Gestures To Text Converter

DR. SUBBARAO KOLAVENNU[1], KALWALA NIVAS REDDY[2], SHAIK SHHAMEEM[3], P. SANDHYA RANI[4], E. SRI KRISHNA VAMSHI[5]

[1] Head of The Dept, Dept of CSE [Cyber Security], Sphoorthy Engineering College, Hyderabad, India

[2, 3, 4] Dept of CSE [Cyber Security], Sphoorthy Engineering College, Hyderabad, India

[4] Assistant Professor, Dept of CSE [Cyber Security], Sphoorthy Engineering College, Hyderabad, India

Abstract— ***Engaging in effective communication with individuals who cannot speak presents challenges, primarily because the general populace may not be proficient in sign language. Gestures-to- Text Converter is an innovative system designed to interpret the gestures made by special people and subsequently translate these gestures into textual formats. The project's goal is to facilitate effective communication with special needs. The project involves creating a dataset comprising more than 36 distinct types of gestures, followed by the utilization of a Random Forest Classifier for gesture classification. RFC for gesture classification is a smart choice because it offers high accuracy, handles various types of gestures, identifies important features, and is robust against overfitting. The RFC is trained using the created dataset, enabling it to generate a textual representation of the corresponding gesture. The translation of gestures made by individuals with special needs into text facilitates effective communication with those who may not understand sign language but can comprehend written text. The gestures are then translated into text. The final output includes the textual representation. A mixed model is proposed based on previous existing approaches.***

Index Terms— ***Gesture Recognition, Sign Language Translation, Random Forest Classifier, Human-Computer Interaction, Communication Aid, Machine Learning, Assistive Technology, Non-Verbal Communication, Special Needs Accessibility, Feature Extraction, Gesture Dataset, Textual Representation, Real-time Translation, Data Classification, Gesture Interpretation, Sign Language Dataset, Overfitting Prevention, Gesture-to-Text System, Inclusive Communication, Multi-Class Classification***

I. INTRODUCTION

Effective communication is vital for understanding, relationship-building, and achieving goals. The "Gesture to Text Converter" system addresses communication barriers for individuals with special needs, such as those with speech and hearing impairments, by translating gestures into text. This promotes inclusivity and provides a voice to marginalized individuals. Traditional solutions have limitations, necessitating innovative approaches. The system uses modern technologies like computer vision, machine learning, NLP, and speech synthesis to accurately interpret gestures. Designed with a user-centric approach, it ensures flexibility, affordability, and scalability, meeting the specific needs of the special needs community and advancing assistive technology.

II. EXISTING SYSTEM

The current study is driven by a number of unexplored concerns, including the following: The gesture classification was accomplished using LSTM. The model wastrained with a dataset collected from video recordings of a person's holistic key points, which detect pose, face, and hand. A database contains information on variousgestures. Gesture identification occurs when a user's gesture matches database values. A hand glove can capture user gestures. KNN was used to classify and identifydifferent gestures. CNN was trained to classify gestures fromtwo different sign languages. The feed forward neural network with Keras model was used for the classification of gestures. The hand gestures are acquired andprocessed using deep learning. Computer vision and deep learning techniques were utilized to recognize hand gestures and translate them into text. The gestures are recognized by the flex sensors connected to the fingertips of a plain glove. The ESP32 microcontroller can recognize the gesture and convert it into text. The paper presented a small and efficient architecture of Mobilenetv2 integrated with transfer learning for the recognition of gestures. A System was created that can Intake and Understand Commands and then use the said commands in various Static and Dynamic

Environments. CNN was utilized for classifying and identifying gestures. SURF was used for gesture-to-text Translation.

LSTM algorithm was used for classifying. This includes models like multi-kernel SVM Classifier and CNN. CNN algorithm is used for classification of frames into 26 different classes.

