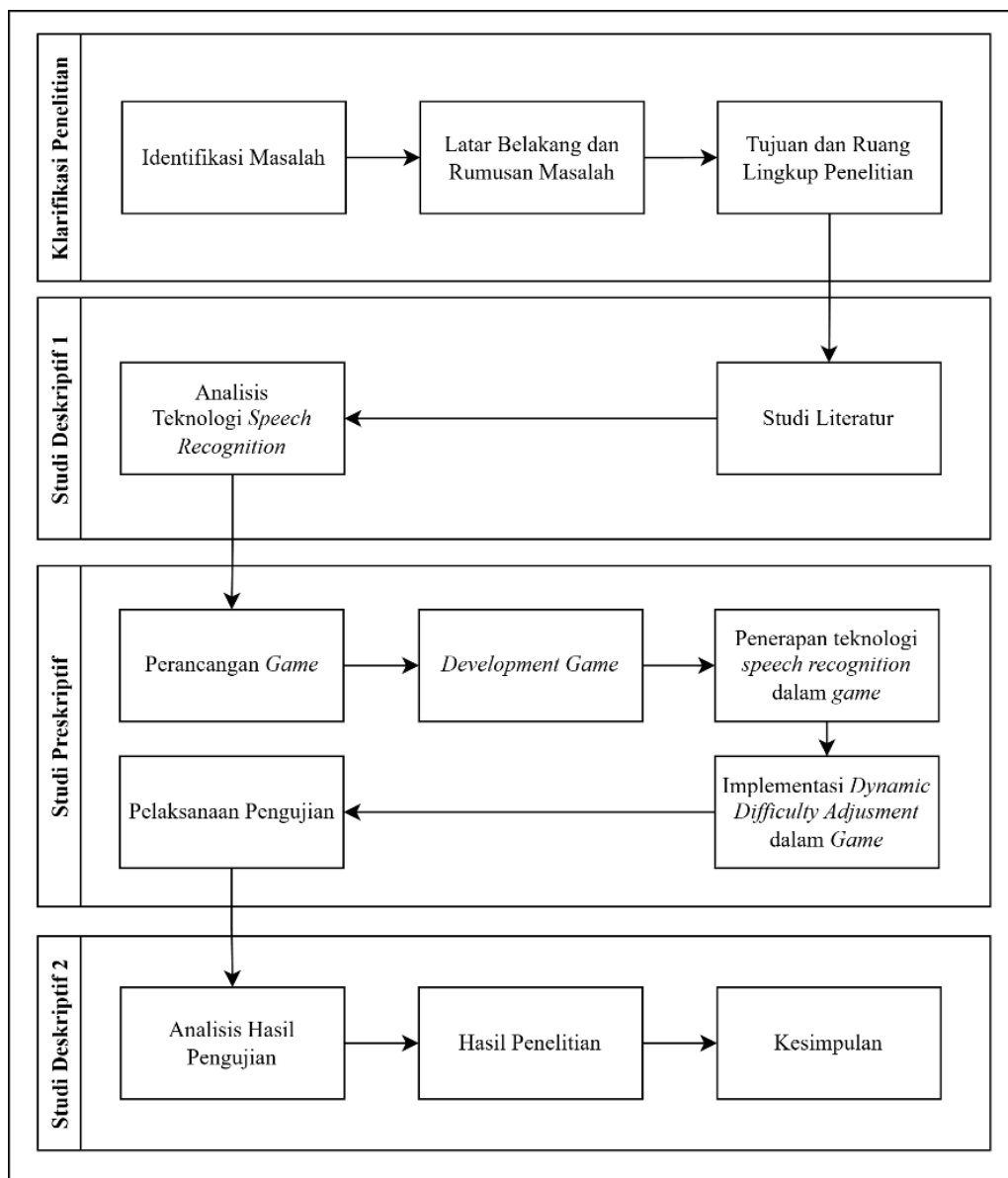


BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

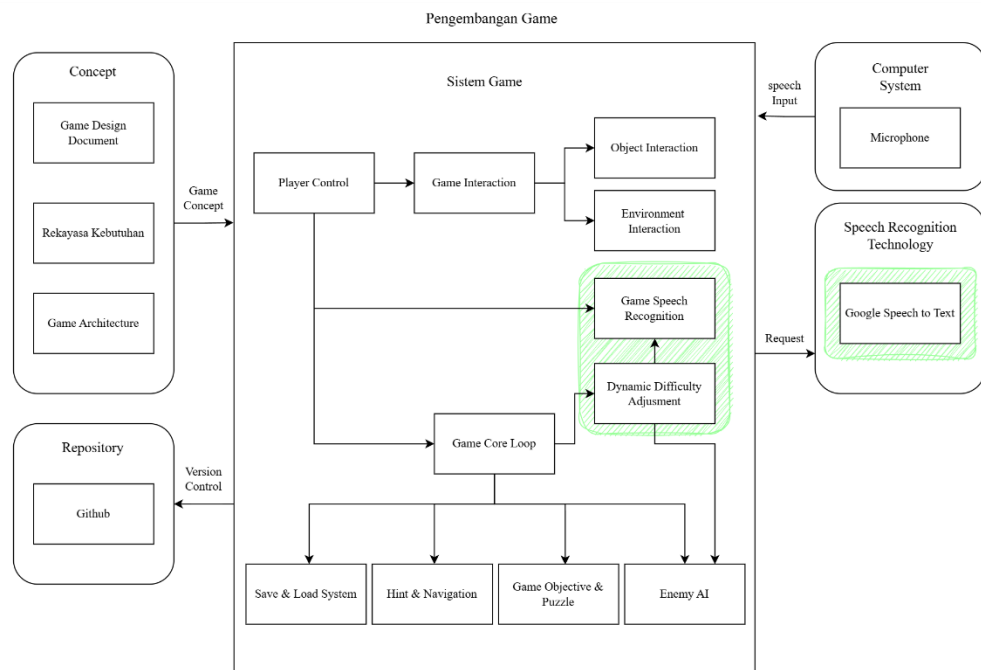
Penelitian ini menggunakan *Design Research Methodology* (DRM) sebagai pendekatan utama dalam proses penelitian dan pengembangan *game*. Model DRM memberikan kerangka kerja yang sistematis dan iteratif sehingga proses penelitian dapat berjalan lebih terstruktur (Sassanelli & Terzi, 2022). Adapun tahapan DRM penelitian ini dijelaskan dalam Gambar 3.1.



Gambar 3.1 *Design Research Methodology*

3.3.1 Klarifikasi Penelitian

Tahapan klarifikasi penelitian dimulai dengan proses identifikasi masalah melalui studi literatur untuk mengenali permasalahan dan solusi yang relevan. Hasil identifikasi dirangkum menjadi poin-poin rumusan masalah sebagai dasar perencanaan penelitian untuk mencapai tujuan penelitian. Kemudian ruang lingkup pengembangan *game* Jaga ini cukup luas, adapun penelitian ini berfokus pada penerapan *speech recognition* menggunakan pendekatan *dynamic difficulty adjustment* dalam *game* Jaga menggunakan teknologi Google Speech-to-Text (GST). Model Google Speech-to-Text dipilih dalam penelitian ini karena menawarkan akurasi tinggi dalam mengenali ucapan serta mampu mengenali berbagai bahasa dan dialek. Kemudian layanan ini mendukung lebih dari 125 bahasa termasuk bahasa Arab dan variasinya sehingga relevan dengan *game* yang dikembangkan. Selain itu, GST juga menyediakan *Application Programming Interface* (API) yang mudah digunakan sehingga mempermudah proses integrasi fitur *speech recognition* ke dalam *game* Jaga (Khairani dkk., 2024). Adapun gambaran desain sistem dan fokus utama penelitian dijelaskan dalam Gambar 3.2.



Gambar 3.2 Desain Sistem *Game* dan Fokus Utama Penelitian

3.3.2 Studi Deskriptif 1

