IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

"Sign Language Gesture Recognition Using Machine Learning"

Vidya R Ghule
Computer Engineering
VPSCET Lonavala
SPPU student India

Sanjivani M Jadhav
Computer Engineering
VPSCET Lonavala
SPPU student India

Gayatri S Wasulkar
Computer Engineering
VPSCET Lonavala
SPPU student India

Aniket T Ghule
Computer Engineering
VPSCET Lonavala
SPPU student India

Prof.Ragho Soni
Department of Computer Eng
Assistant Professor
SPPU India

Abstract:-

Sign language translator is used by deaf people with hearing or speech problems to communicate with each other. Sign language is a form of communication that uses signs and gestures to convey meaning. We propose a new way to convert signed words into text. Our system is designed to help hearing impaired people communicate with others more easily and conveniently. The technique uses computer vision and machine learning to recognize gestures and convert them into appropriate text. Pattern recognition and gesture recognition techniques are used. Gestures, which are an important part of non-verbal communication, play an important role in our daily lives.

Key Words: -

Machine learning, Pattern recognition, Gesture recognition**,**Computer and information processing, Feature extraction, Human Computer Interaction, Image processing, Sign Language Recognition, : Hand Gesture Recognition**.**

1.Introduction:-

Sign language plays a vital role in facilitating communication for individuals who are deaf or mute. In this unique form of language, each gesture carries a precise meaning, allowing for the expression of intricate concepts through the combination of fundamental elements. Sign language is primarily a non-verbal mode of communication, enabling effective interaction between deaf and mute individuals and the hearing community. It adheres to specific rules and grammatical structures to convey messages efficiently. Sign language recognition approaches

can be broadly categorized into two types: image-based and sensor-based. Currently, there is a predominant focus on image-based approaches due to their inherent advantages, such as eliminating the need for complex devices like gloves, helmets or bands which are typically required in sensor-based methods.

Gesture recognition offers us a new, natural and effective way to communicate with computers that humans are more familiar with. The software aims to provide real-time gesture recognition based on the identification of some similar features (e.g. direction, center of mass, and finger state) considering the Similarity of the human hand such as four fingers and thumb. The thumb is in the raised or folded position of the fingers of the hand. Gestures are a method used for non-verbal communication in language. It is often used by deaf people with hearing or speech problems to communicate with each other or with hearing people. Sign language is a form of communication that uses signs and gestures to convey meaning. We propose a new way to convert signed words into text. Our system is designed to help hearing impaired people communicate with others more easily and conveniently. The plan uses computer vision and machine learning to recognize gestures and convert them into appropriate text. Pattern recognition and gesture recognition are developing research areas. Gestures, which are an important part of non-verbal communication, play an important role in our daily lives.

Gesture recognition is gaining increasing significance across various application domains, including human-computer interaction, communication, multimedia, and security. Specifically, sign language recognition is treated as a subset of image understanding, involving two primary phases: sign detection and sign recognition. Sign detection focuses on extracting distinctive features from specific objects based on predefined parameters, while sign recognition involves identifying unique shapes that distinguish one object from others. From a technical perspective, key characteristics of sign language communication encompass its social relevance, technical and technological convenience, and ease of use.

The proposed system utilizes a webcam for image capture, with preprocessing performed using Jupyter Notebook as the integrated development environment (IDE) and Tensorflow & the OpenCV or Scikit library. The captured image sequence undergoes a series of preprocessing steps to eliminate background noise. Additionally, a slope distance-based algorithm, specifically the Fingertip Detection using convex hull algorithm, is employed.

Fig.1:- High Level View of System

The goal here is to implement HCI's ability to nearby human-human interaction by modeling a sign language recognition system that would aid in predicting the context of dialogue between a person to another with the help of an interlocutor, here it is the system. This system makes use of various classifiers for hand detection and uses skin segmentation for recognizing the gestures and empirical tracking method that can dynamically change according to the stage of action. A gesture allows an individual to convey information to another person irrespective of whether they understand the message or not. This system also provides the facility of learning hand sign language with speech recognition that helps all those who want to learn the language. The world is now becoming more disabled friendly making them feel much more normal and the system will make them independent without the need of a third person as a translator.

2.Literature Survey: -

Vision-based hand gesture recognition falls into two primary categories: the 3-D hand model-based method and the appearance-based method. The 3-D hand model-based approach involves comparing input frames to the 2-D appearance projected by a 3D hand model. However, this method necessitates an extensive database to accommodate all potential projections of the 3-D hand model, making it less practical for real-world applications. In contrast, appearance-based techniques extract image features and model them as visual attributes of a hand's posture. They then compare these features to those extracted from live video feeds of users performing gestures, ensuring real-time operation due to the use of 2-D image features. Appearance-based techniques can be further categorized into static hand posture detection and dynamic gesture detection.

