International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 82-88, October-2023

https://doi.org/10.47001/IRJIET/2023.710011

A Multimodal Approach for Real-Time Sinhala Sign Language Translation

[1]Seethoda Gamage, [2]Kavinda Lakshan, [3]Vinodya Wickramarathna, [4]Dr. Dasuni Nawinna, [5]Sathira Hettiarachchi

[1,2,3,4]Department of Computer Systems Engineering, Sri Lanka Institute of Information Technology, Sri Lanka

[5]Department of Information Technology, Sri Lanka Institute of Information Technology, Sri Lanka

Authors E-mail: [1]seethoda123@gmail.com, [2]kavindalakshan1007@gmail.com, [3]vinodyawickramarathna22@gmail.com, [4]dasuni.n@sliit.lk, [5]sathira.h@sliit.lk

***Abstract* - To facilitate communication between individuals with hearing impairments and those without, the use of language translation is necessary. This paper presents the development of a mobile-based solution to assist the communication between individuals having hearing impairments and those without. This study is based in Sri Lanka and the solution is designed to focus on the Sri Lankan Sign Language (SSL). There are three primary functionalities in this app: receiving input in English, converting it to a 3D avatar representing Sri Lankan Sign Language (SSL), and receiving input in Sinhala, converting it to an SSL picture, and receiving output in text form from Sri Lankan Sign Language. The Google Speech-to-Text API is used for translating English and Sinhala voices. The front end of the app was built in Unity and Blender for the Android platform, while the back end was built in Python using the Flask API, TensorFlow, and Kera. Artificial neural network (ANN) and Convolutional Neural Network (CNN) models tailored to the tasks of Natural Language Processing (NLP) and image processing are used during training.**

Keywords: Assistive Technology, Deaf, Hearing Impaired, Natural Language Processing (NLP), Sign Language, Sri Lanka.

I. INTRODUCTION

The deaf community encounters a range of communication obstacles due to their inability to perceive auditory language. Their primary mode of communication involves the utilization of sign language for interaction among themselves as well as with individuals who possess the ability to hear. Nevertheless, it is important to note that sign language is not universally standardized, and various regions have developed their distinct forms of sign language. Deaf individuals encounter notable obstacles when engaging in communication with individuals who possess hearing abilities, which encompasses struggles in comprehending written language. Nevertheless, sign language, comprising manual gestures and bodily movements, functions as the primary mode of communication for individuals in this population. Sign language is a visual-spatial language that employs a system of gestures to communicate words, while sentences are constructed through the amalgamation of multiple gestures. Throughout history, individuals who were deaf have faced significant discrimination and were commonly perceived as incapable of independent learning or cognitive functioning. Nevertheless, in the year 1500, endeavours to provide education for individuals with hearing impairments commenced, leading to the eventual recognition that the ability to hear is not a prerequisite for grasping complex ideas[1].

The prevailing mode of communication utilized by individuals who are deaf and mute in Sri Lanka is Sinhala Sign Language, as widely acknowledged. Sign language exhibits variations across different countries, and a significant number of individuals who are deaf utilize American Sign Language[2]. However, Sri Lanka Sign Language is widely recognized as the prevailing form of sign language in Sri Lanka. In the contemporary era, technology has reached a highly advanced state on a global scale. The endeavour to transform written text into sign language has been explored in various sign languages, yielding a substantial body of productive research. However, a definitive resolution for the conversion of sign language into text (or audio) has yet to be discovered, and a substantial quantity of data is required. The current state of Sinhala sign language utilized in Sri Lanka does not yet possess a comprehensive resolution. The objective of our study is to identify Sinhala sign language in real-time and transform it into a coherent sentence by employing a Time Sequence Neural Network, Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN)[3] machine learning models. To achieve this, we have utilized a limited dataset. Let us examine previous studies conducted to facilitate further investigation[4].

