



PROPOSAL PROYEK AKHIR

**Pengembangan Robot Komunikasi berbasis Ekspresi Wajah
dalam *Block Design Test* untuk Meningkatkan Efikasi Diri**

Putri Cahyani Alfaiqoh
NRP 3223600007

DOSEN PEMBIMBING

Reni Soelistijorini, B.Eng, M.T.
NIP 197104281999032002

Adnan Rachmat Anom Besari, S.ST., M.Sc., Ph.D.
NIP 198509102012121003

**PROGRAM STUDI SARJANA TERAPAN TEKNIK KOMPUTER
DEPARTEMEN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK ELEKTRONIKA NEGERI SURABAYA**

2026

BAB 1

PENDAHULUAN

1.1 LATAR BELAKANG

Block Design Test (BDT) merupakan salah satu instrumen yang digunakan dalam asesmen psikologis untuk mengukur kemampuan kognitif, khususnya kemampuan visual-spasial, penalaran nonverbal, dan pemecahan masalah. Dalam pelaksanaannya, peserta dituntut untuk menyusun balok berdasarkan pola desain tertentu dalam batas waktu yang telah ditentukan. Meskipun BDT berfokus pada pengukuran kemampuan kognitif, performa peserta selama pengerjaan juga dapat dipengaruhi oleh faktor non kognitif, salah satunya adalah efikasi diri. Efikasi diri merujuk pada keyakinan individu terhadap kemampuannya dalam menyelesaikan suatu tugas atau menghadapi tantangan tertentu. Ketika peserta menghadapi pola desain yang semakin kompleks atau tekanan waktu, rendahnya efikasi diri dapat ditunjukkan melalui keraguan, ketegangan, dan menurunnya keyakinan terhadap kemampuan diri. Kondisi tersebut berpotensi mempengaruhi usaha, ketekunan, dan performa peserta selama menyelesaikan tugas BDT [1].

Permasalahan tersebut menunjukkan pentingnya dukungan yang dapat membantu peserta mempertahankan keyakinan terhadap kemampuannya selama menghadapi tantangan dalam pelaksanaan tugas kognitif. Dalam bidang Human-Computer Interaction (HCI) dan Affective Computing, pengembangan sistem interaktif yang mampu merespons kondisi pengguna secara adaptif dapat digunakan untuk memberikan dukungan tersebut. Salah satu pendekatan yang relevan adalah Socially Assistive Robotics (SAR), yaitu penggunaan robot atau agen virtual sebagai pendamping sosial yang memberikan dukungan melalui interaksi verbal maupun nonverbal. Berbeda dengan robot yang berfokus pada bantuan fisik, agen pendamping sosial dapat memberikan penguatan positif, dukungan verbal, dan respons sosial untuk membantu pengguna mempertahankan keterlibatan selama menyelesaikan suatu tugas.

Berbagai penelitian sebelumnya telah menunjukkan bahwa dukungan sosial dan pemberian umpan balik motivasional dapat berperan dalam membantu pengguna menghadapi tugas yang menantang. Pemberian umpan balik yang tepat waktu dapat memberikan penguatan terhadap keberhasilan yang telah dicapai pengguna dan membantu mempertahankan motivasi untuk menghadapi tantangan berikutnya. Namun, sebagian penelitian masih menggunakan robot fisik yang memerlukan perangkat keras khusus, sehingga penerapannya membutuhkan biaya dan kompleksitas yang lebih tinggi. Selain itu, sistem pendamping yang memberikan umpan balik secara berkala atau berdasarkan waktu tertentu belum sepenuhnya memanfaatkan kondisi aktual pengguna dan peristiwa keberhasilan sebagai dasar pemberian intervensi.

Berdasarkan permasalahan tersebut, penelitian ini mengembangkan Robot Komunikasi berbasis ekspresi wajah yang dijalankan melalui platform web pada PC sebagai agen pendamping virtual dalam pelaksanaan *Block Design Test*. Sistem memanfaatkan

webcam, mikrofon, dan speaker yang telah tersedia pada PC sehingga tidak memerlukan perangkat robotik fisik tambahan. Ekspresi wajah pengguna digunakan sebagai sumber informasi utama untuk memperoleh indikasi kondisi pengguna selama mengerjakan BDT. Data wajah yang ditangkap melalui webcam diproses menggunakan *Media Pipe Face Mesh* untuk mengekstraksi landmark wajah, kemudian dianalisis menggunakan model Jaringan Saraf Tiruan (JST) untuk menghasilkan skor probabilitas kondisi visual pengguna. Selain informasi visual, sistem juga memanfaatkan informasi vokal sebagai modalitas pendukung. Ucapan pengguna yang ditangkap melalui mikrofon diproses menggunakan *Speech Recognition* pada *Web Speech API* dan dianalisis berdasarkan kata kunci yang berkaitan dengan keberhasilan, keyakinan, keraguan, atau kesulitan. Informasi visual dan vokal tersebut kemudian diproses menggunakan metode Fuzzy Mamdani untuk menghasilkan skor keputusan pada rentang 0–100%. Pendekatan ini digunakan untuk mengakomodasi ketidakpastian dan dinamika respons manusia selama pengerjaan BDT.

Dalam menentukan waktu pemberian respons, sistem tidak hanya bergantung pada skor kondisi pengguna. Status penyelesaian level BDT digunakan sebagai validasi utama untuk mencegah pemberian respons apresiasi pada kondisi yang tidak sesuai. Integrasi antara kondisi pengguna dan status penyelesaian level diwujudkan melalui mekanisme *Success-Triggered Feedback Loop*. Ketika peserta berhasil menyelesaikan suatu level BDT, sistem mendeteksi peristiwa keberhasilan tersebut dan mengevaluasi kondisi pengguna berdasarkan informasi visual dan vokal yang tersedia. Berdasarkan hasil evaluasi, Robot Komunikasi memberikan respons adaptif berupa motivasi atau apresiasi melalui *Speech Synthesis* pada *Web Speech API* dan speaker PC, serta menampilkan ekspresi wajah pada avatar virtual.

Pemberian respons setelah keberhasilan peserta diharapkan dapat memberikan penguatan positif terhadap pengalaman keberhasilan yang baru dicapai. Dalam konteks efikasi diri, pengalaman keberhasilan merupakan salah satu sumber penting yang dapat memperkuat keyakinan individu terhadap kemampuannya dalam menghadapi tugas. Oleh karena itu, mekanisme *Success-Triggered Feedback Loop* dirancang untuk memberikan dukungan pada momen yang relevan tanpa mengganggu konsentrasi peserta selama proses pengerjaan BDT. Untuk mengetahui efektivitas penggunaan Robot Komunikasi dalam meningkatkan efikasi diri peserta, penelitian ini menggunakan pendekatan eksperimen kuantitatif dengan desain *quasi-experimental*. Pengujian dilakukan melalui perbandingan dua kondisi pengerjaan *Block Design Test*. Pada kondisi pertama, peserta mengerjakan BDT tanpa memperoleh intervensi motivasional dari Robot Komunikasi. Pada kondisi kedua, peserta mengerjakan BDT dengan dukungan Robot Komunikasi yang memberikan respons adaptif berdasarkan ekspresi wajah, respons vokal, dan status penyelesaian level. Hasil dari kedua kondisi tersebut kemudian dibandingkan berdasarkan indikator efikasi diri yang telah ditentukan dalam penelitian.

Dengan demikian, penelitian ini berfokus pada pengembangan Robot Komunikasi berbasis ekspresi wajah dalam pelaksanaan *Block Design Test* serta pengujian efektivitas pemberian umpan balik adaptif dalam meningkatkan efikasi diri peserta. Sistem yang dikembangkan diharapkan dapat menjadi alternatif agen pendamping virtual yang mampu

memberikan dukungan sosial secara adaptif melalui pemanfaatan teknologi pengolahan ekspresi wajah, pengenalan suara, logika fuzzy, dan mekanisme *Success Triggered Feedback Loop* pada platform web.

1.2 PERMASALAHAN

1. Bagaimana merancang Communication Robot berbasis web (client-side) yang mampu mengintegrasikan aliran visual (MediaPipe) dan vokal (Web Speech API) untuk mendeteksi kondisi afektif peserta selama pelaksanaan Block Design Test (BDT)
2. Bagaimana mengimplementasikan metode Logika Fuzzy Mamdani serta Gerbang Pengaman Status Level BDT (*Success/Inprogress*) agar sistem dapat menghasilkan keputusan respons intervensi yang objektif dan adaptif?
3. Bagaimana pengaruh pemberian Motivasi Teks To Speech terhadap tingkat kepercayaan diri dan efikasi diri peserta dalam BDT?

1.3 TUJUAN

1. Merancang sistem Communication Robot berbasis PC yang mampu mendeteksi ekspresi wajah peserta selama pelaksanaan Block Design Test.
2. Mampu menampilkan audio motivasi TTS saat level sukses
3. Menganalisis hasil pengujian dan performa sistem dalam memberikan intervensi sosial guna mendukung peningkatan kepercayaan diri dan efikasi diri peserta selama pelaksanaan BDT.

1.4 BATASAN MASALAH

1. Proyek berfokus pada pengembangan agen virtual (virtual agent) melalui antarmuka layar berbasis web PC.
2. Perilaku robot diwujudkan secara spesifik melalui suara motivasi.
3. Intervensi sosial berupa suara motivasi tidak diberikan secara terus-menerus, melainkan dikendalikan secara adaptif berdasarkan hasil fusi logika fuzzy serta divalidasi oleh gerbang pengaman status keberhasilan level BDT (*Success/Inprogress*).

