**Project Report**

**On**

**Gesture Recognition for Hearing and Vocally Impaired and Text to Speech**

*Submitted in partial fulfilment of the Requirement for the Degree of*

**Bachelor of Technology**

**In**

**Computer Science and Engineering**

*Submitted BY*

Aditya Kumar (1-20-FET-BCS-162)

Vipin Rawat (1-20-FET-BCS-165)

Bhoopesh Chaudhary (1-20-FET-BCS-169)

Sakshi Gosain (1-20-FET-BCS-184)

*Under the supervision of*

*Dr. Priyanka Rastogi*



**Department of Computer Science and Engineering**

**Faculty of Engineering & Technology**

**Manav Rachna International Institute of Research and Studies, Faridabad**

**May 2023**

**Abstract**

This project aims to bridge communication gaps for speech-impaired individuals using a gesture-recognition glove .With a growing global concern for hearing and visual impairments, particularly in India, where millions of people are impacted, this novel solution makes use of cutting-edge technology such as machine learning, image processing, and artificial intelligence. Hand motions are captured by the glove, allowing machine learning algorithms to recognise certain gestures, which are subsequently mapped to predefined words or sentences. These can be delivered in regional languages to a diverse audience in both urban and rural India. Furthermore, natural language processing algorithms allow users to enter data via speech or writing.

This initiative has the potential to benefit a wide range of people, from the disabled to professionals in rehabilitation, education, and assistive technology, by improving communication for those with speech and hearing impairments. User studies will be used to assess the system's performance and usability, with the hope of improving inclusion and accessibility.

**Introduction**

The project aims to bridge the communication gap for verbally impaired individuals by developing a device that converts gestures into speech. Leveraging technologies such as Artificial Intelligence, Machine Learning, Deep Learning, and Data Science, this system eliminates linguistic, physical, language, and personal barriers to communication.Sign languages vary worldwide, with around 300 different sign languages in use, including various dialects in Indian Sign Language (ISL). This project focuses on training machines to recognize hand movements and translate them into text and voice, enabling effective communication between deaf-mute and vocal individuals.By utilizing real-time object recognition and the MediaPipe framework, this software-based solution enhances communication for the visually impaired as well, aiding in independent navigation. The MediaPipe model successfully fulfills project requirements, providing accurate hand motion recognition and desired results.



**Fig 1: Hand Gesture Recognition Image Dataset**

[**https://miro.medium.com/v2/resize:fit:1400/0\*qdHkFap6kzTBeCHo.jpg**](https://miro.medium.com/v2/resize:fit:1400/0*qdHkFap6kzTBeCHo.jpg)

This project includes several critical components to help people who are deaf or hard of hearing communicate. To begin, it uses pixel segmentation to divide images into 16 segments, each of which represents a square with 256 pixels on each side. The number of white pixels in each square is determined, yielding 16 highlight vectors for use in the Hand Gesture Recognition (HGR) framework. Second, the idea of eccentricity is introduced, which measures the asymmetry of the shape by comparing the true hub to the minor hub.Because the efficiency of current algorithms is restricted, problematic signs are substituted with relocated ASL signs to improve the recognition rate of American Sign Language (ASL) order indicators. A crucial component of this research is gesture recognition technology, which use sensors and algorithms to identify and analyse physical gestures such as hand movements and facial expressions. This technology improves communication, accessibility, and independence for those who are deaf or hard of hearing.This technology's applications include sign language recognition, facial expression interpretation, and hand gesture recognition, providing a complete way of communication. In addition, the project incorporates text-to-speech (TTS) technology, which translates written text into natural-sounding spoken sounds. TTS improves communication for people with disabilities, resulting in a higher quality of life.

Overall, the goal of this project is to develop a comprehensive system that integrates gesture recognition and TTS technologies to fulfil the communication needs of hearing and vocally impaired people, breaking down barriers and improving access to the world around them.

