International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8516
***Article Received:** 31 July 2023 **Revised:** 28 September 2023 **Accepted:** 15 October 2023*

Hast Mudra: Hand Sign Gesture Recognition Using LSTM

Riddhi Mirajkar[#a], Pravin Futane[#a], Priya Shelke [#a] , Onkar Parihar [#b], Ujwal Gawali[#b] , Sahil Badgujar[#b] , Rohit Patil[#b],

[a-b] Department of Information Technology, VIIT, Pune-48, India

[a-b]{ riddhi.mirajkar, pravin.futane, priya.shelke, onkar.22120239, ujwal.22120052, sahil.22120193, rohit.22120026, } @viit.ac.in

Abstract—Even though using the most natural way of communication is sign language, deaf and mute people find it challenging to socialize. Alanguage barrier is erected between regular people and D&M individuals due to the structureof sign language, which is distinct from text. Theyconverse by using vision-based communication as a result. The gestures can be easily understoodby others if there is a standard interface thattransforms sign language to visible text. As a result, R&D has been done on a vision-based interface system that will allow D&M persons tocommunicate without understanding one another's languages. In this project, first gathered and acquired data and created a dataset, after which extracting useful data from the images. Keywords After verification and trained data andmodel using the (LSTM) algorithm, TensorFlow,and Keras technology, classified the gestures according to alphabet. Using our own dataset, this system achieved an accuracy of around 86.75% inan experimental test. system uses the (LSTM) algorithm to process images and data.

Keywords-Long-Short term memory (LSTM), TensorFlow, Keras, Sign language, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deaf and mute (D&M).

I. INTRODUCTION

Deaf and mute people communicate most naturally through sign language, but it has been noted that, they have trouble interacting with hearing people. Sign language is a mode of communication that conveys meaning through visual means such as expressions, hand gestures, and body movements.The use of sign language by those who have trouble hearing or speaking can be very beneficial. Sign language incorporates existing spoken language into hand gestures and handexpressions. A sign language recognition systemis made up of a quick, accurate way to translate sign language into written or spoken language spoken language Deaf and mute persons tend to communicate the most naturally using sign language, although it has been noted that they findit challenging to communicate with hearing people.

For communication between regular people and D&M individuals, the structure of sign language,which is different from that of standard text, creates a linguistic barrier. They therefore rely on interaction through vision-based communication.

If there is a standard interface that transforms sign language to text, the movements can be easily understood by others. In order to enable D&M persons to communicate without understanding one another's languages, research has been done on a vision-based interface system.

Because D&M people is only Due to acommunication-related handicap, the only way is through sign language. of communication available to them. The act of communicating ideas and messages through a variety of techniques, including speech, signals, behavior, and writing, is known as communication. People who are deaf and mute (D&M) communicate withothers by making various gestures with their hands. Gestures are nonverbal communicationtools that are interpreted visually. Sign language refers to deaf and dumb people's nonverbal communication.

II. LITERATURE SURVEY

Used static, manual signs, alphabets, and numerals. ASL lexicon video dataset, sign stream,Weizmann face database, Yale b frontal, digital still camera, and specially designed input devices such as CyberGlove®, Sensor glove, and Polhemus FASTRACK were used. In SL Detection, they achieved an accuracy of around 82.20% [1].

Using technologies such as face detection, removal, and HSV thresholding, and for SL detection, they used the CNN method, with an accuracy of around 93% in experimental tests [2].

The author have used multilayer perceptron neural network. In addition, they used marcel static hand posture database and they used two classifiers forthe system. neural network of Raw Features Classifier accepted 3072 specialties as input. Histogram Features Classifier received 512 specialties as input. They trained the model with backpropagation system. With all of these combined they achieved around 70 to 85% positive result [3].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8516
***Article Received:** 31 July 2023 **Revised:** 28 September 2023 **Accepted:** 15 October 2023*

A variety of traditional technologies in conjunction with AI and NLP. With the process of data acquisition, features extraction, gesture classification and verification, they also used CNN and ANN technology for the extraction, cross validation and verification. They did not specified which system they used for training themodule. With the help of all this various methodsauthor achieved an accuracy of 91.1% in their experimental test [4].

