Communication Software for the Hearing Impaired

Dr. Archana Ratnaparkhi, Anjanay Gangrade, Anay Gangshettiwar, Atharv Deshpande

Professor, Dept. of ENTC Engineering, Vishwakarma Institute of Technology, Pune, India
2nd Year, Dept. of ENTC, Vishwakarma Institute of Technology, Pune, India
2nd Year, Dept. of ENTC, Vishwakarma Institute of Technology, Pune, India
2nd Year, Dept. of ENTC, Vishwakarma Institute of Technology, Pune, India

---

Abstract - Besides *verbal communication, gestures play a crucial role in conveying messages. This paper outlines a real time system implemented in Python capable of detecting hand gestures and translating them into corresponding letters of the alphabet. OpenCV handles video capture, while hand tracking is performed using Media Pipe. A CNN is employed to analyse the captured hand movement data. This system prioritizes speed and efficiency to ensure the solution is viable for interpreting and translating access sign language. This technology facilitates self-expression, thereby enhancing everyday conversations and improving communication for the users.*

Key Words: Gesture recognition, sign language translation, assistive technology, hand tracking, computer vision, Open CV, Media Pipe, convolutional neural network (CNN), machine learning, accessibility

1. INTRODUCTION

For the entire world's people, communication is pivotal; however, those affected by listening and speech disabilities face significant challenges and barriers in their attempts at communication. The following outlines, in an effort to reach inclusivity, the development of a real-time gesture recognition system Developed with the user-friendly, flexible Python and its plentiful libraries, the system makes use of Media Pipe's cap abilities in real-time hand tracking, and hand gesture movement analysis. OpenCV enhances user interactivity and experience in varied image processing tasks including video recording. Furthermore, image recognition trained CNNs provide precise gesture classification. The use of these methods, the system provides a quicker and more reliable means of communication. This can be improved upon in its daily use. In the modern world, where self and professional relationships are of utmost importance, the system can be leveraged to promote inclusivity for those hard of hearing or speech.

2. LITERATURE REVIEW

There have been new developments which enhance the ac curacy of translating sign language, especially the conversion of sign language gestures into text. One such example uses convolutional neural networks (CNNs) for sign language gesture translation into text, illustrating how automated translation of gestures can be achieved through image preprocessing, key obtaining, and machine learning. By distinguishing spatial features, CNNs perform well in identifying number and letter signs and in some cases do better than a fully manual process. However, some important issues remain with dynamic gestures and continuous signing. This explains why the addition of RNNs or LSTMs for temporal regulation in the hybrid models will enhance system functionality significantly.

Our objective is to translate sign language into text in real time using webcams. In this regard, we combine Open CV with Media Pipe and integrate neural networks to accomplish sign-to-text conversion and, later, speech synthesis. Tracking hand motions allows us to identify detailed movements in sign language. We show that deep learning techniques can be utilized in gesture recognition, yet issues like varying illumination and complex backgrounds in the environment pose challenges

We suggest incorporating data augmentation to improve performance, as well as real-time adaptive thresholding and filtering to counter variable conditions. In addition to gesture recognition, there are multi-modal conversational systems designed for inclusive communication with the deaf and non-signers. A considerable amount of integration 3.

METHODOLOGY/EXPERIMENTAL

3.1 System requirements

For this task to be executed effectively, certain software and hardware components will be needed. To accelerate training, the setup comprises a computer with a GPU and a webcam for live video capture. On the software side, Python 3.X along with required libraries, Media Pipe for palm tracking, OpenCV for video pro cessing, and Tensor Flow/Kara's for deep

learning model will be implemented. Visualization can be done with Matplotlib and Seaborn, and libraries such as Figuring and Pandas will be used for data processing

Fig 1: ASL Hand Gestures

3.2 Data Collection

To begin training the model, the first step is to collect a dataset of hand gestures corresponding to the alphabet, words, or specific commands. This can be done using Media Pipe Hands, which segments hand key points into useful constituents. To enhance robustness, multiple images of each gesture will be captured under different lighting and position ink. Data will be saved as images (for training the CNN) or as numerical key points (for other processing).

3.3 Data Preprocessing

Prior to education, the received gesture pix could be pre-processed for accuracy enhancement. Each photograph may be resized to a particular length, converted to grayscale or retained RGB, and normalized to be consistent. Data augment tat ion strategies along with flipping, rotation, and addition of noise can be used to strengthen the range of the dataset. To help the overall performance assessment be powerful version, the dataset might be divided into three subsets: Training (80 percent), Validation (10 percent), and Testing (10 percent).