1.5 MANFAAT

1. Memberikan referensi ilmiah mengenai penerapan Affective Computing, pengolahan citra MediaPipe, dan logika fuzzy pada peramban web untuk mengintervensi kondisi psikologis pengguna secara real-time.
2. Menyediakan alat bantu pengujian Block Design Test yang lebih objektif dan interaktif, sehingga beban mental peserta akibat tekanan pengujian dapat ditekan.

3. Membantu meningkatkan rasa percaya diri, mengurangi kecemasan, serta menjaga stabilitas fokus sepanjang pengerjaan tes melalui pendampingan robot virtual yang komunikatif.

1.6 SISTEMATIKA PENULISAN

Sistematika penulisan laporan Proyek Akhir ini disusun secara sistematis ke dalam lima bab sebagai berikut:

Pendahuluan

Bab 1

Bab ini membahas latar belakang penelitian mengenai pentingnya menjaga stabilitas psikologis peserta saat pengerjaan tes kognitif, permasalahan penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, serta sistematika penulisan laporan

Kajian Pustaka

Bab 2

Bab ini membahas teori-teori yang mendukung perancangan sistem yang diusulkan. Cakupan materi meliputi konsep, ekosistem pemrograman menggunakan Framework Flutter dan GetX. Selain itu, bab ini memaparkan tinjauan penelitian terdahulu yang relevan sebagai acuan pengisi celah ilmiah.

Desain Sistem

Bab 3

Bab ini membahas perancangan arsitektur dan metodologi pengembangan sistem komunikasi robot

Eksperimen dan Analisis

Bab 4

Bab ini membahas proses implementasi, pengujian fungsional sistem, pelaksanaan eksperimen, visual dan audio data yang di outputkan menjadi audio motivasi

Penutup

Bab 5

Bab ini menyajikan kesimpulan akhir yang ditarik berdasarkan hasil pengujian sistem, analisis eksperimen, dan pembuktian hipotesis penelitian yang telah dilakukan.

BAB 2

KAJIAN PUSTAKA

2. 1 DESKRIPSI PERMASALAHAN

Pelaksanaan *Block Design Test* konvensional umumnya berfokus pada hasil pengerjaan dan waktu penyelesaian, sehingga belum memberikan dukungan interaktif terhadap kondisi peserta selama menghadapi peningkatan tingkat kesulitan. Ketika peserta mengalami kesulitan atau berhasil menyelesaikan suatu pola desain, belum terdapat sistem pendamping yang mampu merespons kondisi tersebut secara adaptif. Tekanan waktu dan meningkatnya kompleksitas pola dapat mempengaruhi efikasi diri peserta, yang ditunjukkan melalui munculnya keraguan dan menurunnya keyakinan dalam menghadapi tugas berikutnya. Penggunaan robot sosial fisik sebagai pendamping dapat memberikan dukungan interaktif, tetapi penerapannya memiliki keterbatasan berupa biaya, kebutuhan perangkat keras tambahan, dan kompleksitas pemeliharaan. Oleh karena itu, diperlukan alternatif berupa agen pendamping berbasis perangkat lunak yang dapat dijalankan melalui PC.

Berdasarkan permasalahan tersebut, penelitian ini mengembangkan Robot Komunikasi berbasis ekspresi wajah yang menggunakan webcam untuk memperoleh informasi kondisi peserta selama pengerjaan BDT. Informasi tersebut diproses bersama informasi vokal dan status penyelesaian level untuk menentukan pemberian respons. Melalui mekanisme *Success-Triggered Feedback Loop*, sistem memberikan umpan balik berupa suara motivasi atau apresiasi melalui teknologi *Text-to-Speech* setelah peserta berhasil menyelesaikan suatu level. Pemberian umpan balik secara adaptif tersebut diharapkan dapat memberikan penguatan positif dan mendukung peningkatan efikasi diri peserta dalam menghadapi tingkat kesulitan berikutnya.

2.2 TEORI PENUNJANG

2.2.1 Block Design Test (BDT)

Block Design Test (BDT) merupakan salah satu subtes kinerja (performance subtest) dalam asesmen inteligensi yang digunakan untuk mengukur berbagai kemampuan kognitif, terutama penalaran nonverbal, kemampuan visuospasial, analisis dan sintesis visual, serta integrasi visual-motorik. Dalam pelaksanaannya, peserta diminta untuk menyusun sejumlah balok fisik berwarna merah dan putih sehingga membentuk pola geometris yang sesuai dengan contoh yang diberikan. Pengerjaan setiap pola dilakukan dalam batasan waktu tertentu, sehingga peserta dituntut untuk memproses informasi visual dan menyelesaikan permasalahan secara cepat dan tepat. BDT merupakan salah satu subtes dalam Wechsler Intelligence Scale for Children Fifth Edition (WISC-V) yang berkaitan dengan pengukuran kemampuan penalaran visuospasial dan integrasi visual-motorik pada anak (Wechsler, 2014). Seiring meningkatnya tingkat kesulitan pola, peserta perlu menggunakan kemampuan analisis visual dan strategi pemecahan masalah yang lebih kompleks. Selain kemampuan kognitif, kondisi pengerjaan yang melibatkan batasan waktu dan tingkat kesulitan yang meningkat juga dapat memengaruhi pengalaman peserta selama tes, termasuk konsentrasi, ketekunan, dan keyakinan terhadap kemampuan dirinya. Pelaksanaan BDT tidak hanya

melibatkan proses penyelesaian masalah secara kognitif, tetapi juga menuntut peserta untuk mempertahankan fokus dan usaha ketika menghadapi tugas yang semakin sulit.



Gambar 2.1 Block Design Test

2.2.2 Efikasi Diri

Efikasi diri adalah keyakinan seseorang terhadap kemampuan dirinya dalam menyelesaikan suatu tugas dan mencapai tujuan tertentu. Keyakinan ini mempengaruhi cara seseorang menghadapi tugas yang sulit, seberapa besar usaha yang diberikan, serta seberapa lama seseorang mampu bertahan ketika menghadapi kesulitan. Seseorang yang memiliki efikasi diri tinggi cenderung lebih percaya diri dan tidak mudah menyerah ketika menghadapi tantangan. Sebaliknya, efikasi diri yang rendah dapat membuat seseorang lebih mudah meragukan kemampuannya sendiri dan berhenti berusaha ketika mengalami kegagalan atau kesulitan. Dalam pelaksanaan tugas kognitif yang menantang, seperti Block Design Test (BDT), efikasi diri menjadi salah satu faktor yang dapat mempengaruhi respons individu selama proses pengerjaan. BDT menuntut peserta untuk memecahkan masalah, menggunakan kemampuan visuospasial, serta menyelesaikan tugas dalam batasan waktu. Kondisi tersebut dapat menimbulkan tekanan, terutama ketika tingkat kesulitan tugas meningkat atau peserta mengalami kegagalan dalam menyelesaikan suatu pola. Pada situasi tersebut, keyakinan peserta terhadap kemampuannya dapat mempengaruhi keputusan untuk terus mencoba atau mengurangi usaha yang dilakukan.

Bandura menjelaskan bahwa efikasi diri dapat terbentuk melalui beberapa sumber utama, yaitu pengalaman keberhasilan, pengalaman vikarius, persuasi verbal, serta kondisi fisiologis dan emosional individu. Pengalaman keberhasilan memiliki peran yang penting karena keberhasilan dalam menyelesaikan suatu tugas dapat memberikan pengalaman langsung bahwa individu mampu menghadapi tantangan tersebut. Selain itu, dukungan melalui persuasi verbal juga dapat membantu seseorang mempertahankan keyakinan terhadap kemampuannya, terutama ketika sedang menghadapi kesulitan. Dalam konteks tugas yang menantang, umpan balik positif yang diberikan setelah suatu keberhasilan dapat membantu memperkuat persepsi individu terhadap

kemampuannya. Pengalaman keberhasilan yang disertai dengan dukungan positif dapat mendorong individu untuk tetap percaya diri ketika menghadapi tantangan berikutnya.

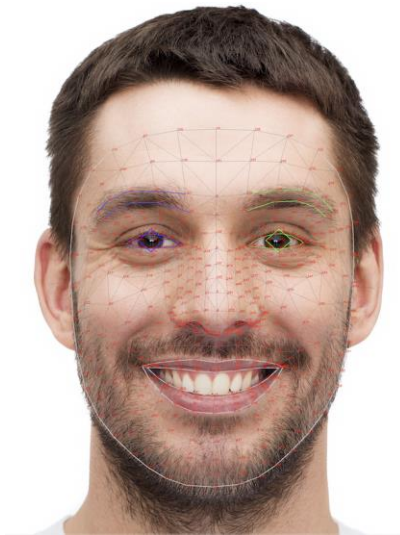
2.2.3 Pengolahan Citra dan Deteksi Wajah (*Media Pipe Face Landmarker*)

Media Pipe Face Landmarker merupakan solusi berbasis pembelajaran mesin yang dikembangkan oleh Google untuk mendeteksi dan melacak karakteristik wajah secara real-time. Teknologi ini memungkinkan sistem memperoleh informasi mengenai struktur wajah dari citra yang berasal dari kamera. Berbeda dengan deteksi wajah yang hanya menentukan keberadaan atau lokasi wajah, Face Landmarker menghasilkan representasi geometris yang lebih rinci melalui titik-titik acuan (facial landmarks) pada berbagai bagian wajah. Proses kerja *Media Pipe Face Landmark* dimulai dengan menerima citra atau frame video sebagai masukan. Sistem kemudian mendeteksi wajah yang terdapat pada frame dan memperkirakan posisi titik-titik penting pada struktur wajah. Titik-titik tersebut mencakup area mata, alis, hidung, bibir, dan kontur wajah. Pada model standar, sistem dapat menghasilkan hingga 478 titik landmark wajah, termasuk informasi tambahan pada area mata, sehingga struktur wajah dapat direpresentasikan secara lebih rinci dalam bentuk data numerik.

