IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Evaluating The Effectiveness Of Sign Language Translation Tools For Deaf And Hard Of Hearing Individuals

[1]Prof. Akshda Dhakade,[2] Ms. Monika Chavhan,[3] Ms. Khushi Bombde,[4] Ms. Suman Rana,

[1]Assistant Professor, [2]Student, [3]Student, [4] Student

[1] Computer Science and Engineering,

[1]Jawaharlal Darda Institude of Engineering and Technology, Yavatmal, India

Abstract: Sign Language is a non-verbal communication system in which people communicate by visually transferring sign patterns to express their meaning. To express their thoughts, sign language is used widely. As a solution to the problem, this project implements a system to assist deaf and speech-impaired people in capturing their sign-based message via the camera and then converting it into text and audio form. Therefore, it will be very easy for them to communicate with any person at any place and get their work could be done very smoothly. The main purpose of the project is to eliminate the communication gap and to improve the interaction between speech-impaired people and the common people. This project consists of main two modules: 1. Hand gesture recognition 2. Speech recognition. For the recognition of gestures, we are making use of a Convolutional neural network (CNN).

Index Terms **– Convolutional Neural Network(CNN), Gesture recognition, Training, Speech recognition,Machine Learning, Image processing, Local Binary Pattern(LBP)**

1.INTRODUCTION

Lately, hand gesture recognition is gaining great significance in human-computer interaction (HCI) and human-robot interaction (HRI). Different approaches have appeared making use of different detectors and biases. Hand-wearable bias was similar to detector gloves have been used although they're generally precious and stone protrusive. Other protrusive wireless biases like the Wii regulator or seeing rings have appeared to overcome these downsides. Cameras and computer vision have proved to be useful tools for this task. In addition, other contact-free detectors have surfaced recently to descry hand motion and interact with a different bias. Still, despite all the formal work, a reasonable result for the gesture recognition problem has not been set up yet.

Utmost complete hand interactive systems can be considered to be comprised of three layers discovery, shadowing, and recognition. The discovery subcaste is responsible for dining and rooting visual features that can be attributed to the presence of hands in the Head of view of the camera(s). The shadowing subcaste is responsible for performing temporal data association between consecutive image frames, so that, at each moment in time, the system may be apprehensive of what is where". also, in model- grounded styles, shadowing also provides a way to maintain estimates of model parameters, variables, and features that aren't directly observable at a certain moment in time. Last, the recognition subcaste is responsible for grouping the spatiotemporal data uprooted in the former layers and assigning the performing groups with markers associated with particular classes of gestures. In this section, exploration on these three depressed subproblems of vision- grounded gesture recognition is reviewed.

The essential end of erecting a hand gesture recognition system is to produce a natural commerce between humans and computers where the honored gestures can be used for controlling a robot or conveying meaningful information. How to form the performing hand gestures to be understood and well interpreted by the computer is considered as the problem of gesture commerce. mortal- computer commerce(HCI) also named Man Machine Interaction(MMI) refers to the relationship between the mortal and the computer or further precisely the machine, since the machine is insignificant without suitable application by the mortal. Two main characteristics should be supposed when designing an HCI system as mentioned functionality and usability. System functionality appertained to the set of functions or services that the system equips to the druggies, while system usability appertained to the position and compass that the system can operate and perform specific stoner purposes efficiently. The system that attains a suitable balance between these generalities is considered as influential performance and important system. Gestures are used for communicating between humans and machines as well as between people using sign language.

Since sign language is used for interpreting and explanations of a certain subject during a discussion, it has entered special attention. A lot of systems have been proposed to fete gestures using different types of sign languages. For illustration(8) honored American subscribe Language ASL using boundary histogram, MLP neural network, and dynamic programming matching. honored Japanese sign language JSL using Recurrent Neural Network, 42 rudiments, and 10 words. honored Arabic subscribe language ArSL using two different types of Neural Networks, incompletely and Completely Recurrent neural Networks.