Technologies such as NLP, CNN methodology, and Sift IN to create models in MATLAB. The author also created their own database. They useddouble layer of algorithm in which they can verifyand forecast which symbols are having close connection with each other. With the help of this technique, they were able to forecast all the symbols with the environment having no noise inthe background and adequate lighting. They achieved an accuracy of 92% in their experimental tests [5].

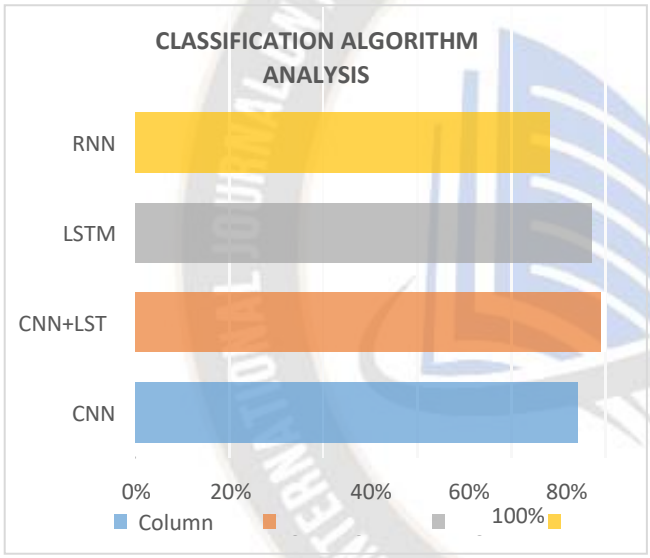

Figure 1. Classification Algorithm analysis

In this project comparing to paper no [1],[2],[3],[4],[5] used (LSTM) algorithm for feature extraction and verification and tensor flow, keras for training the model for detection the language [Figure 1]. Achieved the accuracy ofaround 86.75% with the help of our own created dataset.

Also shown the comparison between all the Paper's algorithm and technologies in the [Figure.2]

Sr. No.	Paper Studied (Year)	Methodology	Limitations
1	**SIGN LANGUAGE RECOGNITION-STATE OF THE ART**:-Ashok K Sahoo, Gouri Sankar Mishra and Kiran Kumar Ravulakollu (2014)	Used static, manual signs, alphabets and numerals in their project. They used ASL lexicon video dataset, sign stream, Weizmann face database, yale b frontal, digital still camera and specially designed input devices like CyberGlove, Sensor glove and polhemus FASTRACK. They got the accuracy of around 82.20% in SL Detection.	There is no standard dataset for various countries. System is slower.
2	**SIGN LANGUAGE RECOGNITION**-Anup Kumar, Karun Thankachan and Mevin M. Dominic(2016)	Used technologies like face detection, removal, HSV thresholding for pre-processing and for SL detection they used CNN method and they got accuracy of around 93% in experimental test.	Can notrun under robust and unfavourable environment (lot of clutter and poor lighting).
3	**SIGN LANGUAGE RECOGNITION**-Tülay Karaylian, Özkan KÖlÖç (2017)	The author have used ANN(Artificial Neural Network) technology and has achieved an accuracy in the range of around 70% to 85%.	The recognition rate can be increased by improving processing an image step.
4	**REVIEW OF THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN SIGN LANGUAGE RECOGNITION SYSTEM** - Oyeniran, Oluwashina A, Oyeniyi, Joshua O, Sotonwa, Kehinde A, Ojo, Adeolu O.(2020)	The author have used various traditional technologies while using AI and NLP with them and achieved accuracy of 91.1% in their experimental test.	Using different algorithms increased time, also author didn't specified which system is used for training the module
5	**A REVIEW PAPER ON SIGN LANGUAGE RECOGNITION FOR THE DEAF AND DUMB**-R Rumana, Reddygari Sandhya Rani, Mrs. R. Prema (2021)	The author have used technologies like NLP, CNN methodology, Sift IN. they have created models using MATLAB and have achieved accuracy of 92% in their experimental test they have conducted.	Tried to include binary threshold and canny edge detection filter but failed to do.

