

BAB I

PENDAHULUAN

1.1 Latar Belakang

Komunikasi merupakan kebutuhan fundamental manusia yang memungkinkan individu untuk mengekspresikan pikiran, perasaan, dan kebutuhan mereka. Menurut Beukelman dan Light (2020), diperkirakan sekitar 5 juta orang Amerika dan 97 juta orang di seluruh dunia dapat memperoleh manfaat dari *Augmentative and Alternative Communication* (AAC) [1]. Di Indonesia, kelompok yang mengalami kesulitan komunikasi verbal mencakup individu dengan tunawicara, autisme, cerebral palsy, dan berbagai kondisi disabilitas lainnya yang memerlukan sistem komunikasi alternatif untuk memfasilitasi interaksi sosial dan pemenuhan kebutuhan sehari-hari.

Tunawicara adalah kondisi dimana seseorang mengalami gangguan atau ketidakmampuan dalam berbicara, yang dapat disebabkan oleh berbagai faktor seperti gangguan pada organ artikulasi, gangguan pendengaran (tunarungu), gangguan neurologis, atau kondisi bawaan seperti autisme [2]. Di Indonesia, data Kementerian Sosial RI (2021) menunjukkan terdapat 251.113 penyandang disabilitas rungu-wicara yang terdaftar secara resmi, namun angka sebenarnya diperkirakan jauh lebih besar karena banyak yang belum teridentifikasi [3]. Individu dengan tunawicara menghadapi tantangan signifikan dalam komunikasi sehari-hari, yang dapat menyebabkan isolasi sosial, kesulitan dalam pendidikan, dan hambatan dalam mengakses layanan publik.

Penelitian oleh Terza Travelancya (2022) menunjukkan bahwa anak tunawicara memiliki kebutuhan khusus dalam pembelajaran dan komunikasi, termasuk kesulitan dalam menyampaikan ide dan gagasan secara langsung [4]. Studi lain oleh Sihotang et al. (2023) mengidentifikasi bahwa anak tunawicara sering melakukan tindakan non-verbal seperti menarik baju atau hijab guru untuk menyampaikan keinginan, menunjukkan urgensi akan sistem komunikasi yang lebih

efektif [5]. Penelitian oleh Nurmansyah et al. (2023) juga mengonfirmasi bahwa permasalahan komunikasi merupakan hambatan utama yang kerap terjadi pada penyandang disabilitas, termasuk tunawicara, yang berdampak pada kualitas hidup dan partisipasi sosial mereka [6].

Untuk mengatasi keterbatasan komunikasi ini, *Augmentative and Alternative Communication* (AAC) telah menjadi solusi yang terbukti efektif. AAC mencakup berbagai metode komunikasi yang dapat melengkapi atau menggantikan kemampuan berbicara atau menulis bagi individu dengan gangguan komunikasi [7]. *Picture Exchange Communication System* (PECS) merupakan salah satu sistem AAC yang paling banyak digunakan dan memiliki bukti ilmiah yang kuat untuk membantu individu dengan gangguan komunikasi dalam mengembangkan keterampilan komunikasi fungsional [8]. Penelitian di Indonesia oleh Christjahja dan Yusuf (2022) mengembangkan aplikasi AAC untuk anak autis di Surabaya yang membantu mengkomunikasikan keinginan dengan lebih mudah dan tepat, menunjukkan potensi besar teknologi AAC dalam konteks Indonesia [9]. Studi oleh Saleh dan Mutahara (2024) juga menunjukkan bahwa AAC menggunakan aplikasi smartphone berisi gambar dan suara efektif mengembangkan kemampuan komunikasi anak dengan kebutuhan khusus [10].

