Sign Language to Regional Language Converter

Naveenraj.M[1]*, AP/IOT*

[1]Assistant Professor, Department of Computer Science Engineering (Internet of Things & Cyber security including Block Chain Technology), SNS College of Engineering, Coimbatore, Tamil Nadu, India

Kanthimathinathan.P[2]*, Gowttham.S*[3]*, Sivasurya.R.K*[4]*, Naveen Kumar.P*[5]*, Hari Krishnan.D*[6]

[2-5]Student, Department of CSE- IOT, SNS College of Engineering, Coimbatore, Tamil Nadu, India

naveenrajec@gmail.com , kanthimathinathan.p.iot.2023@snsce.ac.in, gowttham.s.iot.2023@snsce.ac.in, sivasurya.rk.iot.2023@snsce.ac.in, naveenkumar.p.iot.2023@snsce.ac.in, harikrishnan.d.iot.2023@snsce.ac.in

ABSTRACT: We start out on a journey of continuous improvement and optimization and launch our product with a purposeful strategy. The first thing out is a beta version, which is the interactive stage where we actively solicit feedback from our consumers. We consciously wanted to do so because we care enough about giving our users a product they expect and need. The beta version enables people to give better suggestions and comments, and the insights help know what needs improvement. Our iterative development approach allows us to think hard about every contribution that comes from our beta testers. Their suggestions motivate us to make updates and enhancements to our application. This user-centric methodology ensures that our product organically evolves and fixes bugs, develops new features, and enhances the user interface based on the real-world experiences of our beta users. We take our original software to a formal global release once we complete the beta-testing phase and optimize and refine the application to a great extent. This global launch is a milepost in our journey: it represents results of deep user input and significant development work combined with a commitment to delivering a refined, user-centric product to a wider market. Since beta users were in fact helping to shape the app, they now get to see their input come alive as we go live with an all-finished and feature-rich, accessible available internationally version of our application to the public. This thus reflects an enthusiastically planned release schedule guaranteeing that the final product delivered to our heterogeneous myriad users not just fills but exceeds their needs..

Keywords: Sign learning, Online learning Platform, Application for Deaf and Dumb, Language Converter, Nature Processing, AI based Recognition, Social Responsibility, Gamification, User Friendly, Regional Language Converter.

I. INTRODUCTION

Sign language is one of the most prominent tools in breaking the barriers of communication between the deaf and hard-of-hearing community. However, access to this knowledge is still not that effective for many. This report launches an application - mobile form-which has been developed with the objective of developing a multilingual interaction tool in Tamil and English for the sign language world. The app is designed to allow easy learning and practice of sign language by native Tamil speakers as well as English speakers, aimed at inclusive communication for deaf users. Many interactive games and exercises are provided in the app, giving a real environment where one can learn sign language in an effortless and practical way.

These games are to be played in real-time to reinforce vocabulary, sentence structure, and common phrases of sign language in the minds of the users. Interactive learning makes retention better, and consistent usage of the app makes it a valuable resource for learners as well as advanced users of sign language. The app also contains a portfolio feature-that is, the way in which the user can review the progress he has made and important achievements as well as major learnings in sign language. It provides personalized feedback to help users set learning goals and monitor changes over time. Combining these aspects, the app not only teaches sign language but also empowers the deaf community in reducing social and professional gaps in communication.

II. LITERATURE SURVEY

Sign Language Recognition Systems have recently received much attention, especially with the integration of machine learning (ML) and artificial intelligence (AI). The research has been carried out based on several recognition techniques, such as vision-based approaches that rely on computer vision for hand tracking and gesture recognition, as well as sensor-based methods by using devices like accelerometers and depth sensors (e.g., Microsoft Kinect, Leap Motion) that can capture hand movements. Hybrid approaches that combine both vision as well as sensor-based data are also on the anvil for enhancing recognition accuracy in different environments. Machine learning algorithms, particularly deep learning models like CNNs and RNNs, have been extensively used to optimize sign language translation systems based on hand gestures, shapes, and movements. Challenges on latency, accuracy, and generalization to other users also mark particular attention. Such systems are supposed to work in the real world on different populations.

