19EAC383

Technical Presentation

Final Report

Submitted by

BL.EN.U4EAC22025	K. Tappan Chengappa
BL.EN.U4EAC22026	Krishna P. T.
BL.EN.U4EAC22028	Krithik Ramesh Iyer
Guided by:	
Dr. Sreeja Kochuvila	
Examiner 1	**Examiner 2**
Dr. Pooja Kenchetty	Dr. Lalitha S.

BACHELOR OF TECHNOLOGY

IN

ELECTRONICS AND COMPUTER ENGINEERING AMRITA SCHOOL OF ENGINEERING, BANGALORE AMRITA VISHWA VIDYAPEETHAM

BENGALURU-560035

REPORT

Introduction:

Problem Statement: India's linguistic landscape, comprising over 19,500 languages and dialects, presents significant communication barriers, particularly in rural and tribal regions. Existing machine translation solutions, such as Google Translate, are constrained by their dependence on internet connectivity and text-based outputs, rendering them ineffective in offline scenarios and inaccessible to non-literate populations. This technological gap exacerbates social exclusion, impedes access to essential services, and hinders cross-cultural communication.

Motivation and Significance: The development of an offline, speech-enabled translation system addresses a critical need for equitable communication tools in linguistically diverse, resource-constrained environments. Such technology holds particular relevance for:

- **Healthcare**: Facilitating patient-provider communication in remote clinics
- **Education**: Enabling instruction in native languages
- **Governance**: Improving access to public services for marginalized communities The proposed solution distinguishes itself through real-time audio translation capabilities, portability, and internet independence, making it uniquely suited for deployment in underserved areas.

Alignment with Sustainable Development Goals: This initiative supports the United Nations' 2030 Agenda through direct contributions to:

1. **SDG 10 (Reduced Inequalities)**: Mitigating language-based discrimination by providing universal access to communication technology
2. **SDG 16 (Peace, Justice and Strong Institutions)**: Promoting social cohesion through improved intercultural dialogue

The project embodies the principle of leaving no one behind by specifically targeting populations excluded from digital infrastructure, thereby advancing inclusive development objectives.

Literature survey:

This survey synthesizes advancements and challenges in machine translation (MT), speech-to-text (STT), and text-to-speech (TTS) technologies, with a focus on low-resource languages and hardware integration. The analysis spans six key studies, highlighting conceptual frameworks, methodological innovations, and unresolved gaps relevant to developing an offline, speech-enabled translator.

1) Machine Translation Methodologies:

Hybrid and Neural Approaches:

- **Hybrid Models**: Jayashree Nair et al. (2016) demonstrated a **Hybrid Machine Translation (HMT)** system combining rule-based (RBMT) and statistical methods (SMT) for English-Hindi translation, achieving **96% accuracy** with strong grammatical correctness. However, reliance on rules limited scalability for complex sentences and low-resource languages.
- **Neural Machine Translation (NMT)**: Asmitha M et al. (2024) employed an **encoder-decoder architecture with attention mechanisms** for English-Telugu translation, improving semantic coherence (87.09% accuracy). A critical gap was the absence of real-time translation evaluation, underscoring the need for latency optimization in practical deployments.

Rule-Based Systems:

- Wijerathna et al. (2012) developed a **rule-based Sinhala-English translator** using tokenization and parallel processing. While efficient for structured grammar, it struggled with ambiguity and limited vocabulary, highlighting the necessity of integrating data-driven approaches for robustness.

---

2) Hardware-Software Integration:

- Hirthik T H et al. (2024) pioneered **ML-assisted sign language gloves** using Arduino and Raspberry Pi, achieving **94% accuracy** in recognizing 10 hand signs. Their work validated the feasibility of edge-computing for real-time applications but was restricted to one-way communication (sign-to-speech only). This emphasizes the need for bidirectional systems in spoken language translation.

3) Speech Technologies:

TTS/STT Model Evaluation:

- Surabhi Sudhan et al. (2024) systematically compared **deep learning vs. rule-based TTS/STT models**, noting that most were benchmarked on English datasets, limiting applicability to low-resource languages.
- Nandish et al. (2023) evaluated regional language tweet translators using metrics like BLEU and MOS but faced **limited language coverage**, reinforcing the challenge of scaling speech technologies across India's diverse dialects.