Setiap landmark direpresentasikan menggunakan koordinat tiga dimensi yang terdiri atas koordinat x , y , dan z . Koordinat x dan y menunjukkan posisi relatif titik pada bidang citra, sedangkan koordinat z memberikan informasi mengenai kedalaman relatif terhadap struktur wajah. Kumpulan titik tersebut dapat digunakan untuk merepresentasikan bentuk dan perubahan geometris wajah dari waktu ke waktu. Secara umum, representasi landmark dapat dituliskan sebagai:

dengan L merupakan kumpulan landmark wajah, N merupakan jumlah landmark yang dihasilkan, serta x_i , y_i , dan z_i masing-masing menunjukkan koordinat landmark ke- i .

Media Pipe Face Landmarker juga dapat menyediakan informasi tambahan berupa *face blendshapes*. Informasi ini merepresentasikan perubahan karakteristik ekspresi wajah dalam bentuk nilai numerik, seperti perubahan pada bukaan mata, posisi alis, atau bentuk mulut. Dengan demikian, informasi visual dari wajah dapat diubah menjadi data yang dapat dianalisis oleh sistem secara komputasional. Pemrosesan berbasis landmark menganalisis perubahan wajah dilakukan melalui pendekatan geometris. Sebagai contoh, perubahan ekspresi dapat diamati melalui perubahan jarak atau posisi relatif antara titik-titik tertentu pada wajah. Perubahan tersebut dapat dianalisis dari waktu ke waktu untuk memperoleh informasi mengenai kondisi ekspresi wajah pengguna. Pendekatan ini memungkinkan sistem melakukan analisis wajah tanpa harus menyimpan atau mengirimkan seluruh citra wajah ke server secara terus-menerus.



Gambar 2. 2 Visualisasi Facial Landmarks pada MediaPipe Face Landmarker

Sumber: Google MediaPipe, 2024.

2.2.4 Jaringan Saraf Tiruan (Artificial Neural Network)

Jaringan Saraf Tiruan (JST) atau Artificial Neural Network (ANN) merupakan metode komputasi yang dikembangkan dengan mengambil inspirasi dari cara kerja jaringan saraf pada manusia. JST dapat mempelajari hubungan dan pola yang terdapat pada data sehingga dapat digunakan untuk melakukan pengenalan pola, prediksi, maupun klasifikasi. Proses tersebut dilakukan melalui sejumlah unit pemrosesan yang disebut neuron dan saling terhubung dalam beberapa lapisan. Pada umumnya, JST terdiri atas input layer, hidden layer, dan output layer. Data masukan diterima oleh input layer dan kemudian diproses secara bertahap melalui lapisan-lapisan berikutnya. Setiap neuron melakukan perhitungan berdasarkan nilai masukan, bobot (weight), dan bias sebelum menghasilkan keluaran melalui fungsi aktivasi. Secara sederhana, proses perhitungan pada sebuah neuron dapat dituliskan sebagai:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

Di mana

x_i : Nilai masukan

w_i : bobot

b : bias

f : fungsi aktivasi

Salah satu bentuk JST yang banyak digunakan untuk mengolah data numerik adalah fully connected neural network. Pada struktur ini, setiap neuron pada satu lapisan

terhubung dengan neuron pada lapisan berikutnya. Model seperti ini dapat digunakan ketika data yang akan dianalisis telah diubah menjadi sekumpulan nilai numerik, misalnya fitur yang diperoleh dari hasil ekstraksi data wajah. Dalam pengolahan data wajah, hasil ekstraksi landmark dapat memiliki rentang nilai yang berbeda-beda. Oleh karena itu, data masukan dapat dinormalisasi terlebih dahulu agar perbedaan skala antar fitur tidak terlalu besar. Salah satu metode yang dapat digunakan adalah Min-Max Scaling, dengan rumus:

$$x' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Di mana

x : data awal

x' : data masukan model

$X_{max} - X_{min}$: nilai maximum dan minimum,

Untuk menjalankan model pembelajaran mesin pada aplikasi berbasis web, salah satu pustaka yang dapat digunakan adalah TensorFlow.js. TensorFlow.js merupakan pustaka JavaScript yang memungkinkan model pembelajaran mesin dijalankan secara langsung pada lingkungan web. Dengan pendekatan ini, proses inferensi dapat dilakukan pada perangkat pengguna melalui peramban tanpa harus selalu mengirimkan data ke server. Pemrosesan di sisi klien dapat memberikan keuntungan bagi aplikasi yang membutuhkan respons secara langsung. Data yang diperoleh dari kamera dapat diproses secara lokal sehingga ketergantungan terhadap komunikasi dengan server dapat dikurangi. Hal ini juga dapat membantu menjaga efisiensi proses pemrosesan, khususnya pada aplikasi yang memerlukan analisis data secara berulang dan berkelanjutan. Dalam proses klasifikasi, JST menghasilkan keluaran berupa nilai yang menunjukkan kemungkinan data masukan termasuk ke dalam kelas tertentu. Apabila model memiliki beberapa kelas ekspresi, maka keluaran model dapat direpresentasikan dalam bentuk vektor probabilitas:

$$P = [P_1, P_2, \dots, P_k]$$

Di mana

P_k = menunjukkan nilai probabilitas untuk kelas ke- k . Kelas dengan nilai probabilitas tertinggi kemudian dapat digunakan sebagai hasil klasifikasi.

p = Probabilitas

2.2.5 Pengenalan Suara (*Web Speech API*)

Web Speech API merupakan sekumpulan antarmuka pemrograman yang tersedia pada peramban web untuk mendukung pemrosesan suara. API ini menyediakan dua fungsi utama, yaitu *Speech Recognition* untuk mengubah suara menjadi teks (*speech-to-text*) dan *Speech Synthesis* untuk mengubah teks menjadi suara (*text-to-speech*). Kedua fungsi tersebut memungkinkan aplikasi berbasis web untuk melakukan interaksi dengan pengguna melalui input dan output suara. *Speech Recognition* digunakan untuk menerima ucapan pengguna melalui mikrofon dan mengubahnya menjadi data teks. Proses ini memungkinkan sistem mengenali kata atau kalimat yang diucapkan oleh

pengguna secara langsung. Hasil pengenalan suara dapat digunakan sebagai masukan bagi sistem untuk memahami informasi tertentu dari ucapan pengguna, misalnya respons, perintah, atau kata-kata yang berkaitan dengan kondisi pengguna. Sementara itu, Speech Synthesis mengubah teks menjadi suara yang dapat didengarkan melalui perangkat keluaran audio. Sistem dapat menghasilkan suara berdasarkan teks yang telah ditentukan atau dibuat secara dinamis. Fitur ini memungkinkan aplikasi memberikan respons suara secara langsung dan dapat digunakan untuk menyampaikan informasi, instruksi, maupun pesan tertentu kepada pengguna.

2.2.6 Logika Fuzzy Mamdani

Logika fuzzy merupakan salah satu metode yang dapat digunakan untuk mengolah data yang memiliki nilai tidak pasti atau berada di antara beberapa kondisi. Berbeda dengan logika biner yang hanya menggunakan nilai benar atau salah, logika fuzzy suatu data yang memiliki tingkat keanggotaan tertentu dalam suatu kategori. Pendekatan ini sesuai digunakan pada permasalahan yang melibatkan kondisi manusia, karena perubahan perilaku atau ekspresi seseorang biasanya tidak selalu dapat ditentukan secara tegas. Salah satu metode inferensi fuzzy yang banyak digunakan adalah metode Mamdani. Metode ini menggunakan aturan berbentuk IF-THEN untuk menghubungkan variabel masukan dengan keluaran berdasarkan pengetahuan yang telah ditentukan. Proses inferensi pada metode Mamdani secara umum terdiri atas tiga tahapan, yaitu fuzzifikasi, evaluasi aturan, dan defuzzifikasi.

1. Fuzzifikasi

Fuzzifikasi merupakan tahap untuk mengubah nilai masukan dalam bentuk angka tegas menjadi nilai derajat keanggotaan pada himpunan fuzzy. Nilai tersebut dapat memiliki tingkat keanggotaan antara 0 hingga 1. Sebagai contoh, suatu nilai masukan dapat memiliki tingkat keanggotaan tertentu pada kategori “rendah”, “sedang”, dan “tinggi”. Nilai derajat keanggotaan tersebut menunjukkan seberapa kuat suatu masukan termasuk ke dalam kategori tertentu.

2. Evaluasi Aturan

Setelah proses fuzzifikasi, sistem mengevaluasi aturan fuzzy yang telah ditentukan. Aturan tersebut umumnya ditulis dalam bentuk IF-THEN, misalnya:

IF nilai masukan A tinggi AND nilai masukan B tinggi

THEN keluaran berada pada kategori tertentu.

Pada metode Mamdani, beberapa masukan dapat digunakan secara bersamaan untuk menentukan keluaran. Operator AND digunakan untuk menggabungkan kondisi pada aturan dan menghasilkan nilai aktivasi berdasarkan tingkat keanggotaan masukan.

3. Defuzzifikasi

Tahap terakhir adalah defuzzifikasi, yaitu proses mengubah hasil inferensi fuzzy menjadi satu nilai keluaran yang tegas (crisp value). Salah satu metode yang umum

digunakan adalah metode Centroid atau Center of Area. Metode ini menentukan titik pusat dari area hasil agregasi himpunan fuzzy. Secara umum, nilai keluaran dapat dirumuskan sebagai:

$$z * = \frac{\sum_{i=1}^n \mu(z_i) \times z_i}{\sum_{i=1}^n \mu(z_i)}$$

z^* sebagai nilai keluaran akhir.