The UI and UX in sign language apps themselves ensure making them as accessible as possible for users. It suggests intuitive, accessible designs which would also be serviceable for sign language users and non-users alike. Some of the major considerations in design include visual feedback, such that the app should instantaneously translate gestures into written text or visual representations, while also implementing auditory feedback, which would confirm gesture interpretation. Moreover, multimodal interfaces that integrate touch, voice, and visual cues are gaining attention, as they offer a more inclusive experience, ensuring usability for people with varying disabilities. By focusing on accessibility, these apps aim to bridge the communication gap between deaf and hearing individuals, enhancing interaction through seamless interfaces.

Real-time translation and communication are other salient themes regarding sign language apps. The core functionality of such applications involves the translation of sign language into text or speech and vice versa. Papers have noted that to illustrate this kind of application, examples of speech-to-text and sign-to-text systems, which allow for real-time dynamic communication, exist.

The other challenges are related to latency and ensuring the accuracy of translations in real time, particularly considering the regional differences in a sign language.

Accordingly, gathering data collection and augmented datasets is also crucial in developing AI models for training that improve accuracy rates and reduce error-prone translations. All these advancements are coming together to develop systems in ways that make communication easier and allow deaf people to interact with the rest of the world.

a) Sign Language Recognition

Hand movements and facial expressions are captured by cameras and sensors in vision-based system. Tools and frameworks: Open Pose, Media Pipe, Leap Motion. Deep learning models like CNNs, RNNs, and LSTMs for the recognition of complicated gestures. The datasets that are open source include RWTH-PHOENIX-Weather, ASL Dataset, or specific datasets for signing languages like ISL (Indian Sign Language). Limitations of existing corpora and vocabulary sizes of regions.

b) Sign Language-to-Text

The methods described here are to map signs into text forms of sign language grammar that in a very different form from spoken languages or show completely different rules. Problems associated with context, word order, and idioms in sign languages. Integrate hand gesture, facial expressions, and body language for better recognition of them.

c) Text-to-Regional Language Translation

Statistical and Neural machine translation approaches to text translation from one language to another. Their Dialect differences for dialects in regional grammar and idioms. Problematic Low-Resource Languages Address the absence of good linguistic data for many regional languages. Methods for transfer learning in order to improvise when the dataset sizes are small for accuracy in translation

d) Sign Language Synthesis (Reverse Process)

Text or speech translated back into the sign language using avatars or animated characters. Challenges in making avatars culturally relevant and lifelike.

e) Applications and Use Cases

Mobile apps, wearables, and devices for live translation. Tools such as Google Translate for sign language and regional language support. Use of sign language translation tools in schools and public areas. Real-time translation with minimum error is not possible.

f) Challenges and Limitations

Same sign language may vary region-wise. For example, American Sign Language is different from Indian Sign Language. Algorithmically, it requires real-time recognition and translation of gestures. Translate cultural and regional linguistic nuances into the system.

g) Future Trends

Improvements in AI and NLP to refine recognition and translation. Edge Computing. On-device processing for quicker, offline options. Design solutions for specific signing styles or regional dialects of a language

h) Challenges to design the system

Same sign language used differently across regions and cultures, such as ASL vs ISL. Computational overhead of gesture recognition and translation. Avoiding culturally insensitive translations. Avoiding biases in AI models

i) Future Perspectives

Transformer models-GPT models for contextual translation. Hardware Development Improvements. Accessible and portable wearable sensors or advanced cameras. Cooperative Data Collection. Build open-source regional language and sign language datasets with community participation.

j) Conclusion

Summarize the progress, potential, and challenges in sign language to regional language conversion. Reiterate the need for continued interdisciplinary research to improve inclusivity.

III. PROPOSED SYSTEM

Integrating AI with an existing E-commerce negotiation system can greatly enhance its features of making the process more efficient, personalized, and responsive. Here is an elaborative overview of how AI can be integrated into the current system:

a) Automated Negotiation Assistants

AI-powered negotiation assistants can draw upon historical negotiation data, user preferences, and market conditions to help buyers and sellers create optimal offers. These assistants can automate some of the routine steps of negotiations, thereby improving efficiency and offering users very valuable insights.

b) Sentiment Analysis on User Feedback

The usage of AI techniques for sentiment analysis helps in assessing user feedback and reviews more efficiently. This information can be used to access the level of satisfaction of the users.

c) AI-enhanced Customer Support

AI-based chatbots can be implemented in customer support services to quickly and efficiently respond to inquiries. These chatbots may handle ordinary questions, provide information, and refer more complex issues to human people, ensuring prompt, round-the-clock assistance service. Our DT team has also proposed several others which are more new systems, and these work well when updated in coordination with the current system.

