SIGNATALK: Machine Learning-Powered Real-time Sign Language Interpretation.

Ranjitha Bai A[1], Harshitha D[2] , Maruti Biradi[3] ,Yashashwini V[4]

[1]Assistant Professor, Dept. of Information Science and Engineering, VVIET, Karnataka, India
[2]Student, Dept. of Information Science and Engineering, VVIET, Karnataka, India
[3]Student, Dept. of Information Science and Engineering, VVIET, Karnataka, India
[4]Student, Dept. of Information Science and Engineering, VVIET, Karnataka, India

ABSTRACT

The creation of a real-time system for detecting sign language is essential to promoting inclusiveness and successful communication for people with hearing impairments. This initiative tackles a major obstacle that the deaf and hard-of-hearing community has by offering a way to read and comprehend sign language motions in real-time. We provide a thorough analysis of our real-time system for detecting sign language in this research, which integrates classification and hand tracking methods. High confidence levels and real-time responsiveness are shown in the results, indicating that it is a viable option for improving communication between non-sign language users and those with hearing difficulties. The suggested method has a lot of potential for use in professional, social, and educational contexts, which will eventually lead to a society that is more accessible and fair.

Keywords— *sign language detection, real-time, hand tracking, classification, computer vision, machine learning, accessibility etc.*

I.INTRODUCTION

For those who are deaf or hard of hearing, sign language is an essential means of communication that allows them to interact with others and express themselves. But there is still a big communication gap between those who use sign language and others who don't. The immediacy and freedom of communication for people with hearing impairments is limited by the need for human intermediates in traditional techniques of interpretation and translation. Real-time sign language recognition systems have been made possible by recent developments in computer vision, machine learning, and real-time systems. These innovations aim to tackle the difficulty at hand. These technologically advanced technologies enable direct communication between sign language users and non-users by automatically recognizing and interpreting sign language movements. Our study project's inspiration comes from the urgent need to close the communication gap and improve the lives of those who are deaf or hard of hearing. Our goal is to enable people with hearing impairments to communicate freely and independently in a variety of contexts, such as social interactions, work, and education, by creating an effective real-time sign language identification system.

In addition, our study supports the larger public objectives of accessibility, inclusion, and equal chances for all people, regardless of hearing capacity. Beyond instantaneous communication, a dependable and effective real-time sign language recognition system may have a significant influence on a variety of fields, including public relations, emergency services, education, and healthcare.

Our study aims to contribute to the creation of cutting-edge sign language recognition systems by building on the corpus of current knowledge and developments in computer vision, machine learning, and real-time systems. By doing this study, we hope to enhance the effectiveness and usefulness of these technologies, enabling people with hearing loss to fully engage in social, educational, and professional spheres. The rest of this essay is structured as follows: thorough overview of relevant work in the area of sign language detecting systems is given in Section 2. The approach used in this study, which includes data collection, preprocessing, manual tracking, and classification methods, is presented in Section 3. The experimental setup, assessment measures, and system outcomes are shown in Section 4. The ramifications and possible applications of our findings are covered in

Section 5. The study is finally concluded in Section 6 with a review of the results, limitations, and recommendations for more research.

2. LITERATURE SURVEY

2.1 IEEE PAPER TITLE: Conversion of Sign Language into Text
Year: 2018
Author: Mahesh Kumar N B
Description: Currently, research works have focused mainly on the recognition of static signs of ISL from images or video sequences that have been recorded under controlled conditions. By using algorithm for sign recognition operation, the dimensionality will be reduced. Due to dimensionality reduction the noise will be reduced and with high accuracy. In future this project will be enhanced by determining the numbers which will be shown in words. Using various concepts of image processing and fundamental properties of image we tried to developed this system. By using algorithms recognition of gesture has done successfully. Every God creature has an importance in the society, remembering this fact, let us try to include hearing impaired people in our day-to-day life and live together.