$\mu(z)$ sebagai derajat keanggotaan keluaran pada setiap nilai z .

2.1 PENELITIAN TERKAIT

Dalam penelitian yang dilakukan oleh R. Zhao, J. Kim, dan S. Lee. berjudul “Motivational Feedback Systems in Human Computer Interaction” (2022), penulis mengembangkan sebuah sistem umpan balik motivasi otomatis untuk meningkatkan performa pengguna dalam menyelesaikan tugas-tugas berbasis digital. Pemilihan arsitektur umpan balik reaktif ini didasarkan pada kebutuhan untuk menekan hambatan non-kognitif seperti kecemasan dan tekanan batas waktu. Dengan adanya penguatan positif, pengguna tetap dapat mempertahankan fokus meskipun menghadapi tingkat kesulitan tugas yang terus meningkat. Penelitian ini membahas bagaimana sistem bekerja saat mendeteksi momentum keberhasilan pengguna secara real-time. Pengujian menunjukkan bahwa sistem ini mampu meningkatkan keterlibatan emosional pengguna serta mereduksi tingkat stres secara signifikan dibandingkan sistem konvensional tanpa umpan balik[4].

Sementara itu, A. R. Santos et al. dalam artikelnya yang berjudul “Adaptive Human Robot Interaction Based on User State Detection” (2023), merancang sebuah sistem interaksi manusia dan robot HRI adaptif berbasis deteksi kondisi pengguna. Penulis menggunakan konfigurasi ini karena kemampuannya untuk mendeteksi status pengerjaan tugas secara objektif sehingga robot dapat memberikan intervensi sosial pada waktu yang paling presisi. Penelitian ini menggunakan pendekatan pemantauan indikator perilaku secara real time tanpa mengganggu kenyamanan kognitif pengguna. Hasilnya menunjukkan bahwa sistem berhasil mengoptimalkan waktu intervensi robot pendamping secara reaktif tepat setelah pengguna menyelesaikan target, membuktikan bahwa penentuan yang adaptif sangat efektif untuk menjaga stabilitas konsentrasi pengguna[7].

Penelitian oleh Kim dan Lee yang berjudul “The Effect of Avatar-Based Interaction on User Engagement” (2021), menawarkan pendekatan berbeda dengan memanfaatkan interaksi berbasis avatar virtual untuk memediasi komunikasi dan mempengaruhi psikologis pengguna. Penulis memilih metode representasi grafis digital melalui agen layar gawai ini karena dapat mengurangi kebutuhan komponen mekanis robot fisik yang kompleks serta menekan biaya pembuatan sistem. Dalam proses interaksinya, avatar dilengkapi dengan ekspresi wajah dinamis untuk memicu persepsi positif. Hasil pengujian menunjukkan efektivitas peningkatan rasa percaya diri dan kenyamanan emosional pengguna secara signifikan sesaat setelah menerima respons visual dari avatar tersebut[5].

Adapun M. Tamura, K. Tanaka, dan Y. Sato, dalam artikelnya yang berjudul “Advances in Neural Text-to-Speech Systems for Human Interaction” (2023), menyajikan sebuah studi mengenai pemanfaatan elemen suara yang meliputi perubahan intonasi, volume, dan kecepatan artikulasi dalam komunikasi digital. Penulis menekankan bahwa integrasi antara ekspresi visual dan respons vokal yang positif sangat krusial dalam menyampaikan dukungan emosional secara utuh. Dalam ulasannya, disebutkan bahwa penggunaan penanda asinkron untuk menyelaraskan pemutaran audio dengan pergerakan grafik adalah kunci utama agar pesan motivasi dapat diterima dengan baik oleh psikologis pengguna tanpa mengalami kendala keterlambatan sistem[3].

Terakhir, penelitian dari J. Lazar, J. H. Feng, dan H. Hochheiser. dalam artikel “Human-Computer Interaction Research and Practice” (2021), mengusulkan sebuah panduan perancangan antarmuka visual yang dinamis dan terstruktur untuk menekan beban mental. Penulis memilih pendekatan orientasi layar tertentu karena visualisasi yang luas terbukti efektif memisahkan area interaksi taktil utama dengan area informasi pendukung secara seimbang. Lapisan atas digunakan untuk menampilkan instruksi pengujian, area tengah mengelola pengerjaan tugas, dan area samping dialokasikan untuk visualisasi agen virtual. Hasil pengujian kegunaan menunjukkan bahwa strategi tata letak ini mampu meningkatkan kenyamanan visual dan keterbacaan data sekaligus menjaga fokus pengguna selama proses pengujian kognitif sistem[9].

Penelitian mengenai personalisasi perilaku agen sosial secara otonom juga dikembangkan oleh Maroto-Gómez dkk. (2023). Mereka mengusulkan metode penyesuaian interaksi berbasis HRI untuk melakukan profil pengguna dan personalisasi perilaku robot secara dinamis. Metode Rule-Based Interaction Policy yang diterapkan terbukti efektif menghasilkan respons adaptif yang sesuai dengan karakteristik performa subjek secara personal guna meningkatkan kualitas hubungan interaksi jangka panjang [17].

Sementara itu, dalam konteks asesmen kemampuan kognitif anak, instrumen yang menjadi standar global adalah Wechsler Intelligence Scale for Children Fifth Edition (WISC-V) yang disusun oleh Wechsler (2014). Panduan administrasi dan penilaian ini menyediakan standarisasi batasan waktu yang ketat serta penjenjangan tingkat kesulitan tugas pada sub-tes Block Design Test (BDT). Standardisasi klinis inilah yang menentukan nilai ambang batas (threshold) baku untuk mengukur tingkat efisiensi maupun beban kognitif subjek secara presisi selama proses pengujian berlangsung [18].

Tabel 2.1 Perbandingan penelitian terdahulu

Penelitian	Judul Penelitian	Metode / Pendekatan	Hubungan dengan Proyek Akhir
Zhao(2022)[4]	<i>Motivational Feedback Systems in Human Computer Interaction Real-time motivational</i>	Real-time motivational feedback loop	Diadopsi sebagai landasan teori pemberian penguatan positif sesaat setelah

			sukses.
Santos(2023)[7]	<i>Adaptive Human Robot Interaction Based on User State Detection</i>	Adaptive HRI berbasis pemantauan indikator perilaku reaktif.	Diadopsi untuk menerapkan konsep Silent Monitoring saat berpikir dan memicu respons setelah target terpenuhi.
Kim & Lee(2023)[5]	<i>The Effect of Avatar-Based Interaction on User Engagement</i>	Agen virtual dengan ekspresi wajah dinamis.	Diadopsi sebagai bentuk platform luaran (Avatar berbasis layar PC menggunakan kerangka kerja Flutter).
Tamura dkk. (2023)[3]	<i>Advances in Neural Text-to-Speech Systems for Human Interaction</i>	Sinkronisasi elemen suara (intonasi/volume) dan grafis.	Diadopsi pada Output Layer untuk menyelaraskan facial dan vocal behavior secara simultan melalui pemrograman asinkron Dart.
Lazar (2021)[9]	<i>Human-Computer Interaction Research and Practice</i>	Alokasi tata letak (layout) antarmuka terstruktur dinamis	Diadopsi sebagai referensi pembagian area visual overlay intervensi dan area utama aplikasi.
Maroto Gomez(2023)[17]	<i>Active learning based on computer vision and human–robot interaction for the user profiling and behavior personalization of an autonomous social robot</i>	Rule-Based Interaction Policy untuk personalisasi perilaku.	Diadopsi sebagai metode dasar Decision Engine pada Logic Layer (GetX) untuk menentukan jenis intervensi afektif avatar.
Wechsles(2014)[18]	<i>Wechsler Intelligence Scale</i>	Standarisasi administrasi,	Diadopsi sebagai parameter batasan

	<i>for Children Fifth Edition (WISC-V) Administration and Scoring Manual</i>	penjenjangan level, dan skoring kognitif anak.	durasi maksimal dan rumus penentuan nilai Threshold pada tiap level BDT
--	--	--	---

BAB 3

DESAIN SISTEM

3.1 DESKRIPSI SOLUSI

Sistem yang dirancang dalam penelitian ini merupakan Komunikasi Robot berbasis web yang dijalankan melalui PC. Sistem ini berperan sebagai asisten pendamping virtual dalam pelaksanaan *Block Design Test (BDT)*. Berbeda dengan robot fisik, sistem yang dikembangkan menghadirkan interaksi otomatis melalui browser dengan memanfaatkan perangkat keras yang telah tersedia pada PC, yaitu webcam, mikrofon, dan speaker. Pendekatan sistem memberikan respons adaptif terhadap kondisi peserta tanpa memerlukan perangkat robotik fisik tambahan. Sistem dirancang untuk mendukung interaksi multimodal melalui pendekatan *Multimodal Perception*. Sistem memperoleh informasi pengguna melalui dua sumber data secara paralel dan real-time, yaitu aliran visual dan aliran vokal. Kedua aliran data tersebut diproses secara terpisah sebelum digabungkan pada tahap *Late Fusion*.

