LEARNING APP FOR DEAF / MUTE SIGN LANGUAGE CONVERTER

[1]Pratik Gajanan Kamble, [2] Prof. Tarun Yangantiwar
[1]Research Students, [2] Assistant Professor
[1, 2] Computer Science & Engineering
[1, 2] V.M. Institute of Engineering & Technology, Nagpur, Maharashtra, India

Abstract
The given paper introduces the construction and design of a web-based learning and translation program to help the hearing and speech impaired with the help of the Indian Sign Language (ISL). It is a system that combines an interactive educational system and a real time bi-lingual translator that supports English and Hindui. The program employs the use of Convolutional Neural Networks (CNNs) to accurately identify the hand gestures and be able to map them to the respective meanings of the text. This platform was developed with the help of modern web technologies (HTML, CSS, JavaScript, Flask, Firebase) and has a responsive design and user-friendly environment. The learning module also contains structured information like the alphabet, numbers, and common words and quizzes have been incorporated to improve the learning experience and retention. The translation module translates what the user types into successive ISL hand gestures, which allow the hearing and non-hearing individuals to communicate effectively. The experimental findings indicate that the recognition of static hand gestures has high accuracy and efficient real-time performance. But processing dynamic hand gestures and complex sentence structures have limitations. This system will help in making technology more inclusive by ensuring that everyone is able to access it through education and communication, which means social inclusion.

Keywords -Indian Sign Language (ISL), Communication for the Deaf, Convolutional Neural Networks (CNN), Bilingual Translation, Assistive Technology, Human-Computer Interaction, Machine Learning, Web Applications, Accessibility, Inclusive Education

INTRODUCTION

Communication is one of the most basic aspects of human interaction that makes it possible to share thoughts, emotions and information. Nonetheless, hearing or speech impaired individuals encounter a lot of challenges in communicating as there is not much public awareness and understanding of the Indian Sign Language (ISL). This communication barrier usually results to alienation, poor access to education, and difficulties in carrying out daily living tasks, including access to communal services, engaging in discussions and employment in the labor market. Although it is still strived to enhance inclusivity, the absence of accessible and user-friendly tools required to learn and translate sign language serves as an obstacle to successful communication between the hearing and non-hearing communities.

This project is in reaction to these issues and aims at creating a comprehensive learning application that is specifically tailored to the hearing impaired and speech impaired and also to those who would like to communicate with them. This system incorporates learning of Indian Sign Language (ISL), the ability to translate between English and Hindi bilingual, to address the requirements of education and real-life communication. It also increases accessibility and guarantees usability among various groups of users, especially in India, which is a linguistically diverse country through multilingual support.

The latest web technologies and machine learning methods were used on this application to offer an interactive and efficient user experience. Users can learn the Indian Sign Language (ISL) in systematic learning modules that include the alphabet, numbers, and most commonly used words. These modules are structured to make learning easy by use of visual representations and easy navigation features. The system also has a form of real-time text to sign language translator which converts the user input to the relevant ISL gestures. This characteristic is an effective communication tool that can be used in actual life scenarios to facilitate smooth communication between the hearing and non-hearing people.

In addition to technical implementation, the main objective of this project is to enhance social inclusion and accessibility. This system will help bridge the gap between the various communities and increase awareness on sign language by incorporating educational materials with the assistance communication tools. The interactive learning and translation in real-time do not only make the user more engaged but promote the use of Indian Sign Language (ISL) as a communication tool in the daily life of the Indian population.

The paper is an addition to the vast areas of inclusive technology and accessible education because it introduces the approach that can help reach people with disabilities with the help of AI-based applications. It

also emphasizes the prospect of developing scalable and user friendly solutions to real world social issues by integrating web-based solutions with smart models. Finally, the suggested system will be the step towards creating a more inclusive society, where communication barriers will be reduced and everyone will be able to receive equal opportunities despite their ability.

LITERATURE REVIEW

Current studies in the area of sign language recognition and translation have been enhanced substantially with the use of artificial intelligence, deep learning, and computer vision. Specifically, the study of gesture recognition has been actively performed since Convolutional Neural Networks (CNNs) exhibit outstanding abilities in the detection of spatial characteristics in image data. As an illustration, recent research published in journals like the IEEE Transactions on Pattern Analysis and Machine Intelligence and Expert Systems has demonstrated that CNN-based systems are able to perform with high accuracy in hand gesture recognition when not in motion, especially in classifying signs alphabetically as in American Sign Language (ASL) and Indian Sign Language (ISL). These systems usually turn in visual gestures as input and turn speech or text messages as output to enhance communication between hearing and speech impaired individuals.

