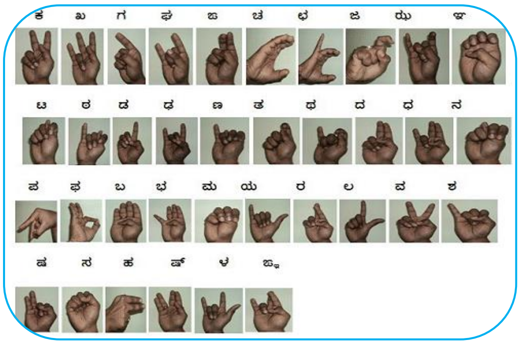
REVIEW OF RESEARCH
ISSN: 2249-894X
IMPACT FACTOR : 5.7631(UIF)
VOLUME - 14 | ISSUE - 8 | MAY - 2025

SIGN LANGUAGE RECOGNITION AND TRANSLATION FOR REGIONAL LANGUAGES IN DAKSHINA KANNADA

Chaitra[1] , Bhuvan Shettigar[2] , Durgaprasad[3] and Gautham Shetty[4]
Students, Department of ISE, AJIET, Masngalore, India.

ABSTRACT :



Machine Learning (ML) and Artificial Intelligence (AI) have made a huge leap in the effectiveness and efficiency of digital recognition systems. While sign language identification has been widely researched academically, real-time systems are still to be made in the context of specialised languages like Kannada. Therefore, through sign language recognition, this study intends to fill the communication gap. The project focuses on building a unified system targeted towards accurately translating sign language gestures into text using state-of-the-art ML techniques. This approach aims at providing assistance to kids with hearing disabilities by enhancing communication, establishing more inclusive educational environments, and facilitating the process of interacting with friends and educators.

KEYWORDS : Machine Learning (ML) and Artificial Intelligence (AI) , educational environments.

INTRODUCTION:

Communication is a fundamental human capacity, but millions of people who struggle with spoken language, especially in social situations, suffer profoundly. For hearing-impaired children in Karnataka, the two main methods of communication are through Indian Sign Language (ISL) and Kannada Sign Language (KSL), which serve as a visual-gestural alternative to oral communication.

To fill this embarrassing gap, Sign Language Recognition (SLR) technologies are well-performing technologies that are able to translate gestures automatically into texts. The emergence of Artificial Intelligence (AI), Machine Learning (ML), and Computer Vision has been a game changer in the precision and efficiency of these systems. Among others, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) networks, have enhanced gesture recognition, improving the accuracy of real-time translation.

Despite these advances, other issues still persist, like the variability in hand types, lighting conditions, and signer-specific factors. As such, this project aims to develop a real-time sign language recognising system that effectively translates 0 to 9 hand gestures in Kannada and Tulu text. which solves the communication issues of hearing-impaired people in the state of Karnataka by making it easier for them to communicate independently with others. By analysing 3D markers in an IMU sensor and enabling social interaction through deep learning-based real-time computer vision, it will be possible to generate more immersive learning environments, promote social inclusion, and increase autonomy.

LITERATURE REVIEW

Recently, Sign Language Recognition (SLR) has attracted a lot of attention and interest because of its potential to close communication with the hearing-impaired community. Numerous studies have examined different approaches that are intended to improve the effectiveness and efficiency of SLR systems. Many researchers have approached this problem from different angles by using various techniques and technologies. The related works that have been done by other researchers are discussed in this section.

A real-time gesture-based sign language detection system is introduced in paper [1], which applies computer vision and deep learning techniques to identify ASL gestures. "For hand and body pose estimation, the system uses Python, OpenCV, and MediaPipe Holistic with LSTM (Long Short-Term Memory) neural network for sequential gesture modelling. With transfer learning and a high-level dataset, it has low latency and high accuracy, allowing smooth communication between signer and non-signer. The system is examined under different lighting conditions, showcasing great promise for real-world applications, with future work being done on the integration of NLP for two-way conversations.

