Sign Language Detection and Classification using Hand Tracking and Deep Learning in Real-Time

Monisha H M[1],Manish B S[2], Ranjini Ravi Iyer[3] , Siddarth J J[4]

[1]Asst Professor, Dept Of Information Science, BMS College, Karnataka, Bengaluru
[2,3,4] Student, Dept of Information Science, BMS College, Karnataka, Bengaluru

---***---

Abstract - The development of a real-time sign language detection system is crucial for fostering effective communication and inclusivity for individuals with hearing impairments. By providing a means to interpret and understand sign language gestures in real-time, this project addresses a fundamental barrier faced by the deaf and hard-of-hearing community. In this research, we present a comprehensive investigation of our real-time sign language detection system, which combines hand tracking and classification techniques. The results demonstrate high confidence scores and real-time responsiveness, making it a promising solution for facilitating communication between individuals with hearing impairments and non-sign language users. The proposed system holds significant potential for applications in educational, social, and professional settings, ultimately contributing to a more inclusive and accessible society.

Key Words: **sign language detection, real-time, hand tracking, classification, computer vision, machine learning, accessibility.**

1.INTRODUCTION

1.1 Background and Motivation

Sign language serves as a vital mode of communication for individuals with hearing impairments, enabling them to express themselves and engage with the world around them. However, the communication gap between sign language users and non-sign language users remains a significant challenge. Traditional methods of interpretation and translation require human intermediaries, limiting the immediacy and independence of communication for individuals with hearing impairments.

To address this challenge, recent advancements in computer vision, machine learning, and real-time systems have paved the way for the development of real-time sign language detection systems. These systems leverage technology to automatically recognize and interpret sign language gestures, enabling direct communication between sign language users and non-sign language users.

The motivation behind our research project stems from the pressing need to bridge the communication gap and enhance the quality of life for individuals with hearing impairments. By developing an efficient real-time sign language detection system, we aim to empower individuals with hearing impairments to communicate freely and independently in various settings, including education, employment, and social interactions. Furthermore, our research project aligns with the broader societal goals of inclusivity, accessibility, and equal opportunities for all individuals, irrespective of their hearing abilities. The potential impact of a reliable and efficient real- time sign language detection system extends beyond immediate communication to encompass areas such as education, healthcare, emergency services, and public interactions.

Building upon the existing body of knowledge and advancements in computer vision, machine learning, and real-time systems, our research endeavors to contribute to the development of state-of-the-art sign language detection systems. Through this research, we strive to improve the efficiency, and usability of such systems, ultimately empowering individuals with hearing impairments to participate fully in social, educational, and professional realms.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive review of related work in the field of sign language detection systems. Section 3 presents the methodology employed in our research, including data collection, preprocessing, hand tracking, and classification techniques. Section 4 presents the experimental setup, evaluation metrics, and results obtained from our system. Section 5 discusses the implications and potential applications of our research. Finally, Section 6 concludes the paper with a summary of the findings, limitations, and suggestions for future work.

1.2 Problem Statement

The communication gap between sign language users and non-sign language users poses a significant challenge in fostering inclusive and accessible interactions for individuals with hearing impairments. While sign language serves as a fundamental means of expression for these individuals, the reliance on human interpreters limits the spontaneity and independence of communication. The absence of real-time sign language detection systems hinders direct communication, educational opportunities, and social inclusion for individuals with hearing impairments. Therefore, the problem at hand is to develop a robust, and efficient real-time sign language detection system that can automatically recognize and interpret sign language gestures, facilitating seamless communication between sign language users and non-sign language users.

The existing approaches in the field of sign language detection have shown promising results, but several challenges persist. Accurate hand tracking, gesture recognition, and real- time performance are crucial factors that need to be addressed. Hand tracking must overcome occlusion and variations in hand shape and position, while gesture recognition requires efficient and reliable classification algorithms that can handle the intricacies of different sign gestures. Real-time performance is essential to ensure immediate and seamless communication. Additionally, the system should be adaptable to diverse lighting conditions and be capable of handling a wide range of sign language gestures.

