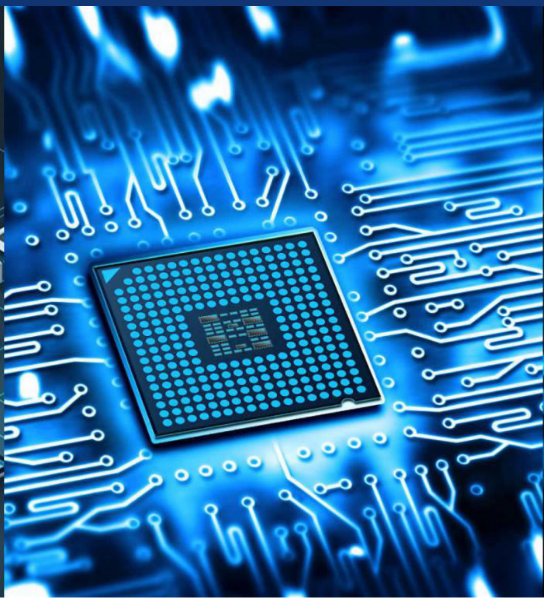
IJIRCCE

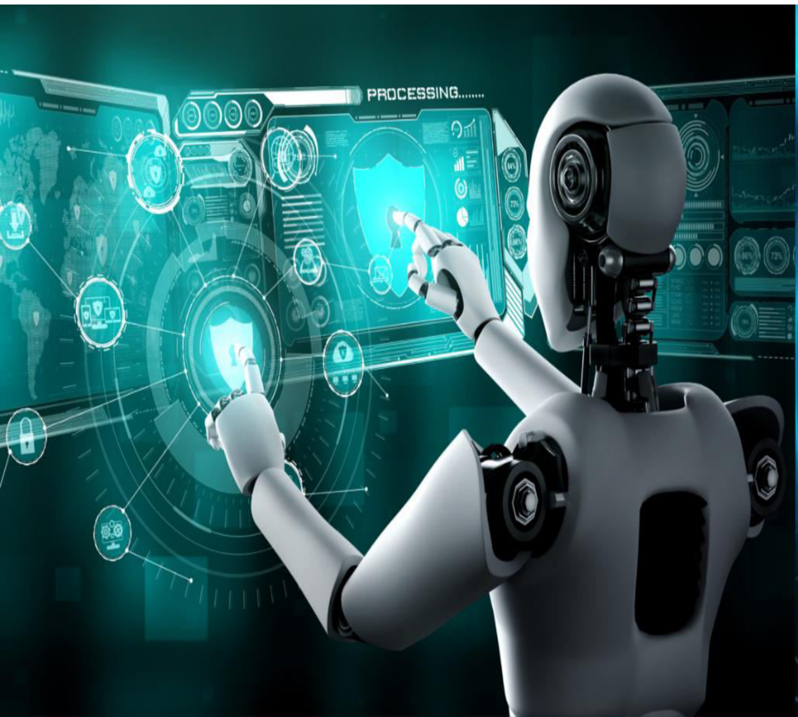
ISSN(O): 2320-9801

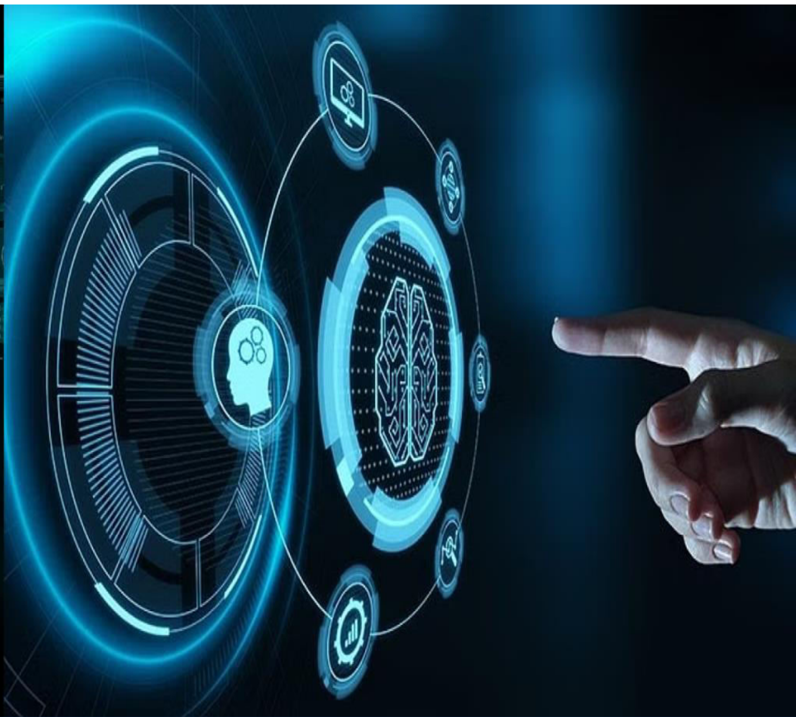
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771 **Volume 13, Issue 4, April 2025**

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304040

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Sign-Language to Text Converter

Melvin Manova L [1], Dr. S. Poorana Seenthilkumar [2]

UG Student, Department of Computer Applications, Dr. N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India[1]

HOD, Department of Computer Applications, Dr. N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India[2]

ABSTRACT: Sign language is one of the oldest and most natural forms of language for communication, but since most people do not know sign language and interpreters are very difficult to come by, we have come up with a real time method using neural networks for fingerspelling based on American sign language. In our method, the hand is first passed through a filter and after the filter is applied the hand is passed through a classifier which predicts the class of the hand gestures. Our method provides 95.7 % accuracy for the 26 letters of the alphabet.