II. LITERATURE REVIEW

American Sign Language (ASL) is widely recognized as the third most prevalent and widely utilized language in the

ISSN (online): 2581-3048
Volume 7, Issue 10, pp 82-88, October-2023
https://doi.org/10.47001/IRJIET/2023.710011

United States. American Sign Language (ASL) encompasses a comprehensive lexicon consisting of more than 4,400 unique signs. According to the information presented in this chapter, Slovenian sign language encompasses approximately 4,000 commonly used signs. In the context of sign language, gestures can be executed using either bilateral or unilateral hand movements, while the interpretation of facial expressions can significantly influence their semantic connotations. Finger spelling is primarily employed for proper nouns and less frequently encountered vocabulary items. However, most sign language conversations rely on complete words, enabling communication to occur at a comparable pace to spoken language[5].

Sign language serves as the primary mode of communication for individuals who are deaf. Efforts aimed at the development of sign language generation systems have the potential to greatly enhance the quality of life for individuals who rely on sign language as their primary means of communication, facilitating a more seamless and efficient communication experience. This study represents a prominent and widely recognized academic literature review on the topic of sign language generation systems. The provided resource offers an academic repository of scholarly literature spanning from 1998 to 2020. Additionally, it proposes a set of classification criteria aimed at organizing and categorizing research studies systematically [6]. A total of 414 research studies were identified and assessed for their direct relevance to the development of sign language generation systems. A total of 162 research studies were subsequently selected, analyzed, and categorized.

The research examines the means of communication employed by individuals who are deaf in Arabic-speaking nations, wherein sign language serves as a gesture-oriented system. Arabic, as an official language, is recognized in 25 countries. However, it is important to note that each of these countries possesses its distinct spoken Arabic dialect. Nevertheless, it is worth noting that all Arabic-speaking nations employ a uniform alphabet, thereby implying that Middle Eastern countries universally adopt the Arabic Sign Language. The Arab deaf communities exhibit a significant degree of isolation from the hearing community, and Arabic Sign Language (ArSL)[7] employs a combination of hand shapes, placement, and movement to communicate meaning, alongside non-manual features including tongue movements and facial expressions. The paragraph underscores the significance of ArSL as a mode of communication for individuals who are deaf and residing in Arabic-speaking nations. The subsequent section outlines the design considerations for a theoretical system aimed at providing a virtual translation of Arabic Sign Language for individuals with hearing impairments. The system will consist of a database comprising recorded 3D gestures in Arabic sign language[8], which were captured using data gloves. Subsequently, the system will proceed to revive the digitized representation of sign language by employing conventional graphical translation techniques, to faithfully reproduce the original gesture with a high degree of precision. To accomplish this task, the Sphinx-4 Speech Recognition Engine will perform semantic translation of commonly used words without the intermediate step of converting them into textual form.

Sign language enables individuals with hearing impairments to engage in visual communication through the utilization of hand and finger gestures. Indian Sign Language (ISL) [9]serves as a highly effective and commonly employed means of communication for individuals who experience deafness or hearing impairment. However, individuals who do not possess knowledge of sign language encounter challenges when attempting to engage in communication with a person who has hearing impairments. The project involved the development of an instructional system for learning ISL sign language, utilizing 3D avatars to convert spoken or typed input into corresponding sign movements. The system is composed of three modules. The initial spoken input is transformed into a sentence in the English language. The Natural Language Processing (NLP)[10] technique is subsequently employed to convert the English sentence into its corresponding ISL statement. Finally, the ISL programming language is employed to precisely define the motions of the three-dimensional avatar. The score for the sign error rate (SER)[11] of the translation module is 10.50 [12].

The research paper discusses a system that has been developed to present Bangladesh Sign Language (BSL)[13] when provided with either a voice in the Bengali language or text originating from Bangladesh. In this study, a dataset comprising 14 Bangladeshi digits was utilized to train a model specifically designed for 2D animation. The system demonstrates an average accuracy of 96.03% for voice input and an average accuracy of 100% for Bangla text input in the form of Bangladeshi composite transliteration[14]. The system incorporates model speech extraction, a restricted sound library for voice recognition, and speaker-dependent training for achieving speaker-independent voice recognition. This research paper discusses a speech recognition system that is extensively employed in various applications. The term being referred to is Mel-Frequency Cepstral Coefficients (MFCCs)[15]. The operational procedure of this system commences with the initial stages of preprocessing and signal conditioning.