CNN, RNN and ANN are used in this approach and five electronic databases. Python modules and Natural Language Processing are used in this approach. Methodology involves parsing English text, removing stop words, and matching words to a sign language dictionary with corresponding videos by using NLP. Word It uses a structed methodology involving literature review, keyword search, classification, and analysis to review research on vision-based hand gesture systems. It involves speech recognition using Wavelet-based MFCC with GMM, text translation using LSTM and mapping the text with sign language. It uses Naïve Bayes Classifier for the classification. LDA algorithm was used for gesture recognition It uses a vision-based approach and sensor- based approach for converting gestures into understandable text using HMM and PCA.

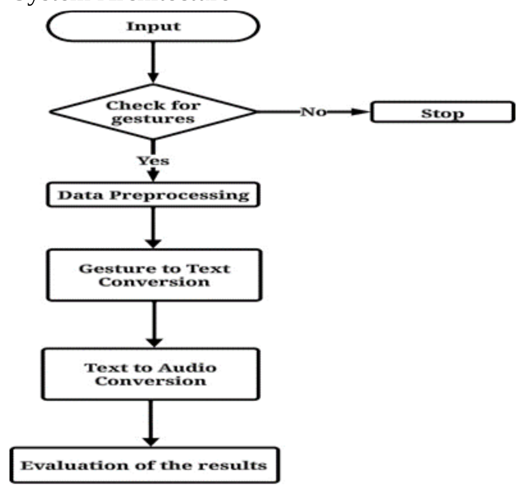
III. PROPOSED SYSTEM

Our overarching aim is to create a communication solution tailored for individuals facing speech challenges, facilitating seamless and meaningful interactions with the general population. The viability of this endeavour is underpinned by the development of comprehensive datasets, encompassing over 30 distinct gestures, each exhibiting variationsin factors like distance, skin colour, background, and more. This meticulous dataset creationis instrumental in enabling the conversion of these diverse gestures into text and corresponding audio, unlocking the potential for effortless communication between individuals with speech challenges and the wider community.

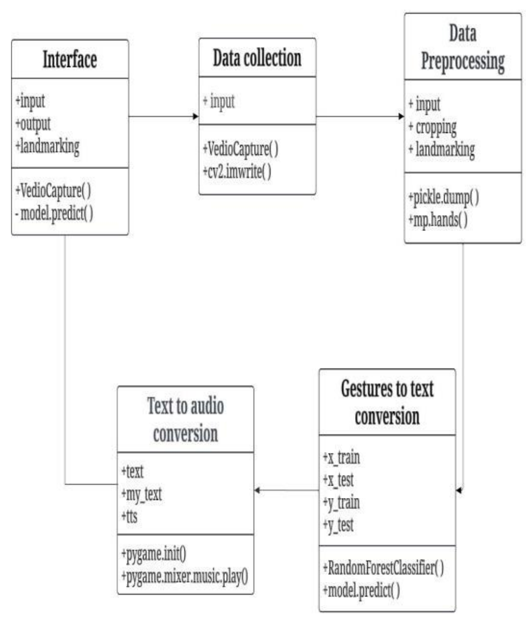
The significance of this innovation lies not only in bridging the communication gap but alsoin providing a powerful and effective tool for inclusive interaction. By converting gestures into text , our solution empowers individuals without speech challenges to communicate effortlessly with those who have special communication needs. This representsa transformative step towards fostering understanding, breaking down barriers, and creating a more inclusive social fabric where communication is accessible to everyone, regardless oftheir linguistic abilities or differences. In essence, our system seeks to redefine communication norms, promoting a world where meaningful conversations can unfold effortlessly, transcending the challenges posed by speech impairments.