They report 99.72% accuracy on detecting ISL and report 99.90% accuracy on ISL images using residual CNNs for detecting ASL with a 9% increment in ASL accuracy compared to previous models [2]. The training process consists of webcam-captured images that undergo preprocessing (resizing, greyscale conversion, normalisation) and data augmentation (rotation, scaling, flipping) to increase the diversity of the dataset. The CNN is composed of convolutional, pooling, and fully connected layers trained with TensorFlow with their training-validation split to reduce overfitting. This system also provides real-time recognition, text- to-speech, and multilingual conversion, making it easier for the hearing-impaired to access information. Future work could broaden coverage to other sign languages, incorporate AR capabilities, and enable more user interactivity.

Paper [3] implements a real-time vision-based ISL recognition system on YOLOv4, which translates Indian Sign Language (ISL) to text. Our proposed model is pre-trained on an extended ISL-CSLTR dataset consisting of 24 signs produced by nine signers along with data augmentations (rotation, zoom, flipping) and manual annotations. Trained on a pre-trained MS COCO network with a batch size of 64, a learning rate of 0.001, and for 11 epochs on a Tesla T4 GPU. With 98.4% mAP accuracy and class precisions greater than 90% for all classes, the system enables real-time efficient word-level recognition. Further enhancement of YOLOv4 is the fast detection capability, which makes it suitable for sign language translation applications.

In paper [4], a survey on image-based ArSLR, including methods for alphabet, isolated word, and continuous recognition techniques, is reviewed. SVM, neuro-fuzzy systems, HMM, and KNN have reached accuracies in the range of 82.2%–98%, while continuous recognition is based on PCNN and graph matching with 70%– 94% accuracy. Video feature extraction methods such as zonal codes, DTC coefficients, and motion chain codes are discussed. The aforementioned reasons highlight the necessity for all improved error correction mechanisms, fusion methods, expanded vocabulary, and fast processing for real-time applications. In future works, exploring bi-directional translation between speech/text and sign language.

Paper [5] proposed a convolutional neural network (CNN)-based system that aims to help disabled individuals by recognising hand gestures and uses a dataset of 6182 images with 9 classes for sign languages. This model accomplished 91.67% accuracy with real-time detection of diseases and was able to outperform existing systems. To boost performance, data augmentation and normalisation were employed. Training was performed with the Adam optimiser on loss to achieve high-priority detection of signs and was tested on SRED in low-ground sign density. The method implemented increases accessibility for disabled people, and future work will involve converting voice to text and solving the issue of speech recognition of different communication styles.

Paper[6] Challenges in Sign Language Recognition (SLR) and Translation (SLT) 312. 314 and a Reversible CNN model for Bi-directional translation from sign language to text/speech and vice versa. The model designed uses CNN as feature extraction to recognise the gestures made by the ISL (Indian Sign Language) and the English alphabets used. The preprocessing step includes background removal, normalisation, and data augmentation. Train this network with the following parameters, which give 70 correct in 100 as well as 46.64%, 75.53%, and 49.42% precision, as well as recall in order to sign to text conversion in order to perform this and F1 score; The paper discusses challenges such as biases in datasets or the limitations of RNN and HMM models. In the future, we plan to extend datasets and refine the translation of expressions at the sentence level.

OBJECTIVE

The goal of a sign language recognition system is to translate sign gestures into text, enabling better communication. It uses computer vision and AI to recognise and interpret hand movements. This technology bridges the gap between sign language users and non-signers.

Purpose

The purpose of a sign language recognition project is to facilitate seamless communication between the deaf and hearing communities by converting sign language gestures into text or speech. It aims to enhance accessibility, inclusivity, and independence for individuals with hearing or speech impairments. By leveraging computer vision and artificial intelligence, the project enables real-time gesture recognition for effective human-computer interaction.

METHODOLOGY

In this part is where we presented our methodological procedures, which we propose. Here in this research work, the primary intention is to identify different sign actions. We will build a CNN model that accurately identifies sign language gestures with an optimal balance of accuracy and low parameter count. The final output of the system is providing a system for recognising and translating the hand movements into sign language. The following subsections provide information about data analysis, data preprocessing, data modelling, etc.