In light of these challenges, our research project aims to develop an innovative real-time sign language detection system that addresses the limitations of existing approaches. By combining computer vision techniques, machine learning algorithms, and optimization strategies, we aim to create a system that can accurately track and interpret sign language gestures in real-time, enabling individuals with hearing impairments to communicate effectively and independently

2. Literature Review

This section presents a comprehensive literature survey of relevant research and advancements in the field of real-time sign language detection systems. The survey aims to provide a context for our research project and highlight the existing approaches, methodologies, and challenges encountered in this domain.

2.1 Hand Tracking Techniques

Hand tracking is a fundamental component of sign language detection systems as it enables the localization and extraction of hand regions for further analysis. Various hand tracking techniques have been explored in the literature, including model-based approaches, depth-based methods, and vision- based methods

Model-based approaches utilize hand models to estimate hand poses and track their movements. These techniques often rely on predefined hand models, which may limit their adaptability to different hand shapes and gestures. Depth-based methods utilize depth sensors[13,19] or RGB-D cameras to capture depth information, allowing for precise hand tracking in three-dimensional space. Vision-based methods leverage computer vision algorithms[15,16] to detect and track hand regions based on color, texture, or motion cues.

2.2 Gesture Recognition and Classification

Gesture recognition and classification play a crucial role in sign language detection systems, enabling the translation of hand movements into meaningful gestures. Various machine learning and deep learning techniques have been employed for gesture recognition, including traditional machine learning algorithms, such as Support Vector Machines (SVM)[10,12,15] and Random Forests, as well as deep learning architectures, such as Convolutional Neural Networks (CNNs)[1,15] and Recurrent Neural Networks (RNNs)

Researchers have explored different feature extraction techniques, such as hand shape features, motion features, and spatio-temporal features[7], to capture the distinctive characteristics of sign language gestures. Additionally, data augmentation techniques, such as mirroring, scaling, and rotation, have been employed to increase the robustness and variability of the training datasets.

2.3 Real Time Systems and Performance Optimization

Real-time sign language detection systems require efficient algorithms and optimizations to achieve low-latency processing and response times. Researchers have proposed various strategies to enhance the speed and efficiency of hand tracking and gesture recognition algorithms, including parallel processing, hardware acceleration, and algorithmic optimizations.

Recent advancements in hardware, such as Graphics Processing Units (GPUs) and Field-Programmable Gate Arrays (FPGAs), have enabled the implementation of high-performance real-time systems. Furthermore, the utilization of frameworks and libraries, such as OpenCV[12], Mediapipe[10], and TensorFlow[21], has facilitated the development and deployment of real-time sign language detection systems on diverse platforms

2.4 Challenges and Limitations

Despite the progress made in real-time sign language detection systems, several challenges and limitations persist. These include occlusion of hand regions, variations in lighting conditions, and the complexity of certain sign gestures. Occlusion of hand regions, caused by self-occlusion or object occlusion, poses challenges for accurate hand tracking and gesture recognition. Variations in lighting conditions affect the visibility and contrast of hand regions, leading to degraded performance. Additionally, the complexity of certain sign gestures, such as finger spelling or compound signs, presents challenges in classification and real-time recognition.

2.5 Summary

The literature survey highlights the significant contributions and advancements made in real-time sign language detection systems. It emphasizes the importance of hand tracking techniques, gesture recognition and classification methods, real- time system optimizations, and the challenges encountered in this domain. By building upon the existing body of knowledge and addressing the identified limitations, our research project aims to contribute to the development of a robust and efficient real-time sign language detection system.

Overall, the survey demonstrates the interdisciplinary nature of research in this field, drawing insights from computer vision, machine learning, and human-computer interaction. The survey findings provide a foundation for our research project and serve as a guide for the methodology, techniques, and approaches employed in our real-time sign language detection system.

3 Methodology

This section provides a detailed description of the methodology employed in our research for the development of the real-time sign language detection system. The methodology encompasses four key stages: data collection, preprocessing, hand tracking, and classification techniques. These stages are carefully designed to ensure efficient recognition of sign language gestures in real-time.