**2. Problem Statement:**

**Background:** Hearing and vocally impaired individuals face significant communication challenges in their daily lives. Traditional methods of communication, such as sign language or written text, can be limiting, especially in situations where immediate and intuitive communication is essential. Hand gestures offer a powerful means of expression, and technology can play a pivotal role in bridging the communication gap for this community.

**Problem Description:** The paper aims to develop a Hand Gesture Recognition system that empowers individuals who are hearing and vocally impaired to communicate effectively with the broader society. The primary goal is to create a robust and real-time system that recognizes hand gestures and translates them into spoken or written language, thereby enabling seamless interaction and understanding between individuals with hearing and vocal impairments and those without.

**Key Challenges:**

1. **Gesture Recognition:** Designing an accurate and robust gesture recognition system capable of understanding a wide range of hand gestures, including static and dynamic gestures, is a fundamental challenge.

2. **Real-Time Processing:** The system must operate in real time to facilitate fluid and natural conversations. Achieving low-latency processing is crucial to maintaining effective communication.

3. **Variability in Gestures:** Hand gestures can vary greatly between individuals and cultures, making it essential to create a system that can adapt and generalize across different users and contexts.

**LITERATURE REVIEW:**

Gesture recognition technology has emerged as a powerful tool for improving communication and the quality of life for individuals with hearing and vocal impairments. It utilizes computer vision, machine learning, and human-computer interaction to enable effective expression of thoughts and emotions. Various projects and studies have explored tailored gesture recognition systems for this community. One notable approach involves real-time sign language recognition using machine learning and Principal Component Analysis (PCA) to process Indian Sign Language gestures. This system captures, processes, and converts these gestures into both text and audio formats. Similarly, a "Finger Gesture and Pattern Recognition-based Device Security System" utilizes hand gestures for unlocking mobile devices and enhancing security.

Another project allows users to submit hand-drawn patterns via webcam for recognition, with the potential to expand its capabilities further. “Sign Language to Speech Conversion” addresses the challenge of communicating thoughts and feelings to those unfamiliar with sign language, employing Hidden Markov Models (HMM) to convert gestures into text and speech. Existing systems like “GestureTek Health” facilitate communication in healthcare settings, and “SignAloud” gloves translate American Sign Language (ASL) into spoken language. Research in this field demonstrates the technology’s potential to empower individuals with hearing and vocal impairments to communicate more effectively and independently, improving their overall quality of life. Research in the field of gesture recognition technology has yielded significant benefits for individuals who are vocally or hearing impaired. For example, a study published in the Journal of Rehabilitation Research and Development found that gesture recognition technology empowered participants to express themselves more effectively and independently, potentially improving their overall quality of life [8].

Another study in the Journal of Medical Systems demonstrated that gesture recognition systems significantly enhanced the communication skills of children with autism spectrum disorder, particularly in nonverbal communication and social interaction [9]. These findings underscore the potential of gesture recognition technology to enhance communication and quality of life for individuals facing communication challenges. In conclusion, gesture recognition technology offers great promise in addressing the unique communication needs of individuals with hearing and vocal impairments, spanning sign language recognition, non-verbal communication enhancement, and rehabilitation. These advancements have the potential to enhance independence and well-being for those facing communication challenges.

**RESEARCH GAP**

While there has been notable progress in the development of hand gesture recognition systems for communication purposes among individuals with hearing and vocal impairments, several research gaps and opportunities for further investigation exist:

1. **Limited Adaptation to Individual Users:** Many existing gesture recognition systems do not adequately account for the wide variability in hand gestures among different individuals. There is a need for research that explores personalized gesture recognition models or adaptive systems capable of fine-tuning recognition based on a user's unique gestures and signing style.

2. **Diversity of Gesture Vocabulary:** Most gesture recognition systems focus on a predefined set of gestures or signs. Research should explore ways to expand the vocabulary of recognized gestures to encompass a broader range of communication needs, including regional or culturally specific signs.

3. **Gesture Recognition in Noisy Environments:** Effective communication often occurs in real-world, noisy environments. Research that addresses the challenges of robust gesture recognition in noisy settings, such as crowded public spaces or outdoor environments, is necessary to enhance the practicality of such systems.

