagai alat bantu komunikasi visual yang praktis bagi penyandang tunarungu.

Penelitian ini memiliki perbedaan dengan penelitian sebelumnya karena mengembangkan alat bantu komunikasi visual berupa kaca mata pintar yang mengintegrasikan teknologi *Speech-to-Text* pada perangkat *wearable*. Sistem

dirancang untuk menampilkan hasil transkripsi suara secara *near real-time* melalui layar OLED sehingga dapat membantu penyandang tunarungu memahami percakapan secara lebih mudah. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan teknologi asistif berbasis *Speech-to-Text* serta menjadi alternatif solusi komunikasi bagi penyandang tunarungu.

2.2 Landasan Teori

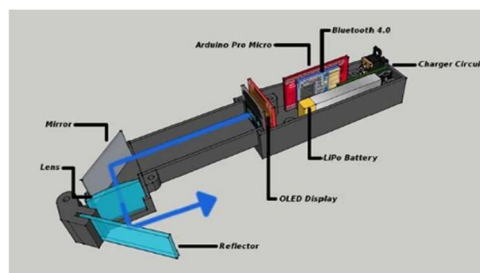
Landasan teori berisi konsep dan teori yang mendukung penelitian mengenai alat bantu komunikasi visual untuk penyandang tuna rungu. Teori yang dibahas meliputi perangkat *wearable* berbasis kacamata, komunikasi visual, *Speech-to-Text*, *Faster-Whisper*, ESP32-S3, mikrofon digital INMP441, OLED, komunikasi data melalui WiFi, serta sistem catu daya yang digunakan dalam perancangan alat.

2.2.1 Perangkat Wearable Berbasis Kacamata

Kacamata Perangkat *wearable* berbasis kacamata merupakan perangkat elektronik yang dirancang untuk dipasang pada atau diintegrasikan dengan frame kacamata sehingga dapat digunakan secara langsung oleh pengguna. Perangkat ini umumnya dilengkapi komponen elektronik seperti layar mini, sensor, mikroprosesor, mikrofon, dan konektivitas nirkabel untuk mendukung penyampaian informasi secara praktis dan *real-time* (Vishal Thorat et al., 2020). Dalam bidang aksesibilitas, perangkat *wearable* berbasis kacamata telah banyak dikembangkan untuk membantu penyandang disabilitas sensorik. Vishal Thorat dkk. (2020) mengembangkan kacamata pintar berbasis Arduino dan sensor ultrasonik untuk membantu penyandang tunanetra mendeteksi hambatan di lingkungan sekitar. Selain itu, (Raju et al., 2025) menunjukkan bahwa penerapan teknologi pengenalan suara pada perangkat berbasis kacamata dapat membantu meningkatkan komunikasi penyandang tuli dalam aktivitas sehari-hari. Pada penelitian ini, perangkat yang dikembangkan berupa modul elektronik yang dipasang pada frame kacamata dan berfungsi menampilkan hasil transkripsi percakapan dalam bentuk teks. Pendekatan ini dipilih karena lebih ringan, sederhana, dan mudah diterapkan dibandingkan kacamata pintar yang terintegrasi penuh.

2.2.2 Desain Casing Modul Kacamata

Desain perangkat tidak menyatu langsung dengan struktur utama kacamata, melainkan dibuat dalam bentuk *casing* modul tambahan yang disangkutkan atau dijepitkan pada *frame* kacamata yang sudah ada, di mana seluruh sistem elektronik termasuk ESP32-S3, OLED, dan *microphone* analog dirakit dalam *casing* mini terpisah yang dapat dilepas-pasang. Desain ini memungkinkan fleksibilitas pemakaian tanpa harus membuat *frame* khusus, serta memudahkan pengguna dalam perawatan atau pengisian daya. Desain fisik dari proyek milik Mr. Vishal Thorat dkk. dapat dilihat pada Gambar 2.1.



Gambar 2. 1 Desain Sistem Milik Vishal Thorat

2.2.3 Komunikasi Visual

Komunikasi visual adalah penyampaian informasi menggunakan media visual seperti teks, ikon, simbol, grafik, dan gambar. Ini merupakan metode penting bagi individu dengan hambatan sensorik, termasuk tuna rungu. Dame dkk. (2024) membuktikan efektivitas sistem komunikasi visual bagi siswa tunarungu menggunakan tombol interaktif dan tampilan gambar dalam lingkungan kelas inklusi. Dalam sistem ini, visualisasi memainkan peran utama sebagai pengganti elemen audio. Oleh karena itu, tampilan hasil pengenalan suara ke layar kecil pada perangkat *wearable* menjadi implementasi langsung dari prinsip komunikasi visual ini.