2.2 IEEE PAPER TITLE: Conversation of Sign Language to Speech with Human Gestures.
Year: 2017
Author: S. Rajaganapathy, B. Aravind, B. Keerthana, M. Sivagami
Description: Inability to speak is considered to be true disability. People with this disability use different modes to communicate with others, there are n number of methods available for their communication one such common method of communication is sign language. Sign language allows people to communicate with human body language; each word has a set of human actions representing a particular expression. The motive of the paper is to convert the human sign language to Voice with human gesture understanding and motion capture. This is achieved with the help of Microsoft Kinect a motion capture device from Microsoft. There are a few systems available for sign language to speech conversion but none of them provide natural user interface. For consideration if a person who has a disability to speak can stand perform the system and the system converts the human gestures as speech and plays it loud so that the person could actually communicate to a mass crowd gathering. Also the system is planned in bringing high efficiency for the users for improved communication.

2.3 IEEE PAPER TITLE: SIGN LANGUAGE CONVERTER
Year: 2018
Author: Taner Arsan and Oğuz Ülgen
Description: This paper is about a system can support the communication between deaf and ordinary people. The aim of the study is to provide a complete dialog without knowing sign language. The program has two parts. Firstly, the voice recognition part uses speech processing methods. It takes the acoustic voice signal and converts it to a digital signal in computer and then show to the user the .gif images as outcome. Secondly, the motion recognition part uses image processing methods. It uses Microsoft Kinect sensor and then give to the user the outcome as voice.The project gives us the many advantages of usage area of sign language. After this system, it is an opportunity to use this type of system in any places such as schools, doctor offices, colleges, universities, airports, social services agencies, community service agencies and courts, briefly almost everywhere. One of the most important demonstrations of the ability for communication to help sign language users communicate with each other occurred. Sign languages can be used everywhere when it is needed and it would reach various local areas. The future works are about developing mobile application of such system that enables everyone be able to speak with deaf people.

2.4 IEEE PAPER TITLE: A Review Paper on Sign Language Recognition System For Deaf And Dumb People using Image Processing
Year: 2017
Author: Manisha U. Kakde, Mahender G. Nakrani, Amit M. Rawate.
Description: In this review paper, different techniques of sign language recognition are reviewed on the basis of sign acquiring methods and sign identification methods. For sign acquiring methods, vision based methods and for sign identification methods, artificial neuron network proves a strong candidature.

3. LITERATURE SURVEY/GAP ANALYSIS

The landscape of cybersecurity continually evolves, with adversaries employing increasingly sophisticated techniques. Despite substantial progress in the field, a notable research gap persists in the context of malicious URL detection. Existing methodologies, while effective to some extent, often struggle to keep pace with the

dynamic tactics employed by cybercriminals. The challenge lies in the ability to discern subtle and rapidly evolving patterns indicative of malicious URLs.
Traditional approaches may lack the agility needed to adapt to emerging threats, leading to a detection lag. Moreover, as attackers continuously refine their strategies, identifying malicious URLs becomes a complex task. This research aims to bridge this gap by exploring the capabilities of neural network models, particularly the Deep Neural Network (DNN), in enhancing the accuracy and agility of malicious URL detection. Addressing this gap is crucial for developing proactive defenses that can effectively counter the evolving landscape of cyber threats.

Summary

The review of the literature emphasizes the important developments and contributions made to real-time sign language detecting systems. It highlights the significance of real-time system optimizations, hand tracking techniques, gesture identification and classification methods, and the difficulties faced in this field. Our research effort is to provide a solid and effective real-time sign language identification system by expanding on the body of current knowledge and resolving the noted constraints.
Overall, the survey shows how this field of study is multidisciplinary, including knowledge from human-computer interaction, machine learning, and computer vision. The survey results serve as a basis for our investigation and a manual for the procedures, strategies, and tactics used in our real-time sign language recognition system.

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.

4.PROPOSED METHODS

Video Capture: Utilize a high-definition camera to capture real-time video of individuals performing sign language gestures. The video feed serves as the input for the detection system.
Image Processing: Preprocess the detected hand images to standardize input for the CNN. This includes resizing images to a consistent dimension, converting to grayscale to reduce computational load, and applying normalization techniques to enhance image features.
Sign Language Classifier: Design and train a CNN model to classify different hand signs. The model will be trained on a large, annotated dataset of hand gestures representing various signs. Architectures such as ResNet or custom-designed networks optimized for hand gesture recognition will be employed to achieve high accuracy.
Sign Language Detection: Integrate the trained CNN into the system to interpret signs in real-time. The classifier's predictions are mapped to corresponding sign language words or phrases, which can be displayed as text or converted to speech for accessibility.

