e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:09/September-2024 Impact Factor- 7.868 www.irjmets.com

A MACHINE LEARNING MODEL TO CONVERT AMERICAN SIGN LANGUAGE TO TEXT

Dr. Sanjay Sharma[*1], Sanchit Sharma[*2]

[*1]Professor, Department Of Electronic And Communication, A.D.G.I.P.S, New Delhi, India.

[*2]ECE, Department Of Electronic And Communication, A.D.G.I.P.S, New Delhi, India.

ABSTRACT

American Sign Language (ASL) is a vital communication method for individuals who are deaf or hard of hearing. However, a significant communication gap exists between ASL users and those who do not understand sign language, leading to challenges in accessibility and inclusivity. This paper presents the development of a real-time, vision-based system capable of recognizing and translating ASL fingerspelling gestures into text. The system leverages Convolutional Neural Networks (CNNs) and various machine learning tools, achieving an accuracy of 98.02% in recognizing the 26 letters of the ASL fingerspelling alphabet. This research offers a cost-effective solution that reduces reliance on interpreters, promoting greater independence and communication for the deaf and mute (D&M) community.

Keywords: American Sign Language, Convolutional Neural Networks, Deaf And Mute Community.

I. INTRODUCTION

American Sign Language (ASL) serves as the primary mode of communication for individuals who are deaf or hard of hearing worldwide. Unlike spoken languages, ASL conveys meaning through hand gestures, facial expressions, and body movements. While effective within the deaf and mute (D&M) community, ASL creates a language barrier when communicating with those unfamiliar with sign language. Traditionally, bridging this gap has relied on skilled interpreters, but their availability and associated costs can limit access to effective communication in various settings.

This project addresses this issue by developing a real-time, vision-based system to recognize and interpret ASL hand gestures, focusing on the fingerspelling technique. Fingerspelling involves representing individual letters of the alphabet through specific hand shapes and movements, allowing users to spell out words and phrases. Despite the challenges encountered in data collection, preprocessing, and resolving ambiguities between similar gestures, our system achieved a 98.02% accuracy rate in recognizing the 26 letters of the ASL fingerspelling alphabet.

The successful implementation of this system has significant implications for promoting accessibility and inclusivity across various sectors, reducing dependence on professional interpreters, and empowering D&M individuals to communicate more effectively.

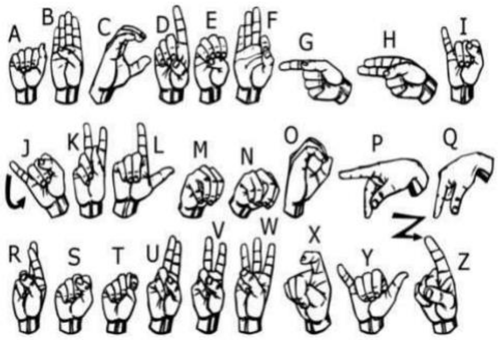

Figure 1: Hand Signs

II. OBJECTIVE

The research project pursued the following key objectives:

1. **Real-time Recognition:** To design a system capable of recognizing ASL fingerspelling gestures in real-time without significant delays.

2. **High Accuracy:** To achieve a high accuracy rate in recognizing and classifying individual letters of the ASL alphabet.
3. **User-friendly Interface:** To create an intuitive and accessible interface for users of varying technical expertise.
4. **Cost-effective Solution:** To provide an affordable alternative to professional interpreters, making ASL translation more widely accessible.
5. **Extensibility and Adaptability:** To ensure the system could be extended to recognize other sign language systems or gestures.
6. **Promote Inclusivity and Accessibility:** To contribute to the broader goal of promoting inclusivity for individuals with disabilities.

III. IMPLEMENTATION

The implementation of the ASL-to-text translation system involved key steps, starting with data collection, where images of ASL gestures were captured under varying conditions to build a diverse dataset. These images were then preprocessed using techniques like grayscale conversion, Gaussian blurring, and background subtraction to enhance gesture recognition. A Convolutional Neural Network (CNN) was developed using TensorFlow and Keras, featuring convolutional and pooling layers to detect and classify the gestures. The model was trained with data augmentation and hyperparameter tuning to achieve high accuracy. Finally, the model was integrated into a real-time system using OpenCV for video processing, enabling live gesture recognition with a user-friendly interface.

IV. MODELING AND ANALYSIS

The Modeling and Analysis of the ASL-to-text translation system focused on designing and evaluating a robust machine learning model capable of accurately recognizing ASL fingerspelling gestures. The core of the system was a Convolutional Neural Network (CNN), chosen for its effectiveness in image classification tasks. The CNN architecture included multiple convolutional layers to extract spatial features, followed by max-pooling layers to reduce dimensionality and focus on key aspects of the gestures. Fully connected layers were used for classification, with a softmax output layer providing probabilities across the 26 ASL letters.