2.2.4 Teknologi Speech-to-Text (STT)

Speech-to-Text (STT) merupakan teknologi yang digunakan untuk mengubah sinyal suara menjadi teks secara otomatis melalui proses pengenalan ucapan (*speech recognition*). Teknologi ini memungkinkan informasi yang disampaikan dalam bentuk suara dapat dikonversi menjadi teks yang dapat dibaca oleh pengguna. Penerapan STT telah banyak digunakan pada berbagai bidang, seperti

transkripsi otomatis, asisten virtual, sistem *caption*, dan alat bantu komunikasi bagi penyandang disabilitas pendengaran karena mampu meningkatkan aksesibilitas terhadap informasi yang disampaikan secara lisan (Irdamurni dkk., 2020).

Secara umum, sistem STT bekerja melalui beberapa tahapan, yaitu akuisisi suara menggunakan mikrofon, pemrosesan sinyal audio, serta proses pengenalan ucapan menggunakan model kecerdasan buatan untuk menghasilkan teks. Perkembangan teknologi *deep learning* meningkatkan kemampuan sistem STT dalam mengenali ucapan dengan tingkat akurasi lebih tinggi dibandingkan metode konvensional (Khoiroh dkk., 2024). Dalam penelitian ini, teknologi STT digunakan untuk mengubah suara percakapan menjadi teks yang kemudian ditampilkan pada layar OLED sebagai media komunikasi visual bagi penyandang tuna rungu.

2.2.5 *Faster Whisper*

Faster-Whisper merupakan implementasi optimasi dari model *Whisper* yang dikembangkan oleh OpenAI untuk tugas pengenalan ucapan (*automatic speech recognition*). *Faster-Whisper* dibangun menggunakan mesin inferensi CTranslate2 yang mampu meningkatkan kecepatan proses transkripsi serta mengurangi penggunaan memori dibandingkan implementasi *Whisper* standar. Meskipun lebih ringan, *Faster-Whisper* tetap mempertahankan tingkat akurasi yang tinggi dalam mengenali berbagai bahasa, termasuk bahasa Indonesia (Guillaume Klein, 2023).

Model *Faster-Whisper* bekerja dengan menerima data audio sebagai masukan, kemudian melakukan proses ekstraksi fitur dan inferensi menggunakan arsitektur Transformer untuk menghasilkan teks transkripsi. Teknologi ini mampu mengenali ucapan pada berbagai kondisi lingkungan dan mendukung penggunaan model dengan ukuran berbeda, mulai dari *tiny*, *base*, *small*, *medium*, hingga *large*, sehingga dapat disesuaikan dengan kebutuhan akurasi dan kemampuan perangkat yang digunakan.

Penelitian oleh Nguyen dkk. (2024) menunjukkan bahwa teknologi *Speech-to-Text* berbasis model *Whisper* dapat diterapkan pada perangkat *wearable* untuk membantu penyandang tuna rungu memahami percakapan secara lebih efektif. Selain itu, Fathony dkk. (2025) menunjukkan bahwa penggunaan *Faster-Whisper* pada arsitektur *client-server* mampu memberikan performa transkripsi yang baik dengan waktu pemrosesan yang relatif cepat pada bahasa Indonesia. Hasil

penelitian tersebut menunjukkan bahwa pemrosesan *Speech-to-Text* dapat dilakukan secara efisien tanpa harus menjalankan model secara langsung pada perangkat *wearable*.

Dalam penelitian ini, *Faster-Whisper* dijalankan pada komputer yang berfungsi sebagai server lokal untuk memproses data audio yang dikirim oleh ESP32-S3 melalui jaringan WiFi. Hasil transkripsi kemudian dikirim kembali ke ESP32-S3 dan ditampilkan pada layar OLED. Penggunaan *Faster-Whisper* dipilih karena memiliki keseimbangan yang baik antara kecepatan pemrosesan, kebutuhan sumber daya komputasi, dan tingkat akurasi transkripsi sehingga sesuai untuk mendukung sistem komunikasi visual yang dikembangkan.