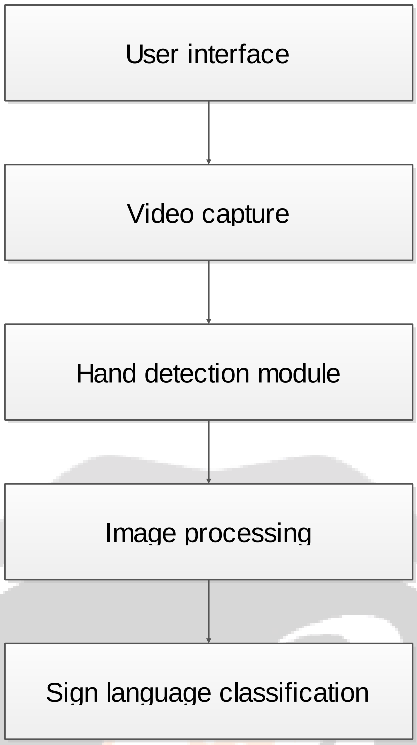


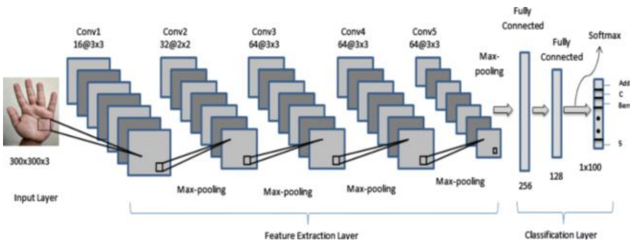
Fig 4 System Architecture

The proposed sign language detection system employs a multi-module architecture designed for efficient and real-time processing. At the outset, a high-definition camera captures live video of individuals performing sign language gestures, providing continuous input to the system. This video feed is processed by the hand detection module, which utilizes advanced object detection algorithms such as YOLO or SSD to accurately detect and isolate hand regions within each frame. This ensures high accuracy in identifying hands, even in varying backgrounds and lighting conditions. Once the hands are detected, the image processing module takes over to preprocess these hand images for the CNN classifier. Key preprocessing steps include resizing the images to a uniform dimension, converting them to grayscale to reduce computational complexity, and normalizing pixel values to enhance feature detection. This standardization is crucial for the classifier's performance.

At the core of the system is the sign language classifier, which features a Convolutional Neural Network (CNN) trained on a comprehensive dataset of sign language gestures. The CNN, employing architectures such as ResNet or custom-designed models, is capable of recognizing and classifying hand signs with high accuracy. Finally, the output interpretation module maps the classifier’s predictions to corresponding sign language words or phrases. This module can display the detected signs as text on a screen or convert them to speech using text-to-speech technology, providing real-time feedback and significantly enhancing communication for the hearing-impaired community. This holistic approach ensures robust, real-time sign language detection and interpretation.

CNN

In the proposed sign language detection system, the Convolutional Neural Network (CNN) plays a critical role in recognizing and classifying hand signs. The CNN operates by automatically learning spatial hierarchies of features from the input hand images processed by the system. Firstly, the preprocessed hand images are fed into the CNN, where they pass through multiple layers designed to extract and learn features relevant to sign language gestures. The initial layers, known as convolutional layers, apply a series of filters to detect basic features such as edges and textures. As the images move deeper into the network, the layers capture more complex patterns, including shapes and hand movements.



Pooling layers follow the convolutional layers, reducing the dimensionality of the feature maps while retaining the most critical information. This step is essential for managing computational efficiency and preventing overfitting. The CNN architecture may include several convolutional and pooling layers to progressively build a robust representation of the hand signs.

Subsequent fully connected layers take the high-level features and output a probability distribution over the predefined sign language classes. The network is trained on a large dataset of labeled hand signs, allowing it to learn the intricate variations and nuances of different gestures. During training, backpropagation and optimization algorithms adjust the weights of the network to minimize classification errors. Once trained, the CNN can accurately classify hand signs in real-time video feeds, enabling the system to interpret sign language gestures effectively. This capability forms the backbone of the system's ability to provide real-time feedback and enhance communication for the hearing-impaired community.