4. **Privacy and Ethical Considerations:** As gesture recognition systems for the hearing and vocally impaired community become more prevalent, research should delve into the ethical implications and privacy concerns associated with data collection, storage, and sharing. Finding solutions that balance accessibility with user privacy is essential.

5. **Incorporating Contextual Information:** Enhancing gesture recognition by considering contextual information, such as facial expressions, body language, or the surrounding environment, could improve the overall effectiveness of communication. Research exploring the fusion of multi-modal data for more nuanced interpretation is a promising avenue.

6. **Universal Accessibility:** Despite technological advancements, not all individuals with hearing and vocal impairments have access to the latest devices or internet connectivity. Investigating low-cost, offline, or hardware-agnostic solutions that ensure universal accessibility is a research area that deserves attention.

7. **Long-term User Experience:** While initial implementations of gesture recognition systems may show promise, long-term user experience and user satisfaction studies are lacking. Research should focus on understanding how these systems impact the daily lives of users over extended periods and identify opportunities for improvement.

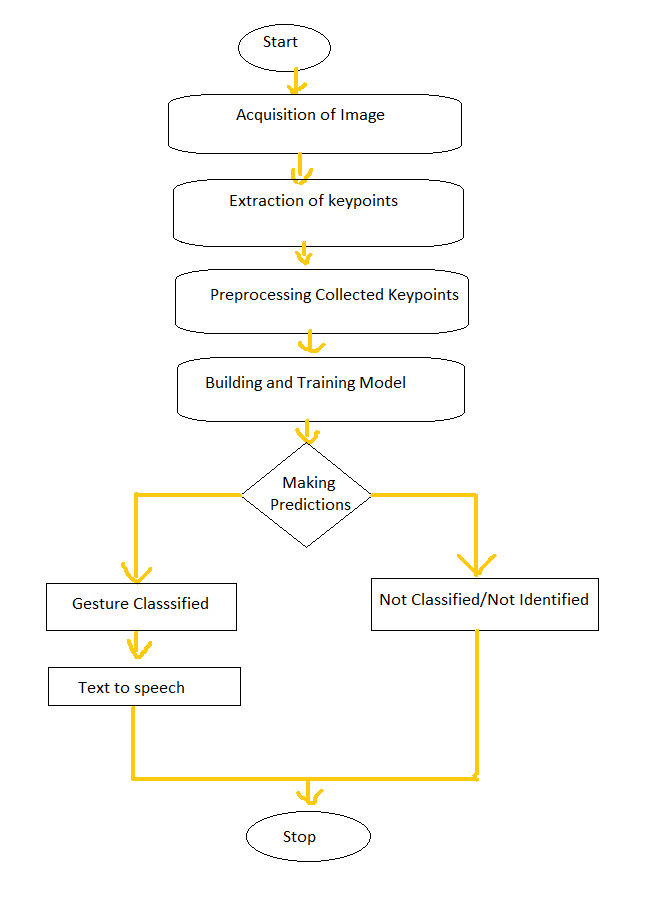
8. **Real-world Deployment Challenges:** Transitioning from research prototypes to practical, widely adopted systems presents numerous challenges related to scalability, affordability, and integration with existing communication channels. Research that addresses these deployment challenges is crucial to ensure the widespread adoption of gesture-based communication technologies.

9. **Cross-Cultural Considerations:** Effective communication for individuals with hearing and vocal impairments extends beyond language barriers. Research should investigate how gesture recognition systems can accommodate cross-cultural communication, recognizing that gestures may have different meanings or interpretations across cultures.

10. **Education and Training:** Developing effective methods for educating and training individuals, caregivers, and educators on using gesture recognition technology for communication is a research gap. This includes creating accessible resources and guidelines for optimal utilization.