Nasser H.D et al. [1] adopt an approach where key features extracted consist of SIFT (Scale Invariant Feature Transform) key-points. They construct a grammar from sequences of hand postures for dynamic gesture detection. Meanwhile, Emil M.P. et al. [2] propose a similar method but use Haar-like features for image description and employ AdaBoost for classification. This approach is valued for its computational efficiency but presents the challenge of requiring a substantial number of features, which complicates the system's training stage when utilizing the AdaBoost classifier.

The recognition and extraction of hands from complex and dynamic backgrounds pose additional challenges. Among various methods, skin color detection is a popular choice. In [3], P.K.Bora et al. introduce the HSV (Hue, Saturation, Value) color space, noting that skin color, regardless of gender or race, consistently falls within specific H and S value ranges. This information allows for the extraction of skin-colored objects from diverse backgrounds.

In their work [4], Macheal V et al. conducted an assessment of different shape descriptors, including Zernike moments, highlighting their effectiveness. Zernike moments have gained popularity in image processing due to their ability to compute higherorder moments independently of lower-order moments, along with their qualities of reconstruction and rotation invariance, which make them an appealing choice for shape description. Additionally, Athira et al. [10] demonstrated the practicality of Zernike Moments by creating a system for recognizing static American Sign Language (ASL) gestures against a uniform background. Their system also incorporates a speech engine for converting speech to gestures.

In reference [5], the groundwork for utilizing Hidden Markov Models (HMM) is laid by drawing a parallel between speech recognition and gesture recognition. HMMs are employed to model time-series data, where hand movement along the coordinate axis is tracked, treating each direction as a state. The paper explores a lexicon of 40 gestures, achieving an impressive 95% accuracy. However, it acknowledges that as the lexicon expands, describing hand configuration alongside hand trajectory becomes more complex and time-consuming. There is a need for a simpler approach to describing dynamic gestures.

In [6], an object tracking algorithm, the Hausdorf object tracker, is employed to extract frames from a live video feed that best encapsulate an object's translation. This is represented as a motion vector, from which a variety of static and dynamic features are extracted for classification.

Emil M.P et al. [7] delve into feature extraction from gesture trajectories for dynamic hand gesture recognition. Their approach involves tracking the Center of Gravity (COG) movement and extracting feature vectors that include parameters like velocity and acceleration. To classify the feature vectors obtained after pre-processing, a prediction method is introduced, allowing classifiers to be trained for predicting unknown feature vectors.

In a comparative study outlined in [8], different classifiers are evaluated, with Multiclass Support Vector Machines (SVM) delivering the best results.

The implementation of Multiclass SVM using the LibSVM library, employing a one-against-one approach, is discussed. The classification process involves a tournament mechanism, and the feature vector is assigned to the class receiving the maximum votes.

3.Problem Statement:-

Developing an effective sign language translator faces several challenges. Firstly, there's a pressing need for better accessibility for the deaf and hard of hearing, as existing solutions often fall short. Convert sign language into spoken or written language. The diversity of sign language dialects and regional variations complicates the creation of a universal translator. Recognizing sign language gestures accurately, especially in noisy environments, presents another problem Achieving high accuracy in converting gestures to text, while considering sign language grammar and syntax, is a central issue. User feedback and real-world testing are necessary to improve system performance, but these require diverse user participation.

4.System Overview: -

The process of gesture recognition typically involves the following steps:

1. Data Acquisition:

Input device: - Devices equipped with sensors capture data related to human movements. Common sensors include cameras, depth sensors, accelerometers, gyroscopes, and more.

Image or Video Input: -In vision-based systems, cameras capture images or videos of the user's movements.

2. Pre-processing: -

Image Processing: -In vision-based systems, the captured images or video frames are processed to extract relevant information. This may involve filtering, noise reduction, and normalization.

Feature Extraction: - Extracting key features from the data is crucial. For instance, identifying key points on a hand, tracking their movement, or calculating angles between joints.

3. Gesture Recognition Algorithm: -

Training the System: To train the system a machine learning model is trained using a dataset of labeled gestures. The model learns to recognize patterns in the data.

Real-time Recognition: - Once trained, the model can predict or classify gestures in real-time. This could involve pattern matching, statistical analysis, or neural networks.

3. Gesture Mapping: -

Mapping Gestures to Command: -The recognized gestures are then mapped to specific commands or actions. For example, a specific hand movement might be associated with "swipe left" or "zoom in."

4. User Interface Integration: -

Interaction with Devices: -The recognized gestures then translate into commands that interact with software or hardware. This could involve controlling a computer, navigating through menus, or interacting with virtual objects.

5. Feedback: -

Feedback to the User: - Providing feedback to the user is essential. This could be Text feedback, confirming that the system has recognized and responded to the gesture.

