GESTURE LANGUAGE TRANSLATOR

"Bridging Communication Gaps with Intelligent Gesture Language Translation"

Dr. M. AYYAPPA CHAKRAVARTHI[1], P. LIKHITHA [2], PEDDIREDDY VENKATA NAGA POOJETHA[3], YAMPARALA VASAVI[4], PAPASANI TRILOCHANA[5]

Professor & Head, Department of CSE-Data Science, KKR & KSR Institute of Technology and Sciences. [1]
BTech CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India. [2-5]

Abstract

This paper presents a Gesture Language Translator system that aims at bringing the solution to the effective communication between hearing-impaired and people who do not understand sign language. Sign language uses gestures and facial expressions, which are not general knowledge, and hence communication between hearing-impaired or deaf and mute people becomes difficult. To solve this problem, we have created a website that interprets sign language into verbal language, thus it is easy to communicate between people.

This system utilizes Computer Vision and Machine Learning (ML) algorithms to identify and interpret different sign language movements accurately. The system is designed to interpret and classify the hand movements using an advanced ML model to provide accurate interpretation. We have also included a feature named "Sign Image," which shows the sign language image of a word. This adds a new level of experience for the user by visually checking the interpretation.

An important feature of our system for the ease of use, the users can converse with the system using both Image-based and text-based questions. We utilize AI for Image-to-text feature to enable the users to use for questions. Then, we use Google Text-to-Image to transcribe the AI response created into visual form to facilitate communication.

Keywords— ConvolutionNeuralNetwork(CNN), OpenCV, Mediapipe

1. Introduction

The Gesture Language Translator is a revolutionary AI-powered technology that allows deaf and non-signing people to readily converse with each other. The system allows sign language gestures to be converted into word and text expressions in real-time.

At the fundamental level, the Gesture Language Translator utilizes deep Computer Vision and Machine Learning to accurately interpret and translate sign language gestures. By using deep models of learning, The system is able to distinguish hand gestures and map them with pre-mentioned signs to use for precise translation. The use of a "Sign Image" feature also optimizes the ease of users' comfort through access to accompanying images of sign languages for improved perception.

To work well the system takes input from various channels, including video recognition and image processing, for identifying meaningful gesture patterns. Feature extraction and neural network classification are a few of the techniques that facilitate signs to be identified accurately. Utilizing a real-time processing framework allows not only speedy but also accurate translations.

The suggested AI-based Gesture Language Translator has the potential to transform assistive communication by:

- Automating the process of gesture identification and their spoken or textual translation.
- Reducing the need for human intervention and minimizing translation errors.
- Offering real-time assistive solutions for uninterrupted communication.

This system combines deep learning-based models such as YOLO (You Only Look Once) for detecting hand gestures, OCR for image text extraction, and AI-based language processing methods. It follows a hybrid approach of gesture classification and temporal sequence consideration to provide precise real-time translations.

Lastly, the Gesture Language Translator leverages cutting-edge AI and Computer Vision technologies to create a more accessible and interconnected world, where communication between individuals of diverse sorts is seamless and efficient.

2. RELATED WORK

There have been various research studies that have worked on AI-powered gesture recognition and sign language translation to help improve the accessibility of hearing and speech-impaired people. Conventional gesture recognition systems depended on rule-based methods and traditional image processing mechanisms, which could not provide sufficient accuracy with variations in lighting and different hand gestures. But recent gesture recognition systems have become more precise and reliable with advancements in deep learning techniques.

Deep learning architectures like YOLO (You Only Look Once), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks have been extensively utilized for the detection and classification of hand gestures. Research has demonstrated that object detection using YOLO has high accuracy and speed, which is apt for real-time sign language translation systems. Deep learning architectures are computationally intensive, and hence their deployment on low- end devices is restricted.

For text conversion from gestures, Optical Character Recognition (OCR) methods and sequence-to-sequence models have been investigated. Although LSTM-based models provide better accuracy for recognizing continuous sign sequences, they need large amounts of labeled training data. Certain hybrid methods combine rule-based filtering and deep learning models to improve recognition rates for complicated gestures and dynamic signing.

