ISSN 2278-3091
Volume 11, No.4, July – August 2022
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse021142022.pdf
https://doi.org/10.30534/ijatcse/2022/021142022

A Review Paper on Real Time Sign Language Detection for the Deaf and Dumb

Anirudh Muppidi, Amar Thodupunoori, Lalitha
1 B.Tech Graduate (IV year), Department of Information Technology, SNIST, Hyderabad
2 B.Tech Graduate (IV year), Department of Information Technology, SNIST, Hyderabad
3 Assistant Professor, Department of Information Technology, SNIST, Hyderabad

Received Date : June 12, 2022 Accepted Date : July 17, 2022 Published Date : August 06, 2022

ABSTRACT

Hand gesture is one of the styles used in sign language for non-verbal communication. It's most generally used by deaf & dumb people who have hearing or speech problems to communicate among themselves or with normal people. Plathora sign language systems had been developed by numerous makers around the world but they're neither flexible nor cost-effective for the end druggies. Hence, it's a software which presents a system prototype that's suitable to automatically recognize sign language to help deaf and dumb people to communicate more effectively with each other or normal people. Dumb people are generally deprived of normal communication with other people in the society, also normal people find it hard to understand and communicate with them. These people have to depend on an interpreter or on some kind of visual communication. An interpreter won't be always available and visual communication is substantially delicate to understand. As a normal person is ignorant of the grammar or meaning of numerous gestures that are part of a sign language, it's primarily limited to their families and/ or deaf and dumb community.

Key words: Hand gesture, Sign language, Communication, ,OpenCV, ANN, CNN

1. INTRODUCTION

Sign language is the mode of communication which uses visual ways like expressions, hand gestures, and body movements to convey meaning. sign language is extremely helpful for people who face difficulty with hearing or speaking. sign language recognition refers to the conversion of these gestures into words or rudiments of being formally spoken languages. Therefore, conversion of sign language into words by an algorithm or a model can help bridge the gap between people with hearing or speaking impairment and the rest of the world.

current exploration in computer vision and machine learning. Being a natural way of human communication, it's an area where numerous experimenters are working on, with the aim of making human computer interaction (HCI) easier and natural, without the need for any additional gadgets. So, the primary aim of gesture recognition research is to build systems, which can identify specific human gestures and use them, for illustration, to convey information. For that, vision- based hand gesture interfaces need fast and extremely robust hand detection, and gesture recognition in real time. Hand gestures are a important human communication modality with lots of implicit operations and in this context, we've sign language recognition, the communication system of deaf people. One of its primary targets is to build systems, which can identify specific gestures and use them to convey information or to control a device. Though, gestures need to be modelled in the spatial and temporal disciplines, where a hand posture is the stationary structure of the hand and a gesture is the dynamic movement of the hand. There are two types of approaches for hand gesture recognition vision- based approaches and data glove approaches. This work main focus is on creating a vision- based system suitable to do real- time sign language recognition. The reason for choosing a system based on vision relates to the fact that it provides a simpler and further intuitive way of communication between a human and a computer. Being hand- gesture one of the most important communication tools in human's day-to-day life, and with the nonstop advances of image and videotape processing ways, research on human- machine communication through gesture recognition led to the use of such sort of technology in a huge broad range of operations, like touch screens, video game consoles, virtual reality, medical applications, and sign language recognition. Although sign language is the most natural way of swapping information among deaf people it has been observed that they're facing difficulties with normal people communication. sign language consists of vocabulary of signs in exactly the same way as spoken language consists of a vocabulary of words. sign languages aren't standard and universal and the grammar differ from country to country.

Sign language is a visual language with three main components:

Table 1: Shows 3 main components of sign language

Fingerspelling	Word level sign vocabulary	Non-manual features
Used to spell words letter by letter .	Used for the majority of communication.	Facial expressions and tongue, mouth and body position.

2. OBJECTIVES

The sign Language Recognition Prototype is a real- time vision- based system whose purpose is to detect the American sign Language given in the alphabet of Fig. 1. The purpose of the prototype was to test the validity of a vision- based system for sign language recognition and at the same time, test and elect hand features that could be used with machine learning algorithms allowing their applications in any real- time sign language recognition systems.