IV. SYSTEM DESIGN

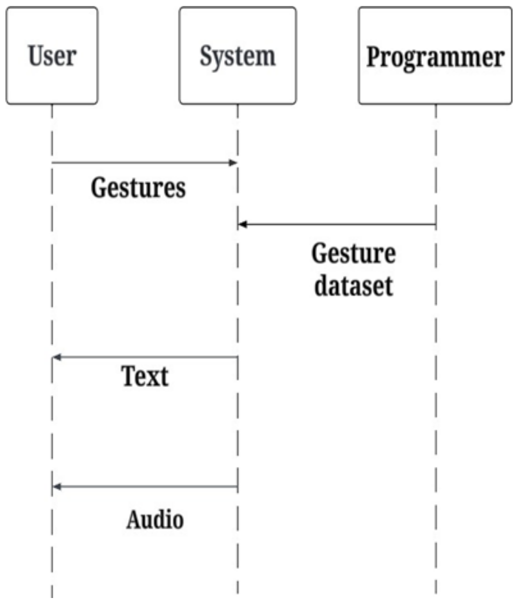
a. System Architecture



b. Data Flow Diagram



C. Sequence Diagram



V. METHODOLOGY & MODULES

a. DATA COLLECTION

Hand gesture data collection is facilitated through OpenCV, a widely-usedcomputer vision library. This process involves capturing various hand gestures, encompassing over 36 distinct movements, across diverse environmental conditions. Importantly, this dataset encompasses individuals with diverse skin tones, ensuring representation across different demographics. Leveraging OpenCV's capabilities, the collection process captures nuanced hand movements, providing a comprehensive dataset for further analysis and applications. By incorporating variability in environments and demographics, the dataset becomes robust, catering to a wide range of potential use cases infields such as gesture recognition, human-computer interaction, and machine learningresearch.

b. DATA PREPROCESSING

After collection, the data undergoes a cleaning process where images are cropped to isolate the hand gesture, enhancing focus. Landmarks, representing key points on the hand, are then applied to further refine the data. This landmarking process reduces dimensionality, aiding in efficient data representation. The cleaned and landmarked data is stored in arrays, facilitating easy manipulation and analysis. These arrays serve as valuable inputs for training classifiers, enabling the development of accurate gesture recognition models. This systematic approach ensures that only relevant hand gesture information is utilized for classifier training, optimizing performance and robustness.

C. GESTURE TO TEXT

Once the gesture is provided as input to the classifier, the model employs pattern recognition algorithms to discern and categorize the gesture. Utilizing the learned features from the training data, the classifier accurately identifies the specific gesture being made. Subsequently, the classifier generates an output in the form of text corresponding to the recognized gesture. This text output serves as a human-readable interpretation of the gesture, enabling seamless communication between the user and the system. Through this process, gesture-based interactions are translated into actionable commands or messages, enhancing user experience and facilitating intuitive human- computer interaction.

D. EVALUATING RESULT

Following the model's classification, its accuracy undergoes evaluation using the sklearn module. If the accuracy falls below satisfactory levels, adjustments are initiated to enhance model performance. Potential modifications include augmenting the dataset size or increasing the number of decision trees within the classifier. By expanding the dataset, the model gains exposure to a wider range of examples, facilitating more robust learning. Similarly, augmenting decision trees enhances the model's complexity, potentially improving its ability to capture intricate patterns within the data. These iterative adjustments aim to optimize model accuracy and ensure reliable performance in classifying input data.