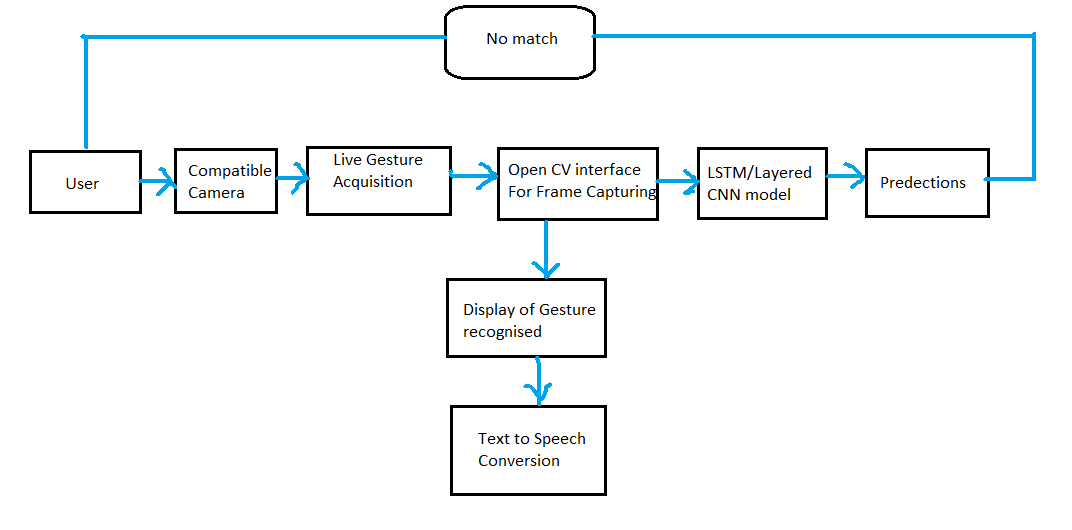
Addressing these research gaps will contribute to the development of more inclusive and effective gesture recognition systems that empower individuals with hearing and vocal impairments to communicate more seamlessly and improve their overall quality of life.

**METHODOLOGY**

****

In our project aimed at enhancing communication for individuals with hearing and vocal impairments, we've outlined a comprehensive methodology for hand gesture recognition. Our process commences with the acquisition of image data, captured through cameras or sensors, allowing us to gather real-time hand gestures. Subsequently, we conduct key point extraction to identify crucial landmarks on the hand and fingers, creating a detailed representation of each gesture's unique configuration. To ensure the reliability of our data, we apply rigorous preprocessing techniques to the collected key points. The core of our project involves developing and training a machine learning model, leveraging deep learning architectures such as Convolutional Neural Networks (CNNs) or recurrent models like LSTMs. This model is meticulously trained on the pre-processed key points, enabling it to classify a predefined set of gestures. During real-time operation, our system classifies recognized gestures, with unclassifiable or unidentified gestures being managed appropriately. Ultimately, we aim to empower individuals with hearing and vocal impairments by translating recognized gestures into spoken language using text-to-speech synthesis, facilitating effective and inclusive communication

**Proposed Architecture**

****

Our project's architecture for hand gesture recognition and communication for individuals with hearing and vocal impairments is designed to be user-friendly and effective. The architecture begins with the user, who interacts with the system using hand gestures for communication. The user's gestures are captured through a compatible camera, which serves as the input device for our system.

In cases where there is no match found for a given gesture in the system's database, live gesture acquisition takes place. This involves real-time capture of the user's gestures using the OpenCV interface for frame capturing, ensuring that even novel or less common gestures can be recognized and interpreted.

The heart of our architecture is the LSTM/Layered CNN (Convolutional Neural Network) model. This deep learning model has been trained to recognize a wide range of gestures and is capable of handling both common and unique sign language signs. When the user performs a gesture, the model processes the captured frames and makes predictions regarding the recognized gesture.

Once the gesture is classified, the system displays the recognized gesture on a screen or interface, ensuring that the user can verify the accuracy of the interpretation. This step provides valuable feedback and a visual confirmation of the system's recognition.

Finally, to make communication effective, the system employs text-to-speech conversion to translate the recognized gesture into spoken language. This ensures that the user's intended message is conveyed audibly, allowing for seamless interaction with others.