1. Dataset Preparation

The first research phase was to create an extensive dataset of hand gesture images depicting numbers 0-9. We took RGB images in a controlled environment with a mobile phone camera. We recorded each numerical gesture several times to obtain a wide variety of hand presentations. Differing lighting conditions were also used to ensure the dataset was more indicative of real-world conditions.

2. Dataset Description

To begin with, we needed to learn what kind of sign languages were used for communicating with disabled people. We have stored images after understanding. We combined images from various individuals on this page. We have 10 classes with a total of 1199 images for each class.

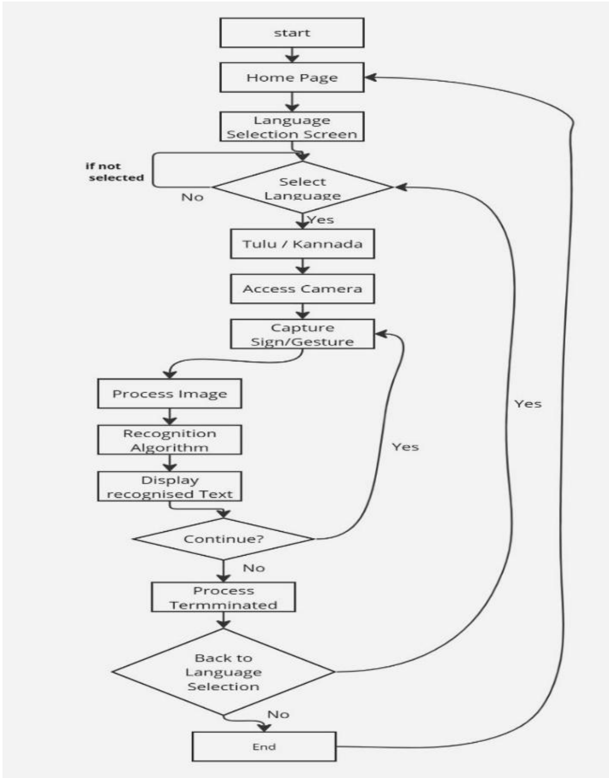


Figure 1: Overall system design

3. Data preprocessing

Data preparation involves capturing webcam frames using OpenCV, flipping them to create a mirror effect, and converting them from BGR to RGB using MediaPipe. MediaPipe detects hand landmarks, which may be accessed using calc_landmark_list(). The pre_process_landmark() method normalises the calculated landmarks by translating their coordinates relative to a base point and scaling them to a defined range.

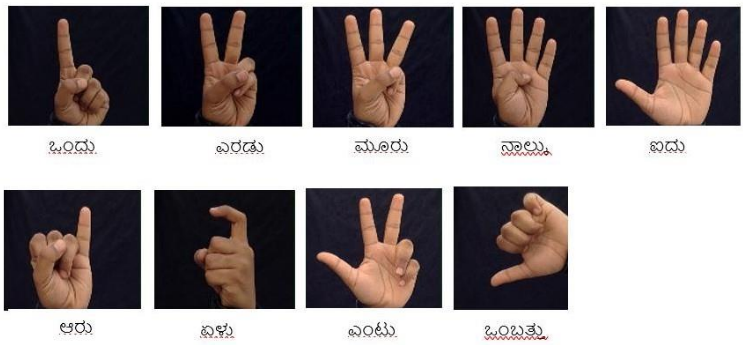


Figure 2: Different Classes of Images from the Dataset

This ensures uniformity independent of location or hand size. After processing, the landmarks are arranged as a DataFrame and sent into the deep learning model for categorisation.