Namun, aplikasi AAC berbasis PECS yang tersedia saat ini masih memiliki keterbatasan signifikan, terutama dalam hal kualitas *text-to-speech* (TTS) berbahasa Indonesia. Sebagian besar aplikasi AAC menggunakan TTS generik yang tidak dirancang khusus untuk bahasa Indonesia, sehingga menghasilkan suara yang kurang natural, intonasi yang kaku, dan pelafalan yang tidak akurat. Penelitian oleh Erlani et al. (2022) menemukan bahwa sistem AAC dalam pembelajaran masih menghadapi kendala komunikasi, dimana sekitar 50% anak dengan gangguan komunikasi tidak mengembangkan kemampuan bicara yang memadai, sehingga alat komunikasi seperti AAC dengan teknologi TTS yang baik menjadi sangat diperlukan [11]. Penelitian oleh Hermawan et al. (2024) dalam pengembangan aplikasi E-Comm juga mengidentifikasi bahwa implementasi AAC memerlukan teknologi yang lebih interaktif dan adaptif untuk kebutuhan komunikasi anak di Indonesia [12].

Perkembangan teknologi TTS telah mengalami kemajuan pesat dalam beberapa tahun terakhir, terutama dengan hadirnya neural TTS yang menggunakan deep learning. IndexTTS2 merupakan model TTS autoregressive berbasis *zero-shot* yang dikembangkan oleh Bilibili, yang menawarkan kemampuan sintesis suara yang sangat natural dengan kontrol durasi yang presisi dan ekspresi emosi yang kaya [13]. Model ini mendukung multiple languages dengan kemampuan voice cloning yang akurat dan pemisahan antara ekspresi emosional dengan identitas pembicara [14]. IndexTTS2 telah menunjukkan performa yang superior dibandingkan model TTS *zero-shot* lainnya dalam hal *Word Error Rate* (WER), kemiripan pembicara, dan kesetiaan emosional, dengan *Mean Opinion Score* (MOS) mencapai 4.2+/5.0 yang mendekati kualitas suara manusia [13].

Namun, IndexTTS2 secara native dilatih untuk bahasa Mandarin Chinese dan Inggris, sehingga memerlukan adaptasi untuk bahasa Indonesia. Penelitian oleh Roosadi et al. (2024) menunjukkan bahwa pengembangan TTS bahasa Indonesia menggunakan Vall-E mencapai naturalness MOS 3.95 dan WER 19.71%, namun masih belum optimal untuk aplikasi real-time dan memerlukan dataset besar [15]. Fine-tuning merupakan pendekatan yang efektif untuk menyesuaikan pre-trained model dengan bahasa atau domain target, namun memerlukan computational resources yang signifikan [16]. *Low-Rank Adaptation* (LoRA) merupakan teknik *parameter-efficient fine-tuning* yang dikembangkan oleh Hu et al. (2021), yang dapat mengurangi trainable parameters hingga $10,000\times$ dan GPU memory hingga $3\times$ tanpa mengorbankan performa [17]. LoRA telah terbukti efektif untuk *Large Language Models* (LLMs), dan memiliki potensi besar untuk aplikasi pada model TTS seperti IndexTTS2 dalam mengadaptasi model untuk bahasa Indonesia dengan resource terbatas.

Integrasi IndexTTS2 yang di-fine-tune menggunakan LoRA untuk bahasa Indonesia ke dalam aplikasi AAC berbasis PECS pada platform Android memiliki potensi besar untuk meningkatkan kualitas komunikasi bagi individu dengan tunawicara di Indonesia. Penelitian oleh Carson et al. (2024) menunjukkan bahwa penggunaan teknologi dalam delivery AAC sama efektifnya dengan face-to-face training, membuka peluang untuk implementasi teknologi advanced seperti TTS neural