2.2.6 ESP32-S3-DevkitC-1

ESP32-S3 merupakan mikrokontroler yang dikembangkan oleh Espressif Systems dengan prosesor *dual-core* Xtensa LX7 berkecepatan hingga 240 MHz serta dilengkapi konektivitas WiFi 2.4 GHz dan Bluetooth Low Energy (BLE). ESP32-S3 dirancang untuk aplikasi *Internet of Things* (IoT), perangkat *wearable*, dan sistem embedded yang membutuhkan komunikasi nirkabel dengan konsumsi daya relatif rendah. Dibandingkan generasi ESP32 sebelumnya, ESP32-S3 memiliki peningkatan kemampuan pemrosesan, dukungan instruksi vektor untuk komputasi AI sederhana, serta antarmuka periferi yang lebih lengkap seperti I2C, SPI, UART, ADC, PWM, dan I2S.



Gambar 2. 2 ESP32-S3

Dalam penelitian ini, ESP32-S3 digunakan sebagai pengendali utama sistem yang bertugas membaca data audio dari mikrofon digital INMP441 melalui antarmuka I2S, mengirimkan data audio ke laptop melalui jaringan WiFi

menggunakan protokol UDP, menerima hasil transkripsi dari sistem *Speech-to-Text* berbasis *Faster-Whisper*, serta menampilkan hasil teks pada layar OLED. Penggunaan ESP32-S3 dipilih karena kemampuan komunikasi nirkabel yang baik, ukuran ringkas, konsumsi daya rendah, serta cukup untuk menangani proses akuisisi dan transmisi data audio secara *near real-time* (System, 2026). Bentuk fisik ESP32-S3 DevKitC-1 yang digunakan dalam penelitian dapat dilihat pada Gambar 2.2, sedangkan spesifikasi teknisnya disajikan pada Tabel 2.2.

Tabel 2. 2 Spesifikasi ESP32-S3

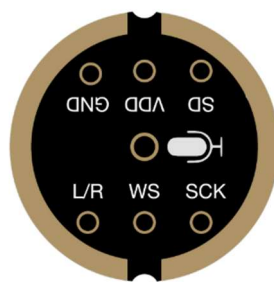
Parameter	Detail
Nama Mikrokontroler	ESP32-S3-DevkitC-1
Catu Daya	3V3-3V6 (via USB C atau GPIO)
SRAM	512 KB
Flash	8 MB
PSRAM	8 MB
Bluetooth	BLE 5.0
WiFi	2.4GHz IEEE 802.11b/g/n
GPIO	50 pin
Mendukung	I2C, SPI, UART, I2S
Camera	CSI-2 Connector
USB	Micro USB On-The-Go (OTG)

ESP32-S3 bekerja sebagai mikrokontroler yang menjalankan *firmware* secara langsung tanpa sistem operasi umum. Pendekatan ini membuat waktu *booting* lebih cepat, konsumsi daya lebih rendah, dan penggunaan memori lebih efisien sehingga sesuai untuk aplikasi *wearable* yang membutuhkan ukuran ringkas dan daya baterai yang terbatas.

2.2.7 Mikrofon Digital

Mikrofon digital INMP441 merupakan mikrofon MEMS *omnidireksional* yang menghasilkan *output* digital melalui antarmuka I2S, sehingga tidak memerlukan konversi analog-ke-digital eksternal. Teknologi ini memungkinkan “transmisi audio digital langsung secara sinkron dengan pemisahan sinyal *clock* dan *data*”, menghasilkan kualitas suara yang lebih bersih dan stabil dibanding

antarmuka analog (Wikipedia, 2025). Pada penelitian ini, INMP441 digunakan sebagai sensor masukan suara yang berfungsi menangkap percakapan dari lawan bicara. Data audio yang diperoleh kemudian dibaca oleh ESP32-S3 melalui antarmuka I2S dan dikirim melalui jaringan WiFi untuk diproses menjadi teks menggunakan sistem *Speech-to-Text*. Pemilihan INMP441 didasarkan pada kualitas audio yang baik, ukuran yang kecil, konsumsi daya yang rendah, serta kemudahan integrasinya dengan ESP32-S3. Mikrofon digital yang digunakan dapat dilihat pada Gambar 2.3, dan spesifikasi dari INMP441 dapat dilihat pada Tabel 2.3.