Pada tahapan ini diawali dengan studi literatur untuk mengetahui landasan penelitian melalui analisis hasil dari penelitian-penelitian sebelumnya dalam bidang yang relevan. Tahapan ini juga melibatkan proses analisis *speech recognition* dari Google Speech-to-Text, batasan serta implementasinya dalam *game*.

3.3.3 Studi Preskriptif

- a. Perancangan *game* mencakup berbagai aspek termasuk aspek desain dan aspek teknis. Pada aspek teknis, *game* dibuat menggunakan *Event Driven Architecture* (EDA) untuk membuatnya menjadi lebih dinamis dengan merespon berbagai peristiwa secara *realtime*. Pada aspek desain, perancangan DDA dilakukan untuk menyesuaikan tingkat kesulitan permainan secara otomatis berdasarkan performa pemain menggunakan pendekatan *Threshold-Based Adjustment* (TBA).
- b. Tujuan utama dari proses pengembangan *game* ini adalah untuk menerjemahkan desain konseptual yang telah dirancang menjadi produk fungsional dalam bentuk *game*. Proses pengembangan *game* ini dilakukan melalui iterasi berulang dalam implementasi mekanika permainan, sistem kontrol, visual, integrasi teknologi *speech recognition* hingga implementasi DDA ke dalam *game*.
- c. Proses selanjutnya dilakukan penerapan *speech recognition* menggunakan model Google Speech-to-Text (GST) pada *game* dalam modul terpisah. Proses ini meliputi banyak tahapan mulai dari konfigurasi GST hingga integrasi *speech recognition* kedalam *game*.
- d. Implementasi DDA pada *game* diterapkan setelah semua modul dalam *game* diimplementasikan. DDA ini berfungsi untuk membantu fitur *speech recognition* melalui penyesuaian tingkat kesulitan berdasarkan kemampuan pemain sehingga dapat berjalan selaras dengan tujuan *game* Jaga sebagai media pembelajaran Al-Quran berbasis *game horror*.
- e. Proses pengujian dilakukan untuk menguji performa *speech recognition* dalam *game* Jaga menggunakan pendekatan DDA serta pengaruhnya terhadap skor bermain pengguna dalam *game* Jaga. Tahap pertama dilakukan pengujian fungsionalitas *game* secara internal untuk