The implemented algorithm uses only one camera, and is grounded on a set of hypotheticals, hereby defined

1. The speaker must be within a defined border area, in front of the camera.
2. The speaker must be within a defined distance range, due to camera limitations.
3. Hand gesture is defined with a bare hand and not clotted by other objects.
4. The system must be used in closed areas, since the named camera doesn't work well under sun light conditions.

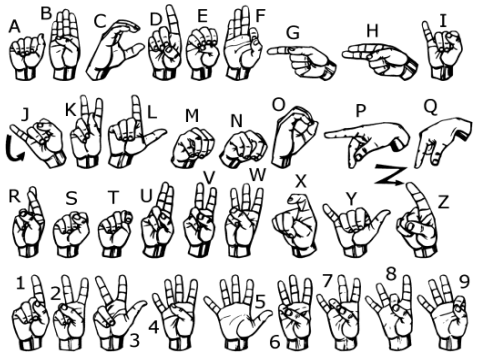

Figure 1: Shows various signs

The proposed system design, which consists of two modules, i.e., data acquisition, pre-processing and feature extraction and sign language gesture classification.

3. LITERATURE SURVEY

The researches done in this field are substantially done using a glove based system. In the glove based system, detectors similar as potentiometer, accelerometers etc. are attached to each of the finger. Grounded on their readings the corresponding alphabet is displayed. Christopher Lee and Yangsheng Xu developed a glove-based gesture recognition system that was suitable to recognize 14 of the letters from the hand alphabet, learn new gestures and suitable to modernize the model of each gesture in the system in online mode. Over the times advanced glove gadgets have been designed similar as the Sayre Glove, Dexterous Hand Master and Power Glove. The main problem faced by this gloved based system is that it has to be recalibrate every time whenever a new speaker on the finger-tips so that the fingertips are linked by the Image Processing unit. We are enforcing our design by using Image Processing. The main advantage of our design is that it is not confined to be used with black background. It can be used with any plane background. Also wearing of color bands isn't needed in our system. By that, securities could authorize an existent's identity depending on "who she is", and not "what she has" and "what she could flash back". Two main classes can be set up in biometrics:

- Physiological – It is associated with the body shape, includes all physical traits, iris, palm print, facial features, Fingerprints, etc.
- Behavioral – Affiliated to the behavioral characteristics of a person. A characteristic extensively used till moment is autographs. Modern methods of behavioral studies are arising similar as keystroke dynamics and voice analysis.

Deaf Mute Communication Interpreter-: This paper aims to cover the numerous prevailing styles of deaf-mute communication interpreter system. The two broad segregation of the communication methodologies used by the deaf –mute people are - Wearable Communication Device and Online Learning System. Under Wearable communication system, there are Glove based system, Keypad system and Handicom Touch-screen. All the over mentioned three sub- divided categories make use of numerous detectors, accelerometer, a suitable micro-controller, a text to speech conversion module, a keypad and a touch-screen. The need for an external device to interpret the communication between a deaf –mute and non-deaf-mute. People can be overcome by the alternate system i.e., online learning system. The Online Learning System has different ways. The five subdivided methods are- SLIM module, TESSA, Wi-See Technology, SWI_PELE System and Web-Sign Technology.

An Effective Framework for American Sign Language Recognition Using Wavelet Tranfigure: The proposed ISLR system is considered as a pattern recognition fashion that has two important modules: feature extraction and classification. The common use of Discrete Wavelet Transform (DWT) grounded feature engineering and nearest neighbor classifier is used to detect the sign language. The experimental results show that the proposed hand gesture recognition system achieves maximum 88.0% classification accuracy while using cosine distance classifier.

Hand Gesture Recognition Using PCA in [3]: In this paper authors presented a scheme using a database driven hand gesture recognition grounded upon skin color model approach and thresholding approach along with an effective template matching with can be effectively used for human robotics operations and such sort of other operations. Originally, hand region is segmented by applying skin color model in YCbCr color space. In the coming stage thresholding is applied to separate focus and background.

Eventually, template based matching fashion is developed using Principal Component Analysis (PCA) for recognition.