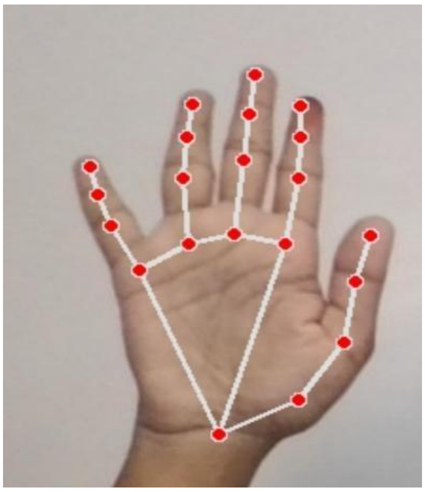


Figure 3: Hand Pose Estimation Using MediaPipe

Model Architecture and Training

The proposed model The architecture of the Convolutional Neural Network (CNN) used for classifying Indian Sign Language (ISL) gestures is mentioned below. The model consists of three convolutional layers with 32 filters, 64 filters, and 64 filters. Next, we add a max pooling layer to the model which reduces the spatial dimension and helps preserve relevant information. After every convolutional layer, a ReLU (Rectified Linear Unit) activation is applied to introduce non-linearity in the model and help it learn complex patterns. After the convolutional layers, a flattening layer converts the 2D feature maps into a 1D vector, which, in turn, is passed to a dense fully-connected layer with 128 neurones that captures complex feature associations. Non-linearity is introduced via ReLU activation.

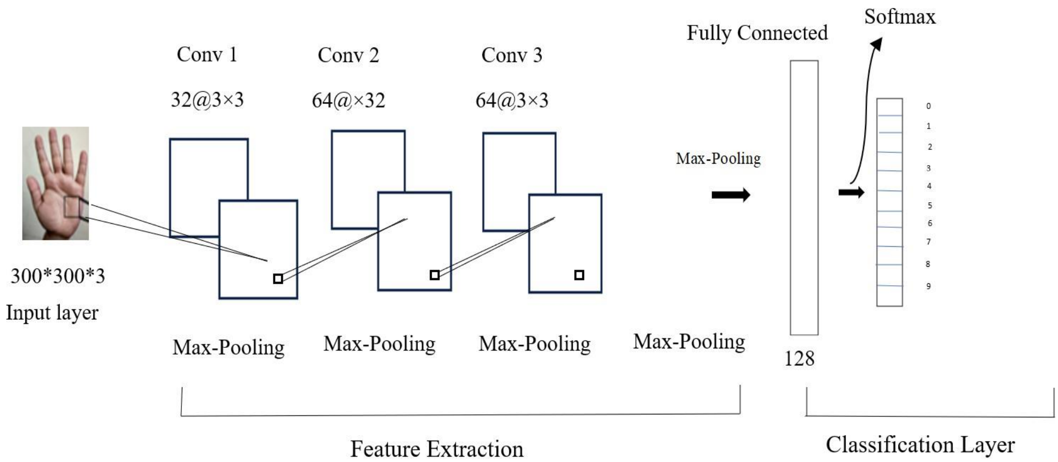


Figure 4: Structure of Convolutional Neural Networks (CNNS's)

Keras and the teaching procedure, this code aims CNN, to prepare the dataset, and then trains against the model. ImageDataGenerator is employed for scaling the pixel values, and augmenting the dataset which is now split into training set, validation set and the test set. The very first steps of a CNN architecture are three convolutional layers, max pooling (the biggest number of the values in the corresponding patch is retained), a flattening layer, and multiple dense layers. Finally, we add a softmax output layer to the model in order to perform multi-class classification. We build the model using the Adam optimiser with categorical cross- entropy loss and accuracy as the metric. The training data undergoes ten epochs of training, with performance validation being done on a different validation dataset. We can visualize the performance of the model throughout the epochs by plotting the accuracy and loss for both training and validation.

