RESEARCH ARTICLE OPEN ACCESS

Hand Gesture Recognition and Voice Conversion For Deaf And Dumb

Jiya Elsa Job [1], Nayina Elsa Aniyan [2], Meera J [3], Mia Dhas[4], Sandeep Chandran[5]

[1], [2], [3], [4] U.G Scholars, Department of Computer Science and Engineering, LBSITW, Trivandrum

[5] Assistant Professor, Department of Computer Science and Engineering, LBSITW, Trivandrum

ABSTRACT

The World Health Organisation (WHO) estimates that 466 million individuals worldwide, or 5% of the total population, are deaf, mute, or suffer from severe hearing loss. There is often a wall of distinction between handicapped people and normal people. We communicate to share our thoughts but for a disabled person (mainly deaf and dumb), it becomes difficult to communicate. In order to facilitate communication for individuals with disabilities in hospitals, this study offers a system that not only automatically detects gestures but also converts them into speech. A camera attached to the computer will capture images of the hand and the contour feature extraction is used to recognize the hand gestures of the person. Based on the recognized gestures, the generated soundtrack will be played. The system is implemented using the concepts of image processing and deep learning. Keras deep learning framework and Long Short-Term Memory (LSTM) networks are used to create a gesture recognition model. On the whole, the solution aims to provide aid to those in need thus ensuring social relevance. The case study will take place at hospitals and clinics. Without going into great detail, this clearly communicates the unique focus of this study. The user-friendly nature of the system ensures that people can use it without any difficulty and complexity. The application is cost-efficient and eliminates the usage of expensive technology.

Keywords: - Deep Learning, image processing, Keras framework, LSTM

I. INTRODUCTION

This study introduces the "Hand Gesture Recognition and Voice Conversion for Deaf and Dumb", aiming to create a robust hand gesture recognition system capable of interpreting various sign language gestures in real-time, specifically within healthcare settings. The proposed real-time translation of sign language into speech has the potential to significantly improve healthcare accessibility for individuals who are deaf or mute. In addition to improving healthcare accessibility, the proposed technology has the potential to enhance patient autonomy and independence. By providing individuals who are deaf or mute with the means to express themselves effectively in medical settings, the system empowers them to actively participate in discussions about their health, treatment options, and care plans. Ultimately, the implementation of this technology can foster a more inclusive healthcare environment where individuals with hearing or speech impairments receive equitable access to healthcare services and experience improved health outcomes. The "Hand Gesture Recognition and Voice Conversion for Deaf and Dumb" project exemplifies a significant step towards achieving healthcare equity and inclusivity for all.

II. RELATED WORKS

[1] Mopidevi, Suneetha & Biradhar, Shivananda & Bobberla, Neha & Buddati, Kiran. (2023). Hand gesture recognition and voice conversion for deaf and Dumb. E3S Web of Conferences 4th International Conference on Design and Manufacturing Aspects for Sustainable Energy (ICMED-ICMPC 2023) .

A hand gesture recognition model designed for real-time applications. By utilizing a webcam, the system enables hand motion detection, allowing users to interact directly through hand movements. Leveraging Machine Learning and Computer Vision techniques, the system employs Python and OpenCV to develop a hand gesture recognizer.For the detection and recognition processes, utilizes the MediaPipe and TensorFlow frameworks, respectively. MediaPipe aids in hand detection and tracking, while TensorFlow facilitates the training and deployment of the gesture recognition model. Overall, demonstrates a practical application of advanced technologies such as CNNs, Machine Learning, and Computer Vision in developing systems that enhance accessibility and interaction for users, particularly those with disabilities.

[2] S. Gupta, R. Thakur, V. Maheshwari and N. Pulgam, "Sign Language Converter Using Hand Gestures, " 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India

A system that enables paralyzed individuals or anyone to convey messages through hand gestures. The system captures hand motions via a webcam, recognizes the gestures, and maps them to corresponding English characters. Messages are delivered in both audio and text formats, with the added functionality of triggering alerts during emergencies. Webcam captures the hand gesture as an image, which undergoes segmentation and background subtraction using SURF features. Points detected are matched with a dataset, and the image with the most matched data points is selected as the output for the gesture. The output is presented in audio and text formats. The system utilizes the free TTS API to convert gesture messages to voice, achieving a quick recognition time of 3 seconds.