memastikan bahwa *game* dapat dimainkan tanpa kendala. Pada tahap kedua dilakukan pengujian akurasi dan performa *speech recognition*. Tahap selanjutnya dilakukan validasi ahli untuk memastikan bahwa *game* layak diteliti dan diuji pada pengguna. Tahap terakhir dilakukan pengujian dengan melibatkan pengguna untuk menguji pengaruh DDA terhadap skor bermain pengguna serta pengujian *System Usability Scale* (SUS) untuk mengukur kemudahan dan keandalan *speech recognition* dan DDA dalam *game* Jaga.

3.3.4 Studi Deskriptif 2

Pada proses ini, hasil dari pengujian kemudian dianalisis secara menyeluruh untuk merumuskan temuan penelitian. Proses ini bertujuan untuk memberikan pemahaman yang mendalam terkait data yang diperoleh sehingga mampu menggambarkan tingkat keberhasilan penelitian. Proses ini diakhiri dengan penyampaian kesimpulan yang mencakup proses dan temuan penelitian secara singkat dalam bentuk kalimat-kalimat penting untuk menjawab pertanyaan-pertanyaan yang diajukan dalam rumusan masalah yang diangkat.

3.2 Partisipan

Partisipan dalam penelitian terbagi atas dua jenis yaitu *end user* dan ahli. Partisipan *end user* yang diuji sebanyak 20 orang dalam rentang umur 13 - 40 dan dianggap sudah memadai dalam penelitian ini mengingat tujuan utama dalam penelitian terletak pada aspek sistem *speech recognition* dan DDA dalam konteks *game* Jaga (Namba, 2021; Anzai dkk., 2021). Namun partisipan yang di uji terbatas pada orang beragama Islam. Adapun karakteristiknya adalah orang yang bisa memainkan *game* desktop, bisa membaca Al-Quran dan sanggup untuk memainkan *game horror*. Kemudian jenis partisipan kedua dalam penelitian merupakan ahli media yang merupakan orang yang memiliki wawasan terkait media *game* atau pengembang *game* profesional melalui sebuah angket kuesioner untuk memberikan validasi media dalam terkait *speech recognition* dan *dynamic difficulty adjustment* (DDA) pada *game* Jaga.

3.3 Metode Pengumpulan Data

Data penelitian dikumpulkan melalui dua metode utama yaitu *log* penggunaan dan kuesioner. Kedua metode ini digunakan untuk mengumpulkan data yang

berbeda namun saling melengkapi untuk memberikan gambaran yang lebih komprehensif mengenai penerapan fitur *speech recognition* dalam *game* menggunakan pendekatan DDA serta dampaknya terhadap skor bermain pengguna.

3.3.1 Log Penggunaan

Log penggunaan merupakan proses untuk mencatat setiap aktifitas pengguna dalam sebuah sistem dan direkam secara otomatis untuk kepentingan tertentu (Marin-Castro & Tello-Leal, 2021). Proses ini digunakan untuk mencatat akurasi dan performa *speech recognition* dalam *game* serta skor bermain pengguna. Data hasil *log* ini kemudian dianalisis untuk menjawab poin-poin rumusan masalah dalam penelitian.

3.3.2 Kuesioner

Kuesioner atau angket adalah teknik pengumpulan data yang dilakukan dengan memberikan serangkaian pertanyaan atau pernyataan tertulis kepada responden untuk dijawab (Mwita, 2022). Metode ini digunakan untuk mengumpulkan data partisipan mengenai kemudahan, keandalan, kompleksitas, *error tolerance*, *learnability* terhadap *speech recognition* dan DDA dalam *game* Jaga.

3.4 Instrumen Penelitian

Penelitian berfokus untuk mengukur kemampuan *speech recognition* dalam mengenali pelafalan Al-Quran serta pengaruh DDA terhadap skor bermain dalam *game*. Instrumen penelitian yang digunakan pada penelitian ini melibatkan sistem *logging*, validasi ahli dan *system usability scale* (SUS).

3.4.1 Sistem Logging

Sistem *logging* digunakan untuk mencatat dan menyimpan data secara otomatis mengenai aktivitas dan performa suatu sistem atau aplikasi. Dengan mencatat berbagai peristiwa seperti permintaan pengguna, kesalahan sistem atau waktu respon, *log* data dapat menjadi sumber informasi yang sangat berharga untuk kebutuhan penelitian, pemeliharaan dan perbaikan berkelanjutan (Ali dkk., 2024). Data hasil *log* kemudian di analisis dan menjawab rumusan penelitian (K dkk., 2024). Adapun data *log* yang digunakan dijelaskan dalam Tabel 3.1.