KEYWORDS: Computer vision, Image processing, Hand gesture recognition, Python, Tensor flow,/flask(for GUI).

I. INTRODUCTION

American sign language is a predominant sign language Since the only disability Deaf and Dumb (hereby referred to as D&M) people have is communication related and since they cannot use spoken languages, the only way for them to communicate is through sign language. Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behavior and visuals. D&M people make use of their hands to express different gestures to express their ideas with other people. Gestures are the non-verbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language. A sign language is a language which uses gestures instead of sound to convey meaning combining hand-shapes, orientation and movement of the hands, arms or body, facial expressions and lip-patterns. Contrary to popular belief, sign language is not international.

II. LITERATURE REVIEW

In recent years there has been tremendous research done on hand gesture recognition.With the help of literature survey, we realized that the basic steps in hand gesture recognition are: -

- Data acquisition
- Data pre-processing
- Feature extraction
- Gesture classification

III. METHODOLOGY

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

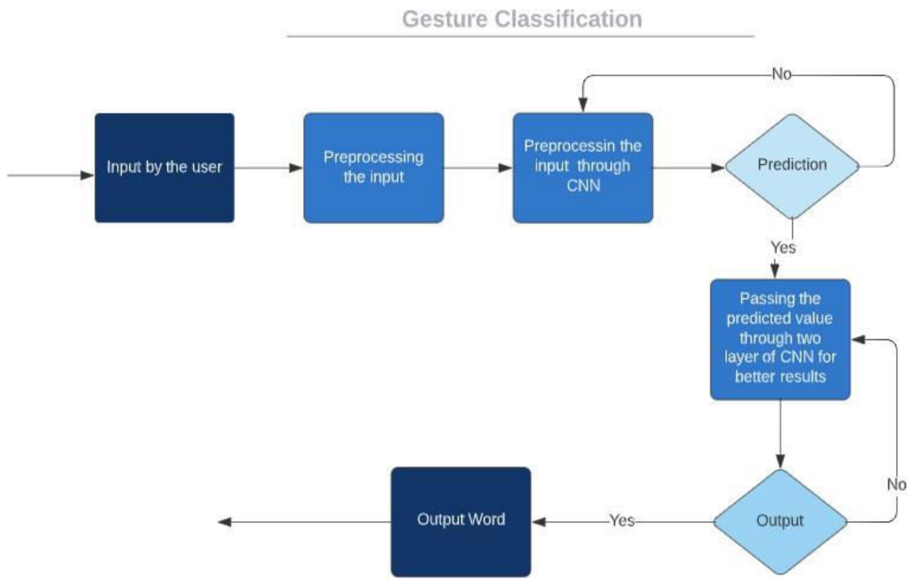
3.1 Data Set Generation:
For the project we tried to find already made datasets but we couldn't find datasets in the form of raw images that matched our requirements. All we could find were the datasets in the form of RGB values.
The overall methodology follows three key phases: Integration, Processing, and Security & Storage, as explained below.

3.2 Gesture Classification:
Our approach uses two layers of algorithms to predict the final symbol of the user.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Algorithm Layer 1:

1. Apply Gaussian Blur filter and threshold to the frame taken with openCV to get the processed image after feature extraction.
2. This processed image is passed to the CNN model for prediction and if a letter is detected for more than 50 frames then the letter is printed and taken into consideration for forming the word.
3. Space between the words is considered using the blank symbol.

Algorithm Layer 2:

1. We detect various sets of symbols which show similar results on getting detected.
2. We then classify between those sets using classifiers made for those sets only.

3.3 AutoCorrect Feature:

A python library Hunspell_suggest is used to suggest correct alternatives for each (incorrect) input word and we display a set of words matching the current word in which the user can select a word to append it to the current sentence. This helps in reducing mistakes committed in spellings and assists in predicting complex words.

IV. IMPLEMENTATION

The implementation of the Sign Language to Text Conversion system involves the integration of multiple modules to ensure smooth and accurate functioning. The system is developed using Python 3.8+ with key libraries such as OpenCV, TensorFlow, Keras, NumPy, and Tkinter for image processing, machine learning, and GUI development.

1. Setting Up the Environment

   To implement the system, the required dependencies are installed using pip. The machine learning model is trained separately and loaded during execution. The system runs on Windows 10/11 with a minimum of Intel i5 processor, 8GB RAM, and a 720p webcam for gesture recognition.

2. Model Training and Integration

   The Convolutional Neural Network (CNN) is trained on a dataset of hand gestures, where each gesture corresponds to a specific letter of the alphabet. After training, the model is saved in HDF5 (.h5) format and is loaded during execution for real-time predictions.