There are different approaches to gesture recognition, including rule-based methods, machine learningbased methods, and hybrid approaches that combine various techniques.

These methods are capable of handling a wide variety of gestures and adapting to different users.

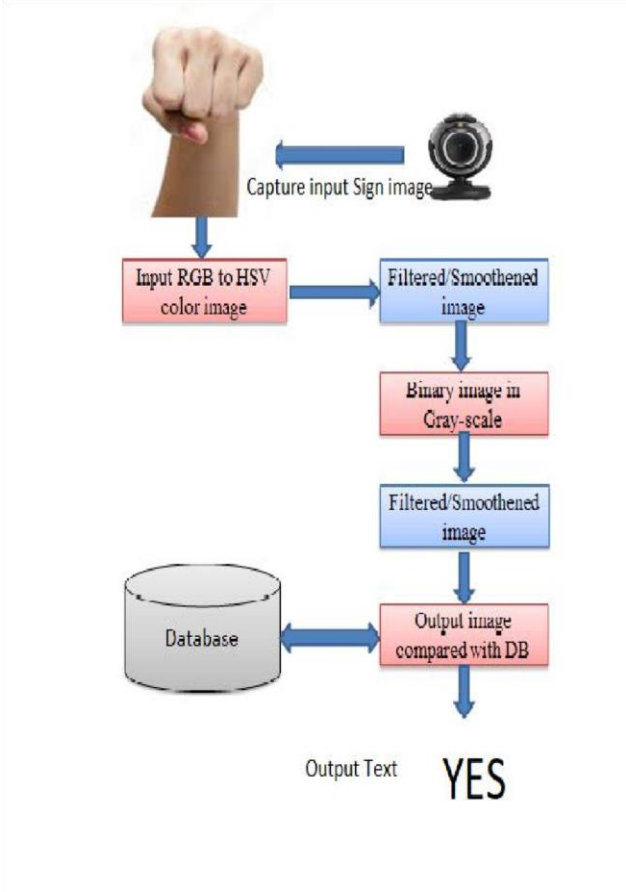

Fig.2:- System Design

The process of recognizing hand gestures relies on a vision-based approach, eliminating the need for external hardware. Users can perform sign language gestures in front of a real-time camera, which detects and simultaneously displays these gestures on the screen. This technique fosters a touch less form of communication between individuals and computers, contributing to the advancement of the HumanComputer Interaction (HCI) field. The general system flow is illustrated in Figure 2.

5.Conclusion & Future Scope: -

The future of sign language translation holds great potential for several exciting developments. Enhancing accuracy and speed in real-time translation remains a priority, and further research into deep learning and AI technologies will likely contribute to these improvements. The adaptation of sign language translation for mobile devices and wearable will increase accessibility. Additionally, addressing the specific needs of regional sign language dialects and cultural variations will be essential for making the technology more inclusive. Integrating sign language translation into educational and professional settings can facilitate better communication. sign language translation is expected to play a significant role in fostering greater inclusivity and accessibility for the deaf and hard of hearing community, transcending barriers in various aspects of their lives.

References:-

[I] Nasser H. Dardas and Nicolas D. Georgians. Realtime hand Gesture detection and recognition using bag-of-features and support vector machine techniques. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, 2011.

[2] Emil M. Petriu Qing Chen, Nicolas D. Georgina's. Real-time visionbased hand gesture recognition using haar-like features, 2007.

[3] P.K. Bora M.K. Bhuyan and D. Ghosh. Trajectory guided recognition of hand gestures having only global motions. International Science Index, 2008. [4] Michael Vorobyov. Shape classification using Zernike moments, 2011.

[5] Thad Eugene Starner. Visual recognition of American Sign Language using hidden

markov models. Master's thesis, Massachusetts Institute of Technology, Cambridge MA, 1995.

[6] P.K. Bora M.K. Bhuyan and D. Ghosh. Trajectory guided recognition of hand gestures having only global motions. International Science Index, 2008.

[7] Emil M. Petriu Qing Chen, Nicolas D. Georganas. Feature extraction from 2d gesture trajectory in dynamic hand gesture recognition, 2006

[8] Manisha M. Ingle Neelam K. Gilorkar. A review on feature extraction for Indian and American Sign Language, 2014.

[9] Chih-Chung Chang and Chih-Jen Lin. A review on feature extraction for Indian and american sign language, 2013.

[10] ATHIRA P K ALEENA K RAJ and DEEPA I K. Sign language conversion software for people with hearing and vocal disabilities, 2013.

[11] Anup Kumar, Karun Thankachan and Mevin M.

[12] Venkatasubramanian S Department of Computer Science and Engineering Saranathan college of engineering, Panjappur, Trichy, India