Tabel 3.1
Data Sistem *Logging*

Data	Deskripsi
Frekuensi pelafalan benar	Frekuensi pengguna melafalkan ayat dengan benar (CER < 10%)
Frekuensi pelafalan kurang tepat	Frekuensi pengguna melafalkan ayat dengan benar (10% < CER < 20%)
Frekuensi pelafalan salah	Jumlah berapa kali pengguna salah dalam melafalkan ayat (CER% > 20)
Frekuensi <i>error</i>	Frekuensi sistem dalam <i>game</i> atau teknologi <i>speech recognition</i> gagal dalam memproses pelafalan pengguna
Durasi bermain	total durasi bermain pengguna
Level DDA	Menghitung capaian level DDA pada <i>gameplay</i> metrik bermain pengguna
Skor	Total skor akhir pengguna
<i>Accuracy rate</i>	Persentase hasil pengenalan suara yang benar dibandingkan dengan total input yang diberikan
<i>Character error rate</i>	Mengukur kesalahan pengenalan karakter dalam <i>speech recognition</i>
<i>Audio duration</i>	Durasi waktu pelafalan
<i>Time process</i>	Waktu yang dibutuhkan untuk memproses suara

3.4.2 Validasi Ahli

Proses validasi ahli berperan krusial untuk menjamin bahwa seluruh elemen media dalam *game* memenuhi standar kualitas yang dibutuhkan baik dari segi teknis, konten maupun fungsionalitas (Yusmaningsih dkk., 2024). Validasi ini dilakukan untuk memastikan bahwa penerapan teknologi *speech recognition* dan DDA dapat diintegrasikan secara efektif ke dalam *game* serta layak diteliti dan diuji pada pengguna. Sangat penting untuk memastikan bahwa spesifikasi, akurasi dan performa *speech recognition* sesuai serta relevan dengan kebutuhan dan tujuan aplikasi yang dikembangkan (Alfadhli dkk., 2024). Kemudian pada aspek DDA juga perlu dipastikan bahwa mekanisme penyesuaian tingkat kesulitan berjalan

sesuai dengan kemampuan pemain, sulit dikenali oleh pemain serta relevan dengan penggunaan *speech recognition* (Martins & Cavaco, 2022). Adapun angket penilaian validasi ahli media dijelaskan dalam Tabel 3.2 berikut.

Tabel 3.2
Angket Penilaian Validasi Ahli Media

No	Indikator	Skor				
		1	2	3	4	5
1	Fitur <i>speech recognition</i> yang dikembangkan relevan dengan konsep gim horor					
2	Fitur <i>speech recognition</i> dalam <i>game</i> berjalan dengan lancar tanpa <i>error</i> atau <i>bug</i>					
3	Teknologi <i>speech recognition</i> dalam <i>game</i> berfungsi sesuai spesifikasi dengan tingkat akurasi yang dapat diterima					
4	<i>Speech recognition</i> dalam <i>game</i> mampu mengenali pelafalan bacaan Al-Quran dengan baik					
5	Performa <i>speech recognition</i> dalam <i>game</i> dapat berjalan dengan konsisten					
6	Penggunaan <i>speech recognition</i> dalam <i>game</i> mudah dipahami dan digunakan					
7	Mekanisme <i>dynamic difficulty adjustment</i> relevan dengan penggunaan teknologi <i>speech recognition</i> dalam <i>game</i>					
8	<i>Dynamic difficulty adjustment</i> berjalan sesuai dengan kemampuan pemain					
9	Sistem <i>dynamic difficulty adjustment</i> merespon dengan cepat terhadap perubahan dalam performa pemain					
10	Perubahan tingkat kesulitan sulit dikenali oleh pemain selama sesi permainan berlangsung					
11	Penggunaan <i>speech recognition</i> dalam <i>game</i> ini membawa elemen baru yang tidak biasa ditemukan dalam <i>game horror</i> lainnya					

12	Sistem <i>speech recognition</i> dan DDA berfungsi dengan baik dalam mendukung dinamika <i>gameplay</i> yang kreatif
13	Fitur <i>speech recognition</i> dan DDA memberikan potensi untuk pengembangan permainan yang lebih adaptif dan responsif

3.4.3 System Usability Scale

Instrumen penelitian berikutnya menggunakan *system usability scale* (SUS) untuk memberikan analisis lebih mendalam terkait interaksi pengguna dengan sistem *game* melalui pengukuran usabilitas *speech recognition* dan DDA dalam *game*. SUS tidak hanya menilai kepuasan pengguna, namun juga memberikan gambaran mengenai performa teknis dan sejauh mana sistem mendukung tugas pengguna secara efisien dan konsisten (Rudianto & Firmansyah, 2024). Pengukuran berdasarkan pengalaman pengguna melalui SUS merupakan langkah penting dalam memastikan bahwa spesifikasi fitur *speech recognition* dan DDA dalam *game* Jaga dapat memenuhi kebutuhan dan ekspektasi pengguna. (Alfadhli dkk., 2024). Adapun kuesioner SUS dijelaskan dalam Tabel 3.3.