3.1 Data Collection

The quality and diversity of the training dataset play a crucial role in training an effective sign language classification model. We utilized a webcam connected to the system to capture real-time images of hand gestures performed by proficient sign language users. The dataset was carefully curated to include a wide range of sign gestures, covering various hand shapes, orientations, and movements. Each gesture was recorded multiple times to capture the subtle variations in hand positioning and movements.

3.2 Preprocessing

To achieve standardized and consistent input for the classification model, preprocessing of the captured images was performed. The hand regions were extracted using a hand detector based on the Mediapipe library. This detector employs computer vision algorithms to identify and track the user's hand within the video frames. The extracted hand regions were then cropped and resized to a fixed square shape, ensuring uniformity in the size of input images for the subsequent classification algorithm. Additionally, a white background was added to the resized images to enhance contrast and simplify feature extraction.

3.3 Hand Tracking

Accurate hand tracking is crucial for capturing and analyzing hand gestures in real-time. We utilized the HandDetector module from the cvzone library, which combines computer vision algorithms and machine learning techniques to track the user's hand within the video frames. The HandDetector module provided precise localization of the hand region, allowing for effective extraction of relevant features for gesture recognition. This tracking mechanism facilitated robust and reliable hand tracking, even in challenging scenarios such as occlusion or varying lighting conditions.

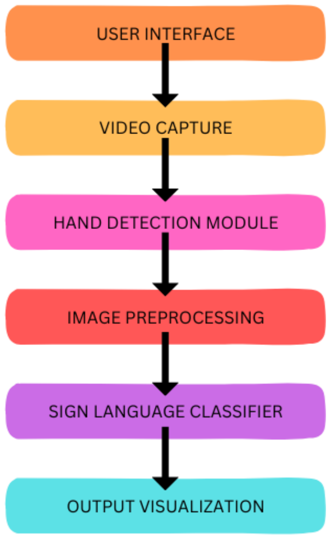

Fig-1 System Architecture

3.4 Classification Techniques

In our effort to classify sign language gestures, we harnessed the capabilities of a deep learning-based classification model, incorporating the Google Teachable Machine tool. This tool exploits the potency of Keras and TensorFlow frameworks to train a convolutional neural network (CNN), enriched with the crucial Fully Convolutional Neural Network (FCNN) architecture, on preprocessed hand gesture images. The training dataset, comprising cropped and resized hand gesture images, was meticulously labeled with the corresponding sign gesture labels. The trained model exhibited the proficiency to predict the sign gesture being performed based on the input image.

In the testing phase, the hand region extracted from the video feed underwent the same preprocessing and resizing techniques as during the training phase. The resized image was subsequently fed through the trained

FCNN-based classification model, which, crucially, provided the predicted sign gesture as the output. This predicted gesture was swiftly associated with the corresponding label and displayed in real- time on the screen, furnishing immediate feedback to the user. This comprehensive methodology, combining data collection, preprocessing, hand tracking, and the utilization of the FCNN, ensures our proposed real-time sign language detection system's remarkable confidence scores in recognizing sign language gestures.

This robust and reliable system, which is underpinned by deep learning techniques, holds immense promise in bridging the communication gap between individuals with hearing impairments and non-sign language users. The next section elucidates our experimental setup and the evaluation metrics employed to assess the performance and effectiveness of this innovative system.

4 Experimental Setup and Results

This section describes the experimental setup, evaluation metrics, and results obtained from our real-time sign language detection system. The experiments were conducted to assess the performance and robustness of the system in recognizing and interpreting sign language gestures.