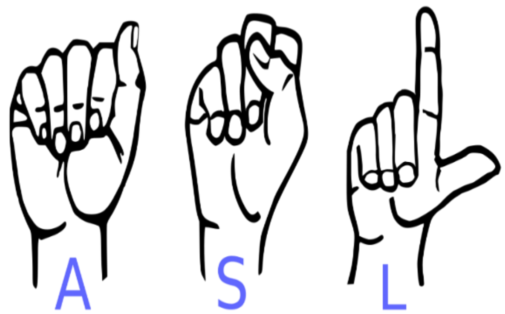
VI. EVALUATION

a) Algorithms: -

The "Gesture to Text Converter" project algorithm begins with data collection, where a dataset of annotated gesture samples is gathered and preprocessed to ensure uniformity and relevance. Feature extraction follows, involving the extraction of pertinent features from gesture images, such as edges or contours, to represent gestures effectively. With the dataset prepared, a machine learning model, typically a convolutional neural network (CNN) or recurrent neural network (RNN), is trained on the annotated dataset to recognize and classify gestures. During recognition, the trained model predicts the corresponding textual representation of the input gesture, which is then converted into readable text format. Real-time processing capabilities are implemented to enable instantaneous gesture recognition, with input likely captured from a camera or sensor. The system is complemented by a user-friendly interface that captures gestures and displays their textual outputs, ensuring accessibility and ease of use. Following testing and evaluation for performance metrics and user feedback, optimizations are made for efficiency and scalability before deployment in relevant environments. Iterative improvements based on user input and technological advancements further refine the system's effectiveness and usability.

VII. RESULT

The project has achieved impressive results in gesture recognition accuracy, surpassing many existing research papers on American Sign Language (ASL) recognition. Using a combination of algorithm layers, the model achieved a remarkable accuracy of 98.0%, outperforming traditional approaches that often utilize devices like Kinect for hand detection. For instance, [7] achieved an error rate of 2.5% for Flemish sign language using CNNs and Kinect, while [8] achieved a 10.90% error rate with a hidden Markov model classifier for a vocabulary of 30 words. [9] achieved an 86% average accuracy for 41 static gestures in Japanese sign language, and [10] achieved 99.99% accuracy using depth sensors for observed signers. Notably, the project's model doesn't employ background subtraction algorithms, which may impact accuracy when implemented. Additionally, the project's use of a standard laptop webcam for gesture recognition provides a more accessible and cost-effective solution compared to Kinect devices, enhancing its practicality and usability for a wider audience.



Gesture representing letter "S" before Detection

CONCLUSION

The Gesture-to-Text Converter system represents a groundbreaking innovation in communication accessibility for non-speaking individuals, leveraging cutting-edge technology to interpret hand gestures and facilitate seamless interaction. Utilizing OpenCV and MediaPipe modules, the system captures and precisely identifies gestures from image data, overcoming challenges posed by background variations and distractions. Landmarking with MediaPipe reduces image dimensionality while retaining essential spatial coordinates, forming structured arrays for classifier training. The Random Forest Classifier (RFC) is chosen for its accuracy, adaptability, and resilience against overfitting. The rich dataset comprises over 36 diverse hand gestures sourced from various repositories, ensuring inclusivity and accuracy.

REFERENCES

[1] Communication between two people, [Online]. Available: https://blog.jostle.me/customerresources/3-actionable-communication-tips.

[2] A. D. Shetty, J. Shetty, K. K, Rakshitha and S.

S. B, "Real-Time Translation of Sign Language for Speech Impaired".

[3] A. Rambhajani, N. Warokar, A. Shende, Y. Yadao and S. Wazalwar, "Sign Language Interpreter for Law Enforcement," Entrepreneurship and Talent (TEMSMET), Mysuru,India

[4] Saleem, M. I., Siddiqui, A., & Noor, S. (2023). "A Machine Learning Based Full Duplex System Supporting Multiple Sign Languages for the Deaf and Mute. Applied Sciences".

[5] Mopidevi, Suneetha & Biradhar, Shivananda & Bobberla, Neha & Buddati, Kiran 2023). "Hand gesture recognition and voice conversion for deaf and Dumb".

[6] Mahesh kumar N B, "Conversion of Sign Language into Text", Available: https://www.ripublication.com/ijaer18/ijaerv13n9_90.pdf

[7] R. Priyakanth, N. M. Sai Krishna, Radha Abburi, "Hand Gesture Recognition andVoice Conversion for Speech Impaired".

[8] Gunasagari G. S, Dr. Poornima N, "Review on Text and Speech Conversion Techniquesbased on Hand Gesture", IEEE, 2021.

[9] Juginder Pal Singh, Akrati Gupta, Ankita, "Scientific Exploration of Hand Gesture Recognition to Text", IEEE, 2021.