[3] M. Agrawal, R. Ainapure, S. Agrawal, S. Bhosale and S. Desai, "Models for Hand Gesture Recognition using Deep Learning, " 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India

A hand gesture recognition system to bridge communication between deaf/mute individuals and others. The system operates in four steps: generating a live stream of hand gestures via webcam, converting video frames into images, preprocessing these images, and recognizing sign language gestures to convert them into text or audio output. Implemented with image processing and neural networks, the system employs Convolutional Neural Networks (CNNs), particularly utilizing transfer learning with the well-known Inception V3 model, which is 48 layers deep. The models trained with Inception V3 exhibit higher accuracies due to its depth, enabling effective interpretation of sign language hand gestures from live webcam streams, facilitating communication for deaf/mute individuals.

[4] G. G. S, P. N, A. Yaji, A. M, A. M. Dsilva and C. S R, "Review on Text and Speech Conversion Techniques based on Hand Gesture, " 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India

Efforts to address the communication challenges faced by deaf-mute individuals have spurred the development of various sensor-based and vision-based solutions. Sensor-based approaches, employing technologies such as flex sensors, gyroscopes, and accelerometers, focus on creating smart gloves capable of recognizing finger movements and orientation to interpret gestures. On the other hand, vision-based techniques leverage image processing and machine learning models to capture, preprocess, and analyze images, enabling accurate gesture recognition and synthesis of text-to-speech signals. These advancements underscore a concerted effort to enhance communication accessibility for the deaf-mute community.

[5] F. Zhan, "Hand Gesture Recognition with Convolution Neural Networks, " 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI), Los Angeles, CA, USA.

The hand gesture recognition system employs 2D convolutional neural networks (CNNs) to extract hand components from images and predict gestures. To enhance generalization and mitigate overfitting, an efficient spatio-temporal data augmentation technique deforms input volumes of hand gestures. Utilizing a CNN classifier, dynamic hand gestures are accurately recognized and classified, leveraging a dataset comprising 500 images across nine distinct gestures.

III. PROBLEM DEFINITION

Communication barriers persist as a significant challenge for individuals who are deaf and dumb within hospital environments, impeding their ability to effectively express their needs and understand medical information. Despite efforts to enhance accessibility, existing solutions inadequately address the complex communication dynamics inherent in healthcare settings. Hospitals often lack dedicated communication tools tailored to the needs of individuals who are deaf and dumb, resulting in reliance on inefficient and ineffective communication methods, such as handwritten notes or lip-reading. Conventional communication methods may prove insufficient in conveying nuanced medical information, leading to misunderstandings between patients and medical professionals. This inadequacy risks compromised diagnosis, treatment, and patient outcomes. Misinterpretation of medical information due to communication barriers poses a substantial risk to patient safety. Inaccurate understanding of symptoms, treatment plans, or medication instructions can result in adverse outcomes, medical errors, or preventable complications. Individuals who are deaf and dumb may experience feelings of marginalization or isolation in hospital settings, exacerbating vulnerability and compromising their sense of dignity and privacy during medical interactions. Healthcare professionals may lack adequate training and awareness regarding effective communication strategies for patients with hearing and speech impairments. This deficiency contributes to communication breakdowns and exacerbates the challenges faced by individuals who are deaf and dumb. By prioritizing accessibility, accuracy, privacy, and ease of integration, a solution like hand gesture recognition and voice conversion can pave the way for improved communication outcomes and enhanced healthcare experiences for individuals who are deaf and dumb in hospital settings.

IV. PROPOSED SOLUTION

The initial step in creating an ML model for hospital situations is gathering a large dataset that includes important gestures used in healthcare environments. This dataset should cover a range of gestures relevant to medical procedures, patient communication, and healthcare workflow. It is necessary to properly annotate and identify each gesture in order to support supervised learning.

Once the dataset is prepared, it undergoes preprocessing to convert it into a numpy array format suitable for ML model training. Data preprocessing includes cleaning, normalization, and feature extraction to ensure the data is in a suitable format for analysis and model building.The ML model architecture is constructed using Keras with an LSTM (Long Short-Term Memory) architecture. LSTM is chosen due to its ability to capture long-term dependencies and patterns in sequence data, making it well-suited for analyzing gestures that unfold over time.

In tandem with model development, MediaPipe is integrated into the pipeline for hand gesture recognition and feature extraction. MediaPipe enables the extraction of relevant features such as hand landmarks, movements, and positions, which are crucial for gesture recognition.The model is trained using the prepared dataset and the extracted features from MediaPipe. The dataset is split into training and validation sets to assess the model's performance accurately. The trained ML model is then deployed in a hospital environment, integrated into a user-friendly interface for real-time gesture recognition. This deployment enables healthcare professionals to utilize gesture-based interactions for improved communication and workflow efficiency.