Hand gesture Discovery systems defy numerous challenges. Variations in the illumination similar as lighting condition affects at the time of rooting the hand skin region. If the hand region is rotated in any direction or any other objects in the scene mingled with the hand to be detected may disrupt the total performance of the system. This section carries a performance evaluative study on the colorful gesture discovery systems bandied in the former sections. Indeed though the accelerometer- grounded and the glove- grounded approach produced good results in some scripts successfully, the Kinect detector(1, 3) grounded approach is set up effective in implanting in the vehicular robotization field. In the first two styles mentioned then, the motorist or the stoner has to wear a glove or a pack of detectors in hand, which may seriously affect one's driving. thus vision- grounded hands-free gesture discovery systems are a better option for vehicular operations and hence the Kinect detector- grounded system is the stylish suit. Kinect detector shows good robustness to cluttered Backgrounds and also deformations and hand variations while acquiring the input image. The glove- grounded system may be stoner dependent as the glove once designed may be unhappy for another stoner that hand size differs from stoner to stoner. Also, it takes time to put on or take off which also may intrude with the driving and reduce driving performance. In an accelerometer- grounded approach, the stoner must attach further no of detectors to the hand or arm thus such a system offers veritably limited movements of the hands. Also, delicacy can be a problem because the variations passed in detector affair due to temperature and the earth's glamorous field. The functional range is set up as some measures in the case of wireless detectors of this type and the range is limited up to the string length in wired grounded approach

2.LITERATURE SURVEY

Since the 1960s, translators for sign language have been in development. The first focus was on creating mechanical or electronic systems that could recognise and transform hand gestures into spoken or written words. The functionality and accuracy of these early devices were frequently constrained, and they frequently needed the user to wear sensors or other equipment in order to record their motions.

Modern sign language translators have become more complex as a result of developments in computer vision and machine learning. These technologies employ cameras to record the user's signing and interpret the user's gestures by observing how their hands and bodies move. They then provide spoken or written language output that hearing people can understand.

The community of the deaf uses Sign Language (SL) as their primary form of communication. The percentage of deaf people worldwide is around 2.5% [3]. Over 10 million individuals in India are hard of hearing, and 4 million people are deaf, according to studies by The All India Federation of the Deaf. Every nation has its own Sign Language that has been established with a variety of grammatical variations.

Indian Sign language is used by more than one million adults and 0.5 million children in India. Starting in 1978, ISL has been studying language. ISL is a wholly natural language with its own morphology, phonetics, grammar, and syntax. ISL uses postures for the head, face, arms, hands, and eyes to convey language information. ISL consists of both manual and automatic parts. A manual component is defined by characteristics like shape, orientation, position, and movement of the hands, whereas a non-manual component is defined by facial expressions, eye contact, and posture of the head and body.

The effectiveness of sign language interpreters has been studied, and the findings have been varied. While some studies have found that these systems are quite accurate and can convert sign language into spoken or written language with great accuracy, others have discovered that they are prone to errors and can have trouble identifying particular motions or signals.

I.DEAF MUTE COMMUNICATION INTERPRETER- A REVIEW[1] :

The purpose of this paper is to discuss the many current approaches of deaf-mute communication translator system. Wearable communication devices and online learning systems are the two primary categories of communication approaches employed by the deaf-mute. There are three types of wearable communication systems: glove-based, keypad-based, and Handicom touch-screen. Each of the three approaches stated above uses a combination of sensors, an accelerometer, an appropriate micro-controller, a text-to-speech module, a keypad, and a touch-screen. The second option, an online learning system, can eliminate the requirement for an external device to translate messages between a deaf-mute and non-deaf-mute person. The Online Learning System employs a variety of techniques. The five subdivided techniques are TESSA, Wi-See Technology, SWI_PELE System, SLIM Module, and Web-Sign Technology.

II. AN EFFICIENT FRAMEWORK FOR INDIAN SIGN LANGUAGE RECOGNITION USING WAVELET TRANSFORM[2] :

The suggested ISLR system is a pattern recognition technique with two major modules: feature extraction and classification. To recognise sign language, discrete wavelet transform (DWT)-based feature extraction and closest neighbour classifier are combined. The experimental findings demonstrate that the suggested hand gesture recognition system, when using a cosine distance classifier, achieves a maximum classification accuracy of 99.23%.

III. HAND GESTURE RECOGNITION SYSTEM FOR DUMB PEOPLE [3]:

By employing digital image processing, authors demonstrated their technique for recognising static hand gestures. SIFT technique is used to the feature vector for hand gestures. The edges that are resistant to scaling, rotation, and noise addition are where the SIFT features have been computed.