This scholarly article, published in the year 2021, discusses the development of an application aimed at

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 82-88, October-2023

https://doi.org/10.47001/IRJIET/2023.710011

mitigating the challenges encountered by individuals with hearing impairments in the country of India. The objective of this application is to transform spoken input into a visual representation in the form of Indian Sign Language. In this context, speech recognition technology is employed to transform auditory input into written text. Word segmentation and root word recognition are tasks that utilize algorithms in the field of natural language processing (NLP)[16]. We also present an analysis of the four main subtasks involved in sign language translation, namely sign2gloss2text, sign2text, sign2(gloss+text), and gloss2text. In this study, we elucidate the fundamental approach to sign language translation (SLT) [17] and present a transformer-based framework for SLT. In this study, we undertake an analysis of the primary obstacles encountered in the field of Speech and Language Technology (SLT) and put forth potential avenues for its advancement[18].

The system is described that can receive voice input, convert it into a textual form, and subsequently presenting corresponding Indian Sign Language images. The graphical user interface (GUI) [19] of this system is developed utilizing the Easy Gui library. The sound input is acquired through the utilization of the microphone, employing the Python pyAudio package. A dependency parser is a linguistic tool that can be employed to analyze the syntactic structure of a sentence and ascertain the dependencies or relationships between its constituent words. The system utilizes the Google Speech API to perform speech recognition. The text is subsequently subjected to preprocessing techniques utilizing Natural Language Processing (NLP).

The research proposes a technique for instantaneously converting 24 fixed Sri Lankan Sign Language alphabet words and numerals into English through the application of artificial neural networks (ANN) and support vector machines (SVM). The proposed methodology aims to address the disparity in communication between individuals proficient in sign language and those who are not. Sign language is a visual mode of communication employed by individuals who experience challenges with speech and hearing, facilitating effective interaction. The methodology involves performing pre-processing on input gestures that are accompanied by a signature, followed by the computation of various region attributes of the pre-processed gesture images. Finally, the signed gestures are transliterated into text and voice using the values obtained from the earlier stages. Python, along with popular libraries such as OpenCV, Keras, and Pickle, is employed in the development of the model.

To foster inclusivity and respect for individuals who are deaf and mute, it is imperative to establish an efficacious means of communication that bridges the gap between this population and the broader society[20]. The Convolutional Neural Network (CNN)[21] based machine learning technique is utilized to process, recognize, and classify these images. The proposed solution demonstrates the capability to accurately detect and classify a total of 26 hand gestures. This is achieved through the utilization of a Convolutional Neural Network (CNN) architecture, resulting in a validation accuracy of 91.23% and a training accuracy of 89.44%.

This article presents a study wherein a sign language interpreter employs the DataGloveTM technology to promptly recognize and interpret sign language gestures. The study focuses on the primary issue of identifying the termination points of gesture input streams. This is accomplished through the utilization of statistical analysis of the four attributes associated with each gesture, namely posture, location, orientation, and movement. The prototype system developed by researchers employed hidden Markov models (HMMs)[22] to facilitate the recognition of 250 phrases in Taiwanese Sign Language (TWL)[23]. These phrases were associated with 51 fundamental postures, six orientations, and eight movement variants. This implies that the system possesses the capability to identify the gesture phrase based on the provided words in real time, with an average recognition rate of 80.4%.

III. METHODOLOGY

Below figure 1 shows 3 main components of the application. Hearing and non-hearing end users. First, the program converts English text or voice commands into a 3D avatar in Sri Lankan sign language. Second, the technology accepts Sinhala speech and text and translates it to Sri Lankan sign language visuals. Sign detection and English word conversion is the final function. According to the user type (Deaf or Hearing user), the app sends API requests. Connecting to the endpoint will be the Python-based machine learning models running on Azure Web Services and the FLASK API.