V. CONCLUSION

The utilization of hand gestures as a means of interaction in hospitals can streamline various processes, from medical procedures to patient communication. Moreover, the integration of gesture recognition technology aligns with the broader trend of incorporating AI and ML solutions in healthcare to enhance patient outcomes and operational efficiency. It reflects a proactive approach towards leveraging technology for social good and addressing real-world challenges faced in medical environments.

Beyond its immediate applications in hospitals, the development of gesture recognition models also has broader implications for accessibility and inclusivity in healthcare. It can enable individuals with physical limitations or communication barriers to interact more effectively with healthcare systems, promoting equal access to healthcare services.

In general, this study is a major advancement towards employing AI and ML for social effect, particularly in the field of healthcare. By fostering innovation in gesture-based interaction, it contributes to creating more responsive, efficient, and inclusive healthcare environments for patients and healthcare professionals alike.

VI. FUTURE SCOPE

The future scope of this project is expansive and holds significant potential for further development and application in healthcare settings. One key area of focus is the continual improvement of gesture recognition accuracy. This can be achieved through the incorporation of more diverse datasets, advanced feature extraction algorithms, and ongoing fine-tuning of model parameters. By enhancing accuracy, the ML model can better interpret and respond to a wider range of gestures, improving overall system performance.

Moreover, gesture recognition can be leveraged for patient monitoring, rehabilitation exercises, and tracking progress in physical therapy programs. This opens up possibilities for personalized care and more effective treatment plans tailored to individual patient needs. The integration of gesture recognition capabilities with Electronic Health Record (EHR) systems holds promise for streamlining documentation, data entry, and workflow automation in healthcare settings. This integration can improve accuracy, reduce administrative burden, and enhance overall productivity within healthcare facilities.

In addition to the aforementioned avenues, another innovative direction for the project's future scope involves the development of a desktop robot to mitigate the discomfort caused by continuous camera surveillance. This desktop robot could serve as an intermediary interface between users and the gesture recognition system, providing a physical embodiment of the technology while also addressing privacy and comfort concerns.

REFERENCES

[1] Mopidevi, Suneetha & Biradhar, Shivananda & Bobberla, Neha & Buddati, Kiran. (2023). Hand gesture recognition and voice conversion for deaf and Dumb. E3S Web of Conferences. 391. 10.1051/e3sconf/202339101060.

[2] S. Gupta, R. Thakur, V. Maheshwari and N. Pulgam, "Sign Language Converter Using Hand Gestures," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp. 251-256, doi: 10.1109/ICISS49785.2020.9315964.

[3] M. Agrawal, R. Ainapure, S. Agrawal, S. Bhosale and S. Desai, "Models for Hand Gesture Recognition using Deep Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2020, pp. 589-594, doi: 10.1109/ICCCA49541.2020.9250846.

[4] G. G. S, P. N, A. Yaji, A. M, A. M. Dsilva and C. S R, "Review on Text and Speech Conversion Techniques based on Hand Gesture," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 1682-1689, doi: 10.1109/ICICCS51141.2021.9432277.

[5] F. Zhan, "Hand Gesture Recognition with Convolution Neural Networks," 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI), Los Angeles, CA, USA, 2019, pp. 295-298, doi: 10.1109/IRI.2019.00054.*FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[6] "HAND GESTURE RECOGNITION AND VOICE CONVERSION SYSTEM FOR DUMB AND DEAF PEOPLE", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.6, Issue 4, page no.152-158, April-2019

[7] F. Shaikh, S. Darunde, N. Wahie and S. Mali, "Sign Language Translation System for Railway Station Announcements," 2019 IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India, 2019, pp. 1-6, doi: 10.1109/IBSSC47189.2019.8973041.

[8] G. Devineau, F. Moutarde, W. Xi and J. Yang, "Deep Learning for Hand Gesture Recognition on Skeletal Data," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018.

[9] Hamed Al-Saedi A.K, Hassin Al-Asadi A, Survey of Hand Gesture Recognition System. IOP Conferences Series: Journal of Physics: Conferences Series 1294 042003. 2019. DOI: https://doi.org/10.1088/1742-6596/4/042003.

[10] S. Jothimani, S. Shruthi, E.D. Tharzanya, S. Hemalatha. “Sign and Machine Language Recognition for PhysicallyImpaired Individuals”, 2022 3rd International Conference on Electron ics and Sustainable Communication System (ICESC), 2022.