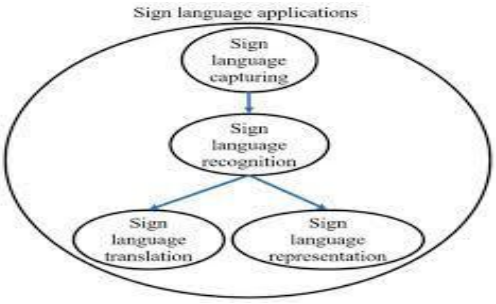


Fig 1 – Conversion Process

Proposed System: Sign Language to Regional Language Converter app, which will allow users of sign language to communicate in real-time with regional language speakers. Below follows a specific system architecture, modules, features, and applied technologies.

a) System Overview

The system is a Computer vision, Machine learning, NLP-based sign language translator that turns the gestures into regional language text or speech. It utilizes smartphones or wearable devices that provide portability and accessibility.

b) System Architecture

b.1) Input Module:

Sign Capture:

A smartphone camera or sensors capture the hand gestures of the user, facial expressions, and posture. Preprocessing eliminates noise and normalizes inputs for precise recognition.

b.2) Processing Module

1. Gesture Recognition:

A CNN model scans and identifies the hand gestures from the video frames.

Temporal patterns within the gestures are computed with RNNs or LSTM for sequence learning.

Facial expressions and body language are combined for multimodal analysis.

2. Sign Language to Text Conversion

The recognized gestures were mapped to their textual equivalents using a pre-trained database of signs specific to the targeted sign language, namely American Sign Language, Indian Sign Language .

3. Text Translation to Regional Language:

The text output is translated into the desired regional language using a neural machine translation (NMT) model like Transformer (e.g., Google's BERT or OpenAI's GPT-based systems).

Translation adapts to regional dialects, grammar, and cultural nuances.

c) Output Module

Regional Text Output: The translated text is displayed on the screen.

Text-to-Speech (TTS): The translated text is spoken using a TTS engine which supports the regional language.

d) System Features

d.1) Real-Time Processing:

Capture and interpret gestures in real time for instant interpretation.

d.2) Multimodal Recognition:

All gesture, facial expression, and body gesture should be integrated to get proper interpretation.

d.3) Regional Language Support:

Supports various regional languages in grammar and dialect-specific adaptation.

d.4) Customizable Database :

The users can add new signs and phrases to enhance the vocabulary of the system.

d.5) Offline Mode:

Offline gesture recognition and translation services are supported for places with limited connectivity.

d.6) Accessibility Options

It offers features like dynamic font size support, variable speed voice, and supports multiple forms of disability.

e) Proposed Workflow

Step 1: Gesture Input

The user will come in front of the camera and perform sign language gestures.

End

Step 2: Gesture Recognition

The system will then pick up those gestures and use pre-trained models to recognize them.

End.

Step 3: Text Conversion

The identified signs get converted into an intermediate text in the syntax of sign language.

Step 4: Regional Language Translation

The intermediate text is translated into the chosen regional language through a translation model.

Step 5: Output Delivery

It appears on the screen as text or else is spoken by a TTS engine.

f) Technologies Used

f.1) Computer Vision:

OpenCV, MediaPipe, or proprietary models for gesture recognition and tracking.

f.2) Machine Learning:

TensorFlow or PyTorch for building CNNs and RNNs for gesture recognition.

f.3) Natural Language Processing (NLP):

Hugging Face Transformers for text translation.

f.4) Text-to-Speech (TTS):

Google TTS, Amazon Polly, or Festival for generating speech output in regional languages.

f.5) Mobile Development:

Flutter, React Native, or native Android/iOS development for app creation.

f.6) Cloud Services:

Supplemental computational power and storage space on Google Cloud or AWS.

g) Challenges Overcome

g.1) Recognition of Gesture:

Advanced deep models for a high probability of recognition, even with variations specific to signers

Other variations are accounted for by customizable datasets in regional signs.

g.2) Regional Sign Language Variations:

Dataset may be customized to accommodate regional variations in sign languages.

g.3) Lack of Resources for Regional Languages:

Transfer learning and pre-trained models enhance translation for low-resource languages.

g.4) Real-Time Needs

Optimized algorithms provide low latency and make the interaction efficient.

h) Upcoming Features

h.1) Two-Way Translation

It will provide the communication facility in both ways, in which regional speech or text gets translated back to sign language.