During the analysis phase, the model was trained on a dataset of preprocessed hand gesture images, with a split for training, validation, and testing to ensure generalizability. The training process involved optimizing the model using the Adam optimizer and minimizing the categorical cross-entropy loss. Performance metrics like accuracy and confusion matrices were used to evaluate the model's effectiveness. The model achieved an accuracy of 98.02%, demonstrating its capability to reliably interpret ASL fingerspelling gestures in real-time applications. Further analysis included testing the model under different lighting conditions and with various users to ensure robustness and adaptability in real-world scenarios.

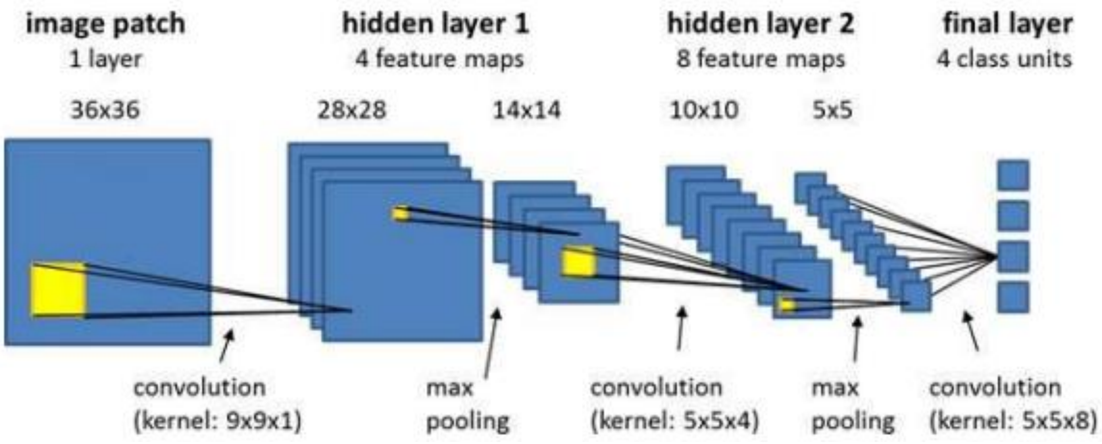

Figure 2: CNN Model

To further validate the model, additional tests were conducted with different users and varying backgrounds to assess its robustness and adaptability. The system consistently maintained high accuracy, confirming its effectiveness across diverse real-world conditions. Future work may involve extending the model to recognize dynamic ASL signs beyond fingerspelling, enhancing its applicability in more complex communication scenarios. The overall success of the model underscores its potential to significantly improve accessibility for the deaf and hard-of-hearing community.

V. CONCLUSION

Using the American Sign Language alphabets, this article describes how a real-time vision-based recognition system was developed to assist Deaf and Hard of Hearing persons. A final accuracy level of 98.02% was achieved on our dataset. Our prediction has been enhanced via the use of two levels of algorithms. In these layers, we have validated and forecasted symbols that are more similar to one another. Once the symbols are shown correctly, background noise is eliminated, and illumination is sufficient, we can detect almost all of them.

VI. REFERENCES

[1] B. Joksimoski et al., "Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities," in IEEE Access, vol. 10, pp. 40979-40998, 2022, doi: 10.1109/ACCESS.2022.3161440.

[2] B. Gupta, P. Shukla and A. Mittal, "K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion," 2016 International Conference on Computer Communication and Informatics (ICCCI), 2016, pp. 1-5, doi:10.1109/ICCCI.2016.7479951.

[3] T. D. Sajanraj and M. Beena, "Indian Sign Language Numeral Recognition Using Region of Interest Convolutional Neural Network," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018, pp. 636-640, doi: 10.1109/ICICCT.2018.8473141.

[4] P. C. Badhe and V. Kulkarni, "Indian sign language translator using gesture recognition algorithm," 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), 2015, pp. 195-200, doi: 10.1109/CGVIS.2015.7449921.

[5] H. Luqman, "An Efficient Two-Stream Network for Isolated Sign Language Recognition Using Accumulative Video Motion," in IEEE Access, vol. 10, pp. 93785-93798, 2022, doi: 10.1109/ACCESS.2022.3204110.

[6] Punekar, Aaditya. (2020). A Translator for Indian Sign Language to Text and Speech. International Journal for Research in Applied Science and Engineering Technology. 8. 1640-1646. 10.22214/ijraset.2020.6267.

[7] dewale, Victoria & Olamiti, Adejoke. (2018). Conversion of Sign Language To Text And Speech Using Machine Learning Techniques. JOURNAL OF RESEARCH AND REVIEW IN SCIENCE. 5. 58-65. 10.36108/jrrslasu/8102/50(0170).