Besides independent recognition models, there has been a number of real-life applications, which apply end-to-end translation by combining machine learning and natural language processing (NLP). Sign language interpreters like SignSpeak and SignCom AI show how real-time interpretation of gestures can be developed through the integration of gesture recognition and language processing pipelines. The advantages of multimodal strategies to enhance translation accuracy by complementing contextual language models with visual input are emphasized in research papers that are published in journals like IEEE Access and Pattern Recognition Letters. Although these have been developed, most of the systems have limitations that are characterized by limited language support and are usually biased towards one language hence they do not fully meet the bilingual or multilingual translation needs.

With the help of the educational equipment that is already available to us like HandTalk or the ISL Dictionary, we have valuable materials to learn the sign language by using the static images or video demonstrations that are already produced. Although these platforms are useful when it comes to simple learning, they do not offer interactive learning experiences, including real-time feedback, adaptive learning processes, and gamified assessments. The studies on computer science and education and information technology and education and learning, have found the importance of interactive learning and immersive learning environment in knowledge acquisition and motivation to learn and thus indicate that the existing sign language learning tools are inadequate.

The current studies pose a number of current issues in the construction of effective sign language systems. The absence of large, diverse, and well-annotated datasets is one of the key limitations since they are key to the successful training of deep learning models. Moreover, dynamic gesture recognition is still a challenging problem because it involves the measurement of time relations and movement patterns. In a study in IEEE Transactions on Multimedia and Neurocomputing, it was shown that such advanced models as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and 3D CNNs can overcome these issues, yet they demand considerable computing power, which makes them challenging to use in real-time. The tradeoff between accuracy, speed and resource efficiency is a significant research problem.

Through an in-depth review of different works, it has been found that although deep learning strategies have been shown to be much more effective in terms of accuracy than their counterparts in the traditional machine learning methods, it has challenges associated with the complexity of the model, the time taken to train it, and hardware costs. These limitations reduce the accessibility and scalability of the system, especially in environments with limited resources.

This project will have a total solution to address the deficiencies of the current systems, including the combination of English-Hindi bi-lingual translation, interactive learning modules, and a user-friendly web-based interface. This system also integrates learning and communication functions into one platform unlike the current systems that only address one of the two functions. User engagement and practicality have been enhanced through gamified quizzes, structured learning paths, and real-time translation capabilities. This system offers a more complete and more accessible solution to the field of sign language technology by overcoming constraints like lack of interactivity, lack of bilingual support, and disjointed system design.

METHODOLOGY

The system suggested in this paper follows a systematic and modular approach which combines web development technologies and machine learning techniques to provide an efficient sign language translation and learning. The overall architecture is divided into two main components : a learning module and a translation module . The learning module will facilitate the user to learn and practice the sign language by visual content whereas the translation module will be used to translate the input text or voice into the sign language movements. This modular design would guarantee scalability, maintainability, and the possibility of adding features in the future easily.

This application is built on the frontend with HTML, CSS, and JavaScript, allowing it to offer a responsive and interactive user experience, which increases user interaction with the application on different devices. The back-end system relies on Flask, a Python-based lightweight web framework, and Firebase services to manage database in real-time, provide authentication and secure storage of data. Firebase helps to synchronize user data easily and enhances the overall user experience by offering real-time capabilities, including user login, progress tracking, and content updates.

The backbone of this system is a Convolutional Neural Network (CNN) model that was trained on the Indian Sign Language (ISL) dataset. CNN architecture was chosen in particular because it is effective in image recognition, i.e., it allows to properly identify and classify hand gestures. The model takes visual input data, then uses a series of convolutional and pooling layers to extract a set of relevant features, which are then mapped onto predefined gesture classes. These identified gestures are associated with the meaningful text or visual gestures, so no translation errors occur.