IV. DESIGN ISSUE AND PROPOSED IMPLEMENTATION OF COMMUNICATION AID FOR DEAF & DUMB PEOPLE [4]:

The author of this research presented a method where hand gestures will be translated into relevant text messages to help deaf and dumb persons communicate using Indian sign language (ISL) with regular people. The main goal is to create an algorithm that can instantly transform dynamic motion to text. The technology will finally be deployed on the Android platform and made available as an application for smart phones and tablet computers after testing is complete.

V. HAND GESTURE RECOGNITION FOR SIGN LANGUAGE RECOGNITION : A REVIEW[5] :

The authors provided a variety of hand gesture and sign language recognition techniques that had been previously developed by a number of scholars. Sign language is the only means of communication for the dumb and deaf. These physically disabled people communicate their feelings and thoughts to others by using sign language.

VI. HAND GESTURE RECOGNITION FOR SIGN LANGUAGE USING 3D CNN [6]:
In this paper, the application of 3DCNN for hand gesture identification is discussed. Six other cutting-edge approaches from the literature were contrasted with the strategies the authors suggested. They outperformed four of them and fared similarly to the other two approaches. It does not, however, function for a live video broadcast.

VII. DEEP CONVOLUTIONAL NEURAL NETWORKS FOR SIGN LANGUAGE RECOGNITION[7] :
The authors proposed a CNN architecture to categorise selfie sign language movements. With the suggested CNN design, there is reduced training and validation loss reported, but the database is not made available to the general population.

VIII. A DEPTH-BASED INDIAN SIGN LANGUAGE RECOGNITION USING MICROSOFT KINECT[8] :
Using this technique, overlapping signs, double-handed signs, and ISL-specific signs may all be recognised. The benefit of this research is that by using this strategy, the average recognition accuracy was raised to 71.85%. For a handful of the signals, the system reached 100% accuracy, but it frequently produces inaccurate translations since it doesn't take the context of gestures into account.

IX. SIGNPRO-AN APPLICATION SUITE FOR DEAF AND DUMB [9]:
The author presented a sign language communication programme that enables the deaf and dumb to interact with the rest of the world. The system's real-time gesture to text conversion is its core component. Gesture extraction, gesture matching, and speech conversion are among the processing phases. Several image processing approaches, including histogram matching, bounding box computation, skin colour segmentation, and region growing, are used in the extraction of gestures. Gesture matching methods include correlation-based matching and feature point matching. Text to gesture conversion and text to voice out of text are some of the other functions in the application.

X. OFFLINE SIGNATURE VERIFICATION USING SURF FEATURE EXTRACTION AND NEURAL NETWORKS APPROACH[10] :
This research proposes off-line signature detection & verification using neural network, where the user is shown the signature as an image.

3.RESEARCH METHODOLOGY

On collected dataset, we divided our approach to attack the bracket problem into three stages. The first stage is to member the skin part from the image, as the remaining part can be regarded as noise w.r.t the character bracket problem. The alternate stage is to prize applicable features from the skin segmented images which can prove importance for the coming stagei.e literacy and bracket. The third stage as mentioned over is to use the uprooted features as input into var- ious supervised literacy models for training and also eventually use the trained models for bracket.

Image Segmentation

Training on skin segmented datasets
We used the skin segmentation dataset which containing about 2,00,000 points for training using learning algorithms like SVM and Random Forest. The trained models are then used to segment out the non-skin classified pixels. As describing our own features may not result in higher efficiency, we started with SIFT(Scale Inverse Feature Transform)features as it computes the key points in the image which is more apt than describing features manually. So, after the skin segmented images were obtrained using the YUV-YIQ model, we used the following approaches for extracting feature vectors.

Machine Learning on Feature Vectors
But before we obtained those best results, we explored the following algorithms on the obtained feature vectors. Multiclass SVM with a linear kernel was used with almost every feature vector. Overall the following approaches were tried.