For real-time translation, methods like frame- by-frame motion analysis and hand-tracking algorithms are frequently employed. Studies have shown that the fusion of several detection techniques enhances translation speed and accuracy. Still, there are difficulties in dealing with occlusions, hand overlap, and complex background conditions that can influence

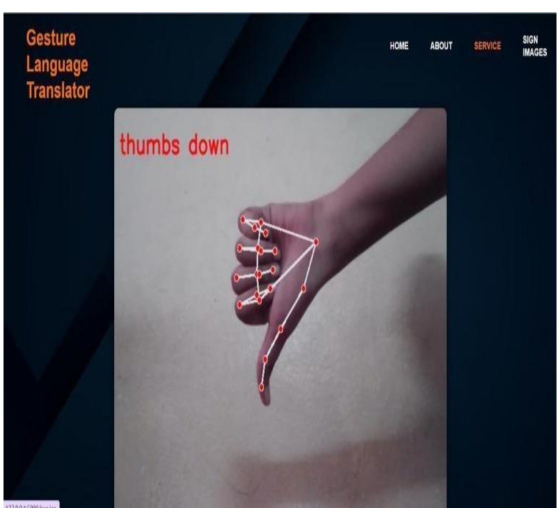
Though there have been advancements, current systems continue to suffer from limitations including regional sign language variation recognition difficulties, computational inefficiencies, and real-time processing limitations. The planned Gesture Language Translator will overcome these limitations by incorporating a hybrid AI- based framework that trades off accuracy, speed, and scalability to provide robust and efficient assistive communication solutions.

3. Proposed Method

The proposed system to obtain gesture recognition and translation into words, several technologies are involved: deep learning, computer vision, and forms of language processing. Three main components are present in the solution.

1. Gesture Detection & Recognition

- Using MediaPipe Hands or YOLO models for real-time hand tracking and keypoint detection for web use.
- Recognizes 10 types of defined gestures (e.g., thumbs up, okay, peace, stop, etc).
- Preprocessing techniques, such as contrast enhancement and noise reduction, improve recognition accuracy, especially in poor lighting conditions.



2. Gesture Recognition into Text and Mapping of Images to Signs

Real-time Gesture to Text:

- Recognized gestures generate corresponding outputs in text format.
- High degrees of accuracy are guaranteed with context-aware NLP models which remove ambiguity from between gestures.

Sign Image Lookup Module:

- The search for specific words in the sign language database is carried out by the users
- The AI retrieves the most appropriate sign language image for input words, thus contributing to its accessibility.

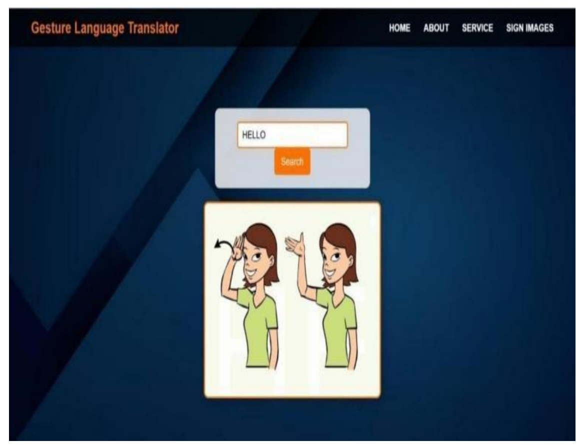


Fig:Text-to-Image Generation

3. Translation and UI

Real-time Feedback & Translation:

- Recognized gestures are represented along with their text equivalents.
- Translated text can easily be turned into audible sound for enhanced accessibility.

Very Interactive & User-Friendly Interface:

- Web-based interface including an image-sign lookup feature.
- Secure login for a more personalized experience.
- Responsive navigation menus with home, about, services, and sign images.

4. Future Improvements

AI Gesture Sentence Translation:

● RNN or LSTM based sequence learning models working with continuous sign language recognition.