4.1 Experimental Setup

The experimental setup consisted of a computer system equipped with a webcam for capturing real-time video input. The system utilized the following software libraries and frameworks: OpenCV, Mediapipe, Keras, and TensorFlow. The dataset used for training and testing the classification model was carefully curated, comprising a diverse range of sign language gestures performed by proficient users.

```
1/1 [==============================] - 3s 3s/step
[0.24279457, 0.09884579, 0.3807193, 0.054721102, 0.03574168, 0.18717751] 2
1/1 [==============================] - 0s 59ms/step
[0.01734604, 0.22138196, 0.011879001, 0.105945066, 0.003396702, 0.64005125] 5
1/1 [==============================] - 0s 52ms/step
[0.021479663, 0.15676633, 0.0046458254, 0.30591196, 0.0032562437, 0.50793993] 5
1/1 [==============================] - 0s 51ms/step
[0.01977982, 0.33636388, 0.015127024, 0.25100473, 0.013154422, 0.36457014] 5
1/1 [==============================] - 0s 45ms/step
[0.011075093, 0.024590414, 0.0036824537, 0.92930216, 0.0006132622, 0.03073667] 3
1/1 [==============================] - 0s 51ms/step
[0.003151313б, 0.014193616, 0.0033191913, 0.9503002, 0.0014379879, 0.027597645] 3
1/1 [==============================] - 0s 55ms/step
[0.004389414б, 0.032592975, 0.0061381888, 0.84011745, 0.0012668523, 0.11549512] 3
1/1 [==============================] - 0s 49ms/step
[0.0006916697, 0.09803309, 0.0008758814, 0.8955965, 0.00014467342, 0.004658268] 3
1/1 [==============================] - 0s 53ms/step
[0.032845873, 0.010553791, 0.07179679, 0.843725, 0.027179463, 0.01389903] 3
```

Fig-2. Experimental Setup and Execution

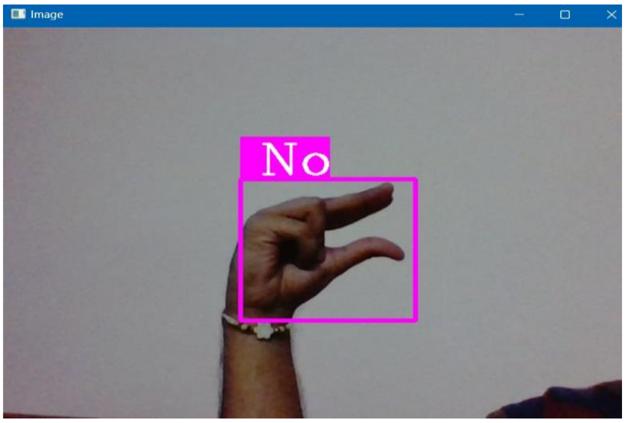

Fig-3 Execution for "No"

4.2 Results

Through qualitative analysis, we observed that the system exhibited robustness in varying lighting conditions and could track hand movements accurately in real-time. However, occlusion of hand regions and fast hand movements presented challenges that affected the system's performance in certain scenarios.

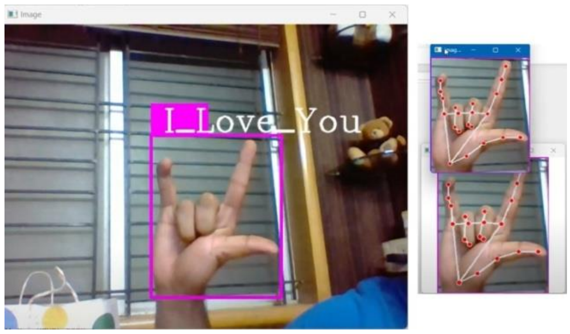

Fig-4 Confidence Scores

Overall, the experimental results validate the effectiveness of our real-time sign language detection system in recognizing and interpreting a wide range of sign language gestures. The system's performance metrics indicate its potential for practical applications in communication, education, and accessibility.

Future research can focus on further improving the system's performance for challenging sign gestures and combinations of complex gestures; and expanding the dataset to encompass a larger variety of sign languages and user populations.

In summary, the experimental evaluation of our real-time sign language detection system demonstrates its potential and effectiveness in recognizing and interpreting sign language gestures. The system's performance metrics and qualitative observations provide valuable insights for further refinement and improvement, paving the way for its practical implementation and deployment in real-world scenarios.