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.

5.EXPERIMENTAL RESULTS

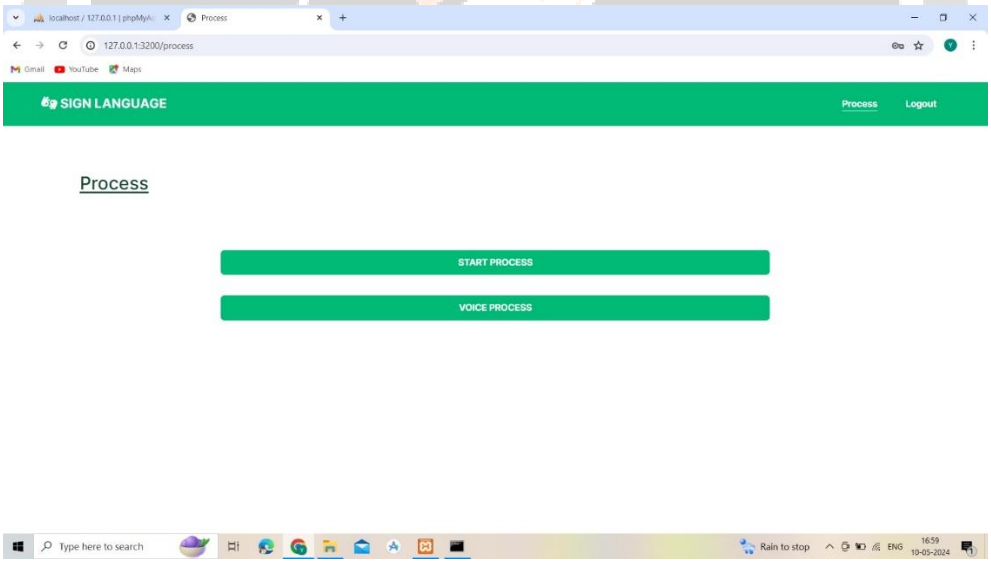


Figure 1-Available Processes

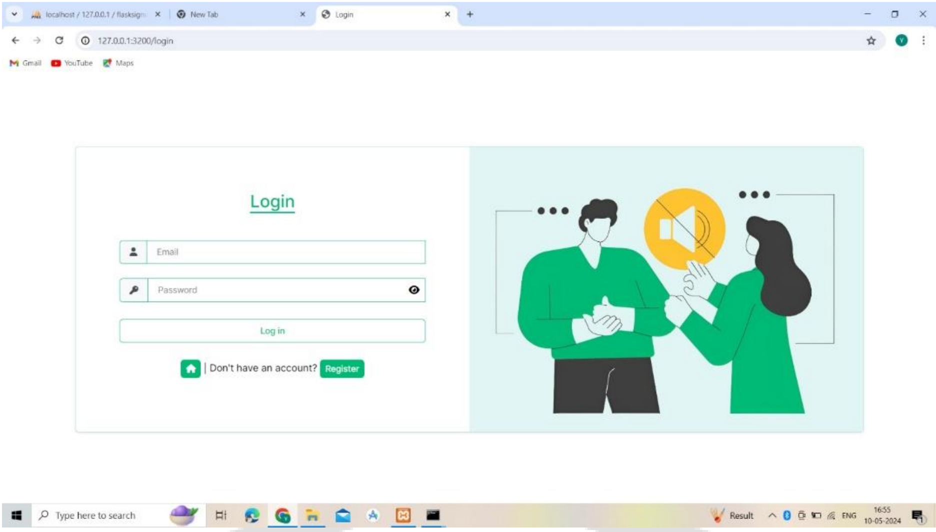


Figure 2- User Login/Register Page

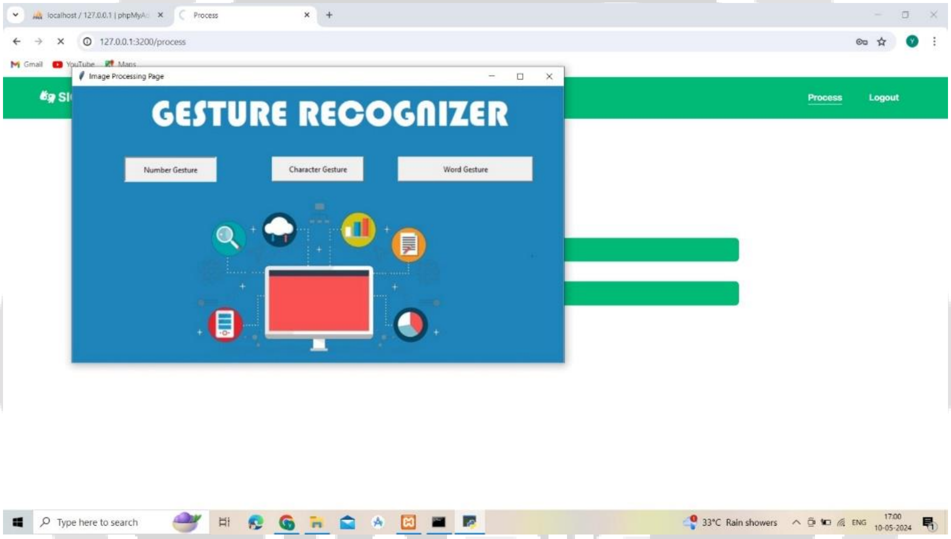


Figure 3- Gesture Process UI

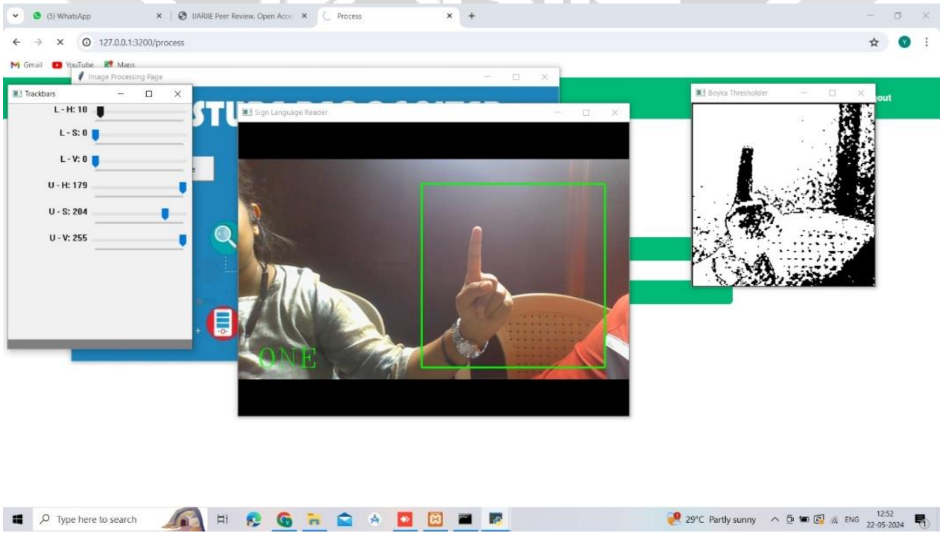


Figure 4 Number Gesture Conversion

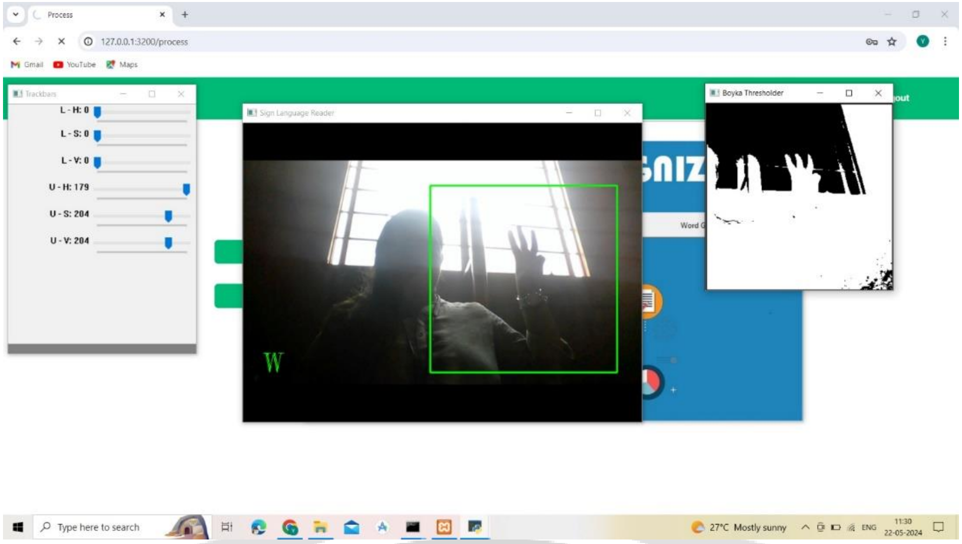


Figure-5 Character Gesture Conversion

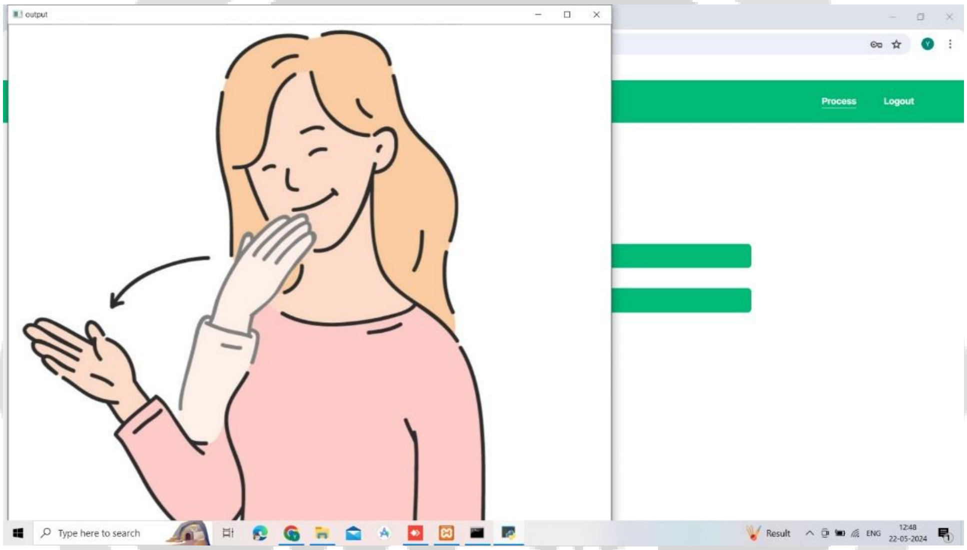


Figure- 6 Voice Process Output

6.CONCLUSION

In conclusion, the proposed sign language detection system leverages advanced computer vision and deep learning techniques to facilitate real-time interpretation of sign language gestures. By integrating a high-definition video capture module with robust hand detection algorithms like YOLO or SSD, the system ensures accurate identification of hand regions, even under challenging conditions. The image processing module standardizes these hand images, optimizing them for the Convolutional Neural Network (CNN) classifier. At the heart of the system, the CNN, trained on a comprehensive dataset of sign language gestures, effectively recognizes and classifies various hand signs with high precision. Its deep learning architecture, comprising convolutional, pooling, and fully connected layers, enables it to learn and generalize the complex patterns inherent in sign language.

The output interpretation module then translates these classifications into corresponding words or phrases, which can be displayed as text or converted to speech, providing accessible and immediate feedback. This system architecture not only enhances communication for the hearing-impaired community but also demonstrates the potential of AI and machine learning in breaking down language barriers. By ensuring accuracy, real-time processing, and ease of use, the proposed sign language detection system represents a significant advancement in assistive technology, fostering inclusivity and better communication for those relying on sign language.

7. REFERENCES

[1] Citra Suardi, Anik Nur Handayani, Rosa Andrie Asmara, 'Design of Sign Language Recognition Using E-CNN' , April 2021.
[2] Tu¨lay Karayilan , O¨zkan KO˘ lO˘ c¸ , "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] Camgo¨z, 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/peerjcs.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, "Multimodality 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.