Tabel 3.3
Kuesioner *System Usability Scale*

No	Pertanyaan	Nilai				
		1	2	3	4	5
1	Saya merasa fitur <i>speech recognition</i> dalam <i>game</i> mudah digunakan.					
2	Saya merasa fitur <i>speech recognition</i> terlalu rumit untuk digunakan dalam situasi horor.					
3	Saya merasa percaya diri pada kemampuan pelafalan bacaan Al-Quran saya saat melalkan ayat <i>speech recognition</i> dalam <i>game</i> .					
4	Saya merasa frustrasi ketika menggunakan fitur <i>speech recognition</i> dalam <i>game</i> ini.					
5	Saya merasa sistem <i>speech recognition</i> dapat mengenali pelafalan saya dengan akurat selama bermain gim.					

-
- Saya merasa ada inkonsistensi atau tidak konsisten
- 6 dalam cara fitur *speech recognition* merespon suara saya.
- 7 Saya merasa bahwa tingkat kesulitan dalam *game* sesuai dengan kemampuan saya.
- 8 Saya merasa hambatan dalam *game* terlalu banyak
- Saya merasa gim ini cocok untuk dimainkan oleh
- 9 berbagai jenis pemain dengan tingkat kemampuan yang berbeda.
- Saya merasa bahwa *game* ini membingungkan dan
- 10 memerlukan waktu lebih lama untuk dapat memahami dan menyelesaikan *game*
-

3.5 Analisis Data

Analisis data dilakukan pada data kuantitatif yang diperoleh melalui kuesioner dan *log* penggunaan. Teknik dan perhitungan analisis berbeda tergantung dari metode pengumpulan dan jenis data yang dikumpulkan.

3.5.1 Metrik Evaluasi Kemampuan *Speech Recognition* Dalam *Game*

Untuk data penelitian yang diperoleh dari *log* penggunaan dianalisis menggunakan beberapa metrik evaluasi diantaranya adalah *Character Error Rate* (CER), *accuracy rate* dan *Real-Time Factor* (RTF).

a. *Character Error Rate* (CER)

CER adalah metrik yang menghitung persentase karakter yang salah pada tingkat karakter. Perhitungan *Character Error Rate* (CER) dilakukan menggunakan Levenshtein Distance, yaitu metrik yang mengukur jumlah minimum operasi penyisipan, penghapusan, dan substitusi karakter yang diperlukan untuk mengubah teks hasil *speech recognition* menjadi teks referensi yang benar. Nilai CER dihitung dengan membagi total jumlah kesalahan dari perhitungan Levenshtein Distance dengan jumlah karakter dalam teks referensi sehingga memberikan gambaran akurat tentang tingkat kesalahan pengenalan dalam sistem yang diuji. Metrik ini digunakan karena secara default *Game Engine* unity tidak bisa memproses huruf Arab yang bersambung.

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

S: Karakter yang salah dikenali (Substitution)

D: Karakter yang hilang atau tidak dikenali (Deletions)

I : Karakter tambahan (Insertions)

N: Jumlah total karakter dalam teks referensi

Adapun kategori penilaian CER yang banyak digunakan sebagai standar kualitasnya terbagi atas beberapa tingkatan. Nilai CER di bawah 5% dikategorikan sebagai sangat baik untuk sistem dapat digunakan secara efektif. Sedangkan nilai 5% - 10% dianggap baik untuk aplikasi umum seperti asisten virtual atau game interaktif dengan pengalaman pengguna yang memadai. Kemudian jika nilai CER berada dalam rentang 10% hingga 20% maka sistem dianggap cukup karena meskipun masih dapat digunakan namun memerlukan koreksi manual atau penggunaan dalam lingkungan yang minim gangguan. Sedangkan nilai CER yang melebihi 20% dikategorikan sebagai buruk dan menunjukkan bahwa sistem membutuhkan peningkatan signifikan baik dalam hal model pengenalan suara maupun kualitas input data yang digunakan (K dkk., 2024).

b. *Accuracy Rate*

Accuracy rate adalah metrik yang digunakan untuk mengukur seberapa akurat hasil pengenalan suara yang dilakukan oleh sistem *speech recognition*. Dalam konteks ini, yaitu mengukur seberapa baik sistem *speech recognition* dalam mengenali pelafalan Al-Quran berdasarkan perbandingan antara hasil yang benar dan seluruh input yang diuji. Sebuah hasil pengenalan suara dapat dikatakan benar jika *Character Error Rate* (CER) yang dihasilkan kurang dari 10% (Yuan dkk., 2021). Kemudian dalam konteks pelafalan Al-Quran, jika persentase kesalahan karakter yang dihasilkan oleh sistem berada di bawah 5% maka pengenalan suara tersebut dapat dianggap sangat akurat dan dapat diterima untuk penggunaan dalam aplikasi dengan standar kualitas tinggi (Harere & Jallad, 2023)

$$Akurasi(\%) = \left(\frac{\text{Jumlah event speech yang berhasil}}{\text{Total speech event yang dijalankan}} \right) \times 100$$

