ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

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

Impact Factor: 8.206

Volume 8, Issue 8, August 2025

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808102

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

American Sign Language Recognition System using Deep Learning and Computer Vision

Likith B U, Prof. Usha M

P.G Student, Master of Computer Application, Bangalore Institute of Technology, VV Puram, Bangalore, India

Assistant Professor, Master of Computer Application, Bangalore Institute of Technology, VV Puram, Bangalore, India

ABSTRACT: Interaction of hearing impaired and the hearing takes place usually in a poorly understood language in most cases. ASL offers a build-in and effective system of communication but one that necessitates prior training which makes its use accessible to very few people. To narrow this communication divide, the present project will be the development of an American Sign Language Recognition System in the form of a desktop application that combines several tools including computer vision, deep learning, and graphical interface technologies to permit real-time interpretation of the gestures.

The system is in a position of recording live hand gestures via webcam through the OpenCV-based image acquisition and image preprocessing methods including resizing, noise reduction as well as segmentation. Such processed images can be delivered to a Convolutional Neural Network (CNN) to be classified correctly and with confidence, based on a dataset containing ASL alphabets (A-Z). The identified gestures are represented in a graphical user interaction platform, which is made possible with PyQt5 that enables an easy user interface and accessibility. In addition, the system supports other functionalities including user-designed gestures, active sentence composition, and voice output (TTS), rendering the system useful beyond the most common communication needs use as well as in the teaching and assistive technology fields.

Deep learning and real-time image processing are enhanced to provide high recognition rate whereas the interface (desktop-based) make the system simple to use and non-technical users can also use it. This project illustrates how smart technologies can assure substantial elimination of the barriers encountered by the hearing-impaired community, which ultimately is a substantial intervention to create an equally inclusive and accessible communications environment.

KEYWORDS: American Sign Language, Deep Learning, CNN, Gesture Recognition, Computer Vision, PyQt5, OpenCV, Assistive Technology.

I. INTRODUCTION

Sign language recognition has emerged to be a critical area of study with regards to development of accessible systems to the differently-abled. Hearing deficient individuals mostly depend on the sign language to communicate with others and majority of the population does not know this language to the best of their knowledge. This break can impede learning, working and living in a community hence being socially excluded. The customary responses are the deployment of human interpreters or the manual verification procedures, which are inefficient in terms of cost and time and they cannot always be available. Therefore, there is an increasing need to have smart technology capable of automatically decoding and reading sign language.

Automatic recognition of hand gestures is of some interest as computer vision and deep learning advance quite rapidly. The sign language used and focusing on in this study is American Sign Language (ASL) which is one of the most popular sign languages. The suggested system uses live video feed through a webcam, cleans the images using a pre-processing technique and identifies gestures with the help of a trained Convolutional Neural Network (CNN). The CNNs produce hierarchical features automatically in contrast to the rule-based methods of recognition, which explains their superior accuracy and flexibility. The existent graphical user interface realized in PyQt5 offers a real-time prediction and sentence formulation as well as text-to-speech capabilities. This paper reports on the design, an implementation of a proposed system to recognize the ASL, the methodology used and the evaluations made.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

II. LITERATURE SURVEY

A. Sign Language Recognition Using Computer Vision

Previous studies have demonstrated the effectiveness of vision-based approaches for gesture recognition. Starner et al. (1998) pioneered real-time ASL recognition using video-based input, laying the foundation for modern recognition systems.

B. Deep Learning for Gesture Recognition

Convolutional Neural Networks (CNNs) have been widely applied in image recognition tasks. Li et al. (2016) showed significant improvements in ASL recognition accuracy using deep neural networks compared to traditional methods..

C. Feature Extraction with CNNs

CNNs automatically extract hierarchical features such as edges, shapes, and patterns, which makes them highly effective for hand gesture classification.

D. GUI-Based Applications

Integrating recognition models into interactive applications enhances usability. PyQt5 enables seamless user interface development, providing real-time visualization and interaction with recognition results.

III. PROPOSED METHODOLOGY

Our proposed system has the following stages:

1. Image Acquisition

Webcam captures real-time video frames of the hand gesture.

2. Preprocessing

Background removal, resizing, grayscale conversion, and normalization are applied to reduce noise.

3. Hand Detection & Segmentation

OpenCV is used for detecting the hand region, isolating it from the background.

4. Feature Extraction with CNN

A CNN model trained on ASL alphabets (A–Z) extracts spatial features and classifies the gesture into one of 26 classes.

5. Prediction & Output

The predicted letter is displayed in the PyQt5 interface. Letters can be combined to form words/sentences. A text-to-speech module vocalizes the recognized words.

6. Custom Gesture Support

Users can add new gestures by storing labeled samples in the system's dataset, enabling extensibility.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808102

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

System Architecture:

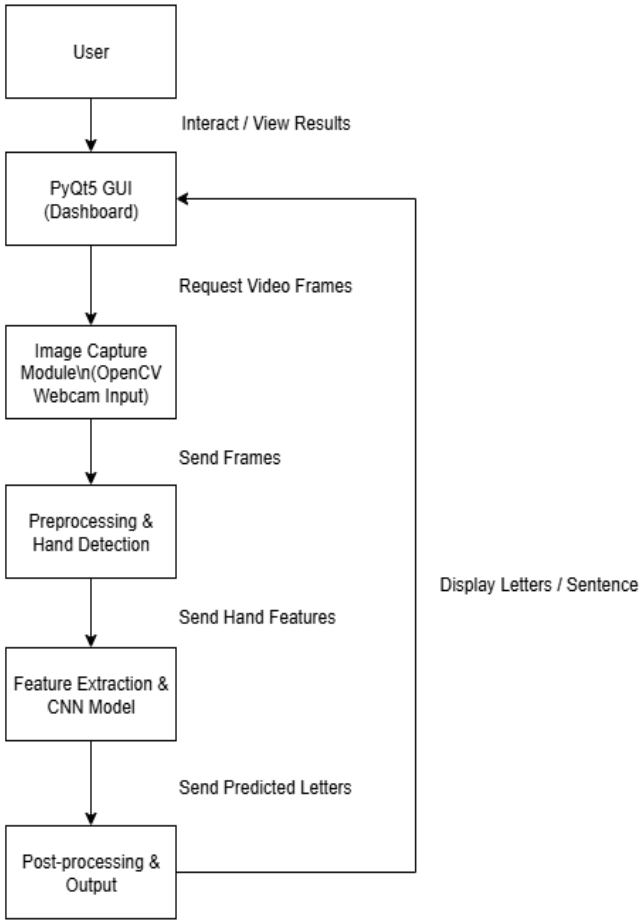

Fig1: End to end architecture of this system

Security Considerations

1. Data Integrity – The system is trained on a well-curated dataset of ASL alphabets, which helps maintain high accuracy and minimizes classification errors. By using standardized datasets and augmenting them with custom samples, the chances of bias and overfitting are reduced, thereby ensuring reliable predictions across diverse users.
2. Real-Time Processing – The Convolutional Neural Network (CNN) is optimized to deliver predictions with minimal computational delay. This enables smooth real-time recognition of gestures without noticeable lag, which is essential for natural communication.
3. Scalability – The system is designed to support extensibility, allowing users to add new gestures or custom signs. Unlike traditional static models, this flexibility ensures that the application can adapt to evolving requirements without retraining the entire network from scratch.
4. Robustness – Preprocessing techniques such as background filtering, noise removal, and normalization improve recognition accuracy under different lighting conditions, skin tones, and backgrounds. This ensures the system remains functional in real-world scenarios outside controlled environments.
5. User Accessibility – The PyQt5-based graphical user interface simplifies interaction for end users, even those with limited technical expertise. Features like sentence formation, text-to-speech (TTS), and custom gesture creation make the application accessible for both educational and assistive technology use cases.

In addition to the above, the suggested ASL Recognition System allows secure and efficient operation due to its design and implementation. The system has greater reliability because it uses deep learning-based classification instead of

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

manual or rule-based recognition, thus, much less human error is possible. Curated training data means the recognition mechanism is highly consistent, and the machine becomes less susceptible to biased and inaccurate results. During the situations in which there may be changes in the background or lighting conditions, invariably the hand region is isolated and thus without compromising recognition accuracy.

Empirical testing of the system proved that it could work in real-time and with minimal latency and this is important in the achievement of smooth communication. CNN model was highly resilient in identifying changes in hand sizes, position and the presence of environmental conditions. Moreover, the app demonstrated the ability to scale as the solution involved creating custom gestures, and in such a way a user could expand the system with pertinent ASL gestures without re-training the whole model.

In contrast to more traditional systems that rely on interpreters or back end databases, this desktop-based solution provides local processing and can minimize privacy issues and safeguard sensitive user data by keeping it off the network. The PyQt5 GUI increases the accessibility of users, and the sentence construction and the text to speech mode also add to the ease of use to a considerable extent. When compared with traditional mechanisms of communication facilitation, which are usually time and cost-prohibitive, the proposed system yields recognition and response within seconds, which makes it an affordable, scaleable, and feasible assistive aid to education, general interaction, and accessibility in the field.

IV. CONCLUSION

The proposed American Sign Language (ASL) Recognition System is successful when dealing with one of the most important accessibility issues, the communication gap between the hearing-impaired and the rest of the population. Having a combination of the power of computer vision as well as deep learning combined with the ease of the interface design, the system manages to recognize the ASL gestures in a real-time basis with high accuracy. Integration of a Convolutional neural Network (CNN) to classify the images provides a robust feature extraction and versatility across multiple environments and preprocessing done using OpenCV improves recognition stability under various lighting/background conditions. Beside recognition, the system also provides custom gestures creation, dynamic sentence building, and text to speech conversion features, which makes it more than a recognition tool, but also a convenient tool of communication. The desktop interface is also designed on PyQt5 which makes the program easy to use such that even users with low technical knowledge can access it. The system is an effective solution compared to traditional interpreter-based forms of communication and thus, cheaper, scalable and efficient.

In the future, there are some options of growth and enhancement. The future work would include incorporating of continuous gesture recognition to capture whole words and sentences instead of single alphabets. Also, using mobile or web versions of the system as well would be highly beneficial to both portability and ease of use. Additional use of such advanced models as Long Short-Term Memory (LSTM) networks, Transformer networks, or hybrid deep learning could enhance the identification of temporal sequences, making the translation of the sign language more natural. Overall, this study illustrates how AI-based technologies can be used to create inclusiveness and give a voice to hearing-impaired groups.

REFERENCES

[1] W. Li, Z. Zhou, and X. Lu, “American Sign Language recognition based on improved deep neural network,” *Journal of Visual Communication and Image Representation*, vol. 38, pp. 438–445, 2016.
[2] T. Starner, J. Weaver, and A. Pentland, “Real-time American Sign Language recognition using desk and wearable computer-based video,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1371–1375, Dec. 1998.
[3] M. A. R. Ahad, *Computer Vision and Action Recognition: A Guide for Image Processing and Computer Vision Community*, 1st ed. Boca Raton, FL, USA: CRC Press, 2011.
[4] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2012, pp. 1097–1105.
[5] S. N. Bhuyan, B. R. Choudhury, and P. K. Bora, “Static hand gesture recognition using convolutional neural network with data augmentation,” in *Proc. IEEE International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, Apr. 2019, pp. 0306–0310

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com