Hand Gesture Recognition System For the Dumb People [4]: Authors presented the stationary hand gesture recognition system using digital image processing. For hand gesture pose feature vector SIFT algorithm is used. The SIFT features have been reckoned at the edges which are steady to scaling, gyration, addition of noise.

An Automated System for American Sign Language Recognition in: In this paper a system for automatic recognition of signs on the base of shape-based features is presented. For segmentation of hand region from the images, Otsu's thresholding algorithm is used, that chooses an optimal threshold to minimize the within-class variance of threshold black and white pixels. Features of segmented hand region are calculated using Hu's steady moments that are fed to Artificial Neural Network for classification. Performance of the system is estimated on the base of Accuracy, perceptivity and particularity.

Hand Gesture Recognition for Sign Language Recognition: A Review in [6]: Authors presented numerous system of hand gesture and sign language recognition proposed in the history by various researchers. For deaf and dumb people, Sign language is the only way of communication. With the help of sign language, these physical disabled people express their feelings and ideas to other.

Design Issue and Proposed implementation of Communication Aid for Deaf & Dumb People in [7]: In this paper author proposed a system to help communication of deaf and dumb people communication using Indian sign language (ISL) with normal people where hand gestures will be converted into suitable text message. Main task is to design an algorithm to convert dynamic gesture to text at real time eventually after testing is done the system will be enforced on android platform and will be available as an application for smart phone and tablet pc. **Real Time Detection and Recognition of Indian and American Sign Language Using Sift In [8]:**

Author proposed a real time vision-based system for hand gesture recognition for human computer communication in numerous applications. The system can detect 35 different hand gestures given by Indian and American Sign Language or ISL and ASL at faster rate with righteous accuracy. RGB-to-GRAY segmentation fashion was used to minimize the chances of false recognition. Authors proposed a system of extemporized Scale Invariant Feature Transform (SIFT) and same was used to engineer features. The system is model using MATLAB. To design and effective user-friendly hand gesture recognition system, a GUI model has been enforced.

A Review on Feature Extraction for American Sign Language in: Paper presented the recent exploration and development of sign language grounded on manual communication and body language. Sign language recognition system generally elaborate three steps preprocessing, feature engineering and classification. Classification methods used for recognition are Neural Network (NN), Support Vector Machine (SVM), Hidden Markov Models (HMM), Scale Invariant Feature Transform (SIFT), etc.

Sign Pro-an Application Suite for Deaf and Dumb. in: Author presented operations that helps the deaf and dumb person to communicate with the rest of the world using sign language. The crucial point in this system is the real time gesture to text conversion. The processing proceeding include: gesture extraction, gesture matching and conversion to speech. Gesture extraction involves use of numerous image processing ways similar as ISSN No: 2454- 2024 (online) International Journal of Technical Research & Science pg. 433 www.ijtrs.com www.ijtrs.org Paper Id: IJTRS-V2-I7-005 Volume 2 Issue VII, August 2017 @2017, IJTRS All Right Reserved histogram matching, bounding box computation, skin color segmentation and region growing. Techniques applicable for Gesture matching include feature point matching and correlation-based matching. The other features in the operation include venting out of text and text to gesture conversion.

Offline Sign Verification Using Surf Feature Extraction and Neural Networks Approach: In this paper, off-line hand recognition & verification using neural network is proposed, where the hand is captured and presented to the speaker in an image format.

Entries and input values at a given position. As we continue this process well produce a 2-Dimensional activation matrix that gives the response of that matrix at every spatial position. That is, the network will learn layers that spark when they see some type of visual feature similar as an edge of some exposure or a blotch of some colour.

4. METHODOLOGY

The system is a vision-based approach. All the signs are represented with bare hands and so it eliminates the problem of using any artificial gadgets for communication.