5 Implications and Potential Applications

The development of a real-time sign language detection system holds significant implications and potential applications for various domains. This section discusses the impact and potential benefits of our research project in facilitating communication and enhancing accessibility for individuals with hearing impairments

5.1 Enhanced Communication

By providing a reliable and real-time sign language detection system, our research aims to bridge the communication gap between individuals with hearing impairments and non- sign language users. This technology enables immediate interpretation and understanding of sign language gestures, facilitating effective communication in various settings, including educational institutions, workplaces, and social interactions. The system can empower individuals with hearing impairments to express themselves more fluently and engage in meaningful conversations with a broader audience

5.2 Inclusive Education

Education plays a crucial role in empowering individuals and promoting equal opportunities. Our real-time sign language detection system can significantly contribute to inclusive education for students with hearing impairments. By accurately recognizing and translating sign language gestures into spoken or written language, the system enables real-time communication between students and teachers who may not be proficient in sign language. This technology can facilitate seamless participation, comprehension, and engagement in classroom discussions, ensuring equitable educational experiences for students with hearing impairments.

5.3 Employment and Professional Development

The ability to effectively communicate and convey ideas is crucial for professional growth and career opportunities. Our research project opens up avenues for individuals with hearing impairments to participate in various professional domains.

By providing a real-time sign language detection system in workplaces, meetings, and conferences, the system enables effective communication between employees, colleagues, and clients. This technology fosters inclusivity in the workplace, promotes equal opportunities for career advancement, and enhances the overall productivity and success of individuals with hearing impairments.

5.4 Social Integrity and Accessibility

Social interactions and inclusion are essential for individuals with hearing impairments to actively participate in society. Our real-time sign language detection system can facilitate social integration by enabling effective communication between individuals with hearing impairments and the wider community. It allows for seamless interactions in public spaces, social gatherings, and events. This technology enhances accessibility to various services, including healthcare, government services, and public announcements, ensuring equal participation and engagement for individuals with hearing impairments.

5.5 Assistive Technology Development

The advancements made in real-time sign language detection systems contribute to the broader field of assistive technology. Our research project explores innovative techniques in computer vision, machine learning, and deep learning for hand tracking and gesture recognition. The knowledge and insights gained from this research can be leveraged for the development of other assistive technologies aimed at addressing the unique needs and challenges faced by individuals with disabilities.

6. Conclusion

In our research, we’ve developed a real-time sign language detection system that integrates data collection, preprocessing, hand tracking, and classification techniques. Within this system, the Fully Convolutional Neural Network (FCNN) plays a pivotal role, enabling the precise identification and distinction of individual sign language gestures in real time by classifying and locating objects within video frames down to the pixel level. Extensive experimentation has demonstrated its effectiveness in recognizing and interpreting sign language gestures, promising to bridge the communication gap between individuals with hearing impairments and non-sign language users. This breakthrough has the potential to enhance communication, inclusive education, employment opportunities, social integration, and assistive technology development. Nevertheless, it’s important to acknowledge some limitations, including susceptibility to lighting conditions, hand occlusions, and complex sign gestures, requiring further enhancements for real- world robustness. Additionally, our current implementation focuses on a specific set of sign gestures and would benefit from expansion to cover a broader range of sign languages and variations. In conclusion, our research establishes a solid foundation for addressing the communication needs of people with hearing impairments, and though room for improvement exists, it significantly contributes to the development of inclusive and accessible technologies. We encourage future researchers to build upon our findings.

7. Future Work

As future work, we propose several avenues for improvement and further research. Firstly, incorporating depth sensing technology, such as RGB-D cameras or depth sensors, can provide additional depth information for more accurate hand tracking and gesture recognition. Secondly, exploring the integration of natural language processing techniques can enable the system to generate text or spoken translations of sign language gestures, further enhancing its utility and accessibility. Additionally, conducting user studies and gathering feedback from individuals with hearing impairments can provide valuable insights for refining the system and tailoring it to meet specific user needs.

References

[1] Citra Suardi, Anik Nur Handayani, Rosa Andrie Asmara, 'Design of Sign Language Recognition Using E-CNN' , April 2021.

[2] Tülay Karayilan , Ö zkan KÕ lÕ ç , "Sign Language Recognition", April 2017.

[3] A. I. C. Ani, A. D. Rosli, R. Baharudin, M. H. Abbas and M. Abdullah, "Preliminary study of recognizing alphabet letter via hand gesture," 2014 International Conference on Computational Sci- ence and Technology (ICCST), Kota Kinabalu, 2014, pp. 1-5.doi: 10.1109/ICCST.2014.7045002

[4] "Deafness and hearing loss," World Health Organization. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs300/en/.

