

Enhancing Sign Language Identification using Key Point-Based Feature Extraction Methods

Shivani Chauhan¹, Sonika Kandari², Kamal Kant Verma³

School of Engineering and Technology, Shri Guru Ram Rai University Dehradun, India,
shivichauhan687@gmail.com 1

School of Engineering and Technology, Shri Guru Ram Rai University Dehradun, India,
dean.set@sgrru.ac.in 2

School of Computer Science and Engineering, IILM University Greater Noida, India,
dr.kamalverma83@gmail.com 3

Article History:

Received:10-10-2025

Revised:27-11-2025

Accepted:10-12-2025

Abstract:

Only the verbal expression of wants, emotions, and thoughts to others is the purpose of human communication. However, people with disabilities, such as the deaf and the dumb, exist on our world. These people are not able to communicate through speech. Only sign language may be used to communicate between deaf and dumb persons. This work applies various feature extraction approaches to static hand gesture photos, with an emphasis on FAST (Features from Accelerated Segment Test), SIFT (Scale Invariant Feature Transform), and ORB (Oriented FAST and Rotated BRIEF). Combinations like FAST+ORB and FAST+SIFT are also studied to enhance performance and resilience. Experiments show that hybrid techniques, such as FAST+ORB and FAST+SIFT, increase computational efficiency and feature matching accuracy, making them suitable for real-time Sign Language Recognition (SLRecog) applications. The best feature extraction methods for gesture-based human-computer interaction systems will be chosen with the aid of this comparative study.

Keywords—sign language recognition, feature extraction, vision based

1.

Introduction

Communication plays a vital role in the human experience, serving as a crucial and efficient means of expressing ideas, emotions, and viewpoints. Nevertheless, a significant portion of the global population does not possess this skill. Numerous individuals experience difficulties with their hearing, speaking abilities, or a combination of both. Hearing loss refers to the partial or total incapacity to hear in one or both ears. Conversely, being mute is a disability that affects speech, rendering individuals unable to speak. When deaf-muteness occurs during childhood, it can hinder language acquisition and lead to language impairment, also known as hearing mutism. These conditions represent some of the most prevalent disabilities on a global scale.

According to a report by the World Health Organization (WHO), there are more than 466 million individuals who experience speech or hearing impairments, with 80% of them being semi-illiterate or illiterate. Non-verbal communication methods, such as sign language, visually express and convey our perspectives, emotions, and ideas. Sign language grammar differs significantly from spoken language, as it involves the use of distinct hand shapes, signals, or gestures performed in specific locations on or near the signer's body, accompanied by particular movements.

Sign Language is a form of communication that relies on visual gestures, involving hand shapes, hand movements, body orientation, facial expressions, and lip patterns to convey meaning instead of relying on sound. Various sign languages are used worldwide, each with its own set of vocabulary and gestures. Examples include American Sign Language (ASL), Indian Sign Language (ISL), German Sign Language (GSL), and British Sign Language (BSL), among others. Sign Language is primarily used within deaf communities, encompassing interpreters, friends, and families of the deaf, as well as individuals who have partial hearing impairments. However, these languages are not

widely known outside of these communities, leading to communication barriers between deaf and hearing individuals.

In Automatic SLRecog system three phases are important: 1) Image preprocessing, 2) Feature extraction and 3) Classification. This work focuses on feature extraction techniques used in vision based Indian Sign Language (ISL) such as SIFT, ORB and FAST.

2. Literature Review

Feature extraction techniques play an important role in SLRecog system to identify the most relevant features from the image. In (I A Adeyanju et al.) different techniques of feature extraction used so far are reviewed such as PCA, Fourier Descriptor, SIFT, SURF and HOG. Further (Wiryananda and Zahra) reviewed the performance feature extraction techniques PCA with different methods are compared and another techniques are also reviewed like hand segmentation, SIFT, Kinect with SDK etc. (Thongtawee et al.) proposed a feature extraction technique for American SLRecog system: NwE, Fcen, AngF and delAng. After feature extraction ANN is used for classification with 95% accuracy.