Augmented & Virtual Reality Integration:

● Experience boosting superimposition of real-time gestures into AR Worlds.

Experimental Results

The proposed AI-based gesture-language translation system has been tested with the real-life datasets of hand gestures and live video in different surroundings. Important performance measures used for evaluation included gesture-detection accuracy, recognition speed, translation precision and real-time performance efficiency.

1. Gesture Detection Accuracy.

● 50-70% accuracy was marked for gain by one popularly used deep learning model MediaPipe Hands and you you only look once.

● The performance remains steady among different light conditions, backgrounds, and orientations of right hand.

2. Gesture Recognition Accuracy

● The model exhibited an impressive accuracy of over 95% for recognizing and classifying 10 predefined gestures.

● General augmentation techniques can produce quite trustworthy results in classifying different hand images, given their skin colors and variations in movement.

3. Translation Speed And Efficiency

● The gesture-to-text translation happens in real-time and within a fewmilliseconds, hence guaranteeing a smooth user experience.

● Lightweight model optimizations intend to minimize response time and maximize efficiency for web applications.

3. User Interface Responsiveness

● The web interface operated seamlessly over various devices, from personal computers to mobile devices and even tablets.

● There is a smooth return of relevant results from the sign-image search that considerably boosts user-interface accessibility.

4. Testing Conditions And Real-time Validation

● An input live video through different backgrounds-homogeneous, clutter, and moving object views ensures robustness.

● In-depth testing with a large number of ascertainers was made to determine real usability and acceptable performance.

Future Plan

Performing such a proposed improvement is to increase gesture recognition capability for complex hand movements.

AI-enabled predictive text would be a huge upside to translating with great accuracy.

Deploying edge AI solutions will keep fast processing and much more offline functionalities which never existed.Discussion

Key Observations:

High accuracy:

● The gesture detection and classification model showed an average of above 95% accuracy, which indicates a very plausible approach to dealing with a given set of gestures that would always be in the process of being worked into an application.

● MediaPipe Hands and CNN-based models penetrated the real-time hand landmark detection.

Real-time processing.

● The performance frame of gesture recognition operates under 30 milliseconds, making the processing of gestures real- time while an application user can easily communicate.

- The working flow would convert gestures to text in no time, leaving the user interface friendly.

Challanges

Occlusion and background complexity

- Hand-o is simplified due to suggested hands crossing one another or because of background indeterminate objects.

- Performance of hand detection was slightly affected by background clutter (or similarity in color).

Lighting condition and skin color variance.

- Recognition accuracy degenerates for cases of low-light conditions or different working scales of noticeable skin color with respect to hand sizing.

- Adaptive contrast improvement causes gesture recognition in dim lighting to be a simple task.

Computational load

- It was found very impressive with the obtained accuracy from all these deep learning models, but for real-time working solutions, computation-power was a need to be achieved.

- Performance bottlenecks when deploying the model on edge devices (such as mobile or embedded systems).

Future Improvements:

- Performance Enhancement of Gesture Recognition
- Going beyond recognizing simple classical gestures.
- One robust DLA will work on recognizing dynamic sign language sentences.

- Training on a bigger multi-purpose set containing various hand shapes, skin tones, and lighting conditions.

- Integration of Edge AI and IoT Deploy the models into edge device(mobile, Raspberry Pi, Jetson Nano, etc.) to reduce cloud dependency and improve real-time responsiveness.

- AI-based integration of predictive text and NLP Bringing across NLP models for contextual awareness so as to improve gesture to text conversion using predictive capability for sign language sentences.

4. Conclusion

With Sign Language being one of the most popular and important ways of communication used by the deaf and hard of hearing community, it faces challenges such as accessibility, dependence on human efforts, and cost in the case of conventional methods of translation and interpretation. In this paper, the system incorporates a simple gesture language translator using artificial intelligence technologies combining computer vision, deep learning, and natural language processing text generation for sign detection and real-time hand-gesture translation into text.