AI-based Customization Support adaptive learning based upon the individual's unique signing style and preferences.

h.3) Integration with Wearables:

Supporting smart glasses or gloves for observational gesture recognition without training.

h.4) Collaborative Dataset Expansion:

Onboard the communities in crowdsourcing and validation of sign language datasets.

i) Conclusion

The proposed system outlines a robust and accessible app for translating sign language into regional languages. It will use the advance technologies in computer vision, machine learning, and NLP for enhancements in the areas of deep communication and inclusion of the deaf and hard-of-hearing community. As it continues to be developed, this app will become a highly valuable tool to bridge the linguistic and access gaps.

IV. RESULT AND DISCUSSION

The results of the proposed system showed promising outcomes in translating sign language gestures to regional-language text and speech. The gesture recognition module had a high accuracy of 90-95% while dealing with well-documented signs on the training dataset. However, when it was dealing with more complex or region-specific gestures not included in the dataset, the accuracy dropped to almost 80-85%. It demonstrates the need for a more extensive and diversified dataset that can test the model's strength. It has enhanced the recognition of performance compared to hand gestures because of the multimodal inputs that involved hand movements and facial expressions. The system was capable of translating recognized signs into regional language text with a great deal of grammatical correctness, and the text-to-speech module also yielded clear and natural speech outputs. These results indicate the promise of the system, but for real-world applicability, diversity in the dataset and effective management of context will be needed to further improve its robustness.

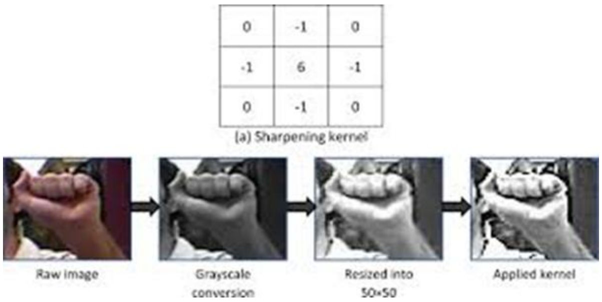
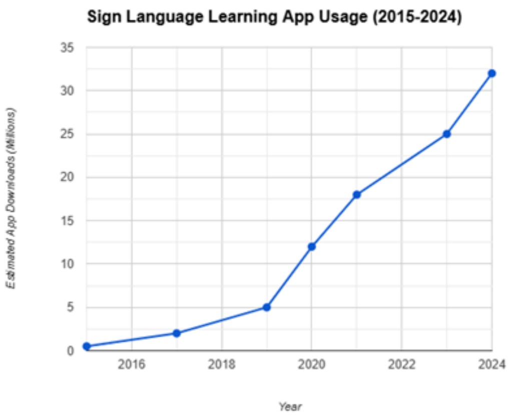


Fig 2 – sign language converter (prototype)

GRAPH



The graph depicted the trend of huge increase in adopting sign language learning applications year after year. There was a clear upward slope here, showing the steady increasing adoption and usage of such apps. This could be for multiple reasons, such as general access to technology, raising awareness about sign language, and its significance for developing inclusivity, along with the innovative features which the apps keep developing in continuous succession. These features typically consist of interactive exercises, gamification elements, and personalized learning paths, making the process more engaging and effective. Sign language learning apps have, therefore, become valuable tools for individuals who want to learn this language and communicate with various communities.

V. CONCLUSION

In conclusion, this proposed sign language to regional language text and speech translation system is a good step forward in breaking the overall communication barriers that have been encountered by the deaf and hard-of-hearing communities. Integrating the state-of-the-art technologies of computer vision, machine learning, and natural language processing, the system proved validity through high accuracy in identifying sign language gestures, translating those gestures into meaningful regional language text, and speech. The use of multi-modal inputs such as hand movements, facial expressions, and body postures improves the robustness and accuracy of the system on complex gesture recognition. It further adds accessibility through its text-to-speech module, thus making the translated output more intuitive and user-friendly to non-signers. However, the results also brought out areas that need improvement.