Furthermore (Tyagi and Bansal) reviewed various feature extraction techniques for vision based Indian SLRecog system such as Zernike Moments, HOG, SURF, SIFT, CNN, ANN, Hu moments. (Ma et al.) Proposed a model for Chinese sign language recognition with 94.8% accuracy for continuous sign and 91.4 % for words. Whereas (Kasireddy et al.) proposed a model for French sign language recognition which based on gloves with 92% accuracy rate. American sign language translator based on neuromorphic sensor and ANN is developed with accuracy 79.58% (Rivera-acosta et al.). (I. A. Adeyanju et al.) Proposed a review a paper in which different machine learning methods used for SLRecog system are discussed. Whereas different

Recognition techniques of hand gestures are discussed in review paper (Cheok et al.). Whereas (Patel et al.) analyzed various image enhancement and restoration techniques like image scaling, color spaces, mean filter, median filter and adaptive filter. (Zhu et al.) proposed an image preprocessing method for gesture recognition. Furthermore, an automatic SLRecog system is proposed in which convex hull and KNN are used for feature extraction and classification with 65% accuracy (Amrutha and Prabu). A “SHAROJAN BRIDGE” is proposed to fill the gap of communication between deaf, dumb and blind people by (Rastogi). Whereas (Pathan et al.) proposed a multi headed CNN model with 98.981% accuracy.

(Tyagi and Bansal) discussed many feature extraction techniques (SIFT, FD, ANN, HOG etc.) in review paper. The “Sign Nevestro Densenet Attention” approached for ASL is proposed by (Hassan et al.). By (Kakizaki et al.) four feature types—angle, motion, distance, and finger detection—are extracted using a Random Forest-based feature classification method. Lastly, the SVM method based on kernels is applied to classification.

3. Approaches

The techniques used to recognize sign language differ according on the type of input device. If a camera is used to capture the gesture, it is a vision-based method; if gloves with sensors are used to track the motion, it is a glove-based method.

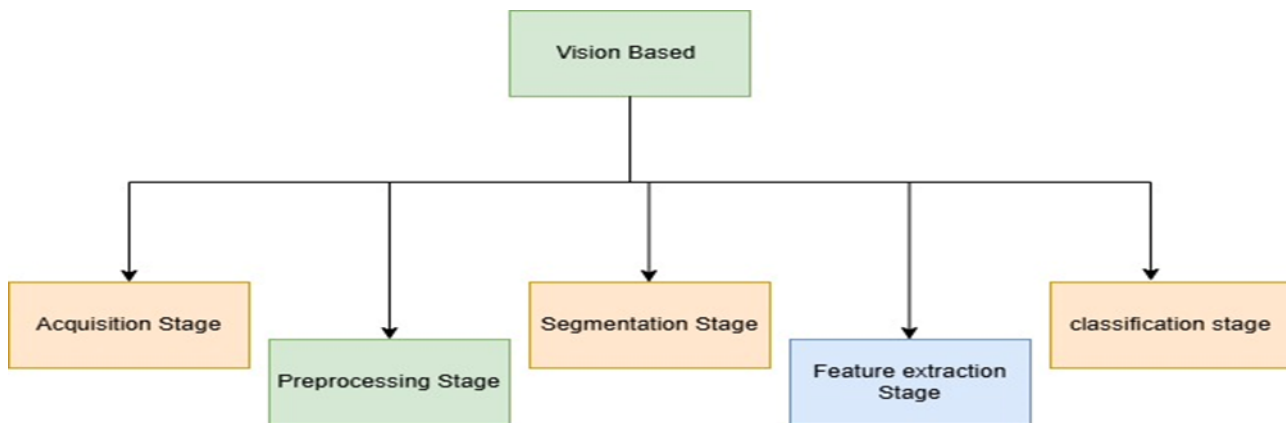


Fig. 1: Stages of Vision Based Approach

3.A Vision Based

In vision-based methods, webcams are used to record gestures. The output's accuracy is impacted by the camera angle. A single sign's meaning will change depending on the perspective from which it is viewed. Other research uses multiple cameras at different angles to record the same move from different perspectives. Fig.1 illustrates the five steps of the vision-based approach to sign language recognition: 1) Acquisition stage; 2) Pre-processing; and 3) Segmentation Step four is feature extraction, while step five is categorization.