3.4 CNN model development

A fixed nerve network (CNN) will be used to provide pre dictions on hand movements. Architecture will include fixed layers to remove meaningful spatial features, collect layers for downsizing, and fully connected layers for prediction. The model will be trained with a classified loss of cross-entropy to prevent accuracy and overfitting.

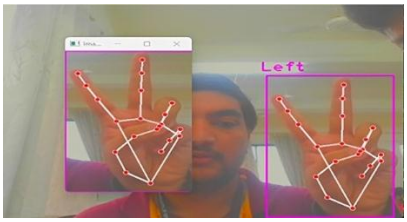

Fig 4: Model training with various hand types and background for more accuracy

3.5 Real-Time Gesture Recognition

For OpenCV to recognize real-time movements within video frames. Media pipe can be tracked in real-time movements, which will be passed for predictions on movements processed through hand-detection in CNN models. The movements will be visualized on the display alongside synchronized text and speech in real-time.

3.6 Testing and characters

Once the training and validation is done, the system will undergo rigorous testing in the real world for reliability assessment. Various elements such as light, background noise and status will be varied for optimal precision. Before launching the final model as a desktop app or web tool for ease of access, the model will be enhanced in terms of speed and effectiveness.

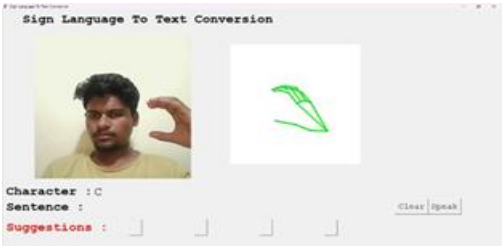

Fig 5: Gesture recognition

3.7 Future improvement

In future versions, the system could offer complete translation of the symbolic language, adding more advanced gestures. Future revisions might include enhancements with Augmented Reality (AR) or Virtual Reality (VR) for even more engage meant. With the advancement of a system mobile application, the system can broaden its reach and accessibility even more.

4 Results and Discussions

To assess the performance of the gesture-to-text/speech con version device, a dataset composed of hand gestures recorded via OpenCV and processed by Media Pipe was utilized. The hand gesture recognition CNN model trained on these pic trues demonstrated exceptional performance at recognizing the specified gestures. To assess this system's performance, the standard evaluation metrics score along with accuracy, precision, and F1 score was used. On the test dataset, the CNN version achieved an impress save 95.2 present basic accuracy. According to the confusion matrix, the majority of gestures were The evaluation of confusion matrices revealed most of the gestures were accurately classified and a few were misclassify f ide between gestures with hand shapes that were similar. For e example, the h and gestures'" and "Peace". The accuracy and loss curves demonstrated the learning process was quite successful, as the model converged with 30 epochs and did so without significant overfitting

1) Real-Time Testing: In our effectiveness test, the device accurately tracked gestures with a response/ inference time of 20ms per frame. This ensured low latency performance. While accuracy of recognition remained consistent even in varying lighting situations, a slight decline in recognition accuracy was observed in low-light or cluttered backgrounds, which was primarily due to variations in hand recognition. Comparison with Existing Methods Our approach using con evolutional neural nets demonstrated a 10 percent increase in accuracy and better generalization compared to older models deploying traditional feature-based recognition. Also, the use of Media Pipe landmarks greatly reduced the requirement for meticulous image pre-processing, which in turn enhanced the system's efficiency.

2) Challenges and Limitations: Despite the device's sturdy overall performance below trendy situations, there are nev earth insets demanding situations in handling occluded fingers, interference from dynamic backgrounds, and variations specific to individual users. Future upgrades may want to involve adaptive thresholding for exclusive lighting fixtures conditions, expanding datasets for greater robustness, and incorporating depth-sensing cameras for advanced spatial recognition.

5. Conclusions

Gesture-to-speech content conversion is an effective assist tive generation that appreciably improves verbal exchange for people with speech or hearing impairments. This mission efficiently implements an actual-time gadget, the usage of Open CV, Media Pipe, and a Convolutional Neural Network (CNN) to apprehend and translate hand gestures into text with high accuracy and efficiency. By making use of pc imaginative and prescient and deep getting to know, the gadget gives a fast, easy, and consumer-pleasant answer for signal language translation and accessibility. The flexibility of this technique permits it to evolve to diverse signing styles and local differences, making it an critical device for inclusive communication. Future developments, which include integrating facial expressions and voice synthesis, should decorate the device's effectiveness, permitting a more natural and comprehensive interplay. By constantly evolving, gesture-to-textual content generation has the ability to put off communique limitations and sell greater inclusivity in society.