3. Hand Gesture Recognition Process

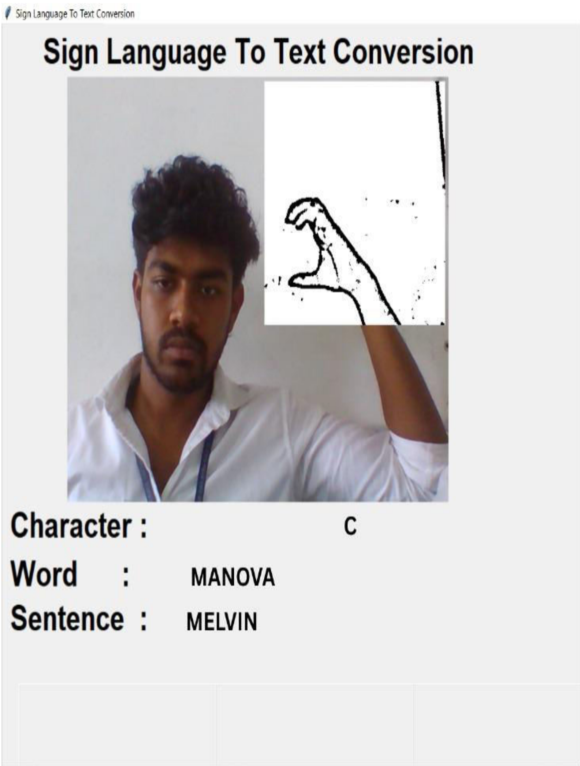
   Once the application is launched, the webcam captures the Region of Interest (ROI) where the user performs hand gestures. The image is preprocessed (grayscale conversion, Gaussian blur, and thresholding) before being passed to the CNN model for classification. The predicted letter is then displayed in the GUI..

© 2025 IJIRCCE | Volume 13, Issue 4, April 2025| DOI: 10.15680/IJIRCCE.2025.1304040
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Spell Correction and Text Formation
   Recognized letters are processed and combined into meaningful words. A spell-checking mechanism (using Enchant or PySpellChecker) corrects any misclassified letters to improve accuracy.
5. User Interface Implementation
   A Tkinter-based GUI is designed to display the live video feed, recognized letters, and final text output. Users can interact with the system and reset the output if needed.

V. RESULT

We have achieved an accuracy of **95.8%** in our model using only layer 1 of our algorithm, and using the combination of **layer 1 and layer 2** we achieve an accuracy of **98.0%**, which is a better accuracy then most of the current research papers on American sign language.
Most of the research papers focus on using devices like Kinect for hand detection.
In [7] they build a recognition system for Flemish sign language using convolutional neural networks and Kinect and achieve an error rate of **2.5%.**
In [8] a recognition model is built using a hidden Markov model classifier and a vocabulary of 30 words and they achieve an error rate of **10.90%**.
In [9] they achieve an average accuracy of **86%** for 41 static gestures in Japanese sign language.



VI. DISCUSSION

Existing messaging applications, such as WhatsApp, Telegram, and Signal, provide robust messaging services, but they have limitations, such as centralized control, limited customizability, and data privacy concerns. The proposed system offers an alternative that enables self-hosting, data control, and modular scalability. The architecture allows developers to integrate additional features such as voice/video calling, artificial intelligence-based chatbots, and enhanced security mechanisms.

One of the significant advantages of this system is its ability to handle real-time chat interactions with low latency,

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

ensuring a smooth user experience. However, challenges were encountered during implementation, including performance optimization of WebSocket connections, database query optimization, and security concerns related to authentication and data transmission. These challenges were mitigated through efficient indexing, token-based authentication, and structured API calls to improve overall system performance.

VII. CONCLUSION

In this report, a functional real time vision based American Sign Language recognition for D&M people have been developed for asl alphabets.
We achieved final accuracy of 98.0% on our data set. We have improved our prediction after implementing two layers of algorithms wherein we have verified and predicted symbols which are more similar to each other.
This gives us the ability to detect almost all the symbols provided that they are shown properly, there is no noise in the background and lighting is adequate.

VIII. FUTURE WORK

We are planning to achieve higher accuracy even in case of complex backgrounds by trying out various background subtraction algorithms.
We are also thinking of improving the Pre Processing to predict gestures in low light conditions with a higher accuracy.
This project can be enhanced by being built as a web/mobile application for the users to conveniently access the project. Also, the existing project only works for ASL; it can be extended to work for other native sign languages with the right amount of data set and training. This project implements a finger spelling translator; however, sign languages are also spoken in a contextual basis where each gesture could represent an object, or verb. So, identifying this kind of a contextual signing would require a higher degree of processing and natural language processing (NLP).

REFERENCES

1. 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.
2. Pujan Ziaie, Thomas Muller, Mary Ellen Foster, and Alois Knoll “A Na ïve Bayes Munich, Dept. of Informatics VI, Robotics and Embedded Systems, Boltzmannstr. 3, DE-85748 Garching, Germany.
3. 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
4. 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).
5. 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

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details