Converting real time sign language into text can be classified into flow of simple steps like:

i. Recognizing a man's or woman's hand motions
ii. Developing a system learning model for picture to textual content translation
iii. Putting words together
iv. Putting sentences together
v. Creating the complete content

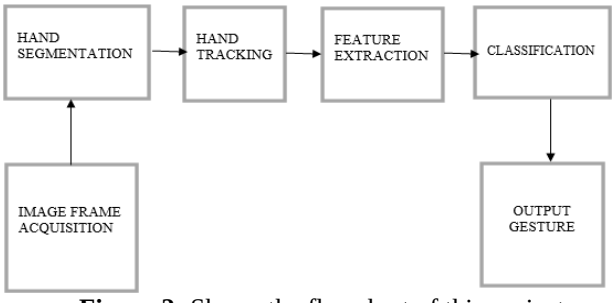

Figure 2: Shows the flowchart of this project

The flow chart of this project can be described as:

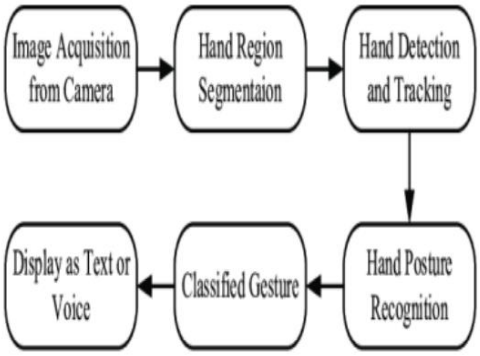

Figure 3: Shows different phases in the project

The figure above shows the phases involved in obtaining the project's goals. The detailed explanation of the above mentioned steps are:

1. Image Acquisition:

The camera is put to use for capturing the motions. The full signing duration is captured with the help of this OpenCV video stream. The frames are taken from the stream and converted to grayscale images with a 50*50 pixel resolution. Because the entire dataset is the same size, this dimension is consistent throughout the project.

2. Hand Region Segmentation & Hand Detection and Tracking:

In the collected photographs, hand gestures are scanned. Before the image is fed into the model for prediction, this is a phase in the preprocessing process. The passages in which gestures are used have been emphasised. This increases the chances of making a successful prediction by a factor of ten.

3. Hand Posture Recognition:

The preprocessed images are sent to the Keras CNN model. The model which was already trained generates the projected label. There is a probability associated with each of the gesture labels. The label which has the highest probability determines the expected label.

4. Display as Text:

The model converts gestures, which are known, into words. The recognized words form a sentence. Therefore, forming the entire context.

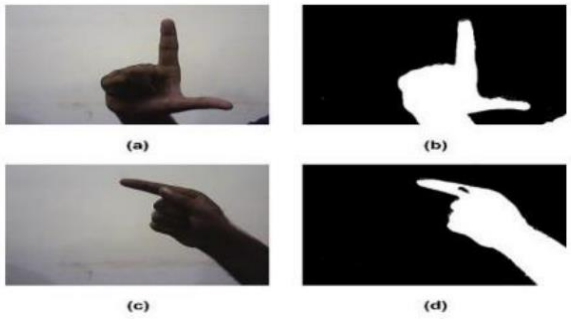

Figure 4: Shows Hand Region Segmentation

Dataset Generation

It is needed to make a proper database of the gestures of the sign language so that the images captured while communicating using this system can be compared. Procedure we followed to produce our data set are as follows. We used Open computer vision (OpenCV) library in order to produce our dataset. Originally, we captured around 13 images of each of the symbol in ASL for training purposes and around 2 images per symbol for testing purpose. First, we capture each frame shown by the webcam of our machine. In each frame we define a region of interest (ROI) which is denoted by a blue bounded forecourt as shown in the image below. From the whole image we uprooted our ROI which is RGB and convert it into labeled Image.

Finally, we apply our single shot detection to our image which helps us extracting various features of our image.

Gesture Classification

The approach which we used for this design is our approach uses two layers of algorithm to predict the final symbol of the speaker.

Labelling Algorithm:

1. Apply labelling made easy for algorithm to recognize the gestures from the pictures.
2. Instead of training whole image algorithm focus mostly on the cropped part of the picture. Hence algorithm will be efficient in duration.
3. And computational power is saved as well.