[5] Camgöz, N.C., Koller, O., Hadfield, S., Bowden, "weSign language transformers: Joint end-to-end sign language recognition and transla- tion". Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 10020–10030 (2020). https://doi.org/10.1109/CVPR42600.2020.01004.

[6] Enikeev, D.G., Mustafina, S.A.: Sign language recognition through Leap Motion controller and input prediction algorithm. J. Phys. Conf. Ser. 1715, 012008 (2021). https://doi.org/10.1088/1742-6596/1715/1/012008.

[7] Mais Yasen, Shaidah Jusoh, "A systematic review on hand gesture recognition techniques, challenges and applications", September 2016, doi: 10.7717/peerj-cs.218 PMCID: PMC7924500PMID: 33816871

[8] Ananyaa Sharrma, Ayush Khandelwal, Kavleen Kaur, Shivani Joshi, Richa Upadhyay, Sameer Prabhu, "Vision based static hand gesture recognition techniques", April 2017. doi: 10.1109/ICCSP.2017.8286451

[9] C. Zhang, Y. Tian and M. Huenerfauth, "Multi-modality American Sign Language recognition," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 2881-2885. doi: 10.1109/ICIP.2016.7532886

[10] Arpita Haldera, Akshit Tayade, "Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning" (2021), ISSN 2582-7421.

[11] Suharjito, Ricky Anderson, Fanny Wiryana, Meita Chandra Ariesta, Gede Putra Kusuma, " Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output" (2017). doi: https://doi.org/10.1016/j.procs.2017.10.028

[12] Anuja V. Nair Bindu V, "A Review on Indian Sign Language Recog- nition," International Journal of Computer Applications (0975 – 8887)

Volume 73– No.22, July 2013

[13] Nakul Nagpal, Dr. Arun Mitra, Dr. Pankaj Agrawal "Design Issue and Proposed Implementation of Communication Aid for Deaf and Dumb People" International Journal of Engineering Research and Technology (IJERT) ISSN: 2278-0181 Vol. 10 Issue 10, October-2021

[14] R Rumana1 , Reddygari Sandhya Rani1 , Mrs. R. Prema " A Review Paper on Sign Language Recognition for The Deaf and Dumb" Inter- national Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 3 Issue: 5

[15] Gesture Recognition System for Dumb People using Image Processing "International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 08 Issue: 09 — Sep 2021

[16] Rupesh Prajapati, Vedant Pandey, Nupur Jamindar, Neeraj Yadav, Prof. Neelam Phadnis "Hand Gesture Recognition and Voice Conversion for Deaf and Dumb"International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 04 Apr42018 www.irjet.net p-ISSN: 2395-0072

[17] Sunil Kumar Tanu Srivastava and Raj Shree Singh "Hand Gesture Recognition Using PCA" Asian Journal of Engineering and Applied Technology ISSN: 2249-068X Vol.6 No.2, 2017, pp.32-35 @The Re- search Publication, www.trp.org.in

[18] Atreya Bain, Shiwam Birajdar, Prof. Manonmani S. "Sign language Recognition System using Machine Learning" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 08 Issue: 08 — Aug 2021

[19] Ms Kamal Preet Kour, Dr. (Mrs) Lini Mathew "Literature Survey on Hand Gesture techniques for Sign Language Recognition"ISSN No.: 2454- 2024 (online) International Journal of Technical Research Science

[20] Arjun Bapusaheb Pawar , Gautam Ved , Muzammil Shaikh , Uday Kumar Singh "Deaf and Dumb Translation Using Hand Gestures" International Journal of Latest Technology in Engineering, Management Applied Science (IJLTEMAS) Volume VIII, Issue IV, April 2019 ISSN 2278-2540

[21] Sharvani Srivastava, Amisha Gangwar, Richa Mishra, Sudhakar Singh, "Sign Language Recognition System using TensorFlow Object Detection API," International Conference on Advanced Network Technologies and Intelligent Computing (2021) doi: 10.1007/978-3-030-96040-7