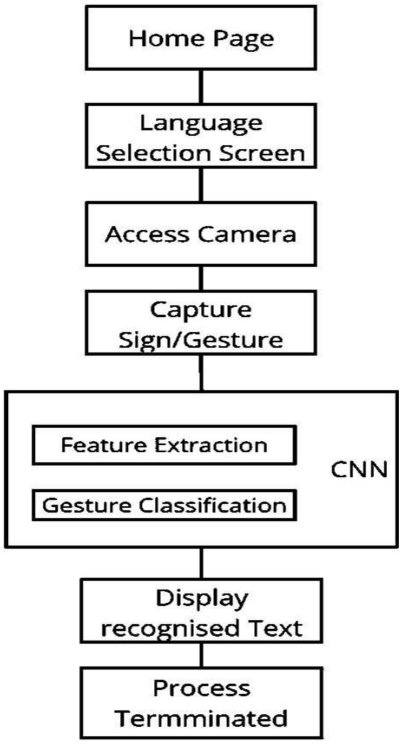


Figure 5: Flow diagram of our Sign Language Recognition System

RESULT DISCUSSION

Result is the outcome of every project. In this section we discuss about the results /output we have obtained after classification of hand gestures. In Figure 6 we can see the hand gestures which are classified into Kannada. Similarly, we have obtained the results for Tulu language also.

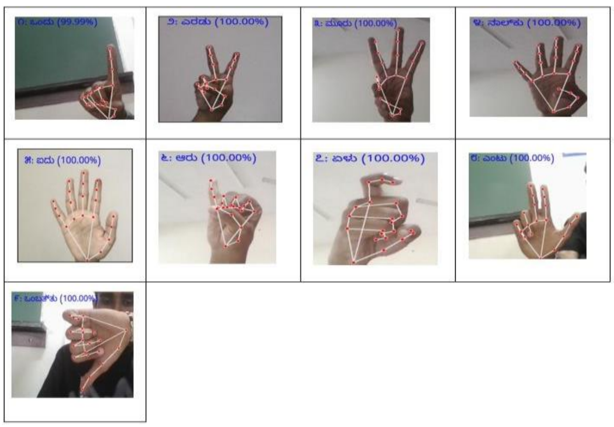


Figure 6: Result snapshots

This figure shows the training and validation graphs offer a comprehensive view of the model's learning process and performance. The training and validation accuracy charts that follow are very high, showing that the model has learnt very well and generalises effectively.

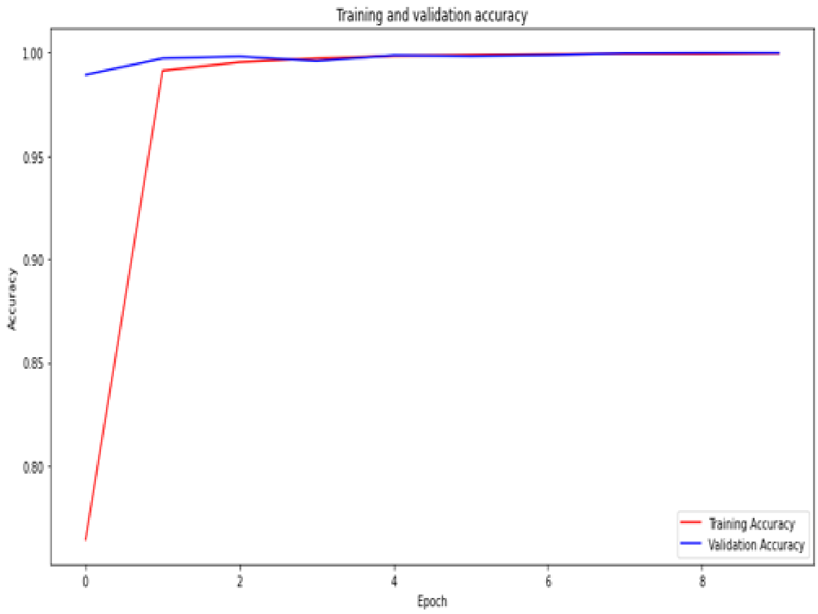


Figure 7. Accuracy Graph

They converge quickly, with training accuracy 77% on initialisation and 99% at the end of the epoch. The validation accuracy starts at 98% and increases slowly to 100%. After Epoch 2, the two measures of the model are stable, suggesting effective convergence. The training and validation curves are very close to each other, which usually is indicative of a strong fit of the model with no overfitting detected. The model learns quickly and reaches accuracy on the training set that is both high and consistent.