Support Vector Machines

Multiclass SVMs were tried on all the feature vectors. Results obtained with linear kernel and four-fold Cross Validated accuracies are reported for all feature vectors. The confusion matrices shown in the results sections below correspond to the different techniques tried using linear kernel Multi Class SVMs. The best accuracies were observed for this algorithm. Or try with the RBF kernel failed miserably on HOG feature vectors as only 4.76% accuracy was observed.

Random Forest

Since this was a 26-class problem, we tried our luck at Random Forest with feature vectors on the compressed images. It fell a little short of the Multiclass SVM with 46.45% 4-fold CV accuracy. Our interpretation is that implementation of decision trees in Python is monothetic which may have led to the formation of rectangular regions whereas the actual boundaries may not have been rectangular.

Hierarchical Classification

One of the major reasons for the comparatively lower accuracy is a large number of classes(26). So one approach we tried was breaking the classification problem into multiple levels or hierarchical classification. First, we train a linear kernel SVM model to classify alphabets as one-handed or two-handed.
This model was performed with 95% accuracy. Then we trained linear kernel Multiclass SVM models to classify the one-handed alphabets(56% accuracy) and two-handed alphabets(60% accuracy) and then put the system together. An alphabet is first classified as one-handed or two-handed, and then depending on the classification is put into the corresponding model and given a label. Even though the individual models performed better than the direct multi-class SVM on HOG features, over all the performance was nearly the same, and four-fold CV accuracy of 53.23% was observed.

4. PROPOSED METHOD

Following is a flowchart of the proposed method for extracting gestures and convert into text

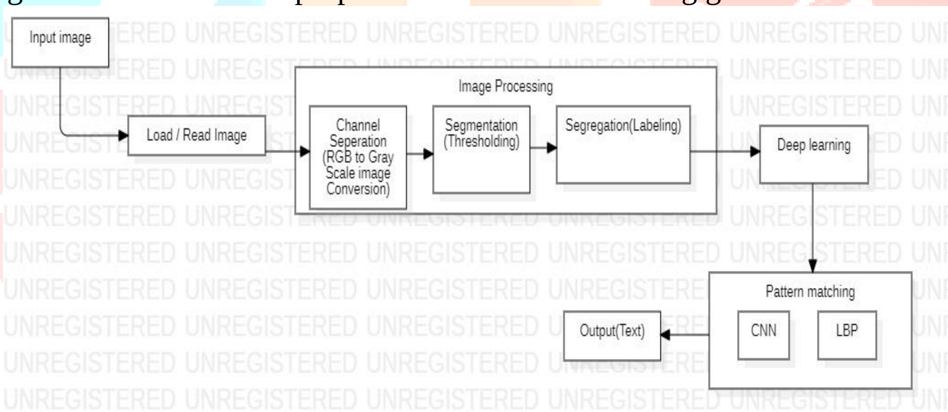

Fig : Block diagram of Proposed Method

A.Image Capture

This is the first step in sign recognition. Camera interfacing is a very critical part. Any web camera is used for capturing the hand gesture. Now web camera is also in built-in laptops & one can use an external camera for interfacing. But captured images need to be in high desciption. So the selection of a good webcam & its interfacing is an important task of this method.

B.Image Preprocessing

Image preprocessing contains cropping, filtering, brightness &contrast adjustment & many more. To do such a process Image enhancement, Image cropping & Image Segmentation methods are used. Originally the captured Images are in the form of RGB. So the first step is to convert RGB images to binary images then cropping of the image is to be done so that unwanted parts of images can be removed. And now enhancement can be done in a certain selected area. In Image segmentation, the Edge detection method is used which can detect the boundary of cropped images which is further used for the feature extraction method.

C.Feature Extraction

Feature extraction is a very useful step to create a database of sign recognition. To characterize the different visual principles of letters in the homemmade ABC efficiently and effectively, both the global visual features and the original visual features are extracted for letter image similarity characterization There are mainly two types of feature extraction methods involved in sign recognition, First is Contour-based shape representation and description methods &I another is Region-based shape representation and description methods. Among those depending on application methods are selected. [1]In this proposed method, the 7Hu moment technique is used & from that 7 moments are found. The database of gesture has been made from those moments.