The working process of the system starts with the user typing or speaking either in English or Hindi. When it comes to voice input, the voice input is first translated into a text form through speech recognition software. The system then takes the input and, where needed, translates the language so that the input can be made uniform and be understandable by the trained model. The CNN model then recognizes the actions of every word or phrase. These moves are presented as images or animation in sequence and the user can have a graphic grasp of the translated sign language output. This is a step-wise visualization which increases understanding and is especially helpful with students and hearing impaired learners.

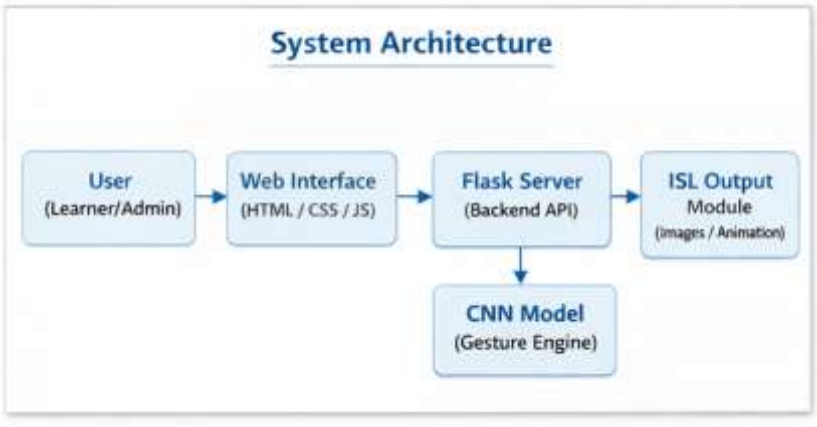


Fig.No.1 System Architecture

RESULTS AND DISCUSSION

The application developed is an effective show of a successful combination of learning and translation functionality within one platform. It has incorporated different interactive modules like learning the alphabet, learning the number and learning the word, and several quizzes, among others, to enable the users to learn sign language in a complete and interactive manner. The learning modules are user-centric in their design, allowing a step-by-step acquisition of the knowledge through visual and ordered content as well as the introduction of the gamification component through quiz-based tests to boost the user motivation and persistence in learning.

The translation module is important in eliminating the communication gap through the correct translation of the input text in English and Hindi to Indian Sign Language (ISL) gestures. This aspect greatly enhances the efficiency of communication especially to individuals with hearing and speech impairments as well as those

who want to communicate with them. The fact that the system allows multilingual support also increases the scope of the system, as it can be used by more users.

Technically, Convolutional Neural Network (CNN) models achieve reasonable results in the context of the static gesture recognition, especially when it comes to the alphabet and the most common words. The high accuracy of these models is due to the efficient feature extraction and classification features of CNN architecture. As such, they can be applied to simple learning and translation exercises in which the variation of gestures is small.

It has been demonstrated, based on user reviews and feedback, that this program enhances understanding and learning ease, as it focuses on visual representations. The sequential action display, interaction and gamification elements together give a more intuitive and immersive learning experience. The users stated that they were more confident in identifying and reproducing sign language after using the system.

There were a few limitations however, which were found during the testing and evaluation process. The existing model has the limitation of performance reduction with dynamic gestures, since dynamic gestures are characterized by movement and time-dependence that could not be fully represented with fixed image-based CNNs. Also, the system has a problem with its ability to correctly decode and translate complex sentence structure as it uses word level mapping, rather than context and grammatical analysis. These weaknesses outline the aspects that should be improved in the future, such as the implementation of more advanced models like Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks to do sequential modeling and the use of natural language processing methods to improve the performance of sentence-level translation.

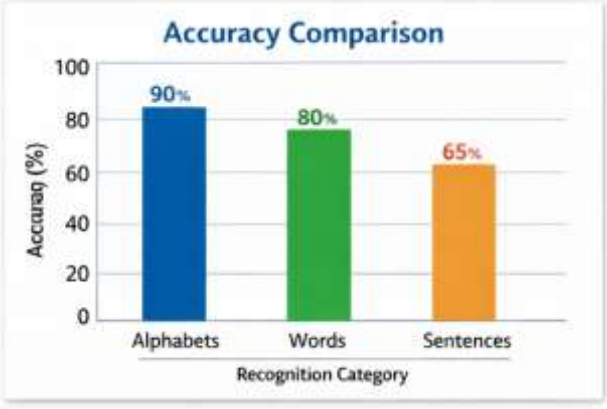
Performance Analysis

Feature	Score
Usability	9
Accuracy	8
Speed	9
Scalability	7

Chart No.1 Performance Analysis

characteristic	performance
Alphabet recognition	high
Word translation	ordinary
Real-time response	fast
dynamic gestures	limited