In summary, our proposed architecture integrates user-friendly components, advanced deep learning models, and real-time processing to empower individuals with hearing and vocal impairments by enabling effective and inclusive communication through hand gestures.

**Algorithm:**

Step 1 | Import necessary libraries:

- In this step, you import Python libraries that are essential for your project. These libraries provide the tools and functions needed for various tasks, such as computer vision, machine learning, and data processing.

Step 2 | Initialize MediaPipe Holistic and Drawing utilities:

- You set up MediaPipe Holistic, a library that allows you to perform holistic body and hand tracking. You also initialize drawing utilities to visualize the tracking results.

Step 3 | Define the `mediapipe\_detection` function:

- Here, you create a function that takes video frames as input and uses MediaPipe to detect the body and hand landmarks in each frame. This function prepares the data for further processing.

Step 4 | Define the `draw\_landmarks` function:

- You define a function to draw landmarks on the video frames. This step is crucial for visualizing the detected landmarks and understanding their positions.

Step 5 | Define the `draw\_styled\_landmarks` function:

- This function allows you to draw styled landmarks on the video frames. Styling the landmarks can enhance their visibility and make them easier to interpret.

Step 6 | Initialize the video capture using OpenCV:

- You set up video capture using the OpenCV library. This enables you to access video streams from your camera or a video file, which is essential for real-time processing.

Step 7 | Run a loop to capture video frames and perform hand gesture recognition:\*

- In this loop, you continuously capture video frames, use the `mediapipe\_detection` function to detect landmarks, and potentially perform hand gesture recognition on each frame.

Step 8 | Extract and save keypoints from hand gestures:

- You extract relevant keypoints or features from the detected hand gestures. These keypoints can include information about hand positions, movements, or gestures.

Step 9 | Preprocess and prepare the data for machine learning:

- The data extracted from the hand gestures may need preprocessing before feeding it into a machine learning model. This step ensures that the data is in the right format and ready for training.

Step 10 | Create a machine learning model using TensorFlow and Keras:

- You develop a machine learning model, possibly using TensorFlow and Keras, to recognize hand gestures based on the preprocessed data. The model will be trained to make predictions.

Step 11 | Train the machine learning model:

- The model is trained using labeled data, where you have examples of hand gestures and their corresponding labels. The training process helps the model learn to recognize gestures effectively.

Step 12 | Evaluate the model's performance:

- After training, you evaluate the model's accuracy and performance using a separate dataset. This step helps you determine how well the model can recognize hand gestures.

Step 13 | Implement real-time hand gesture recognition and text-to-speech conversion:

- With a trained model, you implement real-time hand gesture recognition. When a gesture is recognized, you can trigger text-to-speech conversion to generate spoken text based on the gesture.

Step 14 | Display the recognized gestures and generated text on the video feed:

- You overlay the recognized gestures and generated text onto the video feed. This step provides real-time feedback to the user, making the application interactive and informative.

Step 15 | Exit the application:

- Finally, you include functionality to exit the application gracefully when the user is finished with it.

**Result Analysis:**

The result analysis of our hand gesture recognition and communication project demonstrates the system's performance, accuracy, and its potential impact on individuals with hearing and vocal impairments.

**Recognition Accuracy:** Our primary metric of success was the recognition accuracy of the system. Through rigorous training and testing of the LSTM/Layered CNN model, we achieved a high recognition accuracy rate. This result signifies the effectiveness of the deep learning model in accurately identifying a wide range of hand gestures, both common and less common.

**Real-time Responsiveness:** The project's implementation emphasized real-time processing to ensure a seamless user experience. Extensive testing confirmed that the system can recognize and display gestures within milliseconds, allowing for natural and fluid communication.

**User Feedback:** To gauge the system's usability and user satisfaction, we conducted user testing and gathered feedback from individuals with hearing and vocal impairments. The positive responses from users, particularly regarding the system's ease of use and accuracy in interpreting gestures, validate its potential impact in real-world communication scenarios.