3.A.1 ACQUISITION STAGE

The first stage of the system when raw data representing sign language motions is gathered or recorded is known as the acquisition stage in SLRecog system. The input for later processing stages, such as feature extraction, classification, and recognition, is this data. Webcam, gloves, Kinect, and a Leap Motion Controller are the tools used in the acquisition step.

3.A.2 PRE-PROCESSING

Currently, picture restoration and improvement are being done to increase image quality so that it would be better for both human and computer analysis. Techniques for image enhancement and correction are widely available. There are several image enhancement techniques available, including histogram equalisation, adaptive histogram equalisation, contrast limit adaptive histogram equalization. In order to improve image quality for both human and computer analysis, restoration and enhancement efforts are now underway. There are many methods for improving and correcting images. Histogram equalization, adaptive histogram equalization, contrast limit adaptive histogram equalization (CLAHE), and other methods are among the several image improvement approaches that are accessible. Images can also be restored using a variety of methods, including Gaussian, Mean, and Median filters.

3.A.3 SEGMENTATION

In order to prepare for feature extraction and recognition, the segmentation stage of the SLRecog system is an essential step that entails separating pertinent portions of the input, usually video or picture frames. Separating the signature elements—such as the hands, face, and body—from the background and extraneous details is the aim. There are numerous segmentation techniques available, including edge-based, threshold-based, cluster-based, region-based, and others.

3.A.4 FEATURE EXTRACTION AND CLASSIFICATION

The feature extraction is very important part of any SLRecog system because the accuracy of classification is dependent on the output of feature extraction(I A Adeyanju et al.). In this work various feature extraction techniques are applied on Indian SLRecog datasets containing alphabets [A-Z]. Techniques are used with CNN classifier and The structure is made up of:

- Convolutional Layers: Take input data and extract spatial information.
- MaxPooling Layers: Minimize feature map dimensions without sacrificing important details.
- Flatten Layer: Creates a 1D vector from the 2D feature maps.
- Dense Layers: Completely interconnected categorization layers.
- Class probabilities (26 classes) are output by the Softmax Layer.

The result of various feature extraction techniques is shown in following figures.

3.A.4.i ORB

In computer vision, ORB is a quick and effective feature detector and descriptor that is especially useful for

applications like motion tracking, object recognition, and picture matching. Although ORB is a more computationally efficient option than SIFT and SURF, its accuracy in ISL recognition is limited. In fig.2 ORB is

used with 80 epochs. By epoch 80, training accuracy has steadily improved to about 0.55. For the most of the

training, validation accuracy stays extremely poor and only begins to improve after epoch 55, when it reaches

roughly 0.21. The test accuracy is a very low 21.43%. In fig.3 by epoch 100, the training accuracy has increased

to about 0.85. Compared to the fig.2, Validation Accuracy has improved more, reaching roughly 0.42. The test

accuracy has improved noticeably to 42.86%.

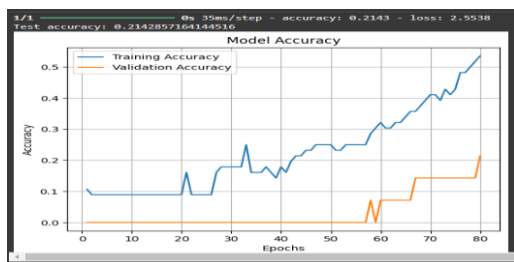


Fig. 2: ORB with 80 epochs

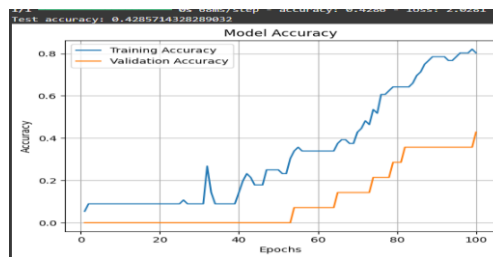


Fig. 3: ORB with 100 epochs

3.A.4.ii FAST + ORB

By enhancing key point identification, the high-speed corner detection technique FAST improves ORB's performance. By the 80th epoch, Training Accuracy has steadily improved to almost 98%. Additionally, validation accuracy rises dramatically to about 75%. With a test accuracy of 0.75, the generalization performance is good as shown in fig.4. In fig.5 training and validation Accuracy both reach 1.0 (100%) with 100 epochs, indicating perfect accuracy on both sets. Test Accuracy is 1.0, showing excellent generalization.