Aliran visual memanfaatkan webcam PC untuk menangkap citra wajah pengguna selama pelaksanaan BDT. Data citra wajah diproses secara langsung pada browser menggunakan pustaka *MediaPipe Face Mesh* untuk mendeteksi dan mengekstrak titik-titik landmark wajah. Informasi geometris dari landmark tersebut kemudian digunakan sebagai fitur masukan bagi model Jaringan Saraf Tiruan yang diimplementasikan menggunakan TensorFlow.js. Model menghasilkan nilai probabilitas yang merepresentasikan kondisi afektif pengguna berdasarkan ekspresi wajah yang terdeteksi. Aliran vokal memanfaatkan mikrofon PC untuk menangkap ucapan spontan pengguna selama pelaksanaan tes. Data suara diproses menggunakan *Web Speech API* melalui fitur *Speech Recognition* yang tersedia pada browser. Hasil pengenalan suara kemudian dianalisis berdasarkan kemunculan dan kepadatan kata kunci tertentu yang berkaitan dengan keberhasilan atau respons positif pengguna. Hasil analisis tersebut menghasilkan skor representasi dari sinyal vokal pengguna.

Nilai yang diperoleh dari kedua aliran sensor kemudian digabungkan menggunakan metode. Setiap input dinormalisasi ke dalam rentang nilai 0 hingga 1 atau 0% hingga 100% agar memiliki skala yang seragam. Selanjutnya, nilai dari aliran visual dan vokal dikombinasikan menggunakan bobot tertentu untuk menghasilkan skor fusi akhir. Skor tersebut merepresentasikan tingkat keyakinan sistem terhadap kondisi afektif atau respons positif pengguna berdasarkan kombinasi informasi visual dan vokal. Hasil fusi kemudian digunakan sebagai dasar pengambilan keputusan interaksi pada sistem. Berdasarkan nilai skor akhir dan aturan keputusan yang telah ditentukan, sistem menentukan apakah diperlukan pemberian respons motivasional kepada pengguna. Respons Komunikasi Robot diwujudkan melalui keluaran audio menggunakan teknologi *Web Speech API* melalui fitur *Speech Synthesis*. Audio tersebut diputar melalui speaker PC sebagai bentuk umpan balik verbal.

3.1.1 Definisi Operasional Aspek *User Confidence*

Dalam Project Akhir ini, kepercayaan diri pengguna didefinisikan secara operasional sebagai tingkat keyakinan pengguna yang direpresentasikan dalam bentuk skor kuantitatif pada rentang 0–100%. Penilaian tersebut diperoleh berdasarkan kombinasi indikator visual dan vokal yang muncul secara spontan selama pengguna mengerjakan *Block Design Test (BDT)* di depan PC. Indikator visual diperoleh melalui pengamatan terhadap ekspresi wajah dan arah tatapan pengguna selama mengerjakan tes. Sementara itu, indikator vokal diperoleh melalui analisis ucapan pengguna berdasarkan kata atau frasa yang menunjukkan keyakinan terhadap hasil pengerjaan.

Kepercayaan Diri Tinggi ditunjukkan melalui kondisi visual berupa fokus tatapan yang relatif stabil pada objek balok, ekspresi wajah yang rileks, atau munculnya senyuman setelah pengguna berhasil menyelesaikan suatu pola. Dari aspek vokal, kondisi ini dapat ditunjukkan melalui ucapan yang mengindikasikan keyakinan atau keberhasilan, seperti “Bisa!”, “Selesai!”, atau “Yes!”. Sebaliknya, Kepercayaan Diri Rendah ditunjukkan melalui indikator visual berupa ekspresi kebingungan atau frustrasi, seperti dahi yang mengerut, ketegangan pada ekspresi wajah, serta arah tatapan yang kurang stabil atau menjauhi objek balok. Dari aspek vokal, kondisi ini dapat ditunjukkan melalui periode tanpa ucapan dalam durasi tertentu atau ucapan yang mengindikasikan keraguan dan kesulitan, seperti “susah”, “bingung”, atau “tidak bisa”. Indikator visual dan vokal tersebut selanjutnya diproses oleh sistem untuk menghasilkan skor indikator kepercayaan diri pengguna dalam rentang 0–100%. Skor ini merepresentasikan tingkat kekuatan indikator perilaku yang berkaitan dengan kepercayaan diri pengguna selama pelaksanaan BDT. Nilai tersebut digunakan sebagai dasar pengambilan keputusan oleh Komunikasi Robot dalam menentukan bentuk dan tingkat respons yang diberikan. Skor yang dihasilkan tidak dimaksudkan sebagai pengukuran psikologis atau diagnosis klinis terhadap tingkat kepercayaan diri pengguna.

3.1.2 Parameter Komunikasi Robot

Parameter pada Komunikasi Robot berbasis web dibagi menjadi tiga bagian, yaitu parameter input, proses dan output. Parameter input digunakan untuk memperoleh informasi mengenai kondisi pengguna selama mengerjakan *Block Design Test (BDT)*, Parameter proses digunakan untuk mengolah data visual dan vokal menjadi indikator numerik yang dapat dianalisis oleh sistem. Sementara itu, parameter output digunakan untuk menentukan bentuk respons yang diberikan oleh Komunikasi Robot.

A. Parameter Input

Parameter input yang digunakan dalam penelitian ini terdiri atas parameter visual dan parameter vokal.

1. Parameter Visual

Parameter visual diperoleh dari data wajah pengguna yang ditangkap melalui webcam PC. Sistem menggunakan *Media Pipe Face Mesh* untuk mendeteksi titik-titik landmark pada wajah. Dari 468 titik landmark yang tersedia, sistem menggunakan titik-titik yang berkaitan dengan area wajah yang menjadi fokus pengamatan, seperti area alis, mata, dan mulut. Perubahan posisi relatif titik-titik tersebut dianalisis setelah dilakukan proses normalisasi. Data ini digunakan untuk memperoleh informasi mengenai perubahan ekspresi wajah dan arah tatapan pengguna selama pengerjaan tes. Hasil analisis tersebut kemudian digunakan sebagai salah satu indikator perilaku yang berkaitan dengan tingkat kepercayaan diri pengguna.

2. Parameter Vokal

Parameter vokal diperoleh dari ucapan pengguna yang ditangkap melalui mikrofon PC. Ucapan tersebut dikonversi menjadi teks menggunakan fitur *Speech Recognition* dari *Web Speech API* yang tersedia pada browser. Teks hasil pengenalan suara kemudian dianalisis menggunakan metode pencocokan kata kunci. Sistem memeriksa kemunculan kata atau frasa tertentu yang telah ditentukan sebelumnya, seperti kata yang menunjukkan keberhasilan, keyakinan, keraguan, atau kesulitan. Frekuensi kemunculan kata kunci dan hasil analisis tersebut digunakan sebagai indikator tambahan untuk merepresentasikan kondisi pengguna selama pelaksanaan tes.

B. Parameter Proses

Parameter proses merupakan hasil pengolahan data visual dan vokal yang digunakan sebagai masukan dalam proses pengambilan keputusan sistem. Data visual yang diperoleh dari webcam diproses melalui *MediaPipe Face Mesh*, normalisasi koordinat landmark, dan model Jaringan Saraf Tiruan (JST) untuk menghasilkan nilai probabilitas visual dalam rentang 0–100%. Sementara itu, data vokal yang diperoleh dari mikrofon diproses melalui *Speech Recognition* dan pencocokan kata kunci untuk menghasilkan nilai probabilitas vokal dalam rentang 0–100%. Kedua nilai tersebut kemudian diproses menggunakan metode inferensi Fuzzy Mamdani. Proses ini menghasilkan sebuah nilai tegas atau crisp value yang disebut sebagai skor indikator kepercayaan diri (z^*) dengan rentang 0–100%.

C. Parameter Output

Parameter output merupakan respons yang diberikan oleh Communication Robot berdasarkan status penyelesaian level BDT dan skor indikator kepercayaan diri. Respons utama sistem dipicu ketika status level berubah menjadi Completed. Selanjutnya, skor z^* digunakan untuk menentukan tingkat respons yang diberikan.

1. Respons Diam

Pada kondisi tertentu, sistem tidak memberikan suara maupun intervensi verbal kepada peserta. Mode ini diterapkan ketika sistem mendeteksi bahwa pengguna sedang berada dalam tahap pengerjaan *InProgress* dengan tingkat keyakinan yang masih rendah, atau ketika tidak terdapat kondisi yang memerlukan pemberian umpan balik. Penerapan Silence Mode bertujuan untuk mencegah robot memberikan intervensi yang tidak diperlukan, sehingga konsentrasi dan fokus pengguna selama menyelesaikan pola balok tetap terjaga dengan baik.

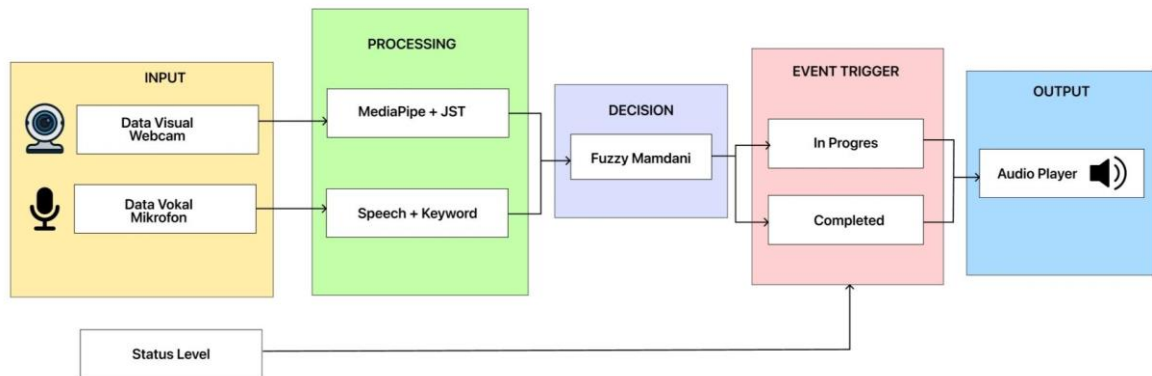
2. Respons Motivasi

Ketika sistem mendeteksi kondisi yang memerlukan dukungan, Komunikasi Robot memberikan respons berupa kalimat motivasi melalui suara. Respons ini dapat berupa dorongan singkat yang bertujuan untuk membantu pengguna mempertahankan keyakinan dan melanjutkan proses pengerjaan tes. Respons suara dihasilkan menggunakan fitur *Speech Synthesis* dari *Web Speech API*, yang mengubah teks motivasi menjadi suara dan memutarkannya melalui speaker PC.