dalam aplikasi mobile [18]. Penelitian di Indonesia juga menunjukkan bahwa teknologi komunikasi interaktif seperti INCOM dapat membantu penyandang disabilitas dalam mengakses pembelajaran lebih efektif, dengan 95% siswa merasa terbantu [19]. Dengan mengimplementasikan teknologi TTS yang lebih natural dan ekspresif yang di-adapt untuk bahasa Indonesia, diharapkan aplikasi TUTUR dapat memberikan pengalaman komunikasi yang lebih baik, meningkatkan motivasi pengguna, dan pada akhirnya meningkatkan kualitas hidup individu dengan tunawicara di Indonesia.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan model text-to-speech IndexTTS2 bahasa Indonesia menggunakan fine-tuning LoRA untuk diintegrasikan ke dalam aplikasi mobile AAC berbasis PECS, khususnya untuk membantu komunikasi individu dengan tunawicara di Indonesia dan pengguna lain yang memerlukan dukungan komunikasi alternatif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana melakukan fine-tuning model IndexTTS2 menggunakan LoRA untuk bahasa Indonesia dan mengintegrasikannya ke dalam aplikasi AAC berbasis PECS pada platform Android?
2. Bagaimana performa dan kualitas output suara model IndexTTS2 yang telah di-fine-tune dalam bahasa Indonesia pada aplikasi TUTUR?
3. Bagaimana tingkat usability dan kepuasan pengguna terhadap aplikasi TUTUR untuk komunikasi individu dengan tunawicara?

1.3 Batasan Masalah

Untuk memfokuskan penelitian dan memastikan hasil yang optimal, penelitian ini memiliki batasan masalah sebagai berikut:

1. Sistem AAC yang diimplementasikan berbasis Picture Exchange Communication System (PECS) dengan fokus pada komunikasi dasar seperti *requesting*, *commenting*, dan *answering*.
2. Model TTS yang digunakan adalah IndexTTS2 dengan fine-tuning menggunakan LoRA untuk bahasa Indonesia.
3. Target pengguna aplikasi adalah individu dengan tunawicara, termasuk anak autisme non-verbal, individu dengan gangguan bicara, dan pengguna umum yang memerlukan dukungan komunikasi alternatif.
4. Evaluasi aplikasi difokuskan pada aspek fungsionalitas, usability (System Usability Scale), kualitas output suara TTS (Mean Opinion Score, Word Error Rate), dan *user satisfaction*.
5. Penelitian ini tidak mencakup pengembangan model TTS baru dari nol, melainkan mengadaptasi model IndexTTS2 yang sudah ada untuk bahasa Indonesia menggunakan LoRA.
6. Dataset bahasa Indonesia untuk fine-tuning menggunakan dataset open-source seperti Common Voice Indonesia atau data yang dikumpulkan khusus untuk penelitian ini dengan jumlah minimal 5 jam audio speech.
7. PECS symbols menggunakan simbol-simbol standar yang tersedia secara open-source atau disesuaikan dengan konteks budaya Indonesia.
8. Evaluasi dilakukan terhadap 15-20 responden yang terdiri dari pengguna tunawicara, caregiver, terapis, dan guru pendamping khusus di wilayah Bandung, Indonesia.

1.4 Tujuan

Tujuan dari penelitian ini adalah:

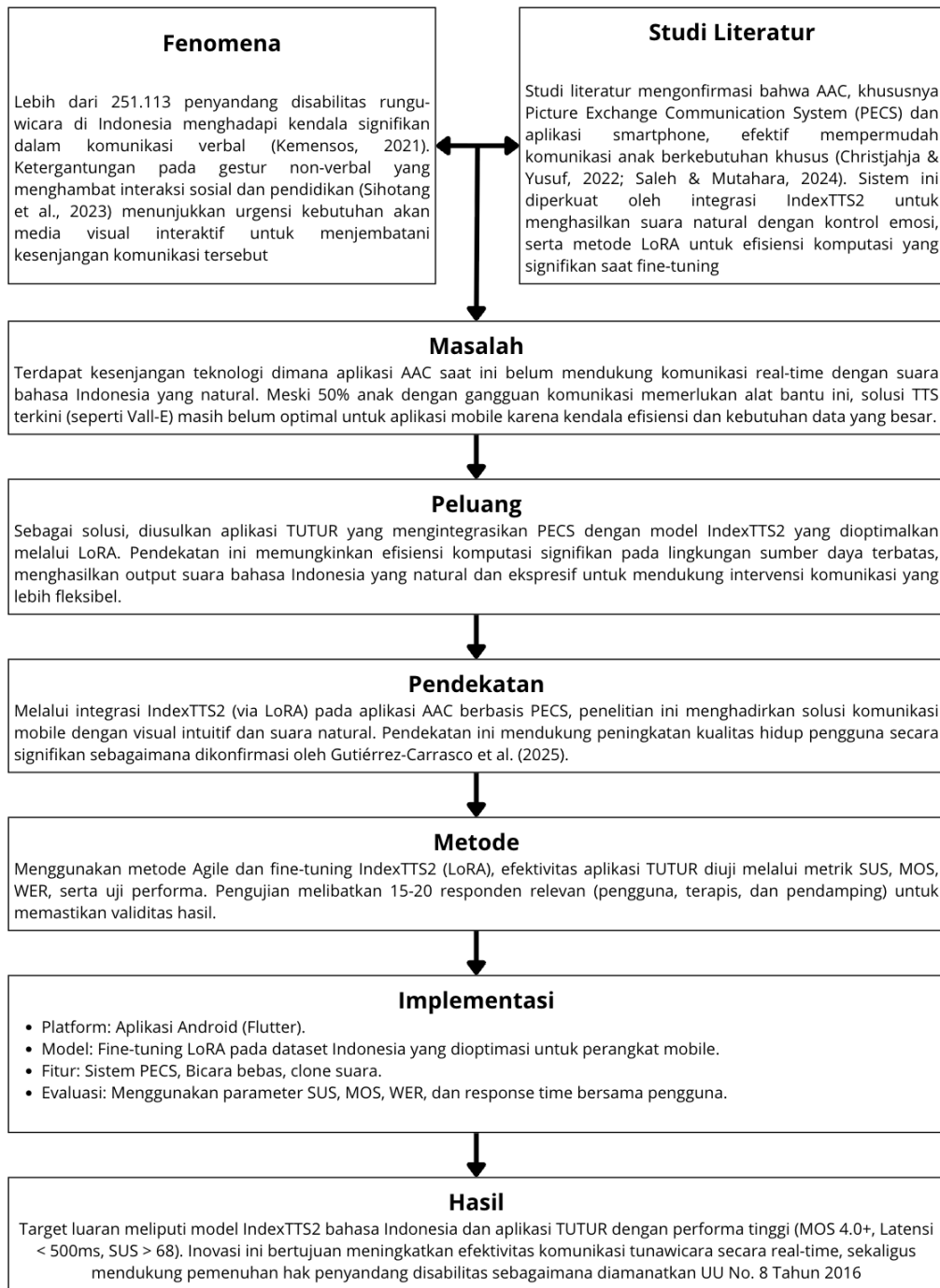
1. Melakukan fine-tuning model IndexTTS2 menggunakan teknik LoRA untuk menghasilkan sintesis suara bahasa Indonesia yang natural dan ekspresif, serta mengintegrasikannya ke dalam aplikasi AAC berbasis PECS pada platform Android.

2. Menganalisis performa dan kualitas output suara model IndexTTS2 yang telah di-fine-tune dalam bahasa Indonesia melalui pengukuran objektif (Word Error Rate, response time, inference latency) dan subjektif (Mean Opinion Score, naturalness, intelligibility).
3. Mengukur tingkat usability menggunakan System Usability Scale (SUS) dan kepuasan pengguna terhadap aplikasi TUTUR untuk komunikasi individu dengan tunawicara di Indonesia.

1.5 Kerangka Pemikiran

Kerangka pemikiran dalam penelitian ini disusun sebagai alur logis yang menjembatani fenomena permasalahan di lapangan dengan solusi berbasis teknologi yang dikembangkan. Secara garis besar, kerangka ini diawali dari permasalahan tingginya jumlah penyandang disabilitas tunawicara di Indonesia yang menghadapi hambatan komunikasi verbal dan masih bergantung pada komunikasi non-verbal yang terbatas. Di sisi lain, aplikasi *Augmentative and Alternative Communication* (AAC) berbasis *Picture Exchange Communication System* (PECS) yang tersedia saat ini umumnya masih menggunakan mesin *Text-to-Speech* (TTS) generik dengan artikulasi yang kaku dan belum teroptimasi untuk bahasa Indonesia. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan pengembangan aplikasi alat bantu komunikasi bernama TUTUR pada platform Android berbasis Flutter yang mengintegrasikan model TTS mutakhir IndexTTS2.