Gambar 2. 3 INMP441

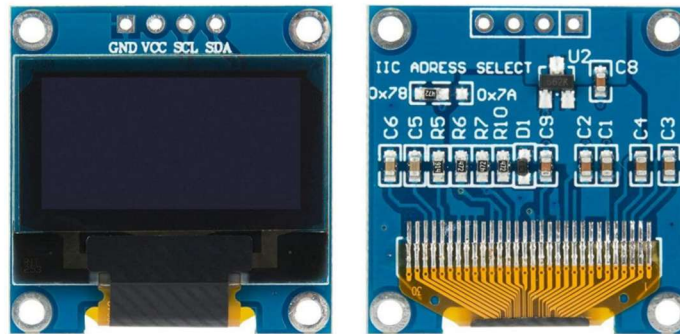
Tabel 2. 3 Spesifikasi INMP441

Parameter	Detail
Nama Modul	INMP441
Rentang Tegangan Operasi	1.8V – 3.3V
Antarmuka	I2S (<i>Inter-IC Sound</i>)
Tipe Mikrofon	<i>Omnidirectional MEMS</i> mikrofon
Konsumsi Daya Rendah	Modul ini dirancang untuk aplikasi dengan efisiensi daya tinggi
Rasio <i>Signal-to-Noise</i> (SNR)	Sekitar 61 dB
Frekuensi Sampling	Mendukung frekuensi sampling audio 60 Hz – 15 kHz

2.2.8 OLED Display

Layar OLED 0.96" beresolusi 128×64 piksel dipilih sebagai media tampilan hasil transkripsi suara. Dengan penggunaan *font* kecil 6×8 piksel, layar ini dapat

menampilkan sekitar 20–22 karakter per baris dan hingga 6 baris teks. Karakteristik OLED yang hemat daya, terang, dan berkontras tinggi membuatnya ideal untuk penggunaan jangka panjang dan tetap terbaca dalam berbagai kondisi pencahayaan. Layar ini dipasang sejajar pandangan mata pengguna untuk memudahkan pembacaan teks secara langsung saat digunakan. OLED Display 0.96" dapat dilihat dari gambar 2.4, sedangkan spesifikasi dari OLED dapat dilihat pada Tabel 2.4



Gambar 2. 4 OLED Display 0.96"

Tabel 2. 4 Spesifikasi OLED Display 0.96"

Parameter	Detail
Nama Modul	OLED Display 0.96"
Rentang Tegangan Operasi	3.3V-5V
<i>Pixel</i>	128*64
<i>Display size</i>	0.96 inch
<i>Interface</i>	I2C Interface
I2C Interface	SSD1306

2.2.9 Catu Daya Portable

Catu daya portabel merupakan sistem penyedia energi listrik yang memungkinkan perangkat elektronik beroperasi tanpa bergantung pada sumber listrik eksternal secara terus-menerus. Pada perangkat *wearable*, penggunaan catu daya portabel menjadi penting karena menentukan mobilitas, kenyamanan, dan lama waktu penggunaan alat. Sistem catu daya umumnya terdiri dari baterai sebagai sumber energi utama, rangkaian pengisian daya, modul pengatur tegangan, serta indikator kapasitas baterai untuk memantau kondisi daya selama perangkat digunakan.

Pada penelitian ini, catu daya portabel dirancang menggunakan baterai Lithium Polymer (LiPo) 3,7 V sebagai sumber energi utama, modul TP4056 sebagai pengisi daya baterai, modul MT3608 sebagai konverter tegangan, dan modul indikator baterai SPBKBS-10 untuk menampilkan status kapasitas baterai. Kombinasi komponen tersebut memungkinkan sistem memperoleh suplai tegangan yang stabil, mudah diisi ulang, serta mendukung penggunaan perangkat secara portabel.

1) ICR-18650 Lithium Ion Battery

Baterai Lithium-Ion tipe 18650 digunakan sebagai sumber daya utama pada sistem alat bantu komunikasi visual yang dikembangkan. Baterai ini dipilih karena memiliki kapasitas yang relatif besar, dapat diisi ulang, serta mampu menyediakan arus yang cukup untuk mendukung pengoperasian ESP32-S3, mikrofon INMP441, layar OLED, dan komponen pendukung lainnya. Selain itu, baterai 18650 memiliki ukuran yang ringkas dan banyak digunakan pada perangkat portabel karena menawarkan keseimbangan antara kapasitas energi, dimensi fisik, dan umur pakai. Pada penelitian ini digunakan baterai Lithium-Ion 18650 dengan tegangan nominal 3,7 V yang terintegrasi dengan rangkaian catu daya portabel. Bentuk fisik baterai yang digunakan dapat dilihat pada Gambar 2.5, sedangkan spesifikasinya disajikan pada Tabel 2.5.