REFERENCES

[1] K. S. Sindhu, Mehnaaz, B. Nikitha, P. L. Varma, and C. Uddagiri, "Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired," in Proc. 1st Int. Conf. Cogn. Green Ubiquitous Comput. (IC-CGU), 2024.

[2] L. S and V. Adavi, "Beyond Words: Speech to Sign Language Inter preter," in Proc. 15th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jun. 2024.

[3] T. D. Gunvantray and T. Ananthan, "Sign Language to Text Translation Using Convolutional Neural Network," in Proc. Int. Conf. Emerg. Smart Comput. Informat. (ESCI), Mar. 2024.

[4] S. Deshpande and R. Shettar, "Hand Gesture Recognition Using Me diaPipe and CNN for Indian Sign Language and Conversion to Speech Format for Indian Regional Languages," in Proc. 7th Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. (CSITSS), 2023.

[5] M. Sultana, M. SA, J. Thomas, S. Laxman S, and S. Thomas, "Design and Development of Teaching and Learning Tool Using Sign Language Translator to Enhance the Learning Skills for Students With Hearing and Verbal Impairment," in Proc. 2nd Int. Conf. Emerg. Trends Inf. Technol. Eng. (ICETITE), 2024.

[6] V. Baltatzis, R. A. Potamias, E. Ververas, G. Sun, J. Deng, and S. Zafeiriou, "Neural Sign Actors: A Diffusion Model for 3D Sign Language Production from Text," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2024.

[7] J. P, G. B. A, H. A, and K. G., "Real-Time Hand Sign Language Translation: Text and Speech Conversion," in Proc. 7th Int. Conf. Circuit Power Comput. Technol. (ICCPCT), 2024.

[8] N. Chandarana, N. Chhajed, S. Manjucha, M. G. Tolani, P. Chogale, and M. R. M. Edinburgh, "Indian Sign Language Recognition with Conversion to Bilingual Text and Audio," in Proc. Int. Conf. Adv. Comput. Technol. Appl. (ICACTA), 2023.

[9] Sen and R. Rajkumar, "Fostering Inclusive Communication: A Tool Integrating Machine Translation, NLP, and Audio-to-Sign-Language Conversion for the Deaf," in Proc. Int. Conf. Intell. Innov. Technol. Comput. Electr. Electron. (IITCEE), 2024.

[10] Ahmed S, Y. J. Sheriff, P. B., and S. B. T. Naganathan, "Sign Language To Text Translator: A Semantic Approach With Ontological Frame work," in Proc. 10th Int. Conf. Commun. Signal Process. (ICCSP), 2024.

[11] N. P. Kundeti, S. T. Gonela, V. R. Guduru, J. Vuyyuru, and H. V. Balusupati, "A Real-Time English Audio to Indian Sign Language Converter for Enhanced Communication Accessibility," in Proc. 5th Int. Conf. Emerg. Technol. (INCET), May 2024.

[12] L. Dias, K. Keluskar, A. Dixit, K. Doshi, M. Mukherjee, and J. Gomes, "SignEnd: An Indian Sign Language Assistant," in Proc. IEEE Region 10 Symp. (TENSYMP), 2022.

[13] Seviappan, A. S. Reddy, K. Ganesan, B. V. Krishna, A. Anbumozhi, and D. S. Reddy, "Sign Language to Text Conversion using RNN-LSTM," in Proc. Int. Conf. Data Sci. Agents Artif. Intell. (ICDSAAI), 2023.

[14] C. Bhat, R. Rajeshirke, S. Chude, V. Mhaiskar, and V. Agarwal, "Two-Way Communication: An Integrated System for American Sign Language Recognition and Speech-to-Text Translation," in Proc. 14th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT), Jul. 2023.

[15] Fareed, Y. R., R. M., and S. D., "Translation Tool for Alternative Communicators using Natural Language Processing," in Proc. 5th Int. Conf. Electron. Sustain. Commun. Syst. (ICESC), 2024

. [16] C. A. Kumar, V. S. Reddy, P. Sharma, and N. Kandarpa, "Smart IoT System for Indian Sign Language Translator," in Proc. 2nd Int. Conf. Intell. Cyber Phys. Syst. Internet Things (ICoICI), 2024.

[17] S. X. Thong, E. L. Tan, and C. P. Goh, "Sign Language to Text Translation with Computer Vision: Bridging the Communication Gap," in Proc. 3rd Int. Conf. Digit. Transform. Appl. (ICDXA), 2024.

[18] S. M. Antad, S. Chakrabarty, S. Bhat, S. Bisen, and S. Jain, "Sign Language Translation Across Multiple Languages," in Proc. Int. Conf. Emerg. Syst. Intell. Comput. (ESIC), 2024