Figure 2. Literature Survey

III. PROPOSED SYSTEM

The important goal of project is to create a modelthat can recognize alphabet-based hand gestures. Because D&M people's only Due to acommunication-related handicap, sign language is the only means of communication available to them. Communication is the act of exchanging thoughts and messages using a variety of behaviors, signs, and written language. People who are deaf and mute (D&M) interact with others by making various hand signals. Nonverbal cues like gestures can be deciphered by visual perception. Sign language is the nonverbalcommunication of the deaf and dumb.

System architecture is consist of 4 stages which isshown in [Figure. 3]

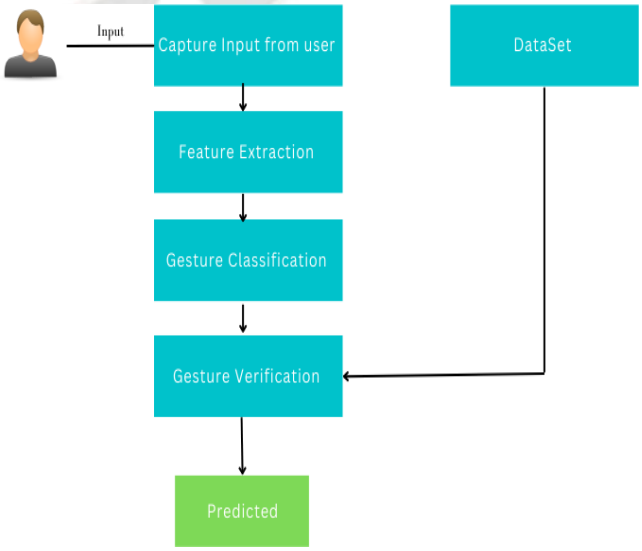

Figure 3. Sign Language Detection Architecture

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8516
***Article Received:** 31 July 2023 **Revised:** 28 September 2023 **Accepted:** 15 October 2023*

a. Data Acquisition

During the data acquisition stage, observing hand information using a computer camera as aninput device. captured around 4500-5000 pictures for all the static alphabets. Because of this data set our sign recognition accuracy has significantly improved. In this stage, created a dataset of alphabets from A to Z using the Python OpenCV library.

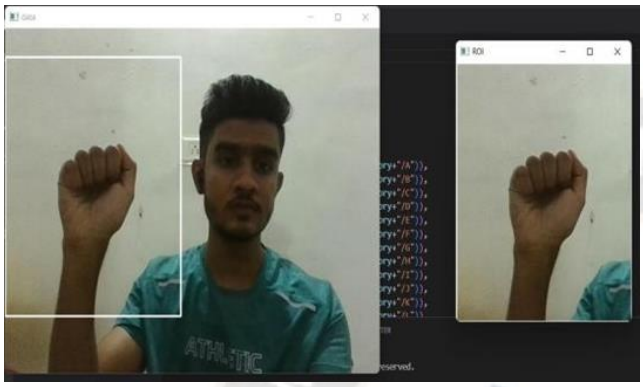

Figure 4. Data acquisition

b. Feature Extraction

Following the creation of a dataset, extract the landmark from the hand in order to recognize thatalphabet. Used the (LSTM) algorithm for extraction.

Following the discovery of a landmark by hand Feature extraction plots, shown in [Figure.5] the hand and then passes it to (LSTM).

Figure 5. Feature Extraction

c. Gesture Classification

The (LSTM) algorithm is used to train the datasetin Gesture classification.

The long short-term memory (LSTM) this type of artificial neural network used in both deep learning and machine intelligence. Shown in [Figure. 6] (LSTM) has feedback connections as compared with conventional cascaded neural networks. This type of recurrent neural network (RNN) can process not only but particular data features (such as images), but also entire data flow (such as speech or video).

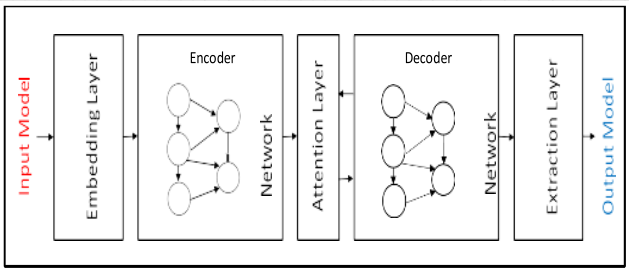

Figure 6. Gesture Classification

d. Verification

Using the dataset created, validated the camera input and hand plotting. Print the output and accuracy based on AML after verification (American Sign Language).