3. Respons Apresiasi

Setelah pengguna berhasil menyelesaikan suatu level BDT, sistem memberikan respons apresiasi sebagai bentuk penguatan positif. Respons ini suara apresiasi yang diputar melalui speaker PC. Pemberian respons tersebut dikendalikan oleh mekanisme *Success Triggered Feedback Loop*. Ketika status penyelesaian level terdeteksi berhasil, sistem memproses informasi level yang telah diselesaikan dan menentukan bentuk respons yang sesuai. Selanjutnya, sistem memberikan umpan balik suara kepada pengguna.

3.2 PERANCANGAN SISTEM



Gambar 3.1 Desain Sistem Communication Robot

3.2.1 Pemrosesan Wajah

Proses pengolahan data wajah dilakukan pada sisi klien pada web browser melalui tahapan berikut:

1. Ekstraksi Landmark Geometri Wajah:

Media Pipe Face Mesh digunakan untuk mendeteksi struktur wajah dan menghasilkan koordinat landmark wajah. Landmark tersebut direpresentasikan dalam bentuk koordinat spasial yang kemudian digunakan sebagai fitur numerik untuk proses analisis

2. Normalisasi Data (*Min-Max Scaling*):

Koordinat piksel mentah yang dihasilkan bersifat sensitif terhadap jarak wajah pengguna terhadap kamera. Oleh karena itu, dilakukan normalisasi ke dalam rentang skala desimal 0.0 - 1.0) untuk mereduksi distorsi jarak:

$$X_{normal} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Di mana

X_{normal} = nilai koordinat setelah normalisasi;

X = nilai koordinat asli;

X_{min} = nilai minimum koordinat;

X_{max} = nilai maksimum koordinat.

3. Klasifikasi JST Ringan

Data koordinat yang telah ternormalisasi selanjutnya diumpankan ke dalam model Jaringan Saraf Tiruan (JST). Arsitektur model ini telah diekspor ke dalam format JSON agar dapat dieksekusi secara lokal pada memori penjelajah web menggunakan pustaka [TensorFlow.js](https://www.tensorflow.org/js). Pada lapisan keluaran (output layer), fungsi aktivasi Softmax diterapkan untuk menghitung

distribusi probabilitas dari tiga kondisi psikologis yang diuji, yaitu Ragu-Ragu, Fokus, dan Percaya Diri. Skor probabilitas wajah tersebut dirumuskan sebagai berikut:

$$P(y = j | x) = \frac{e^{a_j}}{\sum_k^k e^{a_k}} = 1$$

Di mana

$P(y = j | x)$ = probabilitas input termasuk ke dalam kelas (j);

k = jumlah kelas keluaran model;

a_j = nilai keluaran neuron untuk kelas (j);

Nilai probabilitas dari kelas yang merepresentasikan kondisi positif digunakan sebagai basis Skor Visual yang didefinisikan melalui persamaan berikut:

$$P_{\text{visual}} = P_{\text{ragu}} + P_{\text{fokus}} + P_{\text{percaya diri}}$$

Setelah nilai probabilitas visual diperoleh dari hasil akumulasi kelas positif pada tahapan sebelumnya, langkah berikutnya adalah mentransformasikan nilai tersebut ke dalam skala standar yang lebih intuitif untuk interpretasi pengguna. Nilai probabilitas yang berada pada rentang desimal (0.0) hingga (1.0) dikonversi ke dalam bentuk persentase dengan rentang skala (0-100%). Proses ini dirumuskan sebagai berikut:

$$X1 = P_{\text{visual}} \times 100$$

Di mana

x1 = Skor Visual akhir dalam bentuk persentase 0-100

p_{visual} = total probabilitas gabungan dari kelas indikator positif.

3.2.2 Pemrosesan Suara

Pemrosesan sinyal vokal dirancang untuk mendeteksi indikator verbal dari pengguna dengan memanfaatkan fitur bawaan (native Web API) dari peramban modern:

1. Transkripsi STT (*Speech-to-Text*):

Suara yang ditangkap oleh mikrofon diubah menjadi teks secara real-time menggunakan fitur Speech Recognition. Pada implementasi sistem berbasis web, proses ini menggunakan antarmuka SpeechRecognition atau webkitSpeechRecognition sesuai dengan dukungan browser yang digunakan. Hasil dari proses ini berupa teks yang berasal dari ucapan pengguna. Contohnya adalah kata atau kalimat seperti “bisa”, “selesai”, “yakin”, “susah”, atau “bingung”. Teks hasil transkripsi tersebut kemudian digunakan untuk proses pencocokan kata kunci.

2. Kalkulasi Skor Kata Kunci

Untaian teks hasil transkripsi tersebut dicocokkan dengan basis data leksikon lokal di dalam aplikasi web yang berisi kumpulan kata kunci yang merepresentasikan keyakinan, kepastian, atau ekspresi positif. Skor validitas vokal dihitung menggunakan persentase perbandingan:

$$p = \frac{\text{Jumlah keyword}}{\text{Total Kata}} \times 100\%$$

3.1 Tabel Kondisi Pemrosesan Vokal

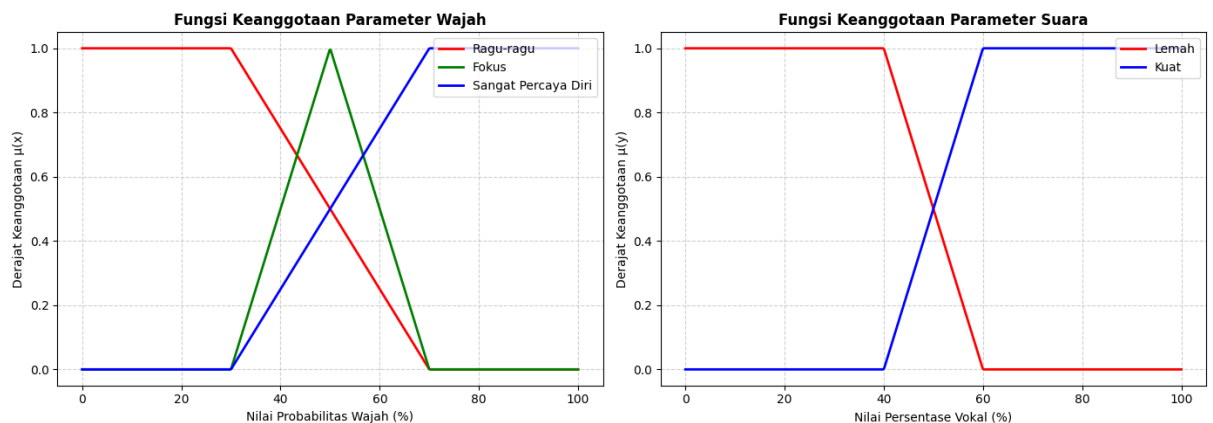
Kondisi Fuzzy	Rentang	Karakteristik Kata Kunci	Kata
Lemah	0-50%	Hanya sedikit kata kunci yang menunjukkan keyakinan / keberhasilan	<i>bingung, susah, sulit, ragu, tidak bisa</i>
Kuat	50-100%	Lebih banyak kata kunci yang menunjukkan keyakinan dan keberhasilan	<i>bisa, yakin, yes, sudah, berhasil, selesai</i>

3.2.3 Fuzzifikasi Nilai Derajat Keanggotaan

Untuk mengakomodasi tingkat ketidakpastian dari ekspresi manusia yang dinamis, nilai numerik mentah dari sensor wajah dan suara dikonversi ke dalam ranah logika kabur *Fuzzy Logic* melalui fungsi derajat keanggotaan :

1. Fungsi Keanggotaan Wajah: Memetakan rentang probabilitas input ke dalam tiga himpunan linguistik utama, yakni Ragu-ragu, Fokus, dan Sangat Percaya Diri dalam rentang skala 0 - 100%.
2. Fungsi Keanggotaan Suara: Memetakan nilai input vokal ke dalam himpunan Lemah dan Kuat.

Ilustrasi visual dari kurva fungsi keanggotaan parameter wajah direpresentasikan melalui diagram berikut:



3.2 Kurva Fungsi Keanggotaan Input Visual dan Vokal

3.2.4 Proses Inferensi Logika Fuzzy (Fuzzy Logic & Decision)

Untuk mengakomodasi ketidakpastian (uncertainty) dari sinyal masukan manusia serta menghasilkan keputusan yang adaptif, sistem mengimplementasikan metode Inferensi Logika Fuzzy Mamdani. Proses penggabungan nilai dari kedua modalitas (wajah dan suara) dilakukan di dalam komponen Web State Controller melalui tiga tahapan matematis berikut:

1. Fuzzifikasi:

Sistem menerima masukan parameter numerik tegas berupa Skor Probabilitas Wajah (X1) dan Skor Kata Kunci Suara (X2). Kedua nilai ini dipetakan ke dalam derajat keanggotaan fuzzy (μ) menggunakan kurva fungsi keanggotaan yang telah ditentukan.