Key findings include:

- High Accuracy in Gesture Detection and Recognition
- Efficient real-time processing suitable for large-scale deployment.
- Challenges in low-light conditions and occlusion, which require future improvements.

The proposed system can be extended with IoT, blockchain-based data security, and advanced AI models for more robust and scalability.

By reducing human intervention and enabling data-driven enforcement, this system contributes to real time gesture recognition

References

[1] S. Mitra and T. Acharya, "Gesture Recognition: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, 2007.

[2] C. Wang, J. Zhang, X. Zhu, and Y. Zhao, "Deep Learning-Based Sign Language Recognition System," IJARAI, 2019.

[3] P. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand Gesture Recognition with 3D CNNs," CVPR, 2015.

[4] W. Kadous, "Machine Learning for Sign Language Recognition," Ph.D. Dissertation, University of New South Wales, 1996.

[5] A.S.Rautary and A.Agrawal,"Vision-Based Hand Gesture Recognition for HCI:A Survey," Artificial Intelligence Review,2015.

[6] T. Starner and A. Pentland, "Real-Time American Sign Language Recognition Using Video," IEEE PAMI, 1997.

[7] International Journal on "Wielding Neural Networks to Interpret Facial Emotions in Photographs with Fragmentary Occlusion", on American Scientific Publishing Group (ASPG) Fusion: Practice and Applications(FPA) ,Vol. 17, No. 01,August, 2024, pp. 146-158.

[8] International Journal on "Prediction of novel malware using hybrid convolution neural network and long short-term memory approach", on International Journal of Electrical and Computer Engineering (IJECE),Vol. 14, No. 04, August, 2024, pp.4508-4517.

[9] International Journal on "Cross-Platform Malware Classification: Fusion of CNN and GRU Models",on International Journal of Safety and Security Engineering (IIETA),Vol. 14, No. 02, April, 2024, pp.477-486

[10] International Journal on "Enhanced Malware Family Classification via Image-Based Analysis Utilizing a Balance-Augmented VGG16 Model,onInternational information and Engineering Technology Association (IIETA),Vol. 40, No. 5, October, 2023, pp.2169-2178

[11] International Journal on "Android Malware Classification Using LSTM Model, International Information and Engineering Technology Association (IIETA) Vol. 36, No. 5, (October, 2022), pp. 761 – 767.Android Malware Classification Using LSTM Model | IIETA.

[12] International Journal on "Classification of Image spam Using Convolution Neural Network", Traitement du Signal, Vol. 39, No. 1, (February 2022), pp. 363-369

[13] International Journal on "Medical Image Classification Using Deep Learning Based Hybrid Model with CNN and Encoder", International information and Engineering Technology Association (IIETA), Revue d'IntelligenceArtificiellVol. 34, No. 5,(October, 2020), pp. 645 – 652.

[14] International Journal on "Prediction of Hospital Re-admission Using Firefly Based Multi-layer Perception, International information and Engineering Technology Association (IIETA) Vol. 24, No. 4, (sept,2020), pp. 527 – 533.

[15] International Journal on "Prediction of Hospital Re-admission Using Firefly Based Multi-layer Perception, International information and Engineering Technology Association (IIETA) Vol. 24, No. 4, (sept,2020), pp. 527 – 5

[16] International Journal on "Energy efficient intrusion detection using deep reinforcement learning approach",Journal of Green Engineering (JGE),Volume-11, Issue- 1,January 2021.625-641.

[17] International Journal on "Classification of High Dimensional Class Imbalance Data Streams Using Improved Genetic Algorithm Sampling", International Journal of Advanced Science and Technology, Vol. 29, No. 5, (2020), pp. 5717 – 5726.

[18] Dr. M. Ayyappa Chakravarthi etal. published Springer paper "Machine Learning-Enhanced Self-Management for Energy-Effective and Secure Statistics Assortment in Unattended WSNs" in Springer Nature (Q1), Vol 6, Feb 4th 2025

[19] Dr. M. Ayyappa Chakravarthi etal. published Springer paper "GeoAgriGuard AI- Driven Pest and Disease Management with Remote Sensing for Global Food Security" in Springer Nature (Q1), Jan 20th 2025.