5. FUTURE WORK :

In order to provide accessibility and communication for those who are deaf or hard of hearing, sign language recognition is a crucial area of study. Future research in this field has a number of potential directions, including:

AIMING TO INCREASE RECOGNITION ACCURACY: Complex or nuanced signs can still be difficult for current sign language recognition algorithms to correctly identify. Future research could concentrate on increasing recognition accuracy by creating more complex algorithms and training models on larger and more varied datasets.

EXPANDING THE SET OF SIGNS THAT CAN BE RECOGNISED: Many existing sign language recognition systems can only now identify a small set of signs, which can limit their applicability in real-world scenarios. The goal of future research might be to increase the number of signs that can be recognised, potentially by creating systems that can distinguish between signs in many sign languages or regional variants of a single sign language.

Real-time sign language recognition is necessary to allow sign language users and non-signers to communicate in real time. Future research might concentrate on creating systems that can instantly translate signs into text or speech and recognise them in real-time.

Communication between sign language users and sign language recognition systems can be significantly hampered by sign language recognition systems' difficulty in recognising signs in noisy surroundings. The development of systems that can recognise signs in noisy surroundings may be the focus of future research, maybe using more sophisticated sensors or noise reduction techniques.

Sign language recognition in specialised fields: specialised fields like healthcare or education can benefit greatly from sign language recognition. The development of sign language recognition systems that are customised to certain domains, potentially by including domain-specific signals or terminology, could be the main goal of future work.

6. CONCLUSION :

In conclusion, sign language translation is a crucial field of study and development with the potential to significantly enhance communication for the deaf and hard-of-hearing populations. Technology has made tremendous advancements in sign language translation and recognition systems, which can aid in bridging the communication gap between hearing and deaf people.

The complexity of sign language makes it challenging to effectively capture and translate all of its intricacies, thus there is still considerable work to be done in this area. Sign languages vary considerably across different countries and civilizations. Therefore, to further enhance the accuracy and utility of sign language translation technology, continued research and development will be required.

Overall, sign language translation is a crucial tool for promoting inclusivity and accessibility for the deaf and hard-of-hearing communities and has the potential to significantly enhance communication and understanding amongst persons of various abilities.

REFERENCES :

[1] Sunitha K. A, Anitha Saraswathi.P, Aarthi.M, Jayapriya. K, Lingam Sunny, "Deaf Mute Communication Interpreter- A Review", International Journal of Applied Engineering Research ,Volume 11, pp 290-296 , 2016.

[2] Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, " An Efficient Framework for Indian SignLanguage Recognition Using Wavelet Transform" Circuits and Systems, Volume 7, pp 1874- 1883, 2016.

[3] Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people",International Journal of Science and Research (IJSR).

[4] Nakul Nagpal,Dr. Arun Mitra.,Dr. Pankaj Agrawal, "Design Issue and Proposed Implementation of Communication Aid for Deaf & Dumb People", International Journal on Recent and Innovation Trends in Computing and Communication ,Volume: 3 Issue: 5,pp- 147 – 149.

[5] Pratibha Pandey, Vinay Jain, "Hand Gesture Recognition for Sign Language Recognition: A Review", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 3, March 2015 .

[6] M. Al-Hammadi, G. Muhammad, w. Abdul, M. Alsulaiman, M. A. Bencherif, M. A. Mekhtiche, "Hand Gesture Recognition for Sign Language Using 3DCNN", 2020.https://ieeexplore.ieee.org/abstract/document/9078786.

[7] G. A. Rao, K. Syamala, P. V. V. Kishore, A. S. C. S. Sastry, "Deep Convolutional Neural Networks for Sign Language Recognition", 2018. https://ieeexplore.ieee.org/abstract/document/8316344.

[8] T. Raghuveera, R. Deepthi, R. Mangalashri, R. Akshaya, "A depth-based Indian Sign Language recognition using Microsoft Kinect", 2020.https://link.springer.com/article/10.1007/s12046-019-1250-6.

[9] Ashish Sethi, Hemanth ,Kuldeep Kumar,Bhaskara Rao ,Krishnan R, "Sign Pro-An Application Suite for Deaf and Dumb", IJCSET , Volume 2, Issue 5, pp-1203-1206, May 2012.

[10] Priyanka Sharma,"Offline Signature Verification Using Surf Feature Extraction and Neural Networks Approach", International Journal of Computer Science and Information Technologies, Volume 5 (3) , pp 3539-3541, 2014.