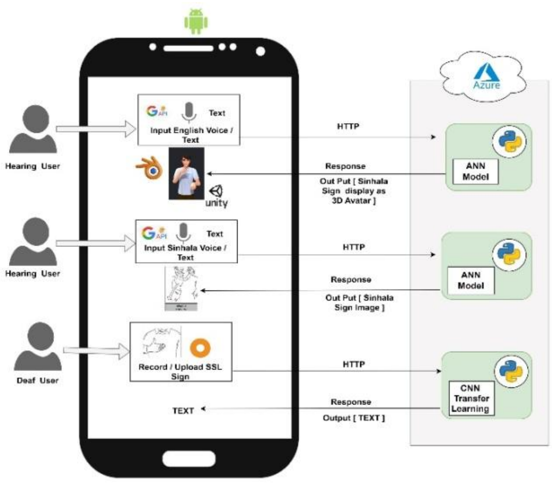

Figure 1: SSL Talk Overall Diagram

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 82-88, October-2023

https://doi.org/10.47001/IRJIET/2023.710011

Our Mobile Application has 3 main features, Such as

1) English Voice or text to 3D Avatar Sign Language Conversion using ANN

Manually collect the dataset, then make a 3D avatar using the blender tool, then program 200 sign movements showing common Sri Lankan sign language gestures. Employ the UNITY framework to Incorporate the GIFs' sign art into the 3D avatar, making sure that every gesture is represented accurately. Before utilizing, the English voice should be cleaned up and normalized. The Google API was used to transcribe the English-language recordings. Data derived from animated Gifs of Sri Lankan sign language will be matched with English text. Artificial neural network (ANN)-based models are your best bet. The learned ANN model should be included in the 3D avatar system that can interpret English speech. For real-time English speech-to-text conversion, you need Google's Speech-to-Text API. Apply the ANN model you just trained to predict the correct sign language gestures for the English text you converted.

2) Sinhala Voice or text to Sri Lankan Sign Language Image Conversion using ANN

Acquire a collection of 500 photos depicting various signals and motions in Sri Lankan sign language. Resized all the images to the same height and width. Clean up and normalize the Sinhala voice before using it. The Sinhala audio files are translated into text using the Google API. Also, text preprocessing by removing unwanted special characters.

Gather the Sinhala text and sign language visuals that will be used to train the ANN model. The system able to process Sinhala voice inputs should be updated to include the trained ANN model. Use the Google Speech-to-Text API to instantly transcribe spoken Sinhala into text. The transformed Sinhala text may then be used to forecast the appropriate sign language visuals using the ANN model.

3) Real-time Sri Lankan Sign Language to Text Conversion using VGG19 and Transfer Learning

100-Sign Limited SSL Video Capacity Record between a thousand and a hundred Sri Lankan sign language gestures on camera. All videos are acquired by hand, and it is suggested that 10 films per sign be used as training data. Getting rid of background noise might improve the quality of recorded movies. During training, you should reduce the video's width and height to ease the strain on the computer. Divide pixel values by 255 to get them into the range [0, 1] for faster convergence during model training. Images were taken from the movie using an image processing technique, and then CNN was used to analyse the data. Transfer learning for image categorization using VGG19 applied to sign language recognition. The CNN model VGG19 might benefit from having its training video segmented into several pictures. Make use of the pre-trained VGG19 models' feature extraction skills with the help of transfer learning. Overfitting may be prevented by freezing certain layers and optimizing others for the sign language recognition task at hand.

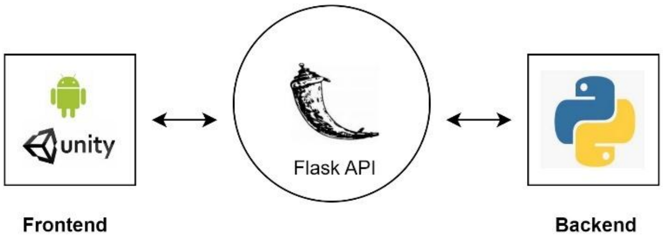

Figure 2: Technology used for the development.