Algorithm:

1. We used tensorflow object detection algorithm.
2. And we used single shot detector algorithm.
3. Due to labelling it doesn't need a huge dataset. So the training duration will be deduced and space will be used effectively.
4. Collect images for deep learning using your webcam and OpenCV.
5. Label images for sign language detection using LabelImg.
6. Setup Tensorflow Object Detection pipeline configuration.
7. Use transfer learning to train a deep learning model.
8. Detect sign language in real time using OpenCV

Training and Testing

We label our input images (RGB) and create a new xml file which holds details about the labeled part of the image. We apply adaptive threshold to extract our hand from the background and resize our images to 128 x 128. We feed the input images after pre- processing to our model for training and testing after applying all the operations mentioned over.

Challenges Faced

There were numerous challenges faced by us during the project. The foremost issue we faced was of dataset. We wanted to deal with raw images and that too square image as Object detection in Tensorflow as it was a lot more convenient working with only square images. We couldn't find any being dataset for that hence we decided to make our own dataset.

Alternate issue was to select a filter which we could apply on our images so that proper features of the images could be attained and hence also we could give that image as input for SSD model. We tried number of filter including binary threshold, canny edge detection, gaussian blur etc. but finally we settled with single shot detection. Further issues were faced relating to the accuracy of the model we trained in earlier phases which we ultimately bettered by adding the input image size and also by perfecting the dataset.

5. CONCLUSION

In this report, a functional real time vision based American sign language recognition for Deaf and Dumb people have been developed for asl rudiments. We achieved final accuracy of 92.0 on our dataset. We're suitable to better our predictions after enforcing two layers of algorithm in which we corroborate and predict symbols which are more analogous to each other, This way we're suitable to detect nearly all symbols handed that they're shown duly, there's no noise in the background and lighting is acceptable.

6. FUTURE ENHANCEMENTS

We intend to test a variety of background subtraction techniques in order to improve accuracy even while dealing with complex backgrounds. We're also working on improving the preprocessing so that we can better predict gestures in low-light situations.

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] Mandeep Kaur Ahuja, Amardeep Singh, "Hand Gesture Recognition Using PCA", International Journal of Computer Science Engineering and Technology (IJCSET), Volume 5, Issue 7, pp. 267-27, July 2015.

[4] Sagar P.More, Prof. Abdul Sattar, "Hand gesture recognition system for dumb people",

[5] International Journal of Science and Research (IJSR)

[6] Chandandeep Kaur, Nivit Gill, "An Automated System for Indian Sign Language Recognition", International Journal of Advanced Research in Computer Science and Software Engineering.

[7] 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.

[8] 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,pp147 – 149.

[9] S. Shirbhate1, Mr. Vedant D. Shinde2, Ms. Sanam A. Metkari3, Ms. Pooja U. Borkar4, Ms. Mayuri A. Khandge/Sign-LanguageRecognition-System.2020 IRJET Vol3 March,2020.

[10] T. Yang, Y. Xu, and "A. , Hidden Markov Model for Gesture Recognition", CMU-RI-TR-94 10, Robotics Institute, Carnegie Mellon Univ.,Pittsburgh,PA, May 1994.

[11] Pujan Ziaie, Thomas M¨uller , Mary Ellen Foster , and Alois Knoll"A Na¨ive Bayes Munich,Dept. of Informatics VI, Robotics and Embedded Systems,Boltzmannstr. 3, DE-85748 Garching, Germany.

[12]https://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_bilateral_filter/gausian_median_blur_bilateral_filter.html

[13] Mohammed Waleed Kalous, Machine recognition of Auslan signs using PowerGloves: Towards large-lexicon recognition of sign language.

[14]aeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-ConvolutionalNeural-Networks-Part-2/

[15]http://www-i6.informatik.rwth-aachen.de/~dreuw/database.php

[16] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cham

[17] Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision based features. Pattern Recognition Letters 32(4),572–577 (2011) 46

[20] N. Mukai, N. Harada and Y. Chang, "Japanese Fingerspelling Recognition Based on Classification Tree and Machine Learning," 2017 Nicograph International (NicoInt), Kyoto, Japan, 2017, pp. 19-24. doi:10.1109/NICOInt.2017.9

[21] Byeongkeun Kang , Subarna Tripathi , Truong Q. Nguyen "Real-time sign language fingerspelling recognition using convolutional neural networks from depth map" 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)