International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 08 | August - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

SIGN LANGUAGE TRANSLATOR USING PYTHON

BAGATH S, DHANA VARSHINI S, JERSIL G M

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

SRI SHAKTHI INSTITUTE OF ENGINEERING AND TECHNOLOGY

COIMBATORE 641062

MS.M.BUVANA [AP/IT]

DEPARTMENT OF INFORMATION TECHNOLOGY

SRI SHAKTHI INSTITUTE OF ENGINEERING AND TECHNOLOGY

COIMBATORE 641062

ABSTRACT

The complex and ancient form of sign language, which is largely used by the deaf and hard of hearing, struggles with both a lack of qualified interpreters and low comprehension. Convolutional Neural Networks (CNNs), a cutting-edge technology, are being used to decode American Sign Language (ASL) motions in response and are particularly effective at fingerspelling. Pre-processing is used in this advanced method to clean up captured motions, which are then fed into a CNN that has been painstakingly trained on a variety of datasets. This makes it possible for the network to extract critical patterns and features for exact recognition. The CNN's classification layer effectively converts these movements into ASL letters or phrases. The ramifications are wide-ranging and touch on accessibility, communication, and education. The use of this technology might improve interactions between hearing people and deaf people. Additionally, it could help close societal disparities by promoting diversity in fields like healthcare and education. CNN-based ASL recognition represents significant progress toward seamless communication and understanding across hearing and non-hearing cultures, despite obstacles including illumination variability and gesture variation. These issues are expected to be resolved as methods and technology develop, improving the performance of CNN-based ASL identification systems and changing how people who are deaf or hard of hearing communicate with the outside world, ultimately resulting in a more inclusive society.

KEYWORDS

Open CV, TensorFlow, gTTS, NLP, Python, Recognition, Conversion, Translation.

INTRODUCTION

The sign language used by the Deaf and Hard of Hearing community is a sophisticated and complicated system of communication that allows people to communicate themselves through a dynamic combination of hand gestures, facial emotions, and body movements. The communication gap between those who use sign languages and those who do not, despite its elegance, nevertheless remains. In a world that is growing more connected and varied, closing this gap is crucial because it creates the foundation for inclusive communication and mutual understanding.

An important step in solving this issue has been the creation of sign language translators, a technological marvel that connects sign language users with the broader population. This device can recognize and translate sign language motions into spoken or written English in real-time thanks to the utilization of cutting-edge technologies such as computer vision, machine learning, and natural language processing. By removing linguistic barriers, the sign language translator enables people who use sign language to engage freely and meaningfully in a number of contexts, including educational and clinical settings, normal social interactions, and critical public services.

The goal of this paper is to study the creation and application of sign language translation as a transformative instrument for fostering successful communication between sign language users and non-signers. It sets off on an enlightening trip into the realm of sign language translation. The main capabilities of the sign language translator take advantage of developments in computer vision methods, machine learning algorithms, and real-time processing. Our suggested translation aims to provide a practical and useful solution that improves accessibility and advances the inclusion ideal by combining these technologies[2].

The following sections of this paper go into further detail into the intricate technological foundation of the sign language translator. This includes a thorough explanation of the techniques used to identify and analyse complex hand movements, the sophisticated algorithms that power recognition, and the translation mechanisms that turn these signals into understandable spoken or written language. We give results from a series of thorough experimental tests to support the effectiveness of our method, demonstrating the precision and effectiveness of our translator in faithfully translating a wide variety of sign language motions.

The real-world applications of our sign language translator are also explored in this research, demonstrating its potential to fundamentally alter communication dynamics across linguistic and cultural boundaries. The translator emerges as a change agent in bridging gaps and advancing cross-cultural understanding by encouraging meaningful contacts between disparate populations.

The development of reliable and effective sign language interpreters holds great potential for enhancing accessibility and removing the obstacles to productive communication as technology advances. By embracing innovation, we help to build an inclusive society where language barriers do not prevent dialogue. Our commitment to creating a more peaceful and connected world, where everyone can appreciate and understand the beauty of sign languages, is demonstrated by our pursuit of sign language translation. We make considerable progress toward a future in which communication barriers are eliminated and the idea of an inclusive society is fulfilled through our constant commitment to enhancing sign language translation.