Adaptasi bahasa Indonesia dilakukan secara efisien menggunakan teknik *Parameter-Efficient Fine-Tuning* (PEFT) melalui *Low-Rank Adaptation* (LoRA) agar menghasilkan artikulasi suara yang natural dan ekspresif tanpa membebani komputasi perangkat bergerak melalui arsitektur cloud API. Efektivitas sistem diuji secara komprehensif melalui evaluasi kualitas suara (*Mean Opinion Score* dan *Word Error Rate*), latensi komputasi, serta tingkat kebergunaan (*System Usability Scale*) bersama para pengguna sasaran dan pendamping di Sekolah Luar Biasa. Alur keterkaitan antara fenomena, masalah, peluang, pendekatan, implementasi metode, hingga target luaran yang diharapkan digambarkan secara terstruktur pada Gambar 1.1 berikut



Gambar 1.1: Kerangka Pemikiran

1.6 Sistematika Penulisan

Sistematika penulisan ini disusun untuk memberikan gambaran umum mengenai penelitian yang dilakukan. penelitian ini terdiri dari lima bab adapun struktur dan penjelasan dari setiap bab adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, kerangka pemikiran dan sistematika penulisan mengenai pengembangan model Text-to-Speech (TTS) berbahasa indonesia untuk aplikasi alat bantu komunikasi penyandang disabilitas tunawicara.

BAB II KAJIAN LITERATUR

Bab ini memuat penelitian terdahulu, kesenjangan penelitian, serta landasan teori yang mencakup tunawicara, *Augmentative and Alternative Communication* (AAC), *Picture Exchange Communication System* (PECS), *Text-to-Speech* (TTS), arsitektur model *IndexTTS2*, *Low-Rank Adaptation* (LoRA), *Android application development*, *Flutter*, *Deployment model*, *System Usability Scale* (SUS), dan metodologi *CRISP-DM* sebagai dasar dalam pengembangan sistem yang diteliti.

BAB III METODOLOGI

Bab ini menjelaskan metode penelitian berbasis kerangka kerja *CRISP-DM*, mencakup *business understanding*, *data understanding* (dataset *Common Voice*, *FLEURS*, dan *ID2*), *data preparation*, *modeling* (*fine-tuning LoRA* pada *IndexTTS2*), *evaluation model* (*Word Error Rate*, *inference latency*, dan *Mean Opinion Score*), *deployment* menggunakan *Modal.com*, *usability testing* menggunakan kuesioner *System Usability Scale* (SUS), dan terakhir analisis hasil *usability*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil implementasi sistem, termasuk pelatihan dan evaluasi performa model *IndexTTS2* dengan eksperimen variasi *rank LoRA* (32, 64, 128) untuk sintesis suara bahasa Indonesia, analisis metrik evaluasi seperti *Word Error Rate*

(WER), *inference latency*, dan *Mean Opinion Score* (MOS), serta hasil *usability testing* (SUS) terhadap aplikasi TUTUR melibatkan pengguna tunawicara, *caregiver*, dan terapis. Bab ini juga membahas temuan, interpretasi hasil, dan perbandingan dengan penelitian terdahulu.

BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan dari seluruh penelitian sebagai jawaban atas rumusan masalah serta saran untuk pengembangan lebih lanjut, baik dalam peningkatan performa model TTS, perluasan dataset audio bahasa Indonesia, pengayaan fitur aplikasi AAC-PECS TUTUR, maupun implementasi skala lebih luas untuk mendukung kemandirian komunikasi bagi penyandang disabilitas tunawicara.