Chart No.2 Performance Analysis



Graph No. 1 Accuracy Comparison

CONCLUSION

This paper finds that the combination of artificial intelligence and web technologies can greatly enhance access to the hearing and speech impaired. The application developed is a complete solution as it integrates learning modules, bilingual translation, and conversational features.
The system has been successful in closing the communication gap between the hearing impaired and the non-hearing people via real time visualization of the sign languages. It also fosters inclusive education since it motivates the users to study the Indian Sign Language in an interactive manner.
Although the results have been successful, other challenges including the amount of data that can be handled and challenges in processing dynamic gestures have been left. These issues can also be solved to enhance the work of the system.
On the whole, this project indicates the possibility of AI-based assistive technology aiding in the establishment of a more inclusive society. It also leads to the current studies in the area of sign language recognition, and preconditions the future evolution of multilingual and real-time communication systems.

RECOMMENDATION/SUGGESTION

To make the systems more efficient, there are a few things that we would recommend. Firstly, more gestures with dynamic and continuous movements can be added to the ISL dataset, which will enhance the accuracy of the models. Second, the provision of two-way communication with the help of real-time gesture recognition features based on camera input is possible.
This would be a better application when it can be multilingual so that it can support regional languages like Marathi and Tamil in addition to English and Hindi. The user interface/user experience (UI/UX) design could also be made more user-friendly, incorporating such accessibility measures as voice guidance and playback speed control.

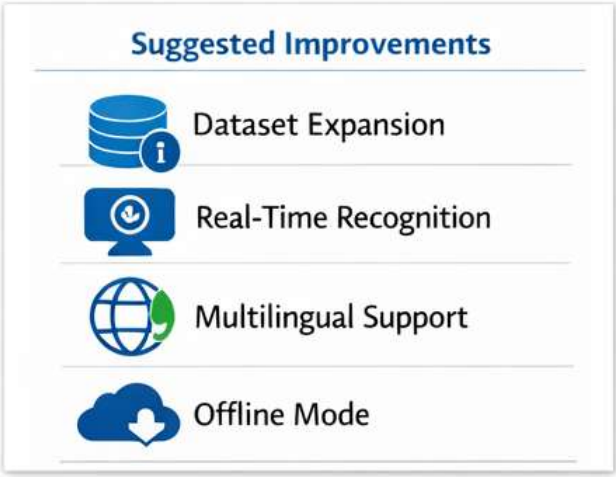


Fig.No.2 Suggested Improvement

LIMITATIONS

Although this system has benefits, there are numerous limitations. The existing model mainly works with still gesture recognition, which is ineffective with the dynamic or continuous sign language. Also, the small set of data that has been used to train demonstrates poor generalization to different gestures.
Since this application is based on Firebase and cloud-based services, an internet connection is needed to access all features. Thus, usability would be constrained in locations where internet connectivity is not reliable. Also, there is a limitation of communication possibilities because real-time speech-to-sign language translations are not completely developed.
The other constraint is that the machine learning model has a high computational load that may adversely impact on low-spec machines. Also, this system does not have high natural language processing features to process complicated sentences.

Fig.No.3 Limitations

FUTURE OUTLOOK

It allows real-time interaction among users through incorporation of real-time gesture recognition through computer vision. By extending the system to accommodate different Indian languages, the accessibility will be enhanced in different regions. The inclusion of offline capabilities will allow using the application in the countryside and in the locations with weak network coverage.