To improve the accuracy of model, all of these steps necessitate the use of additional technologies.

The following are some of the technologies used in this project,

- **Tensor Flow**

Tensor Flow is a cost free and available for all numerical computation library. After defining the nodes of the computation graph, the 13 actual computations occur within a session. Tensor Flow is a well-known machine-learning framework.

- **Keras**

Keras is a superior level neural network library written in Python that serves as a wrapper for TensorFlow. It is used when need to quickly construct and test a neural network using only a few lines of code.

- **OpenCV (Open-Source ComputerVision)**

Real time Machine vision programming functions in an open source library called OpenCV (Open Source Computer Vision). It largely practiced for face and object identification, image processing, and video collection and analysis.

- **Long short-term memory (LSTM)**

This type of artificial neural network used in both deep learning and machine intelligence. (LSTM) has feedback connections as compared withconventional cascaded neural networks.This type of recurrent neural network (RNN) is capable of processing not only particular data features (such as images), but also entire data flow (such as speech or video). For example, Video games, healthcare, robotics control, natural language processing, machine translation,and unsegmented, handwritten word analysis can all gain from (LSTM).

The (LSTM) neural network is the most frequently mentioned neural network of the twentieth century**.** A recurrent neural

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8516
Article Received: *31 July 2023* ***Revised:*** *28 September 2023* ***Accepted:*** *15 October 2023*

network is long short-term memory. The previous step's output is incorporated into the current phase of RNN. It dealt with the difficulty of RNN long term dependencies, which occur when the RNN could not forecast words accumulated in extendedterm memory but can make more perfect forecasting based on recent information. As the gap length increases, RNN is not capable to provide efficient performance. (LSTM) can save data for a prolonged period by default. It is used to process data from a time series, predict and classify it.

In order to compare the photographs taken while using this technology for communication, a proper database of sign language motions mandatory to create. Below are the steps followed to produce the data set used.

IV. DATASET GENERATION

To create dataset, used an open computer vision library. For testing purposes, first took roughly 4500–5000 photos of every alphabet out of which used approximately 4000 images and 180–200 images of each letter. First, record every frame that the webcam on computer produces. As seen in the image below, designate a space of interest in each frame, which indicated by a blue outlined box.

V. RESULT AND ANALYSIS

This study has built an effective USA Sign Language detection based on eyesight in real time system for D & M users. By employing two algorithms layers that more symbols to validate and forecast similar to one another, can improve forecasting. With this technique, can virtually always identify symbols as long as they correctly displayed, there is no background noise, and the lighting is adequate.

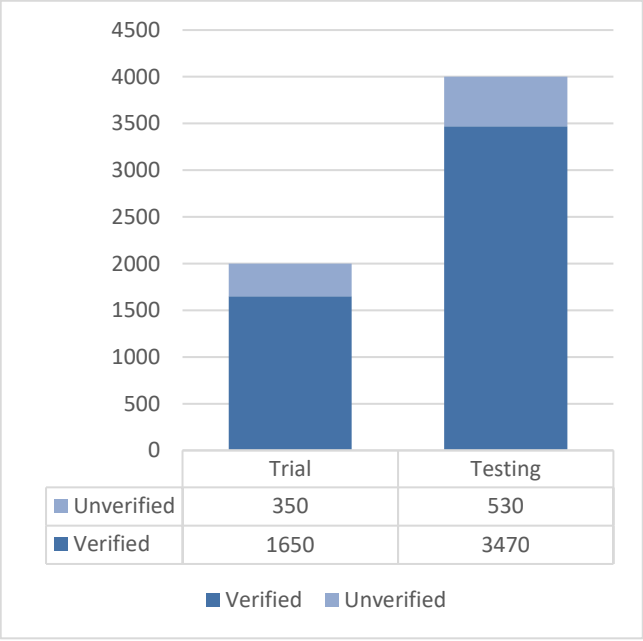

Figure 7 Trial And Testing

The system was tested on a 2000 image dataset and gave an accuracy of around 82%, which is approximately 1649 images matched with the input.