Synthesis and Relevance to the Current Project - The reviewed studies collectively underscore three critical insights:

1. **Hybrid Architectures**: Combining NMT's semantic flexibility with rule-based grammatical rigor could enhance translation quality for complex sentences.
2. **Edge Computing**: Raspberry Pi's proven role in hardware prototypes (e.g., sign language gloves) supports its use for portable, offline translation.
3. **Low-Resource Language Gaps**: Most TTS/STT models lack multilingual training data, necessitating curated datasets for tribal/regional languages.

By addressing these gaps—through optimized NMT models, bidirectional speech processing, and expanded language coverage—this project advances toward a scalable, inclusive communication tool aligned with SDGs 10 and 16.

Conceptual Breadth: Encompasses MT methodologies (hybrid, neural, rule-based), hardware integration, and speech technologies.

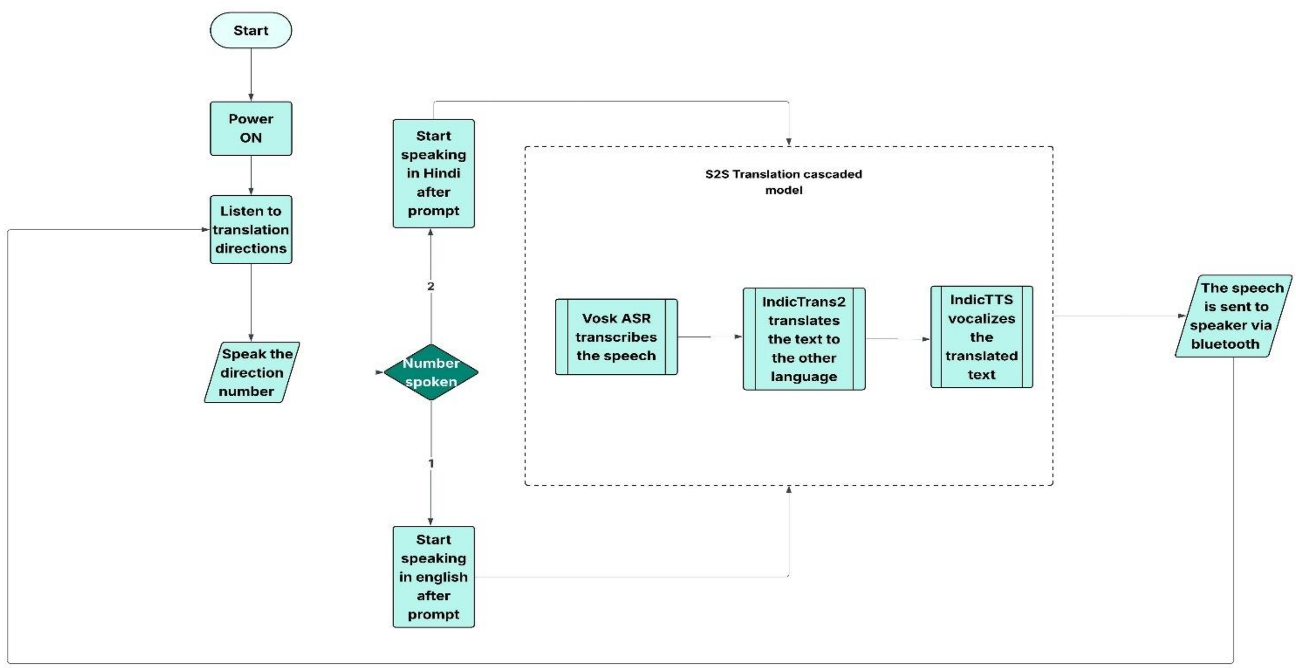
Depth: Critically evaluates accuracy metrics, scalability constraints, and sociotechnical applicability to marginalized communities