Table 1: Component and Model used for SSL talk

Component	Model
English to 3D Avatar	ANN
Sinhala to SSL Image	
SSL to Text	CNN

IV. RESULTS AND DISCUSSION

The RNN model was used first. However, the RNN did not support video data sets. Secondly, the Long-term Recurrent Convolutional Network (LRCN) and ConvLRCN, were used. The ConvLRCN was able to achieve better accuracy than LRCN to get 53% accuracy, finally trying CNN + VGG19 [22]. Then we achieved 93% accuracy it represented Figures 6 and 7 (component 3 (a),(b)) represent the loss and accuracy functions of Sri Lanka sign language converted to the text which displays 93% accuracy after 10 epochs and using Softmax layers activation function for the multiclass.

The following diagrams show how precise and efficient certain parts are. Figure 4 represents the accuracy and loss functions of English voice or text converted to the 3d avatar which displays 99% accuracy after 600 epochs, using Selu, Relu, and Softmax layers for activation function, SGD optimizer. Figure 5 represents the accuracy and loss functions of Sinhala voice or text converted to the SSL image which displays 99% accuracy after 600 epochs using Selu, Relu, and Softmax layers for activation function, SGD optimizer.

ISSN (online): 2581-3048
Volume 7, Issue 10, pp 82-88, October-2023
https://doi.org/10.47001/IRJIET/2023.710011

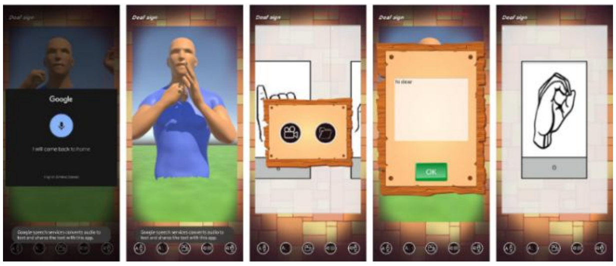

Figure 3: User interfaces

The following diagrams show how precise and efficient certain parts are.

Component 1

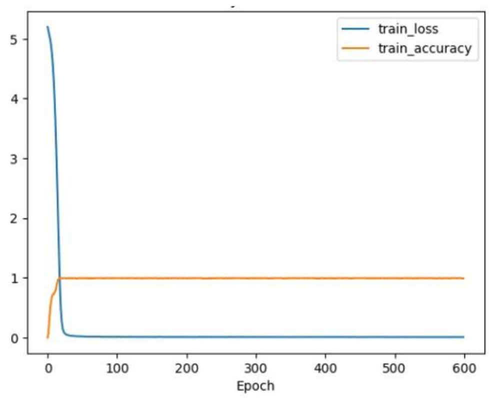

Figure 4: Accuracy for 3D avatar

Table 2: Accuracy and loss for 3D avatar

Epoch	Loss	Accuracy
595/600	0.0082	0.9892
596/600	0.0077	0.9946
597/600	0.0082	0.9892
598/600	0.0083	0.9946
599/600	0.004	0.9946
600/600	0.0079	0.9946

Component 2

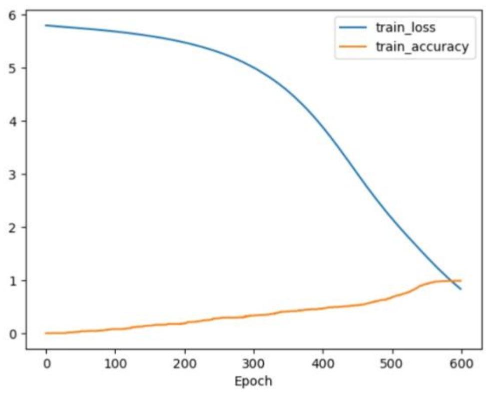

Figure 5: Sinhala voice or Sinhala text to Sri Lankan sign language images using ANN

Table 3: Accuracy and Loss of Sinhala voice or Sinhala text to Sri Lankan sign language images using ANN