Kualitas *accuracy rate* dalam *speech recognition* dikategorikan berdasarkan tingkat akurasi yang biasa dipakai dalam standar industri. Nilai akurasi lebih dari 95% ($\geq 95\%$) menunjukkan kualitas sangat baik dalam sistem dengan kesalahan minimal dan cocok untuk sistem yang menuntut akurasi tinggi. Nilai akurasi 90-95% dikategorikan baik untuk sebagian besar sistem. Kemudian nilai akurasi 80-89% dikategorikan cukup baik meski memerlukan sedikit koreksi. Sementara nilai akurasi kurang dari 80% ($< 80\%$) memiliki tingkat kesalahan yang sering dan kurang direkomendasikan untuk sistem digunakan dalam industri khususnya dalam *game*.

c. *Real-Time Factor* (RTF)

Standar kualitas dalam *speech recognition* tidak hanya mengacu pada metrik *character error rate* dan *accuracy rate* tetapi juga *Real-Time Factor* (RTF), yaitu rasio antara waktu pemrosesan dan durasi audio. Sistem dengan $RTF \leq 1.0$ dianggap ideal karena mampu memproses suara secepat atau lebih cepat dari durasi audio asli. RTF 1.0 - 2.0 masih dapat diterima tetapi sedikit lebih lambat, sementara $RTF > 2.0$ dianggap kurang optimal karena membutuhkan waktu lebih lama dari durasi audio yang dapat menghambat respons sistem. Standar industri umumnya menetapkan bahwa $RTF \leq 1.0$ diperlukan untuk pengalaman pengguna yang optimal (Arriaga dkk., 2024).

$$RTF = \frac{Waktu Proses}{Durasi Audio}$$

3.5.2 Skala Likert

Skala Likert adalah metode pengukuran yang digunakan untuk mengukur pendapat atau persepsi seseorang terhadap suatu topik atau pernyataan. Metode ini digunakan pada kuesioner validasi ahli dan SUS. Adapun skala penilaian dijelaskan pada Tabel 3.4 berikut.

Tabel 3.4
Penilaian Skala Likert

Jawaban	Nilai Jawaban
Sangat Tidak Setuju	1
Tidak Setuju	2
Netral	3
Setuju	4
Sangat Setuju	5

Pada instrumen penilaian validasi ahli, total skor penilaian dihitung menggunakan rumus dibawah ini untuk mendapatkan persentase kualitas yang *speech recognition* dan *dynamic difficulty adjustment* dalam konteks *game* Jaga yang sudah dibuat.

$$\text{Persentase} = \frac{\text{Total Skor}}{\text{Nilai Maksimal}} \times 100\%$$

Setelah menghitung skor kualitas media oleh validasi ahli, kemudian hasil persentase tersebut dikategorikan berdasarkan nilai persentase pada Tabel 3.5 (Wahyuliana & Andrian, 2022)

Tabel 3.5
Kategori Kualitas Media

Persentase Nilai	Kriteria
85% - 100%	Sangat layak, dapat digunakan tanpa revisi.
70% - 84%	Layak, dapat digunakan dengan revisi minor
50% - 69%	Cukup layak, dapat digunakan dengan revisi mayor
0% – 49%	Tidak layak, tidak dapat digunakan

3.5.3 Analisis Regresi

Regresi adalah metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan satu atau lebih variabel independen (Muhammad Bangkit Riksa Utama & Hajarisman, 2021). Tujuan utama regresi adalah untuk memahami, memodelkan, dan memprediksi hubungan antara variabel-variabel tersebut. Adapun penelitian ini menggunakan analisis regresi linear sederhana untuk menganalisis hubungan antara audio rekaman pelafalan

pengguna terhadap waktu proses pengolahan suara dan pengaruh DDA terhadap skor bermain pengguna menggunakan rumus berikut.

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Hasil dari regresi linear mencakup sejumlah parameter dan statistik yang dapat digunakan untuk mengevaluasi hubungan antara variabel independen dan variabel dependen.

- a. Koefisien Regresi (β_i) menunjukkan pengaruh satu unit perubahan pada variabel independen terhadap variabel dependen dengan asumsi variabel lain tetap konstan. Jika nilai $\beta_i > 0$ maka menunjukkan hubungan positif, sebaliknya jika $\beta_i < 0$ menunjukkan hubungan negatif dan $\beta_i = 0$ berarti tidak memiliki hubungan.
- b. *intercept* adalah nilai prediksi variabel dependen ketika semua variabel independen bernilai nol.
- c. Nilai signifikansi (P-value) menunjukkan apakah koefisien regresi (β_i) secara statistik signifikan. Jika $p \leq 0.05$ maka koefisien berpengaruh signifikan dan sebaliknya jika $p > 0.05$ maka koefisien tidak berpengaruh signifikan.
- d. Koefisien determinasi (R^2) mengukur proporsi variasi dalam variabel dependen yang dijelaskan oleh semua variabel independen.