METHODOLOGY

The suggested method is made up of a number of crucial steps that are carefully organized to produce accurate hand gesture translation from American Sign Language (ASL) into text and speech. In order to provide clear and correct communication for sign language users, each stage is important.

IMAGE CAPTURING FROM CAMERA

Camera Setup And Calibration

Achieving precise image capture entails meticulous camera setup, accounting for crucial elements like camera angle, distance from the subject, and lighting conditions. Accurate camera calibration may involve adjusting intrinsic parameters (focal length, optical centre) and extrinsic parameters (position and orientation) for optimal performance. Calibration techniques, such as chessboard calibration patterns, are utilized to fine-tune these parameters, enhancing the camera's ability to capture hand gestures with fidelity.

FRAME EXTRACTION AND PROCESSING

In this stage, frames are systematically extracted from the continuous video stream recorded by the camera. These individual frames encapsulate specific moments in the signing sequence. Pre-processing steps are then applied to these frames to enhance their quality for subsequent analysis. Techniques such as noise reduction diminish unwanted visual artifacts, contrast enhancement improves visibility, and frame alignment ensures consistency. These preparatory measures lay the groundwork for accurate gesture recognition and translation, shaping the overall effectiveness of the system.

HAND GESTURE SCAN

Region Of Interest Detection

This stage recognizes the crucial regions of interest within each frame, particularly the hands and gestures, using advanced techniques like skin colour segmentation or motion detection. The system effectively decreases computing effort and focuses on the crucial aspects for analysis by identifying these critical components. This separation guarantees that subsequent operations are precisely adjusted to the relevant visual information, improving the precision of gesture detection and system effectiveness.

BACKGROUND REMOVAL

In this step, immobile or obtrusive components in the frame are skilfully removed using modelling or background subtraction methods. The technology speeds further analysis by extracting the dynamic hand motions and removing extraneous background components. By eliminating unnecessary visual components, the input for recognition algorithms is improved, interference is reduced, and only the most important motions are submitted to additional processing, improving the system's overall performance and accuracy[3].

HAND POSTURE RECOGNITION

Convolutional Neural Network (CNN) Architecture Design

Carefully crafted to excel in hand gesture detection, CNN is the central component of the recognition process. Layer count, kernel dimensions, and activation function choices are purposefully selected to enable the network to extract crucial elements from gesture input. The network can recognize complex patterns, which are essential for exact recognition, thanks to the architecture's intelligent design, which also maximizes computing efficiency.

Transfer Learning

This method adapts the network for the goal of hand gesture detection by drawing on pre-trained CNN models, frequently generated from sizable datasets like Image Net. The network gains expertise from multiple domains by fine-tuning the model using specific gesture data. As a result of adapting this information to recognize movements reliably, learnt patterns from various visual data sources are effectively used to improve recognition accuracy. The method considerably cuts down on training time while improving the system's ability to correctly understand motions[4].

GESTURE CLASSIFICATION

Identification Of Similar Motions

Using methods like clustering or feature analysis, this stage locates groups of visually related motions. The algorithm develops a sophisticated comprehension of these gesture sets by classifying motions with similar visual properties. This procedure helps in the development of customized classifiers that can distinguish between gestures that may look similar but have different semantic implications. It improves identification accuracy and ensures that even minute changes are correctly interpreted when visually similar motions are segmented into relevant groups.

Custom Classifiers

Carefully crafted and trained classifiers are used for each cluster of visually comparable motions. Employed methods include Support Vector Machines (SVMs), Random Forests, and Neural Networks, all of which concentrate on identifying minute differences within the cluster. The system improves its capacity to effectively interpret and classify visually comparable movements, enhancing the overall recognition accuracy and precision. This is accomplished by training classifiers customized to the distinctive characteristics of each gesture category[5].