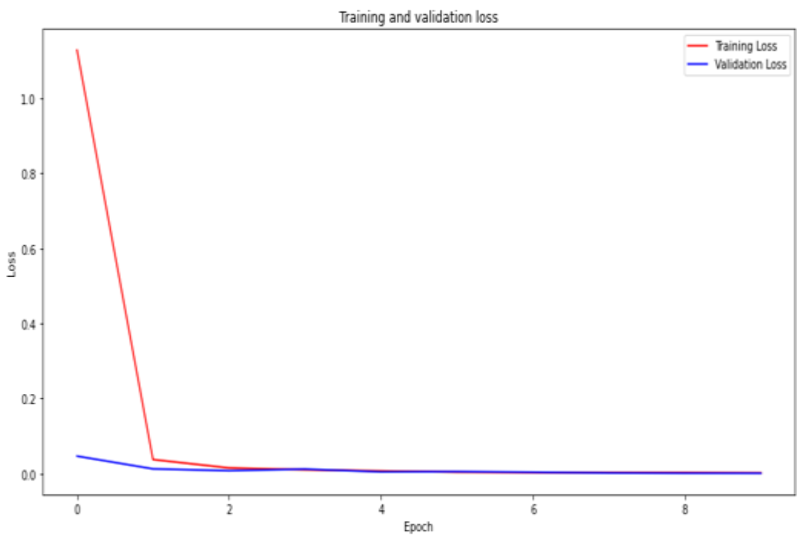


Fig 8: Loss graph

The loss graph shown above here shows the training and validation loss for 8 epochs to get a better picture of how the model learnt. Training loss (red) starts out high at around 1.1 and falls very sharply during the first epoch, indicating fast learning. By contrast, the validation loss (blue) starts low, roughly at 0.05, and stays relatively unchanged throughout the training process. By the second epoch, the two trajectories had converged and stabilised at zero, denoting effective optimisation. The low and steady loss values are in line with the great accuracy of the previous graph, showing that the model has been trained well and has good generalisation capability without any overfitting.

Advantages of the Proposed Model

The proposed model offers several unique features, making it useful for sign language recognition and translation. It can detect hand motions in real time and with high accuracy, enabling users to communicate seamlessly. The approach underscores the spirit of inclusion and removes barriers for native speakers by integrating regional languages—Kannada and Tulu. Its high accuracy and dependable generalisation make the system perform excellently on both known and unknown data. Its easy-to-use interface provides guidance through the steps in the experiment, from selecting a language to recognising gestures in real-time, and targets users of various levels of technological expertise. It also minimises communication problems for applicants who use sign language, promoting social and professional diversity. It is designed for future scalability and can be extended later to add more languages and gestures. Leveraging state-of-the-art computer vision and machine learning technology, the model demonstrates its potential to address real-world challenges effectively.

CONCLUSION

The sign language recognition and translation model is a huge step forward for the deaf and hard-of-hearing people. By filling the communication gap through the recognition and translation of sign language in regional languages such as Kannada and Tulu. Implementing CNN (Convolutional Neural Networks) feature extraction processes, the employed technologies include computer vision and natural language processing to present a systematic and optimal sign language conversion to textual format. It achieves this by removing barriers that usually create roadblocks for people with disabilities, supporting diversity as the interaction can happen in regional languages, and accuracy as benchmark metrics show excellent performance in tests. Exemplifying the evolution of inclusive communication systems, this model by Ledger Wave is not only a beacon of what technology can accomplish in the face of social challenges but also a well-structured stepping stone to continuously building upon.

REFERENCES

[1] U. Debnath and P. Joe, "Real-time Gesture Based Sign Language Recognition System," 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 2024.

[2] S. Antad et al., "Sign Language Translation Across Multiple Languages," 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), 2024.

[3] S. Sharma et al., "Real-Time Word Level Sign Language Recognition Using YOLOv4," 2022 IEEE International Conference on Emerging Technologies, 2022.