Epoch	Loss	Accuracy
595/600	0.8945	0.9931
596/600	0.8837	0.9931
597/600	0.8728	0.9931
598/600	0.8622	0.9931
599/600	0.8517	0.9931
600/600	0.8411	0.9931

Component 3 (a)

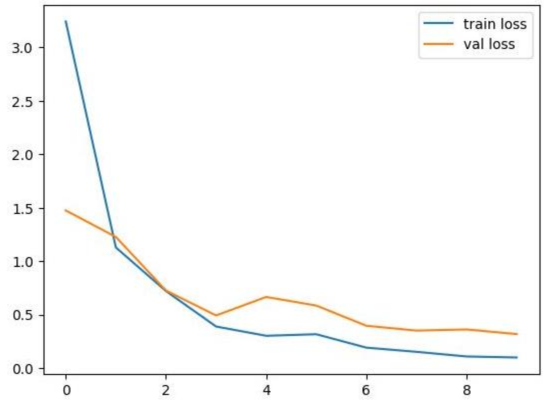

Figure 6: SSL to Text Loss

Component 3 (b)

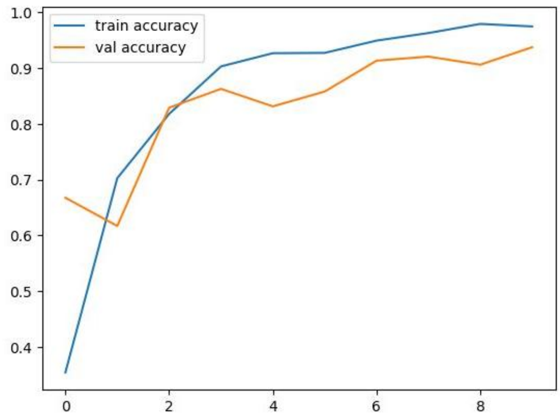

Figure 7: SSL to Text Accuracy

Table 4: SSL to Text Accuracy and Loss

Epoch	Loss	Accuracy	Val_Loss	Val_Accuracy
5/10	0.3014	0.9266	0.6647	0.8313
6/10	0.3160	0.9272	0.5842	0.8578
7/10	0.1904	0.9492	0.3951	0.9133
8/10	0.1513	0.9628	0.3498	0.9205
9/10	0.1079	0.9791	0.3606	0.9060
10/10	0.0985	0.9746	0.3174	0.9373

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 82-88, October-2023

https://doi.org/10.47001/IRJIET/2023.710011

The precision of the 3D avatar is seen in Figure 4. In Figure 5, we see how ANN is used to translate from Sinhala speech or text to Sri Lankan sign language graphics. Component 3 was mentioned in the previous figure 6 and 7. The sign has been texted in Sri Lankan. The precision and efficiency losses of each component were summarized in Figure 8.

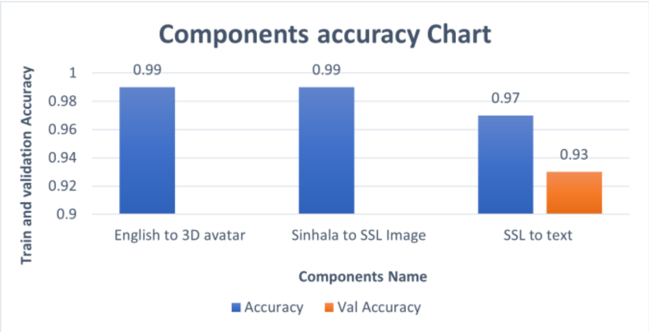

Figure 8: Comparison of accuracy

V. CONCLUSION

The process of communication between individuals who can hear and those who are deaf can present difficulties, particularly in situations where there is a scarcity of proficient speakers in specific languages. An illustrative instance pertains to the employment of sign language within the deaf community in Sri Lanka, which diverges from the sign language commonly employed by the broader populace.

The main objective of this primary research is to develop an effective translation service for the deaf community. There are main 3 features in the mobile application such as English Voice or text to 3D Avatar, Sinhala Voice or text to Sri Lankan Sign Language Image Conversion using ANN, and Sri Lankan Sign Language to Text Conversion using CNN.