3.5.4 System Usability Scale

Data penelitian hasil kuesioner dianalisis menggunakan metode SUS. Analisis ini dilakukan dengan mengumpulkan respon dari pengguna menggunakan skala Likert 1–5 dengan mengacu pada prinsip penilaian SUS dengan ketentuan berikut.

- a. Untuk setiap pernyataan ganjil

$$\text{Skor Ganjil} = \text{Nilai Responden} - 1$$

- b. Untuk setiap pernyataan genap

$$\text{Skor Genap} = 5 - \text{Nilai Responden}$$

- c. Hasil dari kedua nilai ini kemudian dijumlahkan untuk mendapatkan total skor dengan nilai maksimum 40

$$\text{Total Score} = \text{Score Ganjil} + \text{Skor Genap}$$

- d. Total skor kemudian dikonversi ke skala 0–100 dengan cara mengalikan total skor dengan 2.5

$$SUS\ Score = (Total\ Skor) \times 2.5$$

Langkah berikutnya adalah menginterpretasikan skor SUS tersebut berdasarkan kategori yang telah ditetapkan dalam kerangka evaluasi SUS pada Tabel 3.6 Nilai SUS berada dalam rentang 0 – 100 dengan nilai 68 sebagai rata-rata standar yang digunakan sebagian besar organisasi sebagai indikator bahwa sistem tersebut memiliki kegunaan yang cukup baik untuk digunakan oleh pengguna. Sementara Skor di bawah 68 menunjukkan bahwa ada area yang perlu perbaikan (Vlachogianni & Tselios, 2022).

Tabel 3.6
Grade Score SUS

Skor SUS	Grade	Keterangan
85 – 100	A	<i>Excellent usability</i> (Sangat baik)
70 – 84	B	<i>Good usability</i> (Baik)
50 – 69	C	<i>Marginal usability</i> (Cukup baik)
30 – 49	D	<i>Poor usability</i> (Buruk)
0 – 29	F	<i>Unacceptable usability</i> (Tidak dapat diterima)

3.6 Prosedur Pengujian

Pengujian dalam penelitian ini terbagi menjadi beberapa tahapan yaitu pengujian kemampuan *speech recognition* pada *game* Jaga melalui pengukuran akurasi dan kecepatan sistem dalam memproses audio pelafalan dan pengujian pada pengguna.

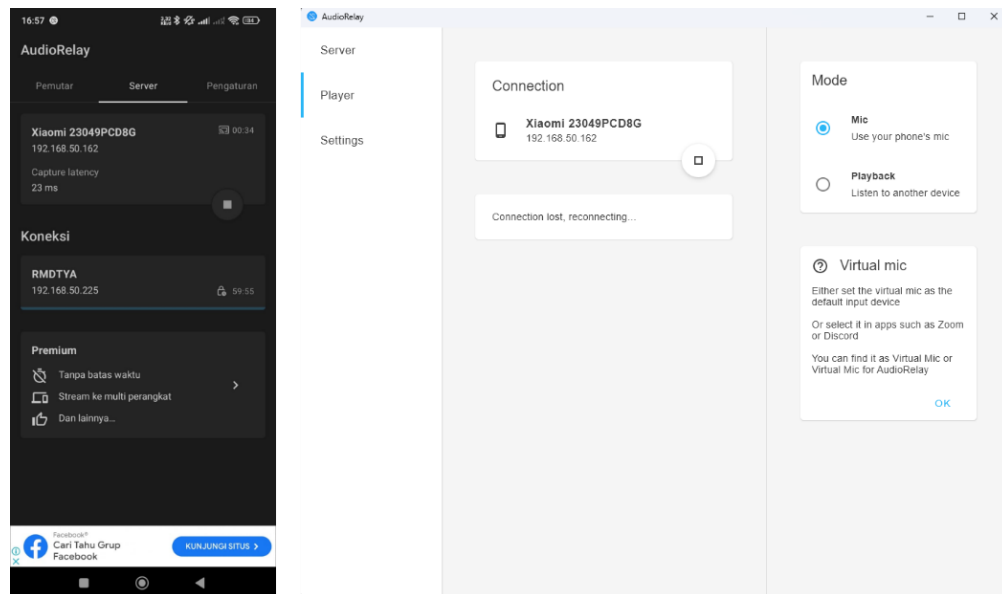
3.6.1 Pengujian Kemampuan *Speech Recognition*

Tahapan pertama adalah pengujian kemampuan *speech recognition* dalam kondisi yang paling optimal menggunakan metrik CER, akurasi dan RTF dengan pendekatan *single-pass evaluation* dimana setiap *speech event* diuji hanya sekali pada setiap ayat tanpa pengulangan atau koreksi. Pendekatan *single-pass evaluation* memungkinkan evaluasi dilakukan secara cepat dan efisien serta memberikan gambaran yang jelas tentang kemampuan sistem dalam mengenali ucapan secara

langsung (Li dkk., 2024). Untuk mendapatkan hasil yang lebih akurat, MuslimPro digunakan sebagai media yang membantu untuk mengukur akurasi dan performa *speech recognition*. MuslimPro menyediakan fitur bacaan Al-Quran digital oleh beberapa qari terpercaya yang dapat dijadikan acuan standar untuk diteladani atau sebagai rujukan bacaan yang representatif bagi umat Muslim. Qari adalah istilah dalam bahasa Arab yang merujuk pada seseorang yang memiliki kemampuan membaca atau melantunkan ayat-ayat Al-Quran dengan indah dan sesuai dengan kaidah tajwid (Alfadhli dkk., 2024). Seorang qari memiliki pelatihan khusus dalam melafalkan Al-Quran secara benar, termasuk pemahaman mendalam tentang tajwid, makhraj dan irama (Kulsoom, 2024).