2. Evaluasi Basis Aturan

Setelah seluruh nilai parameter masukan tegas berhasil dikonversi ke dalam ranah linguistik melalui proses fuzzifikasi, tahapan berikutnya adalah mengevaluasi hubungan antar variabel menggunakan sekumpulan aturan yang dibentuk dalam bentuk implikasi IF-THEN. Pada sistem ini, evaluasi aturan dirancang untuk menggabungkan dua premis masukan yakni derajat keanggotaan ekspresi wajah dan derajat keanggotaan suara untuk menentukan arah respons sistem. Metode inferensi Mamdani menggunakan fungsi MIN untuk menghitung tingkat kekuatan aktivasi atau yang biasa disebut dengan nilai a dari setiap aturan ke i . Secara matematis, proses pencarian nilai dirumuskan sebagai berikut:

$$a_i = \min(\mu_{Ai}(x_1), \mu_{Bi}(x_2))$$

Di mana

a_i : menyatakan kekuatan atau bobot aktivasi dari aturan ke i ;
 i : 1, 2,... (total keseluruhan aturan yang ada di dalam matriks);
 $\mu_{Ai}(x_1)$: derajat keanggotaan input visual;
 $\mu_{Bi}(x_2)$: derajat keanggotaan input vokal.

Setelah nilai dari seluruh aturan berhasil dihitung, langkah berikutnya adalah menggabungkan atau mengagregasi seluruh daerah fuzzy dari tiap-tiap aturan yang aktif. Proses komposisi antar aturan menggunakan fungsi maksimum, dimana sistem mengambil nilai tertinggi dari irisan daerah keluaran aturan-aturan tersebut. Persamaan agregasi seluruhnya dinyatakan sebagai berikut:

$$\mu_{solusi}(z) = \max(a_1, a_2, \dots, a_n)$$

Di mana

$\max$: Fungsi matematika untuk mencari nilai tertinggi di himpunan a predikat yang aktif.

Dari persamaan di atas, terbentuklah satu fungsi keanggotaan daerah hasil akhir yang merepresentasikan gabungan seluruh keputusan linguistik yang kemudian dibawa ke tahap akhir defuzzifikasi untuk diterjemahkan kembali menjadi satu angka tegas (z^*).

3. Tabel Matriks Basis Aturan

Untuk memetakan hubungan antara kondisi probabilitas wajah dan validitas kata kunci suara yang ditangkap, sistem menggunakan sekumpulan aturan logika. Aturan ini dirancang menyerupai pola pengambilan keputusan manusia, di mana intervensi suara (TTS) akan disesuaikan berdasarkan tingkat kesesuaian kedua modalitas masukan tersebut. Himpunan keluaran pada proses inferensi ini dibagi menjadi tiga variabel linguistik, yaitu Rendah, Sedang, dan Tinggi. Matriks basis aturan yang digunakan terdiri 6 kombinasi aturan IF-THEN yang dijabarkan pada Tabel 3.1 berikut:

Tabel 3.2 Matriks Basis Aturan (Rule Base) Inferensi Fuzzy

Aturan	Wajah(X1)	Operator	Suara(x2)	Keluaran(z) Fuzzy
R1	Ragu-Ragu	AND	Lemah	Rendah
R2	Ragu-Ragu	AND	Kuat	Sedang
R3	Fokus	AND	Lemah	Sedang
R4	Fokus	AND	Kuat	Tinggi
R5	Percaya Diri	AND	Lemah	Tinggi
R6	Percaya diri	AND	Kuat	Tinggi

Berdasarkan matriks aturan di atas, mekanisme pembacaannya dibagi menjadi tiga pola utama: jika ekspresi wajah pengguna terlihat ragu-ragu dan suara vokal lemah R1, sistem memberikan keluaran target Rendah agar robot menyalurkan kalimat dorongan semangat yang menenangkan. Jika salah satu modalitas menunjukkan indikator positif R2 dan R3, sistem mengambil nilai tengah Sedang untuk memberikan penguatan normal. Sebaliknya, jika ekspresi wajah sudah dominan pada tingkat Fokus atau Sangat Percaya Diri R4, R5, dan R6, keluaran otomatis didorong menjadi Tinggi guna memberikan apresiasi suara selebratif kepada pengguna

4. Defuzzifikasi

Tahap terakhir dari keseluruhan proses inferensi logika fuzzy adalah defuzzifikasi, yaitu proses mengonversi daerah hasil gabungan yang diperoleh dari tahap agregasi sebelumnya kembali menjadi satu nilai angka skalar yang tegas. Nilai tegas inilah yang nantinya akan digunakan oleh

sistem untuk menentukan jenis respons suara yang akan dikeluarkan. perancangan sistem ini menggunakan metode Centroid atau yang sering disebut Center of Area (CoA). Berdasarkan metode ini, nilai tegas dicari dengan cara menghitung titik pusat atau titik berat dari seluruh luasan kurva daerah hasil agregasi. Karena komputasi dilakukan secara digital di dalam program, persamaan yang digunakan adalah bentuk diskrit. Secara matematis, rumus perhitungan defuzzifikasi metode Centroid dinyatakan sebagai berikut:

$$z^* = \frac{\sum_{i=1}^n \mu(z_i) \times z_i}{\sum_{i=1}^n \mu(z_i)}$$

Di mana

z^* : nilai tegas akhir.

z_i : nilai titik diskrit pada himpunan keluaran.

n : jumlah total keseluruhan titik diskrit pada domain keluaran.

3.3 Tabel Rentang Nilai Keputusan

Output	Rentang Nilai	Interpretasi
Rendah	0% - 30%	Tingkat keputusan rendah
Sedang	30% - 70%	Tingkat keputusan sedang
Tinggi	70% - 100%	Tingkat keputusan Tinggi

3.2.5 Logika Keputusan Berbasis Status Level dan Output TTS

Sistem menerapkan status penyelesaian level BDT sebagai salah satu parameter utama dalam menentukan waktu pemberian umpan balik kepada pengguna. Status level digunakan untuk membedakan antara kondisi ketika pengguna masih mengerjakan suatu level dan kondisi ketika pengguna telah menyelesaikan level tersebut. status level dibagi menjadi dua kondisi, yaitu InProgress dan Completed. Status InProgress menunjukkan bahwa pengguna masih berada dalam proses pengerjaan pola balok. Sementara itu, status Completed menunjukkan bahwa pengguna telah menyelesaikan level BDT sesuai dengan kondisi penyelesaian yang ditentukan oleh sistem.

Penggunaan status level sebagai pemicu utama bertujuan untuk mencegah Communication Robot memberikan respons apresiasi sebelum pengguna menyelesaikan tugas. Hal ini penting karena indikator visual dan vokal yang

menunjukkan ekspresi positif atau keyakinan tidak secara langsung menunjukkan bahwa suatu level BDT telah selesai. Oleh karena itu, perubahan status level menjadi Completed digunakan sebagai pemicu mekanisme Success-Triggered Feedback Loop

1. Kondisi InProgress *Silence Mode*

Ketika status level berada pada kondisi InProgress, pengguna masih mengerjakan pola balok. Pada kondisi ini, Komunikasi Robot berada dalam Silence Mode dan tidak memberikan intervensi suara secara terus-menerus. Skor indikator kepercayaan diri yang dihasilkan dari pemrosesan visual dan vokal tetap dapat dihitung oleh sistem sebagai informasi kondisi pengguna, tetapi nilai tersebut tidak langsung digunakan untuk memicu respons suara selama pengguna masih berada dalam tahap pengerjaan. Silence Mode bertujuan untuk menjaga konsentrasi pengguna dan mencegah pemberian intervensi yang tidak diperlukan selama proses pengerjaan BDT.

2. Kondisi Completed dengan Skor Indikator Rendah

Ketika status level berubah menjadi Completed dan skor indikator kepercayaan diri berada pada rentang:

$$0\% \leq z^* \leq 30\%$$

sistem menginterpretasikan kondisi tersebut sebagai tingkat indikator kepercayaan diri yang rendah. Communication Robot memberikan respons berupa penguatan positif dengan intensitas ringan melalui suara.

3. Kondisi Completed dengan Skor Indikator Sedang

Ketika status level berada pada kondisi Completed dan skor indikator kepercayaan diri berada pada rentang:

$$30\% \leq z^* \leq 70\%$$

sistem menginterpretasikan kondisi tersebut sebagai tingkat indikator kepercayaan diri sedang. Communication Robot memberikan respons apresiasi dengan intensitas normal.

4. Kondisi Completed dengan Skor Indikator Tinggi

Ketika status level berada pada kondisi Completed dan skor indikator kepercayaan diri berada pada rentang:

$$70\% \leq z^* \leq 100\%$$

sistem menginterpretasikan kondisi tersebut sebagai tingkat indikator kepercayaan diri yang tinggi. Communication Robot memberikan respons apresiasi dengan intensitas lebih tinggi.

3.6 Pengujian dan Analisa

Hasil pengujian fungsionalitas sistem serta analisis performa dari keseluruhan arsitektur Robot Komunikasi Berbasis Ekspresi Wajah yang dikembangkan untuk mendampingi pelaksanaan Block Design Test (BDT) dalam rangka meningkatkan efikasi diri pengguna. Pengujian ini bertujuan untuk memastikan seluruh komponen mulai dari Input Layer deteksi visual via *Media Pipe* dan vokal via *Web Speech API*, proses inferensi Fuzzy

Logic Mamdani, hingga Output Layer pemberian intervensi motivasional melalui *Text-to-Speech*.