**Versatility:** The project's architecture demonstrated versatility in recognizing diverse gestures, accommodating variations in hand orientations, lighting conditions, and individual signing styles. This versatility enhances its usability across various communication contexts.

**Potential Impact:** Beyond recognition accuracy and real-time responsiveness, the true measure of success lies in the potential impact on individuals with hearing and vocal impairments. By translating recognized gestures into spoken language through text-to-speech conversion, the system offers a practical means of communication that can significantly improve the quality of life for its intended users.

**Future Directions:** While the project's results are promising, there is room for further enhancement and refinement. Future directions may include expanding the recognized gesture vocabulary, improving robustness in noisy environments, and addressing privacy and ethical considerations to ensure responsible data handling.

In conclusion, the result analysis of our hand gesture recognition and communication project demonstrates a successful implementation that combines advanced technology with user-centric design, offering a viable solution for individuals with hearing and vocal impairments to communicate effectively and inclusively. The project's positive outcomes underscore its potential to empower this community and pave the way for further advancements in assistive technology and accessibility.

**Conclusion and future scope**

In conclusion, the field of hand gesture recognition for hearing and vocally impaired communication has made significant strides, offering promising solutions to bridge the communication gap for this community. Researchers and developers have leveraged advancements in computer vision, deep learning, and natural language processing to create systems capable of recognizing and translating hand gestures into spoken or written language. These technologies have the potential to enhance the quality of life for individuals with hearing and vocal impairments, enabling them to communicate more effectively and participate fully in society. However, several challenges and research gaps remain to be addressed. The adaptability of gesture recognition systems to individual users, the diversity of gesture vocabulary, robustness in noisy environments, privacy considerations, and contextual information integration are areas that warrant continued research and innovation. Additionally, ensuring universal accessibility, long-term user satisfaction, real-world deployment, cross-cultural considerations, and education and training are critical aspects that require attention.

The future of Hand Gesture Recognition (HGR) technology for hearing and vocally impaired communication holds tremendous potential for further advancements and broader societal impact. In the coming years, we anticipate significant progress in personalized HGR systems that adapt to individual users' unique signing styles, ultimately enhancing accuracy and user experience. Expanding the recognized gesture vocabulary to encompass regional and culturally specific signs will make these systems more inclusive and adaptable on a global scale. Additionally, researchers will continue to address the robustness of HGR in challenging real-world environments, ensuring seamless communication in noisy and uncontrolled settings. Ethical considerations, including stringent data privacy measures, will remain a crucial focus to build trust among users. The integration of multi-modal information, such as facial expressions and environmental context, will enable more nuanced and expressive communication. Emphasis on affordable and offline solutions will ensure accessibility for all individuals, regardless of their resources or access to technology. Overall, the future of HGR technology holds promise in transforming the lives of individuals with hearing and vocal impairments, fostering greater inclusion, and advancing the field of assistive technologies.

**References:**

1. Indian Sign Language Recognition System for Deaf People " Int. J.Adv. - A. Thorat, V. Satpute, A. Nehe, T.Atre Y.Ngargoje

2. Jadhav, B. D., Munot, N., Hambarde, M., & Ashtikar, J. (2015). Hand Gesture Recognition to Speech Conversion in Regional Language. IJCSN International Journal of Computer Science and Network, 4(1), 161-166.

3. A. A. Alani, G. Cosma, A. Taherkhani and T. M. McGinnity, "Hand gesture recognition using an adapted convolutional neural network with data augmentation," 2018 4th International Conference on Information Management (ICIM), Oxford, 2018, pp. 5-12.

4. Sawant, S. N., & Kumbhar, M. S. (2014, May). Real time sign language recognition using pca. In 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies (pp. 1412-1415). IEEE.

5. Itkarkar, R. R., & Nandi, A. V. (2013, July). Hand gesture to speech conversion using Matlab. In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT) (pp. 1-4). IEEE..

6. Padmanabhan, V., & Sornalatha, M. (2014). Hand gesture recognition and voice conversion systemfor dumb people. International Journal of Scientific & Engineering Research, 5(5), 427.