S.No	Authors name(s)	Full title of the paper	Inference from the paper	Gaps
1	Jayashree Nair, Amrutha Krishnan K, Deetha R.	An Efficient English to Hindi Machine Translation System Using Hybrid Mechanism	Hybrid Machine Translation (HMT) combining Rule-Based (RBMT) and Statistical Machine Translation (SMT). Achieves **96% accuracy** and enhances **grammatical correctness**.	-Limited Handling of Complex Sentences -Dependency on Rule-Based Approaches -Limited Lexical Resources
2	Asmitha M, Kavitha C.R.	Bridging the Language Gap: Enhancing English-to-Telugu Translation using NMT and Encoding-Decoding Techniques	**Neural Machine Translation (NMT)** using Encoder-Decoder with Attention mechanism. Achieves **87.09% accuracy** and improves **semantic translation**.	-Challenges with Low-Resource Language Data -Lack of Comparison with Other Models:* -Real-Time Translation Considerations Missing
3	Hrithik T H, Rhethika S, K H Akhil, K Deepa	ML Assisted Sign Language to Speech Conversion Gloves for the Differently Abled	-Arduino Mega processes sensor data and transmits it to Raspberry Pi -Raspberry Pi runs a Random Forest ML model for sign recognition -Converts detected signs into speech using text-to-speech synthesis Achieves **94% accuracy** for detecting 10 hand signs	-Limited dataset (only 10 alphabets tested) -Limited to one-way communication (no speech-to-sign conversion)
4	H D Nandish, Rahul R, T K Luchingba, Neelima N	A Novel Implementation of a Cohesive Regional Language Tweet Translator	Covers **comparative analysis of TTS and STT models**, focusing on **speech-based translation** and evaluation using **BLEU, MOS, and MSE**.	Limited Language Coverage: The system primarily focuses on Indian regional languages and English, potentially excluding other languages and dialects that are not covered by the Indic-Transliteration library.
5	Surabhi Sudhan, Parvathy P Nair, MG Thushara	Text-to-Speech and Speech-to-Text Models: A Systematic Examination of Diverse Approaches	Reviews **TTS/STT models** for **assistive tech and language translation**, comparing **deep learning vs. rule-based approaches**.	-Limited Language Coverage -Many models are primarily evaluated on English datasets, limiting their applicability to other languages, especially low-resource languages.
6	L. Wijerathna, W.L.S.L. Somaweera, S.L. Kaduruwana, Y.V. Wijesinghe, D.I. De Silva, K. Pulasinghe, S. Thellijjagoda	A Translator from Sinhala to English and English to Sinhala (SEES)	Rule-based machine translation system Tokenization for **breaking sentences** into components. Rule-based translation engine using a knowledge base. **Parallel processing** with threading for performance optimization. Handles active and passive voice. Uses **C#.NET**, **SQL** 2005	-Cannot handle ambiguous words perfectly without human input . - Limited vocabulary and grammar support. -Focused only on Sinhala and English

Methodology:

The project employs a three-stage modular architecture to achieve offline, real-time speech translation: **Speech-to-Text (STT)**, **Machine Translation (MT)**, and **Text-to-Speech (TTS)**. Using a Raspberry Pi 5 as the core processing unit, the system captures audio input via a microphone, processes it through the Vosk STT model for transcription, translates the text using the IndicTransV2 transformer-based model (fine-tuned for low-resource Indic languages), and synthesizes the translated text into natural speech via the A14BITS TTS engine. The translated audio is output through a Bluetooth speaker, enabling seamless bidirectional communication (e.g., Hindi ↔ English). Key technical optimizations include model quantization for edge deployment on Raspberry Pi, parallel processing to reduce latency, and offline functionality by pre-loading all models.

Block Diagram:



Comparing with existing methods:

The following analysis evaluates existing translation technologies and hardware solutions against the proposed offline speech translator, focusing on cost, efficiency, scalability, and key usability factors.

1. Google Translate:

- **Cost**: Free for users but requires continuous internet access, which incurs data costs in remote areas.
- **Efficiency**:
  - BLEU Score: 34.5–42.8 (high accuracy for mainstream languages).
  - Latency: Dependent on internet speed; unsuitable for real-time use in low-connectivity regions.
- **Scalability**: Supports 100+ languages but performs poorly on low-resource dialects (e.g., tribal languages).
- **Key Limitations**: Internet dependency, text-only output, no audio support for marginalized populations.