The proposed methodology entails the training of neural network models to accurately identify and interpret signs by utilizing a real-time database of SSL sequences. The optimization techniques such as NLP and Image processing will be employed to further enhance the performance of these models. The system will be assessed in comparison to that of an interpreter.

The limited exchange of information between individuals who have hearing impairments and those who possess normal hearing abilities poses challenges in terms of accessing diverse services and information. By expanding the data collection, it is possible to enhance the accuracy of the results. The data collection should be modified to accommodate multiple sign transactions since the program provides support for such transactions.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude, appreciation, and acknowledgment to Mr Kasun Chandana, president of the Ruhuna Sumaga Deaf Organization, and Miss Jayamini, a staff member at Rohana Special School Matara.

REFERENCES

[1] M. Rahaman, Md. P. Hossain, M. M. Rana, M. Rahman, and T. Akter, *A Rule-Based System for Bangla Voice and Text to Bangla Sign Language Interpretation*. 2020.

[2] V. Mueller, "American Sign Language (ASL)," 2013, pp. 133–136. doi 10.1007/978-1-4419-1698-3_1654.

[3] D. Patel and S. Naskar, "Image Classification Using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN): A Review," 2020, pp. 367–381. doi: 10.1007/978-981-15-1884-3_34.

[4] K. Saija, S. Sangeetha, and V. Shah, "WordNet Based Sign Language Machine Translation: from English Voice to ISL Gloss," *2019 IEEE 16th India Council International Conference (INDICON)*, pp. 1–4, Dec. 2019, doi: 10.1109/INDICON47234.2019.9029074.

[5] A.Mehta, K. Solanki, and T. Rathod, "Development of an Automatic Translate Real-Time Voice to Sign Language Conversion for Deaf and Dumb People," in *1st International Conference on Multidisciplinary Industry and Academic Research 2020*, Institute of Industry and Academic Research Incorporated, Dec. 2020. Accessed: Jul. 05, 2023. [Online]. Available: https://www.neliti.com/publications/358242/

[6] R. Kumar Attar, V. Goyal, and L. Goyal, "State of the Art of Automation in Sign Language: A Systematic Review," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, vol. 22, no. 4, p. 94:1-94:80, Apr. 2023, doi: 10.1145/3564769.

[7] Z. Alsaadi, E. Alshamani, M. Alrehaili, A. A. D. Alrashdi, S. Albelwi, and A. O. Elfaki, "A Real Time Arabic Sign Language Alphabets (ArSLA) Recognition Model Using Deep Learning Architecture," *Computers*, vol. 11, no. 5, Art. no. 5, May 2022, doi: 10.3390/computers11050078.

[8] A.Boukdir, M. Benaddy, A. Ellahyani, O. E. Meslouhi, and M. Kardouchi, "3D gesture segmentation for word-level Arabic sign language using large-scale RGB video sequences and autoencoder convolutional networks," *Signal, Image and Video Processing*, vol. 16, pp. 1–8, Feb. 2022, doi: 10.1007/s11760-022-02167-6.

[9] K. Shenoy, T. Dastane, V. Rao, and D. Vyavaharkar, "Real-time Indian Sign Language (ISL) Recognition," in *2018 9th International Conference on Computing,*

ISSN (online): 2581-3048
Volume 7, Issue 10, pp 82-88, October-2023
https://doi.org/10.47001/IRJIET/2023.710011

Communication and Networking Technologies (ICCCNT), Jul. 2018, pp. 1–9. doi: 10.1109/ICCCNT.2018.8493808.

[10] S. Katoch, V. Singh, and U. S. Tiwary, "Indian Sign Language recognition system using SURF with SVM and CNN," *Array*, vol. 14, p. 100141, Jul. 2022, doi: 10.1016/j.array.2022.100141.