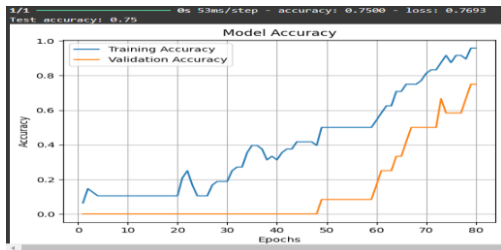


Fig.4: FAST+ ORB with 80 epochs

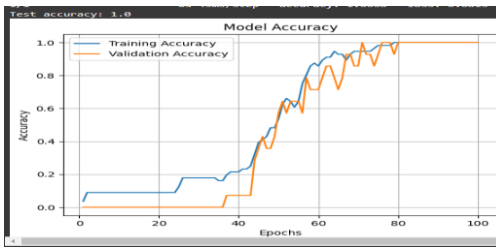


Fig .5: FAST+ORB with 100 epochs

3.A.4.iii FAST+SIFT

SIFT is quite good at extracting characteristics that are invariant to rotation and scaling. In conjunction with FAST, it accomplishes: with 80 epochs by about epoch 60, training accuracy has increased gradually to 100%. Additionally, validation accuracy steadily increases, hitting 100% about the same time. With a reported test accuracy of 1.0, perfect generalization is indicated as shown in fig.6. In fig.7 with 100 epochs Validation and Training Around epoch 40, accuracy both reach 100% even more quickly. For the remaining epochs, the model remains perfectly accurate and the little loss (0.0280) further supports the excellent performance. Additionally, excellent generalization is confirmed by the test accuracy, which is once again 1.0.

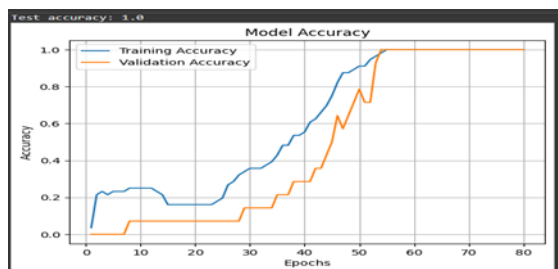


Fig. 6: FAST+SIFT with 80 epochs

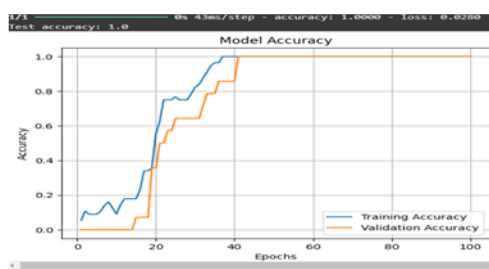


Fig. 7: FAST+SIFT with 100 epochs

Table.1: Summary of results

Variant of the Model	Training Epochs	Training of Accuracy	Validation of Accuracy	Test Accuracy	Remarks
ORB (Fig. 2)	80	0.55	0.21	0.2143	Inadequate to generalize
ORB (Fig. 3)	100	0.85	0.42	0.4286	Moderately improved
FAST + ORB (Fig. 4)	80	0.98	0.75	0.75	Strong generalization
FAST + ORB (Fig. 5)	100	1.0	1.0	1.0	Accurate precision
FAST + SIFT (Fig. 6)	80	1.0	1.0	1.0	Excellent generalization
FAST + SIFT (Fig. 7)	100	1.0	1.0	1.0	Excellent performance, very low loss (0.0280)

3.B GLOVE-BASED

Glove-based methods use sensors to monitor the indication. On a glove, there are five sensors on each finger. Some researchers have used LCD in order to correctly combine the accurate symbol. This method's disadvantage is that gloves are required for use. Gloves cannot be carried about constantly, and a user's unintentional finger movement is likewise interpreted as a sign.

4. Conclusion and Future work

According to the results, FAST + ORB, with optimized hyperparameters (100 epochs, batch size 32), is a good option for feature extraction since it strikes a balance between execution time and accuracy. Nevertheless, even with somewhat higher computing requirements, FAST + SIFT is still a reliable technique for guaranteeing high accuracy. In future SLRecog systems should be deployed on portable or low-power devices, such as smartphones, embedded systems, and wearable technologies, is made practical by increasing real-time performance and decreasing computing latency. This will improve these technologies' usability and accessibility for regular communication.