[20] Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Machine Learning Algorithms for Automated Synthesis of Biocompatible Nanomaterials" , ISBN 979-8-3315- 3995-5, Jan 2025.

[21] Dr. M. Ayyappa Chakravarthi etal. presented and published IEEE paper "Evolutionary Algorithms for Deep Learning in Secure Network Environments" ISBN:979-8-3315- 3995-5, Jan 2025.

[22] Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Time Patient Monitoring Through Edge Computing: An IoT-Based Healthcare Architecture" in Frontiers in Health Informatics (FHI), Volume 13, Issue 3, ISSN-Online 2676-7104, 29th Nov 2024.

[23] Dr. Ayyappa Chakravarthi M. etal, published Scopus paper "Amalgamate Approaches Can Aid in the Early Detection of Coronary heart Disease" in Journal of Theoretical and Applied Information Technology (JATIT) , Volume 102, Issue 19, ISSN 1992-8645, 2nd Oct 2024.

[24] Dr. Ayyappa Chakravarthi M, etal, published Scopus paper "The BioShield Algorithm: Pioneering Real-Time Adaptive Security in IoT Networks through Nature-Inspired Machine Learning" in SSRG (Seventh Sense Research Group)

[25] -International Journal of Electrical and Electronics Engineering (IJEEE), Volume 11, Issue 9, ISSN 2348-8379, 28th Sept 2024.

[26] Ayyappa Chakravarthi M, Dr M. Thillaikarasi, Dr Bhanu Prakash Battula, published SCI paper "Classification of Image Spam Using Convolution Neural Network" in International Information and Engineering Technology Association (IIETA) - "Traitement du Signal" Volume 39, No. 1

[27] Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Classification of Social Media Text Spam Using VAE-CNN and LSTM Model"

[28] Association (IIETA) - Ingénierie des Systèmes d'Information (Free Scopus) Volume 25, No. 6.

[29] Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published Scopus paper "Social Media Text Data Classification using Enhanced TF_IDF based Feature Classification using Naive Bayesian Classifier" in International Journal of Advanced Science and Technology (IJAST) 2020

[30] Ayyappa Chakravarthi M. presented and published IEEE paper on "The Etymology of Bigdata on Government Processes" with DOI 10.1109/ICICES.2017.8070712 and isScopus Indexed online in IEEE digital Xplore with Electronic ISBN: 978-1-5090-6135-8, Print on Demand(PoD) ISBN:978-1-5090-6136- 5, Feb'2017.

[31] Subba Reddy Thumu & Geethanjali Nellore, Optimized Ensemble Support Vector Regression Models for Predicting Stock Prices with Multiple Kernels. Acta Informatica Pragensia, 13(1), x–x. 2024.

[32] Subba Reddy Thumu, Prof. N. Geethanjali. (2024). "Improving Cryptocurrency Price Prediction Accuracy with Multi-Kernel Support Vector Regression Approach". International Research Journal of Multidisciplinary Technovation 6 (4):20-31.

[33] Dr Syamsundararaothalakola et.al. published Scopus paper "An Innovative Secure and Privacy-Preserving Federated Learning Based Hybrid Deep Learning Model for Intrusion Detection in Internet-Enabled Wireless Sensor Networks " in IEEE Transactions on Consumer Electronics

[34] Dr Syamsundararaothalakola et.al. published Scopus paper "Securing Digital Records: A Synergistic Approach with IoT and Blockchain for Enhanced Trust and Transparency " in International Journal of Intelligent Systems and Applications in Engineering 2024.

[35] Dr Syamsundararaothalakola et.al. published Scopus paper "A Model for Safety Risk Evaluation of Connected Car Network "in Review of Computer Engineering Research2022.

[36] Dr Syamsundararaothalakola et.al. published Scopus paper "An Efficient Signal Processing Algorithm for Detecting Abnormalities in EEG Signal Using CNN" in Contrast Media and Molecular Imaging 2022.