[11] R. Rajashekar, K. V. S. Hari, and L. Hanzo, "Antenna Selection in Spatial Modulation Systems," *Communications Letters, IEEE*, vol. 17, pp. 521–524, Mar. 2013, doi: 10.1109/LCOMM.2013.012213.122650.

[12] R. M. Rishan, S. Jayalal, and T. Wijayasiriwardhane, *Translation of Sri Lankan Sign Language to Sinhala Text: A Leap Motion Technology-based Approach*. 2022, p. 223. doi: 10.1109/ICARC54489.2022.9754050.

[13] S. Das, Md. S. Imtiaz, N. H. Neom, N. Siddique, and H. Wang, "A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier," *Expert Systems with Applications*, vol. 213, p. 118914, Mar. 2023, doi: 10.1016/j.eswa.2022.118914.

[14] S. K. Youme, T. A. Chowdhury, H. Ahamed, Md. S. Abid, L. Chowdhury, and N. Mohammed, "Generalization of Bangla Sign Language Recognition Using Angular Loss Functions," *IEEE Access*, vol. 9, pp. 165351–165365, 2021, doi: 10.1109/ACCESS.2021.3134903.

[15] S. Majeed, H. HUSAIN, S. Samad, and T. Idbeaa, "Mel frequency cepstral coefficients (Mfcc) feature extraction enhancement in the application of speech recognition: A comparison study," *Journal of Theoretical and Applied Information Technology*, vol. 79, pp. 38–56, Sep. 2015.

[16] K. Jain and V. Prajapati, "NLP/Deep Learning Techniques in Healthcare for Decision Making," *Primary Health Care*, no. 3, p. 4, 2021.

[17] A.Núñez-Marcos, O. Perez-de-Viñaspre, and G. Labaka, "A survey on Sign Language machine translation," *Expert Systems with Applications*, vol. 213, p. 118993, Mar. 2023, doi: 10.1016/j.eswa.2022.118993.

[18] Z. Liang, H. Li, and J. Chai, "Sign Language Translation: A Survey of Approaches and Techniques," *Electronics*, vol. 12, no. 12, Art. no. 12, Jan. 2023, doi: 10.3390/electronics12122678.

[19] D. V. G and K. Goyal, "Indian Sign Language Recognition Using Mediapipe Holistic." arXiv, Apr. 20, 2023. doi: 10.48550/arXiv.2304.10256.

[20] S. Hettiarachchi, "Machine Learning Approach for Real Time Translation of Sinhala Sign Language into Text," 2020. Accessed: Aug. 03, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Machine-Learning-Approach-for-Real-Time-Translation-Hettiarachchi/63e58eb09460d5028ebc42ca3e976a427cf13241

[21] F. M. J. M. Shamrat and S. Chakraborty, "Face Mask Detection using Convolutional Neural Network (CNN) to reduce the spread of Covid-19," *IEEE*, Jan. 2021, Accessed: May 01, 2022. [Online]. Available: https://www.academia.edu/49065936/Face_Mask_Detection_using_Convolutional_Neural_Network_CNN_to_reduce_the_spread_of_Covid_19

[22] B.-J. Yoon, "Hidden Markov Models and their Applications in Biological Sequence Analysis," *Curr Genomics*, vol. 10, no. 6, pp. 402–415, Sep. 2009, doi: 10.2174/138920209789177575.

[23] S. Duncan, "Gesture in Signing: A Case Study from Taiwan Sign Language," *Language and Linguistics*, vol. 6, Jan. 2005.

[24] V. Sudha and D. Ganeshbabu, "A Convolutional Neural Network Classifier VGG-19 Architecture for Lesion Detection and Grading in Diabetic Retinopathy Based on Deep Learning," *Computers, Materials & Continua*, vol. 66, pp. 827–842, Jan. 2020, doi: 10.32604/cmc.2020.012008.

Citation of this Article:

Seethoda Gamage, Kavinda Lakshan, Vinodya Wickramarathna, Dr. Dasuni Nawinna, Sathira Hettiarachchi, "A Multimodal Approach for Real-Time Sinhala Sign Language Translation" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 10, pp 82-88, October 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.710011

****