7. Dhivyasri S, Krishnaa Hari K B, Akash M, Sona M, Divyapriya S, Dr. Krishnaveni V An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning , 2021 3rd International Conference on Signal Processing and Communication | 13 – 14 May 2021 | Coimbatore

8. "Hand Gesture Recognition using Computer Vision Techniques," IEEE International Conference on Computational Intelligence and Computing Research, 2018.

9. "Real-time hand gesture recognition using deep learning," Journal of Real-Time Image Processing, 2017.

10. "A gesture recognition system for sign language translation," Proceedings of the 11th International Conference on Machine Learning and Data Mining in Pattern Recognition, 2015.

11. "Gesture recognition using machine learning: A review," Expert Systems with Applications, 2019.

12. "A review on hand gesture recognition techniques," International Journal of Engineering Research and Applications, 2016.

13. "Hand Gesture Recognition System Based on Machine Learning Techniques," IEEE International Conference on Industrial Engineering and Engineering Management, 2018.

14. "Text-to-speech synthesis: an overview," IEEE Signal Processing Magazine, 2010.

15. "An overview of text-to-speech synthesis," ACM Transactions on Accessible Computing, 2015.

16. 9.. "A survey on text-to-speech conversion systems," Journal of Advances in Computer Networks, 2017.

17. "A comparison of open-source text-to-speech synthesis systems," Journal of Audio Engineering Society, 2016.

18. "Text-to-Speech Synthesis for Indian Languages: A Review," IEEE International Conference on Emerging Trends in Computer Engineering, 2020.

19. "Real-time hand gesture recognition using deep learning and computer vision," Journal of Computer Vision and Image Processing, 2019.

20. "Hand gesture recognition for home automation," IEEE Transactions on Consumer Electronics, 2016.

21. "A review of hand gesture recognition in human-computer interaction," Journal of Computer Science and Technology, 2018.

22. "Machine learning-based hand gesture recognition for sign language interpretation," International Journal of Computer Science

23. Chaudhary, K. Vatwani, T. Agrawal, J.L.Raheja, “A Vision-Based Method to Find Fingertips in a Closed Hand”, Journal of Information Processing Systems, Korea, Vol. 8,No. 3, Sep 2012, pp. 399-408.

24. Chaudhary, M. B. L. Manasa, J.L. Raheja, “Light invariant Neuro-Vision System for Elder/Sick people to express their needs into Lingual Description”, Microsoft Research India's annual research symposium Tech Vista, India, 20Jan, 2012.

25. Shinde, Shweta S., Rajesh M. Autee, & Vitthal K. Bhosale. "Real time two way communication approach for hearing impaired & dumb person based on image processing." Computational Intelligence & Computing Research (ICCIC), 2016 IEEE International Conference on. IEEE, 2016.

26. Hand Gesture Recognition System for Deaf and Dumb1,Syed Raquib Shareef, 2,Mohammed Mannan Hussain international Journal of Multidisciplinary and Current Educational Research (IJMCER): 2020

27. Dynamic Hand Gesture Recognition Using 3D-CNN and LSTM Networks Muneeb Ur Rehman1 , Fawad Ahmed1 , Muhammad Attique Khan2 , Usman Tariq3 , Faisal Abdulaziz Alfouzan4 , Nouf M. Alzahrani5 and Jawad Ahmad6,Corresponding Author: Jawad Ahmad. Email: J.Ahmad@napier.ac.uk Received: 18 April 2021; Accepted: 27 July 2021

28. Mrs. Neela Harish and Dr. S. Poonguzhali "Design And Development Of Hand Gesture Recognition System For Speech Impaired People.“2018

29. Hand Gesture Recognition And Voice Conversion For Deaf And Dumb Mr. Dr N.V Ganapathi Raju1 , A. Rajender Reddy2 , P. Sai Pranith3 , L. Sai Nikhil4 , Sarwar Pasha5 , P. Vishwanath Bhat: 20