TEXT AND SPEECH TRANSLATION

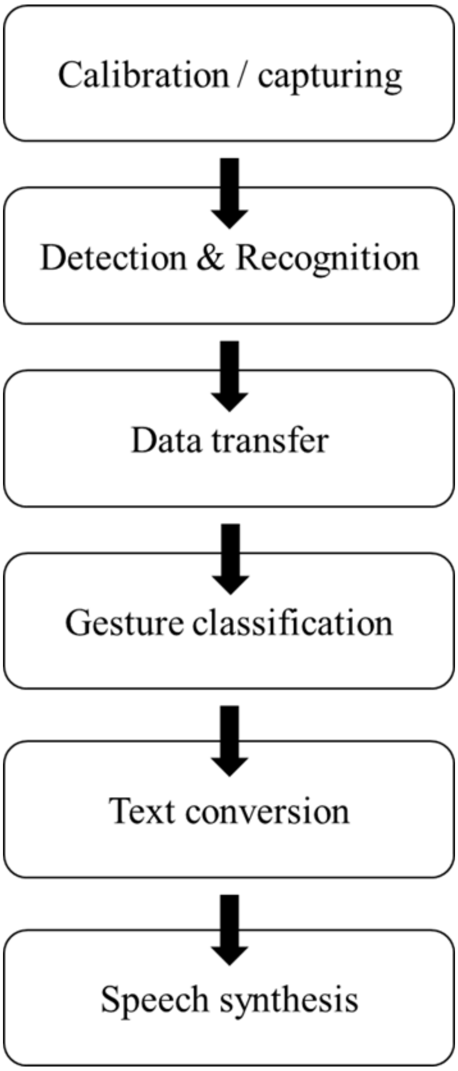
Gesture-to-Text Conversion

In this stage of transformation, identified motions are given linguistic meaning by being mapped to relevant text. The system converts gestures into meaningful textual representations by using either a predetermined vocabulary or a machine learning-based technique that takes use of learnt associations. This translation establishes a foundation for understandable communication between sign language users and non-signers by tying the visual language of gestures to the written word[1].

Integration Of Speech Synthesis

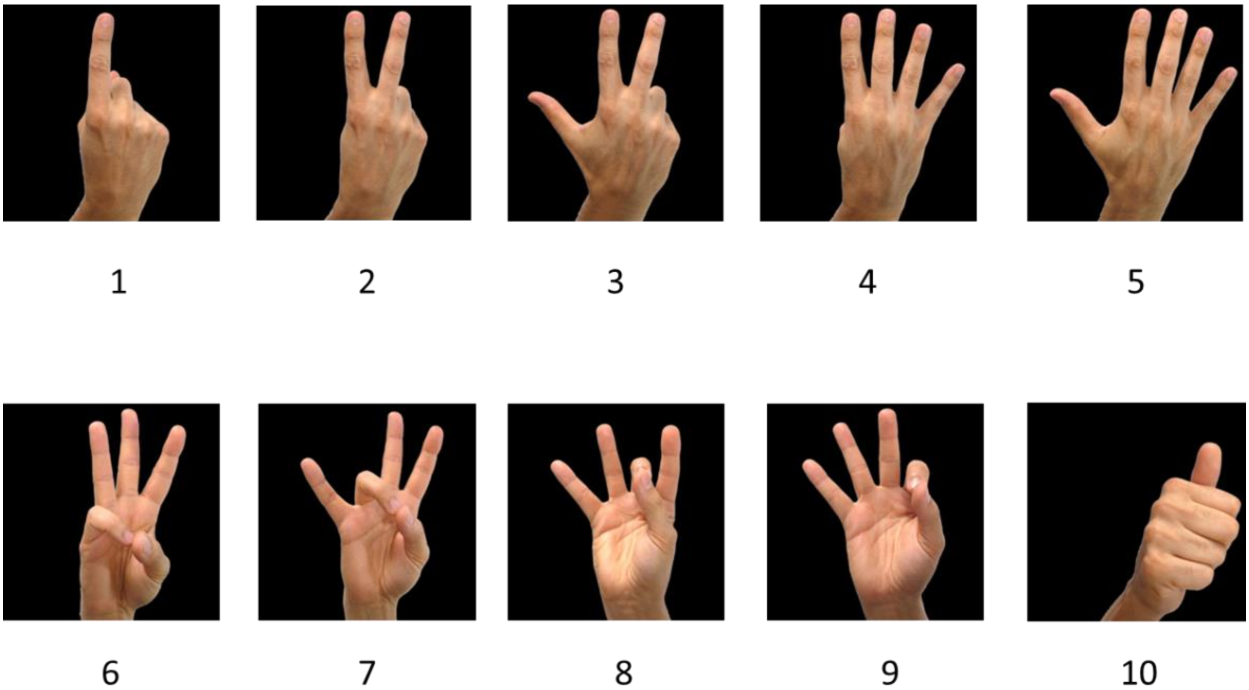
The "gTTS" library or comparable tools are integrated to bring recognized text into the auditory dimension. This synthesis entails choosing appropriate voices, accents, and pacing, and it results in the production of output that sounds natural. The system improves the whole immersive communication experience and eliminates language barriers by imitating spoken language, making interactions more interesting and relatable.