Gambar 2. 5 Baterai Lithium-Ion 18650

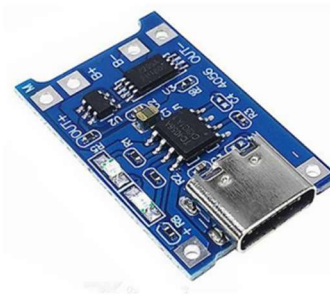
Tabel 2. 5 Spesifikasi Baterai Lithium-Ion 18650

Parameter	Detail
Tipe	<i>Lithium-Ion (18650)</i>
Tegangan	3.7V

Parameter	Detail
Kapasitas	2000mAh
Nominal Voltage	3.7V
<i>Maximum Charge (Voltage)</i>	4.2V
<i>Maximum Charge (Current)</i>	500mAh

2) Modul Charger TP4056 (Type-C)

Modul TP4056 digunakan sebagai rangkaian pengisian daya baterai LiPo 3,7 V. Modul ini mendukung pengisian ulang melalui konektor USB Type-C sehingga memudahkan pengguna dalam melakukan pengisian daya menggunakan adaptor maupun *power bank*. TP4056 menerapkan metode pengisian *Constant Current/Constant Voltage (CC/CV)* yang umum digunakan pada baterai lithium untuk menjaga keamanan dan memperpanjang umur baterai. Selain itu, modul ini juga dilengkapi fitur proteksi terhadap *overcharge*, *overdischarge*, dan *overcurrent* sehingga dapat meningkatkan keamanan sistem selama proses pengisian maupun penggunaan. Modul TP4056 yang digunakan dapat dilihat pada Gambar 2.6, sedangkan spesifikasinya ditunjukkan pada Tabel 2.6.



Gambar 2. 6 TP4056

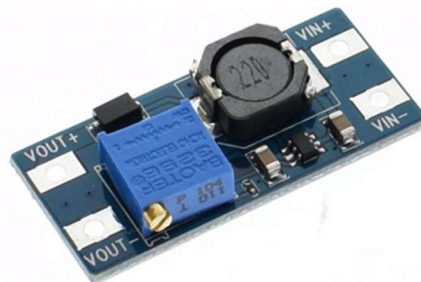
Tabel 2. 6 Spesifikasi TP4056

Parameter	Detail
Nama Modul	TP4056
<i>Input voltage</i>	5V
<i>Charging cut-off voltage</i>	4.2V $\pm$ 1%
<i>Maximum charging current</i>	1000mA

Parameter	Detail
<i>Battery over-discharge protection voltage</i>	2.4V $\pm$ 100mV
<i>Battery overcurrent protection current</i>	3A
<i>Board size</i>	2.6*1.7cm

3) Step-Up Converter (MT3608)

Tegangan dari baterai LiPo (3.7V) tidak cukup untuk langsung digunakan oleh Raspberry Pi yang membutuhkan 5V. Oleh karena itu, digunakan modul step-up converter MT3608 untuk menaikkan tegangan dari 3.7V menjadi 5V yang stabil. MT3608 memiliki ukuran yang ringkas dan efisiensi tinggi, cocok untuk perangkat portabel dan *wearable* seperti proyek ini. *Step-Up Converter (MT3608)* dapat dilihat pada Gambar 2.7, dan spesifikasi dapat dilihat pada Tabel 2.7.



Gambar 2. 7 Module Step Up MT3608

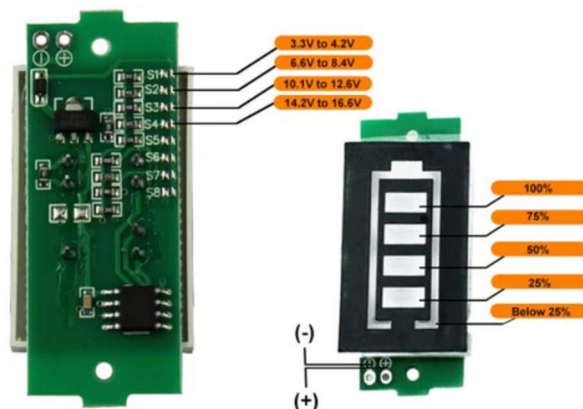
Tabel 2. 7 Spesifikasi Step Up MT3608

Parameter	Detail
Nama Modul	MT3608
<i>Maximum output current</i>	2A
<i>Input voltage</i>	2 v - 24 v
<i>Maximum output voltage</i>	> 5V-28 v
<i>Efficiency</i>	> 93%