REFERENCES

1. Adeyanju, I A, et al. "Intelligent Systems with Applications Machine Learning Methods for Sign Language Recognition: A Critical Review and Analysis." *Intelligent Systems with Applications*, vol. 12, 2021, p. 200056, <https://doi.org/10.1016/j.iswa.2021.200056>.
2. Adeyanju, I. A., et al. "Machine Learning Methods for Sign Language Recognition: A Critical Review and Analysis." *Intelligent Systems with Applications*, vol. 12, Elsevier B.V., 1 Nov. 2021, <https://doi.org/10.1016/j.iswa.2021.200056>.
3. Amrutha, K., and P. Prabu. "ML Based Sign Language Recognition System." 2021 International Conference on Innovative Trends in Information Technology, ICITIIT 2021, Institute of Electrical and Electronics Engineers Inc., 2021, <https://doi.org/10.1109/ICITIIT51526.2021.9399594>.
4. Cheok, Ming Jin, et al. "A Review of Hand Gesture and Sign Language Recognition Techniques." *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 1, Jan. 2019, pp. 131–53, <https://doi.org/10.1007/s13042-017-0705-5>.
5. Hassan, Esraa, et al. "A Novel Model for Expanding Horizons in Sign Language Recognition." *Scientific Reports*, vol. 15, no. 1, 2025, pp. 1–21, <https://doi.org/10.1038/s41598-025-09643-2>.
6. Kakizaki, Manato, et al. "Dynamic Japanese Sign Language Recognition Throw Hand Pose Estimation Using Effective Feature Extraction and Classification Approach." *Sensors*, vol. 24, no. 3, 2024, <https://doi.org/10.3390/s24030826>.
7. Kasireddy, Shivaji, et al. Gestures with an IMU-Based Glove †. 2018, pp. 1–18, <https://doi.org/10.3390/informatics5020028>.
8. Ma, Jiyong, et al. Stream State-Tying for Sign Language Recognition.
9. Patel, Isha, et al. "Analysis of Various Image Preprocessing Techniques for Denoising of Flower Images Digital ScareCrow Using Iot View Project CHARUSAT Apps (Mobile Application) View Project Analysis of Various Image Preprocessing Techniques for Denoising of Flower Images." Article in INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING, 2018, www.ijcseonline.org.
10. Pathan, Refat Khan, et al. "Sign Language Recognition Using the Fusion of Image and Hand Landmarks through Multi-Headed Convolutional Neural Network." *Scientific Reports*, vol. 13, no. 1, 2023, <https://doi.org/10.1038/s41598-023-43852-x>.
11. Rastogi, Rohit. A Novel Approach for Communication among Blind, Deaf and Dumb People Role and Efficacy of Positive Thinking on Stress Management and Creative Problem Solving for Adolescents View Project Audio Visual EMG & GSR Biofeedback Analysis for Effect of

- Spiritual . 2015, <https://www.researchgate.net/publication/283507754>.
12. Rivera-acosta, Miguel, et al. Neural Network. 2017, <https://doi.org/10.3390/s17102176>.
 13. Thongtawee, Ariya, et al. "A Novel Feature Extraction for American Sign Language Recognition Using Webcam." 2018 11th Biomedical Engineering International Conference (BMEiCON), 2018, pp. 1–5.
 14. Tyagi, Akansha, and Sandhya Bansal. Feature Extraction Technique for Vision-Based Indian Sign Language Recognition System : A Review for Vision-Based Indian Sign Language Recognition System : A Review. no. August 2020, Springer Singapore, 2024, <https://doi.org/10.1007/978-981-15-6876-3>.
 15. Wiryana, Fanny, and Amalia Zahra. "Feature Extraction Methods in Sign Language Recognition System : A Literature Review." 2018 Indonesian Association for Pattern Recognition International Conference (INAPR), 2018, pp. 11–15.
 16. Zhu, Yiding, et al. "Image-Preprocessing Method for near-Wall Particle Image Velocimetry (PIV) Image Interrogation with Very Large in-Plane Displacement." Measurement Science and Technology, vol. 24, no. 12, 2013, <https://doi.org/10.1088/0957-0233/24/12/125302>.