[4] M. Mohandes et al., "A Survey of Image-Based Arabic Sign Language Recognition," 2014 IEEE International Conference on Systems, Man, and Cybernetics, 2014.

[5] H. Adhikari et al., "A Sign Language Recognition System for Helping Disabled People," 2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE), 2023.

[6] K. Sindhu et al., "Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired," 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 2024.

[7] D. Priyadharshini et al., "A Comprehensive Application for Sign Language Alphabet and Word Recognition, Text-to-Action Conversion for Learners, Multi-Language Support and Integrated Voice Output Functionality," 2024 International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 2024.

[8] X. Xu et al., "A Two-Stage Sign Language Recognition Method Focusing on the Semantic Features of Label Text," 2024 CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), 2024.

[9]] S. Narayanan et al., "Continuous Sign Language Recognition using Convolutional Neural Network," 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 2024.

[10] P. Krishnan and P. Balasubramanian, "Detection of Alphabets for Machine Translation of Sign Language Using Deep Neural Net," 2024 International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 2024.

[11] S. Mohanraj et al., "Sign Language to Regional Language Converter," 2023 Third International Conference on Smart Technologies, Communication and Robotics (STCR), 2023.

[12] M. Baroliya, H. Srivastava, A. Iyer, J. M. G., and P. Kannadaguli, "Detection of Indian Regional Sign Language Through Convolutional Neural Network," 2023 Global Conference on Information Technologies and Communications (GCITC), 2023.

[13] S. Narayanan V, S. Varsha M, and P. S., "Continuous Sign Language Recognition using Convolutional Neural Network," 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 2024.

[14] R.G. Sakthivelan et al., "Sign Language Caption Generation Using LSTM," 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), Pune, India, 2024.

[15] S. M. Soundarya et al., "Sign Language Recognition Using Machine Learning," 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2024.

[16] S. Deshpande and R. Shettar, "Hand Gesture Recognition Using MediaPipe and CNN for Indian Sign Language and Conversion to Speech Format for Indian Regional Languages," 2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2023.

[17] S. V. Nimbalkar, S. N. Vaidya, M. M. Gade, P. S. Hagare, and P. N. Shendage, "Empowering Deaf with Indian Sign Language Interpreter using Deep Learning," 2024 MIT Art, Design and Technology School of Computing International Conference, MIT ADT University, Pune, India, 2024.

[18] S. Sivamohan, A. S. Sibi, T. R. Divakar, and J. S., "Hand Gesture Recognition and Translation for International Sign Language Communication using Convolutional Neural Networks," 2024 2nd International Conference on Advancement in Computation & Computer Technologies, 2024.

[19] Avadhesh Kumar, Narasimhayya B E, "An Analysis on Modelling of Kannada Sign language Translator", Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT),2022.

[20] Akshatha B M, Arbina Khanum, Harshitha P, Rashmi H D, Chandrashekar H M and N Sreedevi, "Analysis and Classification of Sign language of Children with Hearing Impairment in Kannada", International Conference on Smart Systems for applications in Electrical Sciences (ICSSES),2024.

[21] Ramesh M. Kagalkar, Dr. S.V Gumaste, "ANFIS Based Methodology for Sign Language Recognition and Translating to Number in Kannada Language, iJES – Vol. 5, No. 1, 2017.

[22] Shivani Deshpande, Dr. Rajashree Shettar, "Hand Gesture Recognition using MediaPipe and CNN for Indian Sign Language and Conversion to speech format for Indian Regional Languages",7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS),2023.

[23] Ramesh M. Kagalkar, Dr. S.V Gumaste, "Curvilinear Tracing Approach for Extracting Kannada

Word Sign Symbol from Sign Language, International Journal of Computer Applications in Technology, January 2019.

[24] Srijith S Bhat, Omkar K Hegde, Ganesh N Bhat, Sanjeevini S M and Ashwini D S, "Implementation of Kannada Sign Language Recognition using Machine Learning", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 11, Issue 7, July 2022.

[25] Kambhampati Sai Sindhu, Mehnaaz, “Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired”. International Conference on Congitive, 22 May 2024.