3.6.1 Pengujian Fungsionalitas Modul Komunikasi Robot

Pengujian fungsionalitas dilakukan untuk memvalidasi apakah setiap modul di dalam sistem reaktif terhadap dinamika ekspresi dan ucapan pengguna selama berinteraksi dengan robot pendamping

3.4 Tabel Pengujian Fungsionalitas

No	Modul Pengujian	Skenario Uji	Hasil yang Diharapkan
1	Ekstraksi Wajah (MediaPipe)	Pengguna duduk di depan <i>webcam</i> pada jarak optimal (40–60 cm) sembari mengerjakan BDT.	Sistem mendeteksi dan melacak 468 titik <i>landmark</i> wajah secara <i>real-time</i> di dalam peramban web.
2	Transkripsi Vokal (STT)	Pengguna merespons tes secara verbal (" <i>Selesai!</i> ", " <i>Susah...</i> ").	Web Speech API mengonversi suara menjadi teks secara akurat dan <i>real-time</i> .
3	Inferensi Logika Fuzzy	Memasukkan parameter probabilitas wajah (X_1) dan skor kata kunci vokal (X_2) ke sistem.	Sistem mengonversi nilai tegas ke ranah <i>fuzzy</i> , mengevaluasi aturan Mamdani, dan menghasilkan nilai <i>crisp value</i> (z^*) melalui defuzzifikasi <i>CoA</i> .
4	Gerbang Pengaman Status Level	Mensimulasikan kondisi di mana skor fusi tinggi, namun status pengerjaan BDT bernilai InProgress.	Sistem mengunci respons selebratif dan mengalihkan ke mode diam atau kalimat korektif yang menenangkan.
5	Intervensi Motivasi (TTS)	Sistem memicu <i>State Response Sukses</i> setelah level berhasil diselesaikan dengan keyakinan tinggi.	<i>Speaker</i> PC mengeluarkan suara apresiasi berenergi tinggi untuk memperkuat efikasi diri pengguna.

3.6.2 Analisis Logika Keputusan Fuzzy Mamdani dan Respon Sistem

Analisis dilakukan berdasarkan enam kombinasi aturan yang terdapat pada basis aturan Fuzzy Mamdani. Sistem menggabungkan informasi dari kondisi wajah dan indikator vokal untuk menentukan tingkat respons yang sesuai. Nilai keluaran sistem

kemudian digunakan sebagai dasar untuk menentukan bentuk respons Communication Robot.

1. Analisis Kondisi Ragu-Ragu dan Vokal Lemah

Pada kondisi R1, sistem menerima masukan berupa kondisi wajah Ragu-Ragu dan indikator vokal Lemah. Kombinasi tersebut menghasilkan keluaran fuzzy Rendah. Kondisi ini menunjukkan bahwa sistem belum memperoleh indikator yang cukup kuat untuk memberikan respons apresiasi. Apabila status level BDT masih berada pada kondisi InProgress, sistem mempertahankan kondisi diam atau memberikan dorongan singkat yang menenangkan.

2. Analisis Kondisi dengan Satu Indikator Positif

Pada kondisi R2 dan R3, salah satu modalitas menunjukkan indikator yang lebih positif, sedangkan modalitas lainnya masih berada pada tingkat yang lebih rendah. Kondisi tersebut menghasilkan keluaran fuzzy Sedang. Pada kondisi ini, sistem dapat memberikan respons motivasional normal. Ketika ekspresi fokus atau percaya diri terdeteksi dan dikonfirmasi dengan vokal yang kuat, sistem menaikkan skor keluaran menjadi Tinggi

3. Analisis Kondisi dengan Indikator Keyakinan Tinggi

Pada kondisi R4, R5, dan R6, sistem mendeteksi indikator visual dan/atau vokal yang menunjukkan tingkat keyakinan lebih tinggi. Kombinasi tersebut menghasilkan keluaran fuzzy Tinggi. Namun, nilai keluaran Tinggi tidak secara langsung menyebabkan sistem memberikan apresiasi keberhasilan. Pemberian respons apresiasi tetap dikendalikan oleh status penyelesaian level BDT.

4. Analisis Integrasi Fuzzy dengan Status Level BDT

Apabila status level masih InProgress, sistem tidak memberikan apresiasi keberhasilan. Apabila status level telah berubah menjadi Success, sistem menjalankan mekanisme Success-Triggered Feedback Loop dan memberikan respons apresiasi melalui TTS.

3.6.3 Pengujian Respons Adaptif terhadap Efikasi Diri Pengguna

Pengujian terhadap efikasi diri dilakukan untuk mengetahui perbedaan kondisi pengguna sebelum dan setelah memperoleh dukungan dari Communication Robot. Pengujian dilakukan melalui dua sesi pelaksanaan BDT, yaitu sesi pertama tanpa Communication Robot dan sesi kedua dengan Communication Robot.

1. Sesi 1: BDT Tanpa Communication Robot

Pengguna mengerjakan BDT tanpa memperoleh respons motivasional atau umpan balik verbal dari sistem. Sesi ini digunakan sebagai kondisi pembandingan untuk mengetahui kondisi efikasi diri pengguna tanpa adanya intervensi Communication Robot.

2. Sesi 2: BDT dengan Communication Robot

Pengguna mengerjakan BDT dengan dukungan Communication Robot. Sistem memberikan respons adaptif berdasarkan kombinasi informasi visual, vokal, serta status penyelesaian level BDT. Respons diberikan melalui mekanisme *Success-Triggered Feedback Loop* dan keluaran suara menggunakan TTS.

3.3 Pembuatan Laporan

Pada tahap ini dilakukan pembuatan laporan sebagai bukti kegiatan penelitian yang telah dikerjakan selama kurun waktu yang telah ditentukan. Pembuatan laporan dilakukan berdasarkan hasil pengujian dan analisa yang telah didapatkan. Dari laporan ini akan diketahui apakah hasil proyek akhir ini sesuai dengan target atau tidak, sehingga dapat diambil kesimpulan untuk proyek akhir yang sudah dilakukan.

3.4 Perencanaan Jadwal

[illegible]

DAFTAR PUSTAKA

1. [1] D. Wechsler, Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV) Technical and Interpretive Manual. San Antonio, TX: Pearson, hlm. 1–450, 2021.
2. [2] M. Jung dan A. Paiva, “Social Robots and Human Interaction: A Review of Recent Advances,” *International Journal of Social Robotics*, vol. 13, no. 8, hlm. 1723–1740, 2021.
3. [3] M. Tamura, K. Tanaka, dan Y. Sato, “Advances in Neural Text-to-Speech Systems for Human Interaction,” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, hlm. 1–5, 2023.
4. [4] R. Zhao, J. Kim, dan S. Lee, “Motivational Feedback Systems in Human Computer Interaction,” *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*, hlm. 1–12, 2022.
5. [5] S. H. Kim dan K. Lee, “The Effect of Avatar-Based Interaction on User Engagement,” *Computers in Human Behavior*, vol. 116, hlm. 106620, 2021.
6. [6] Y. Li dan X. Chen, “Emotion-Aware Systems in Human-Computer Interaction: A Review,” *IEEE Access*, vol. 9, hlm. 134567–134580, 2021.
7. [7] A. R. Santos et al., “Adaptive Human–Robot Interaction Based on User State Detection,” *Robotics and Autonomous Systems*, vol. 170, hlm. 104547, 2023.
8. [8] T. Nguyen dan P. K. Singh, “Real-Time Emotion Recognition for Interactive Systems,” *IEEE Access*, vol. 12, hlm. 44521–44535, 2024.
9. [9] J. Lazar, J. H. Feng, dan H. Hochheiser, “Human-Computer Interaction Research and Practice,” *ACM Comput. Surv.*, vol. 53, no. 6, hlm. 1–38, Nov. 2021.
10. [10] R. M. Thompson, “Mastery experiences and psychological resilience,” *Journal of Applied Psychology*, vol. 14, no. 2, hlm. 55–72, 2026.
11. [11] S. K. Gupta, *Foundations of Affective Behavior in Human-Centered Systems*. Singapore: Springer Nature, hlm. 210–235, 2022.
12. [12] L. M. Fernandez dan K. T. Zhang, “The influence of confidence on user engagement in complex tasks,” *International Journal of Social Sciences*, vol. 16, no. 3, hlm. 400–415, 2024.
13. [13] Google Developers, “Flutter Documentation,” *Official Flutter Docs*, <https://docs.flutter.dev>, Diakses tanggal 12 Mei 2026.
14. [14] Dart Project, “Dart Language Overview,” *Dart Official Portal*, <https://dart.dev/overview>, Diakses tanggal 12 Mei 2026.
15. [15] J. Aguilar, “Reactive State Management in Flutter using GetX,” *Medium - Flutter Community Blog*, <https://medium.com/flutter-community/getx>, Diakses tanggal 14 Mei 2026.
16. [16] Blue Fire Team, “Audioplayers Package for Flutter,” *Pub.dev Documentation*, <https://pub.dev/packages/audioplayers>, Diakses tanggal 15 Mei 2026.
17. [17] M. Maroto-Gómez, J. C. Castillo, J. Alonso-Martin, S. Marqués-Villaroya, and Á. Castro-González, “Active learning based on computer vision and human–robot interaction for the user profiling and behavior personalization of an autonomous social robot,” *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105631, Jan. 2023,

18. Wechsler, D. (2014). Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) Administration and Scoring Manual. San Antonio, TX: Pearson Assessment.
19. [19] T. K. Park and H. J. Kim, "The Impact of Countdown Distractions on Block Design Test Performance in School-Aged Children," *Journal of Intelligence and Child Development*, vol. 29, no. 2, pp. 115–128, Nov. 2021.
20. [20] S. J. Vaughan, M. D. Coleman, and L. R. Johnston, "Digitizing Clinical Psychometrics: Evaluation of Time Constraints and Cognitive Load in Web-Based Block Design Test Alternatives," *Computers in Human Behavior*, vol. 140, p. 107592, Mar. 2023.