During the testing phase, a fresh dataset of over 2000 photos updated, and with the previous dataset included, the system delivered around 86.75% accuracy, which is roughly 3470 images matched with input [Figure.7].

Successfully developed a model with an accuracy of about 86.75% as shown in testing phase using the (LSTM) (Long Short Term Memory) methodology, as well as several image categorization and verification method.

Figure 8. Concluding Output

Comparing to [1] they used manual signs and pre developed dataset for the experimental test and they achieved around 82.20% accuracy whereas compared to [2] they have used face detection, removal and HSV thresholding and for the SL they used CNN algorithm and have achieved accuracy of 93%. In [3] the authors have used artificial neural network technology and have output of around 70 to 85%. In the paper[4] the used various contemporary technologies with the help of Al and NLP and resulted 91.1% accuracy.[5] In this they used NLP,CNN Methodology and sifting to create model in MATLAB after that they got 92% positive result.

VI. CONCLUSION

In this project at the beginning captured around 4000-4500 images for the dataset. After that used (LSTM) algorithm for feature extraction where points on the hands are cited to detect the hand symbols. After that gesture classification step occurs where (LSTM) is also used and the classification between different various symbols happens.

In last step of the sign detection verification step takes place where system have used tensor flow and keras for training the model for detection the symbols. In addition, these two technologies have also used for cross validation and verification.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8516
***Article Received:** 31 July 2023 **Revised:** 28 September 2023 **Accepted:** 15 October 2023*

Successfully developed a model with an accuracy of about 86.75% in experimental test using the (LSTM) (Long Short Term Memory) methodology, as well as several image categorization and verification method Using the (LSTM) algorithm, can incorporate handwriting recognition and text-to-speech conversion in future updates. This algorithm also allows us to use video capturing technology to create dynamic signs. Also intend to translate the results into local languages and integrate this technology into mobile devices.

REFERENCES

[1] Ashok K Sahoo, Gouri Sankar Mishra and Kiran Kumar Ravulakollu (2014)- Sign language recognition-State of the art
[2] Anup Kumar, Karun Thankachan and Mevin M. Dominic(2016)- Sign Language Recognition
[3] Tülay Karaylian, Özkan KÕlÕç (2017)- Sign Language Recognition
[3] Oyeniran, Oluwashina A, Oyeniyi, Joshua O, Sotonwa, Kehinde A,Ojo, Adeolu O.(2020)- Review Of The Application Of Artificial Intelligence In Sign Language Recognition System
[4] R Rumana, Reddygari Sandhya Rani, Mrs. R. Prema (2021)- A Review Paper on Sign Language Recognition for The Deaf and Dumb
[5] V. N. T. Truong, C. K. Yang and Q. V. Tran (2016)- A Translater For American Sign Language To Text And Speech
[6] M. M. Islam, S. Siddiqua and J. Afnan (2017)- Real Time Hand Gesture Recognition Using Different Algorithms Based On American Sign Language
[7] I. C. Ani, A. D. Rosli, R. Baharudin, M. H. Abbas and M. F. Abdullah(2014)-
[8] Preliminary Study Of Recognising Alphabet Letter Via Hand Gesture
[9] Rajshri Bhandra & Subhajit Kar- Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network. (2021)
[10] Vishwa Hariharan Iyer, U.M Prakash, Aashrut Vijay, P.Sathishkumar – Sign Language Detection using Action Recognition(2022)
[11] Tengfei Li Yongmeng Yan; Wenqing Du - Sign Language Recognition Based on Computer Vision(2022)
[12] Rahib Abiyev; John Bush Idoko; Murat Arslan - Reconstruction of Convolutional Neural Network for Sign Language Recognition (2020)
[13] Priyanka Pankajakshan; Thilagavathi B-Sign language recognition system(2015)
[14] S.A.M.A.S Senanayaka; R.A.D.B.S Perera; W. Rankothge; S.S. Usgalhewa; H.D Hettihewa; P.K.W. Abeygunawardhana-Continuous American Sign Language Recognition Using Computer Vision And Deep Learning Technologies (2022)
[15] Nipun Jindal; Nilesh Yadav; Nishant Nirvan; Dinesh Kumar-Sign Language Detection using Convolutional Neural Network (CNN),(2022)