Finally, the suggested system's integrated methodology—from camera-based picture capture to gesture identification, classification, and translation—unveils its potential to provide sign language users with sophisticated communication capacities.

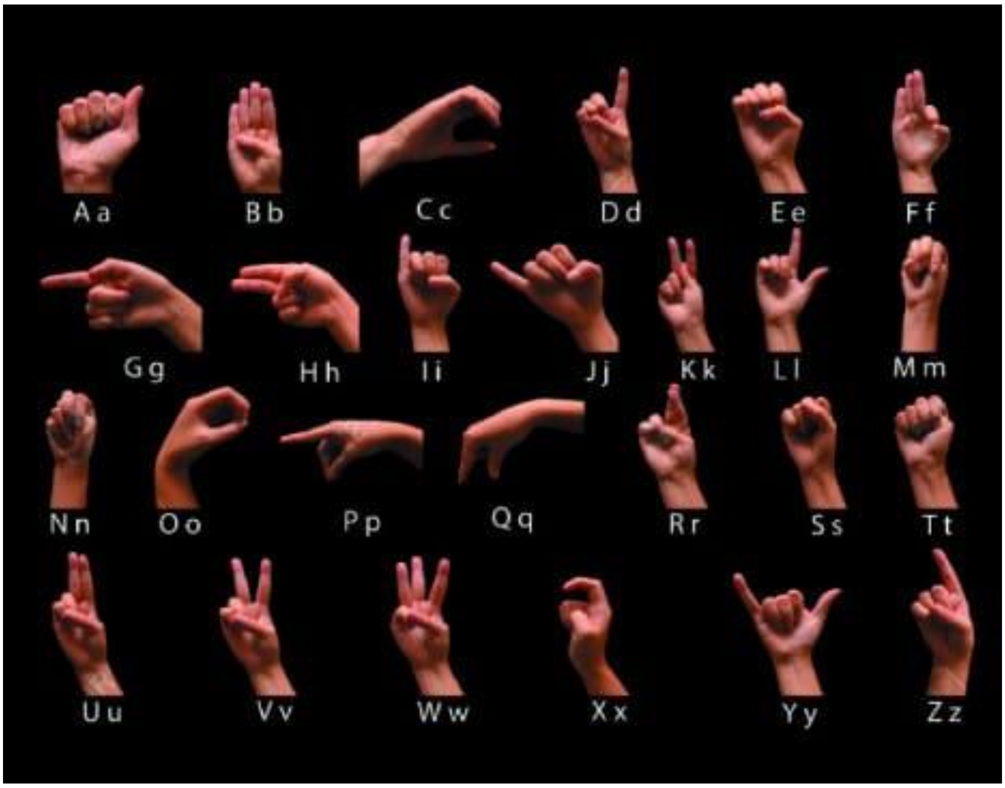


Proposed flowchart for ASL convertor

Volume: 07 Issue: 08 | August - 2023 SJIF Rating: 8.176 ISSN: 2582-3930



Numbers from 1-10 in ASL



Alphabets in ASL

SYSTEM SPECIFICATIONS

Software Requirements

Python is a versatile, high-level programming language known for its simplicity and flexibility. Developed by Guido van Rossum in 1991, it emphasizes readable code and offers a concise syntax. Python supports various programming paradigms, including procedural, object-oriented, and functional programming. Its extensive standard library empowers developers across domains such as web development, data analysis, scientific computing, automation, and artificial intelligence (AI). TensorFlow, an essential library, aids in creating machine learning models for gesture recognition. OpenCV is crucial for real-time hand gesture interpretation, enabling efficient communication for sign language users. Natural Language Processing (NLP) plays a pivotal role by translating gestures into text or speech, bridging the gap between sign language and spoken language. Through these technologies, sign language translators facilitate smoother communication and inclusivity, enhancing accessibility for all.

WORKING

Capturing Input

The system records input as video footage, usually from a camera or webcam. The user's hand motions, which are the foundation of sign language communication, are captured in this video stream.

Hand Tracking and Gesture Recognition

It can be done with the help of OpenCV which has Python bindings. This library offer resources for finding and locating hands in video frames. The system recognizes numerous hand configurations that indicate distinct sign language symbols using predefined motions or machine learning models[7].

Feature Extraction

Extract pertinent features from the recorded hand motions, such as the form of the hand and the movements of the fingers. The features act as the succeeding phases' input data.

Machine Learning

To recognize movements, this system uses machine learning, which involves building a model from a collection of annotated sign language gestures. These models can be created and trained with the use of TensorFlow or other machine learning libraries. Based on the features gathered, the trained model can then forecast the sign language gestures[6].

Translation and Text Generation