2. Hybrid Machine Translation (HMT):

- **Cost**: Moderate (combines rule-based and statistical models; requires initial setup).
- **Efficiency**:
  - Accuracy: 96% for simple sentences (e.g., English-Hindi).
  - Fails with complex grammar or slang due to rigid rule-based components.
- **Scalability**: Limited by manual rule creation; adding new languages is labor-intensive.
- **Key Limitations**: Struggles with ambiguity and low-resource languages.

3. Neural Machine Translation (NMT):

- **Cost**: High (computationally intensive training; GPU resources required).
- **Efficiency**:
  - Accuracy: 87.09% (semantically coherent but untested for real-time use).
  - BLEU Score: ~37.6 for Indic languages (lower than Google Translate).
- **Scalability**: Adaptable to new languages but requires large datasets, which are scarce for tribal dialects.
- **Key Limitations**: High latency in edge deployments, energy-intensive.

4. Rule-Based Systems (e.g., Sinhala-English Translator):

- **Cost**: Low post-development but high upfront investment in linguistic expertise.
- **Efficiency**:
  - Works well for structured grammar but fails with ambiguous phrases.
  - Limited vocabulary and grammar support.
- **Scalability**: Not scalable due to manual rule updates.
- **Key Limitations**: Inflexible, excludes dialects outside predefined rules.

5. Sign Language Gloves:

- **Cost**: Moderate (100–100–200 for Arduino/Raspberry Pi components).
- **Efficiency**:
  - Accuracy: 94% for 10 hand signs.
  - One-way communication (sign-to-speech only).
- **Scalability**: Limited to tested sign languages; not applicable for spoken languages.
- **Key Limitations**: Niche use case, no bidirectional interaction.

Proposed Offline Speech Translator:

- **Cost**: Affordable (Rs 9000 for Raspberry Pi 5, microphone, and speaker).
- **Efficiency**:
    - BLEU Score: ~22.25 (lower than Google but optimized for Indic languages).
    - Real-time processing: <2 minutes latency via parallel STT-TTS pipelines.
- **Scalability**:
    - Modular design supports adding tribal languages with curated datasets.
    - Edge-compatible models (IndicTransV2, A14BITS) reduce cloud dependency.
- **Key Advantages**:
    - **Offline Functionality**: Eliminates internet costs and connectivity barriers.
    - **Audio-Centric Design**: Prioritizes non-literate users with natural speech output.
    - **Targeted Inclusion**: Focuses on underserved languages (e.g., Kannada, Hindi) excluded by mainstream tools.

Summary:

Approach	Cost	Efficiency
Google Translate	Free (data)	High (internet)
HMT	Moderate	Rigid accuracy
NMT	High	Semantic coherence
Rule-Based Systems	Low (post-dev)	Structured grammar
Sign Language Gloves	Moderate	Niche accuracy
Proposed Translator	Rs 9000	Real-time, audio

Approach	Scalability	Key Differentiator
Google Translate	Broad but shallow	Internet dependency
HMT	Low	Rule-based constraints
NMT	Moderate	Data hunger
Rule-Based Systems	None	Inflexibility
Sign Language Gloves	Limited	Unidirectional
Proposed Translator	High (modular)	Offline, inclusive, low-resource focus

Conclusion:

The proposed **offline speech translator**, integrating edge computing with IndicTransV2 and A14BITS models, emerges as the optimal solution for bridging language divides in low-resource settings. By prioritizing offline functionality, real-time audio translation, and affordability (~10000), it overcomes the internet dependency of Google Translate, the rigidity of rule-based systems, and the high costs of NMT frameworks. While its BLEU scores (22.25) trail behind cloud-based tools, its modular design allows scalable inclusion of tribal languages, directly advancing SDG 10 (reduced inequalities) and SDG 16 (inclusive governance). Future enhancements, such as hybrid AI architectures and Coral TPU acceleration, could further refine latency and accuracy, solidifying its role as a transformative tool for marginalized communities—proving advanced technology can be both accessible and equitable.