Qari yang dipilih yaitu Saad Al-Ghamdi dan Ali al Hudhaify yang keduanya memiliki karakteristik bacaan yang berbeda. Saad Al-Ghamdi menggunakan metode Mujawwad dengan irama yang merdu, mendayu dan cenderung mendengung serta tempo yang lambat. Sebaliknya, Ali Al-Hudhaify menggunakan metode Murattal, dengan bacaan yang jelas, tegas dan cenderung lebih cepat (Arliana Arlin dkk., 2024). Perbedaan karakteristik membuat pengujian menjadi lebih komprehensif dan akurat. Variasi dalam gaya dan intonasi kedua qari memberikan cakupan yang lebih luas dalam menguji kemampuan sistem mengenali pelafalan yang berbeda. Hal ini membantu memastikan bahwa sistem tidak hanya mampu mengenali satu pola bacaan, namun juga dapat beradaptasi dengan variasi pelafalan sehingga meningkatkan keandalan dan performa sistem secara keseluruhan.

Untuk meminimalisir gangguan external serta hasil audio yang lebih optimal, *software* AudioRelay digunakan sebagai media yang menyalurkan audio pelafalan MuslimPro dari *smartphone* ke *virtual microphone* dalam komputer. Adapun mekanismenya, AudioRelay pada *smartphone* diatur sebagai *server* sedangkan pada komputer diatur sebagai *player*. Adapun konfigurasi AudioRelay dijelaskan pada Gambar 4.2 berikut.



Gambar 3.3 Konfigurasi AudioRelay

3.6.2 Pengujian Pada Pengguna

Prosedur pengujian ini dilakukan oleh partisipan end-user melalui pengujian secara langsung dengan memainkan *game* Jaga. Adapun prosedurnya terdiri atas dua tahapan yaitu tahapan bermain *game* Jaga dan pada tahapan kedua adalah mengisi angket System Usability Scale (SUS). Pada tahap pertama partisipan dapat memainkan langsung dari perangkat yang disediakan pengembang atau secara daring melalui platform Discord, namun keduanya dilakukan sambil dipantau oleh tim pengembang. Selama pengguna bermain, sistem logging dalam *game* akan mencatat secara otomatis untuk mengumpulkan data terkait performa sistem dan *gameplay* metrik pengguna. Tahap kedua dilakukan setelah partisipan menyelesaikan sesi bermain yang kemudian diarahkan untuk mengisi 10 butir soal kuesioner *system usability scale*. Data hasil pengujian kemudian dianalisis untuk menunjukkan tingkat efektivitas dari *speech recognition* menggunakan pendekatan dynamic difficulty adjustment dalam *game* Jaga.

3.7 Alat dan Bahan

Penelitian ini menggunakan beberapa alat dan bahan berupa *software* dan *hardware*. Perangkat utama yang digunakan adalah laptop dengan prosesor Ryzen 5 4600H 6 CPU @~3.0 GHz dan TP-Link TL-WN725N USB adaptor WiFi 4 (802.11n) dengan paket wifi IndiHome 10Mbps digunakan agar *game* dapat meminta request ke API Google Speech-to-Text. Perangkat lainnya adalah

Ramaditya Firdaus, 2025

PENERAPAN SPEECH RECOGNITION PADA GIM HOROR MENGGUNAKAN PENDEKATAN DYNAMIC DIFFICULTY ADJUSTMENT

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

smartphone Android yang dilengkapi dengan aplikasi MuslimPro dan AudioRelay sebagai penghubung antara *smartphone* dan komputer.

3.8 Alat Pendukung

Terdapat beberapa alat pendukung yang digunakan untuk membantu dan membuat proses pelaksanaan penelitian ini lebih mudah. Adapun rincian alat – alat pendukungnya sebagai berikut.

1. Microsoft Excel

Microsoft Excel digunakan untuk mengelola berbagai jenis data termasuk analisis data hasil dari pengujian.

2. Unity Asset Store

Unity Asset Store menyediakan berbagai model 3D, tekstur hingga *plugin* yang diperlukan untuk mempercepat proses pengembangan dan meningkatkan fungsionalitas *game*.

3. *Documentation*

Documentation menyediakan berbagai informasi komprehensif tentang bagaimana penggunaan teknologi yang digunakan.