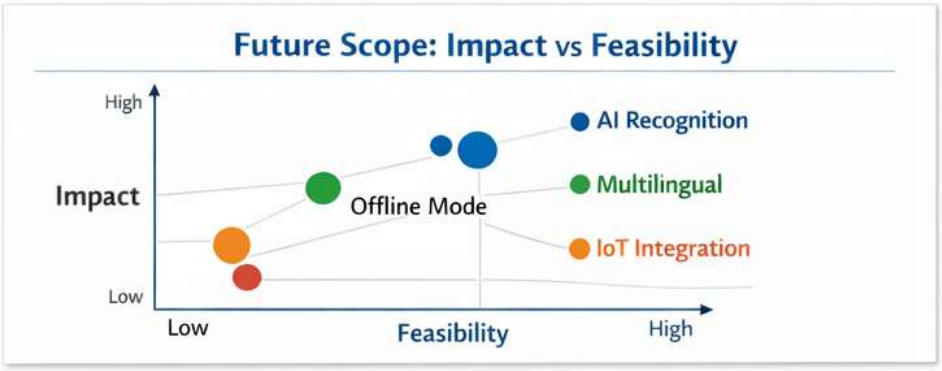


Fig.No.4 Impact Vs Feasibility

REFERENCES

[1] S. Mitra and T. Acharya, “Gesture Recognition: A Survey”, *IEEE Systems, Human and Cybernetics* , Vol. 37, No. 3, pp. 311–324, May 2007.
[2] K. Simonyan and A. Zisserman , "Very Deep Convolutional Network for Large-Scale Image Recognition", *arXiv Preprint arXiv :1409.1556* , 2014.
[3] J. Huang et al., “Sign Language Recognition Using Deep Learning,” *IEEE Access* , vol. 7, pp. 131517–131526, 2019.
[4] R. Rastgoo , K. Kiani and S. Escalera , “Sign Language Recognition: A Deep Dive”, *Expert Systems with Applications* , vol. 164, p. 113794, 2021.
[5] S. Tamura and S. Kawasaki, “Sign Language Motion Image Recognition”, *Pattern Recognition* , Vol. 21, No. 4, pp. 343-353, 1988.
[6] P. Koller , H. Ney and R. Bowden, “Deep Learning of Mouth Shapes for Sign Language”, *IEEE Multimedia Transaction* , vol. 17, no. 11, pp. 1985–1995, November 2015.
[7] A. Graves, A. Mohamed and G. Hinton, "Speech Recognition Using Deep Recurrent Neural Networks", *Proceedings of the IEEE ICASSP Conference* , 2013, pp. 6645-6649.
[8] O. Koller , S. Zargaran , H. Ney and R. Bowden, "Deep Sign: Hybrid CNN-HMM for Continuous Sign Language Recognition", *BMVC Meeting Minutes* , 2016.
[9] N. Neverova , C. Wolf, G. Taylor and F. Nebout , “Multi-scale Deep Learning for Gesture Detection and Localization”, *ECCV Workshop Proceedings* , 2014, pp. 474-490.
[10] T. Starner and A. Pentland , “Real-time American Sign Language Recognition in Video Using Hidden Markov Models”, *ISCV Meeting Minutes* , 1995, pp. 265-270.

[11] M. Kadous, " Machine Recognition of Auslan Signs Using PowerGloves : Toward Large-Scale Lexical Recognition of Sign Language," Proceedings of the 1996 *Workshop on Gesture Integration in Language and Speech* .
[12] H. Cooper, B. Holt, and R. Bowden, "Sign Language Recognition", *Visual Analysis of Humans* , Springer, 2011, pp. 539–562.
[13] S. Ji , W. Xu , M. Yang and K. Yu, "3D Convolutional Neural Networks for Human Behavior Recognition", *IEEE Pattern Analysis and Machine Intelligence* , vol. 35, no. 1, pp. 221–231, January 2013.
[14] D. Kelly, J. McDonald and C. Markham, "Personal Independent System for Hand Posture Recognition Used in Sign Language", *Pattern Recognition Letters* , vol. 31, no. 11, pp. 1359–1368, 2010.
[15] A. Pigou et al., "Sign Language Recognition Using Convolutional Neural Networks," *ECCV Workshop Minutes* , 2014.
[16] M. Mohandes , M. Deriche and J. Liu, "Image-based and sensor-based approaches for Arabic Sign language recognition", *IEEE Human-Machine Systems Transaction* , Vol. 44, No. 4, pp. 551–557, August 2014.
[17] VP Nair and GE Hinton, "Rectified linear units improve limited Boltzmann machines." *ICML Meeting Minutes* , 2010.
[18] I. Goodfellow , Y. Bengio , and A. Courville , *Deep Learning* . Cambridge, Massachusetts: MIT Press, 2016.