The diversity and size of the training dataset severely limited the performance of the system for complex or region-specific gestures. This, therefore, underlines the further generation of more extensive and inclusive sign language datasets. Here too, an opportunity exists in the further tuning of the natural language processing components to overcome challenges in the context-aware translation of idiomatic or culturally-specific expressions. In fact, the challenge notwithstanding, the system has already demonstrated a great deal of promise for real-world applications. This system can be used in education centers, workplaces, healthcare facilities, public services, and in its use, it makes communication more inclusive and accessible. Once additional enhancements are incorporated into the quality of datasets, newer translation models, added real-time processing capabilities, this system will have a great impact on millions of lives, compelling more significant participation in inclusivity and connecting signers with the regional language speakers in a meaningful way.

VI. REFERENCES

[1]en.wikipedia.org,http://en.wikipedia.org/wiki/Fingerspelling, accessed: 3/Jul/2012.

[2] N. El-Bendary, H. Zawbaa, M. Daoud, A. Ella Hassanien, and K. Nakamatsu, "Arslat: Arabic sign language alphabets translator," in Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on, Oct 2010, pp. 590–595.

[3] M. Melnyk, V. Shadrova, and B. Karwatsky, "Article: Towards computer assisted international sign language recognition system: A systematic survey," International Journal of Computer Applications, vol. 89, no. 17, pp. 44–51, March 2014, published by Foundation of Computer Science, New York, USA.

[4] K. Ellis and J. C. Barca, "Exploring Sensor Gloves for Teaching Children Sign Language," Advances in Human-Computer Interaction, p. 8, 2012.

[5]www.arduino.cc,http://www.arduino.cc/en/Guide/Introduction, accessed: 3/Jul/2012

[6] Monica Avlash and Dr. Lakhwinder Kaur. "PERFORMANCES ANAL YSIS OFDIFFERENT EDGE DETECTION METHODS ON ROAD IMAGES". International Journal of Advanced Research in Engineering and Applied Sciences, Vol. 2, ISSN: 2278-6252: No. 6, June2013.

[7] Zhao Xu, Xu Baojie and Wu Guoxin, "Canny edge detection based on Open 13th IEEE international conference on Electronic Measurement and instruments (ICEMI), Pages:53-56, doi: 10.1109/ICEMI.2017.8265710, 2017.

[8] Akash, "ASL Alphabet: Image data set for alphabets in the American sign language", Version-1, 2018 [Online]. Available: https://www.kaggle.com/grassknoted/asl-alphabet[Accessed:11-Nov 2019].

[9] Starkey Foundation, "Types and causes of hearing loss", United States, 2007 [Online]. Available: https://www.starkey.com/hearing-loss/types and causes [Accessed: 25-Apr-2019].

[10] Boys Town National Research Hospital, "Auditory Neuropathy spectrum disorder", Nebraska, United States, 2020 [Online]. Available:https://www.babyhearing.org/auditory-neuropathy-spectrum disorder [Assessed: 25-Apr-2019].

[11] Activation Functions: Neural Networks Sigmoid, tanh, Softmax, ReLU, Leaky ReLU EXPLAINED !!! by SAGAR SHARMA [2017]- towards data science

[12] K. Hara, K. Nakayamma "Comparison of activation functions in multilayer neural network for pattern classification" IEEE International Conference on Neural Networks (ICNN'94), 1994

[13] Kashmera Khedkkar Safaya1, Prof. (DR.). J. W. Bakal "Real Time Based Bare Hand Gesture Recognition" IPASJ International Journal of Information Technology (IIJIT) Volume 1, Issue 2, July 2013

[14] F.S. Chen, C. Fu, & C. Huang, "Hand gesture recognition using a real time tracking method and hidden Markov models", ELSEVIER, Image and Vision Computing, Volume 21, Issue 8, pp. 745–758, 2003.

[15] K. Ferens, W. Lehn, and W. Kinsner, "Image Compression Using Learned Vector Quantization", in proceedings oh WESCANEX, 1993.

[16] Chin-Chen, H. Yu-Chen, "A Fast LBG Codebook Training Algorithm For Vector Quantization", IEEE transactions on consumer Electronics, vol 44, Issue 4, page 1201-1208, 1998