4) Indikator Baterai

Modul SPBKBS-10 merupakan indikator kapasitas baterai yang digunakan untuk memantau kondisi daya baterai secara sederhana. Modul ini menampilkan tingkat kapasitas baterai melalui indikator LED sehingga

pengguna dapat mengetahui sisa daya baterai tanpa perlu melakukan pengukuran menggunakan alat tambahan. Pada penelitian ini, SPBKBS-10 digunakan untuk memberikan informasi kondisi baterai 18650 yang menjadi sumber daya utama sistem. Penggunaan modul ini membantu pengguna dalam memantau kebutuhan pengisian ulang baterai sehingga perangkat dapat digunakan secara lebih praktis dan andal. Bentuk fisik modul SPBKBS-10 dapat dilihat pada Gambar 2.8, sedangkan spesifikasinya ditunjukkan pada Tabel 2.8.



Gambar 2. 8 *Battery Indicator*

Tabel 2. 8 Spesifikasi *Battery Indicator*

Parameter	Detail
Nama Modul	Battery Indikator (SPBKBS-10)
Tegangan Input	3.3 v - 30 v
Metode Indikasi	LED Bar Indikator
Aplikasi	Li-ion & LiPo

2.2.10 WiFi (*Wireless Fidelity*)

WiFi (Wireless Fidelity) merupakan teknologi komunikasi nirkabel yang memungkinkan perangkat untuk melakukan pertukaran data melalui jaringan tanpa menggunakan koneksi kabel. Teknologi ini bekerja berdasarkan standar IEEE 802.11 yang menggunakan gelombang radio sebagai media transmisi data.

Pada sistem alat bantu komunikasi visual ini, WiFi digunakan sebagai media komunikasi antara ESP32-S3 dan laptop yang menjalankan model *Whisper Speech*

Recognition. Data audio hasil tangkapan mikrofon digital INMP441 dikirimkan dari ESP32-S3 menuju Laptop melalui jaringan WiFi untuk dilakukan proses transkripsi suara menjadi teks.

Penggunaan WiFi dipilih karena ESP32-S3 telah memiliki modul WiFi bawaan sehingga tidak memerlukan perangkat komunikasi tambahan. Selain itu, WiFi memiliki kemampuan transmisi data yang cukup untuk mendukung pengiriman audio secara *real-time* sehingga sistem dapat bekerja secara responsif pada perangkat *wearable*.

2.2.11 WiFi dan User Datagram Protocol (UDP)

Teknologi komunikasi nirkabel berbasis jaringan lokal digunakan sebagai media transmisi data utama pada sistem ini, memanfaatkan Wi-Fi pada lapisan fisik dan *User Datagram Protocol* (UDP) pada lapisan transport.

1. Teknologi WiFi

Wi-Fi merupakan teknologi jaringan nirkabel yang bekerja berdasarkan standar IEEE 802.11. Pada mikrokontroler, modul Wi-Fi bawaan memungkinkan perangkat bekerja dalam *Station Mode* untuk terhubung ke jaringan lokal. Teknologi ini dipilih karena menyediakan *bandwidth* yang besar, yang sangat dibutuhkan untuk melakukan aliran data audio digital secara berkelanjutan (*continuous streaming*) dengan jangkauan area operasional lokal yang memadai.

2. User Datagram Protocol (UDP)

User Datagram Protocol (UDP) adalah protokol lapisan transport yang bersifat *connectionless* dan *unreliable*. UDP tidak menerapkan mekanisme jabat tangan (*three-way handshake*) maupun pengecekan ulang paket data yang hilang (*no acknowledgment*). Karakteristik utama UDP meliputi *overhead header* yang sangat minimal (8 byte) sehingga proses enkapsulasi data berlangsung sangat cepat. Karakteristik ini membuat UDP sangat ideal untuk aplikasi aliran data (*streaming*) kontinu yang mengutamakan latensi sangat rendah (*near real-time*) dibandingkan keandalan mutlak pengiriman paket.