It's time to convert the gestures into text once the system has recognized them. Rule-based or machine learning-driven approaches in Natural Language Processing (NLP) techniques can translate identified gestures into textual representations of the related words or phrases.

Text-to-Speech Conversion

Text-to-speech (TTS) engines can further translate the translated text into spoken language to complete the dialogue. This procedure can be facilitated by Python modules like gTTS (Google Text-to-Speech).

Output Display

To make the message understandable to users who do not utilize sign language, the translated text or spoken output is shown on a screen or transmitted over speakers.

FUTURE SCOPE

Future systems might develop to understand not just the signs' physical characteristics but also the feelings and sentiments they represent, deepening communication. Collaborative translation, which enables several users to participate in the translation process in real time, may be made possible using sign language translators. Sign language interpreters might materialize in actual settings thanks to holographic and augmented reality (AR) technologies, enabling more realistic and immersive encounters. In order to enable the deaf and hard-of-hearing community to participate more actively in a variety of professions, sign language translators could play a crucial role in inclusive education and employment integration

REFERENCES

1. Starner.T, & Pentland, A. (1997). Real-time American Sign Language recognition from video using hidden Markov models. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 265-271).

2. Han, J., Martinez, J. M., & Dahmoush, H. (2013). Real-time sign language recognition using a consumer depth camera. In Proceedings of the 15th ACM on International Conference on Multimodal Interaction (ICMI) (pp. 235-238).

3. Li, H., & Hu, J. (2013). Real-time sign language recognition based on RGB-D. In Proceedings of the 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) (pp. 2320-2325).

4. Stein, G., Tyagi, A., Lohan, K. S., Pitsch, K., Gehre, F., & Sagerer, G. (2010). The Bielefeld 2000/2001 (BiNVis) American Sign Language corpus. In Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC) (pp. 3150-3155).

5. Koller, O., & Bowden, R. (2008). Learning sign language by imitation: An evolutionary robotics approach. IEEE Transactions on Evolutionary Computation, 12(6), 711-731.

6. Buehler, P., Feiten, W., Sagerer, G., & Ritter, H. (1995). Dynamic recognition of gestures with a multi-cue system. In Proceedings of the 5th International Conference on Automatic Face and Gesture Recognition (pp. 50-55).

7. Pfister, T., Charles, J., Hanke, J., & Van Gool, L. (2014). Seeing the Signs: Hand Segmentation in Real-Time Depth Images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops.