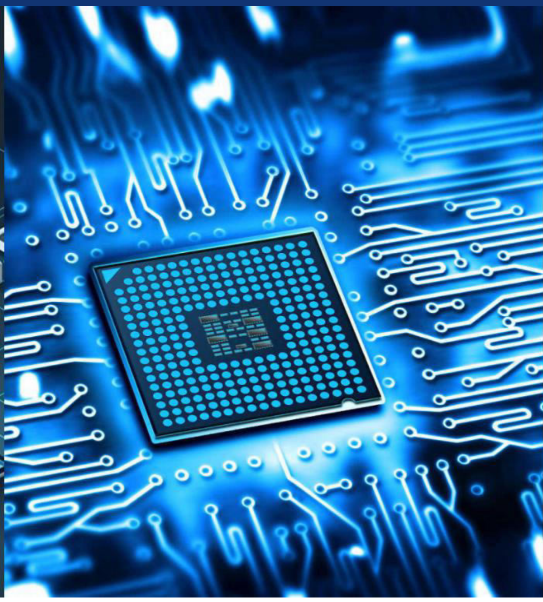
IJIRCCE

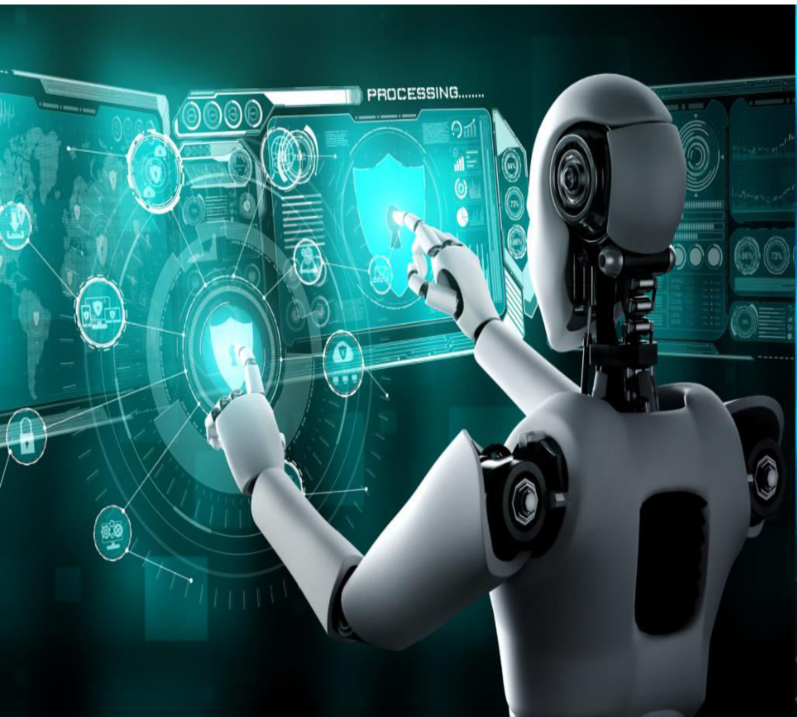
ISSN(O): 2320-9801

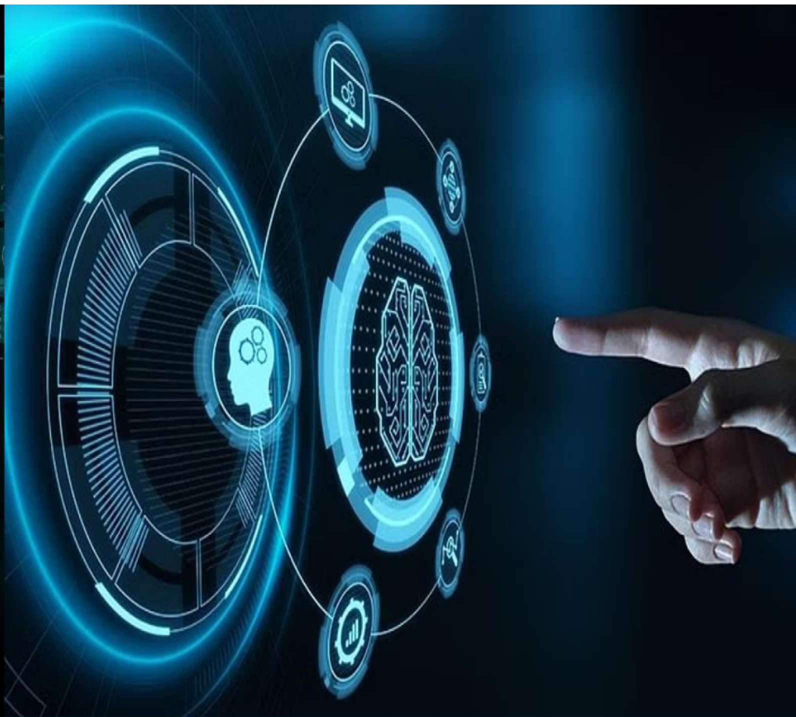
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

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





Impact Factor: 8.625

Volume 13, Issue 1, January 2025

www.ijircce.com | ijircce@gmail.com | +91-9940572462 | +91 63819 07438

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

Scale Invariant Feature Transform and Neural Network for Sign Language to Text Detection

Aman Kumar Sinha[1], Shrikrishna[2], Praveen N[3], Shafiya Khanum[4], Ajay Prakash B V[5], Kushala M V[6]

UG Student, Department of ISE, Dr. Ambedkar Institute of Technology, Bangalore, India[1,2,3,4]

Professor, Department of AIML, Dr. Ambedkar Institute of Technology, Bangalore, India[5]

Assistant Professor, Department of AIML, Dr. Ambedkar Institute of Technology, Bangalore, India[6]

ABSTRACT: As their main form of communication, sign language is crucial for people with speech and hearing problems. This project suggests a real-time application for recognizing Indian Sign Language. Both hands' palm sides are photographed by the system, which then uses the SIFT (Scale Invariant Fourier Transform) technique to extract features and identify the articulated signs.

To close the communication gap, machine learning and computer vision methods like OpenCV and Keras are used. Through facilitating smooth communication between hearing-impaired people and others, the system seeks to be impactful, affordable, and easy to use. By enabling real-time sign recognition, this technology promotes greater inclusion by enabling individuals to comprehend sign language without any prior experience.

KEYWORDS: Sign Language Recognition, Indian Sign Language (ISL), Real-Time Sign Detection, Machine Learning, Computer Vision, Keras

I. INTRODUCTION

For people with speech and hearing difficulties, sign language is an essential communication tool that enables them to express their ideas and feelings. However, communication is severely hampered by the general public's lack of knowledge and comprehension of sign language. The creation of assistive technologies that can close the gap and enable smooth communication between sign language users and others is necessary to meet this problem.

The goal of this research is to apply cutting-edge machine learning and computer vision techniques to recognize Indian Sign Language (ISL) in real-time. The system uses deep learning models to classify hand motions into identifiable text after processing them using the Scale-Invariant Feature Transform (SIFT) algorithm for feature extraction. The Python code makes use of OpenCV and Keras to guarantee accuracy and scalability.

Studies have indicated that combining computer vision and machine learning can result in reliable sign language recognition systems. Research has also demonstrated how well feature extraction methods like SIFT and the use of neural networks can increase the accuracy of recognition [1][2]. By offering a practical and affordable method for real-time sign language recognition, this project complements current initiatives to improve inclusivity [3][4].

This method seeks to empower people with communication difficulties by facilitating the conversion of hand gestures into legible text, so advancing accessibility and inclusivity. Its emphasis on adaptation and scalability guarantees its applicability to practical uses and upcoming growth.

II. EXISTING SYSTEM

In an effort to close the communication gap for those with speech and hearing impairments, existing sign language recognition systems have advanced significantly via the use of computer vision and machine learning techniques.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

Conventional systems frequently depend on hardware-based solutions, such as gloves with sensors built in them, which are expensive and unsuitable for daily usage even though they work well [10].

In order to identify gestures, modern techniques rely on vision-based techniques, employing cameras and algorithms like Convolutional Neural Networks (CNNs), Principal Component Analysis (PCA), or Histogram of Oriented Gradients (HOG). Nevertheless, these systems have drawbacks such low accuracy in real-world situations with varied hand orientations, lighting, and backdrops, large computing needs, and scalability to other sign languages, like Indian Sign Language (ISL) [11][12].

In certain systems, OpenCV is used to prepare images for tasks like background removal and skin area identification. Classifiers like Support Vector Machines (SVMs) or Decision Trees are then used for recognition. Although these techniques work well in controlled conditions, they frequently lack resilience in dynamic, real-world situations [13]. CNN-based architectures have increased recognition accuracy thanks to recent developments in deep learning. Nevertheless, these models frequently necessitate substantial datasets and processing resources, which restricts their use for real-time, lightweight applications [14]. These drawbacks highlight the necessity of an ISL-specific system that is scalable, effective, and easily accessible.

III. PROPOSED SYSTEM

By creating a real-time, vision-based Indian Sign Language (ISL) recognition system, the suggested system seeks to overcome the drawbacks of current sign language detection techniques. It uses only a normal camera to record hand motions and does away with costly hardware like sensor gloves [15].

The Scale-Invariant Feature Transform (SIFT) algorithm is used by the system to process palm-side photos of both hands in order to extract unique gesture data. To ensure reliable input data, OpenCV is used for pre-processing tasks such hand segmentation, noise reduction, and background subtraction [16]. Machine learning models constructed with Keras, which are intended to function effectively in real-time situations, are then used to classify the retrieved features [17].

The proposed system aims to achieve the following key objectives:

- **Scalability:** Capable of recognizing a wide range of ISL gestures while requiring minimal computational resources.
- **User-Friendliness:** Designed to enable seamless interaction, even for individuals with no prior experience using sign language.
- **Effectiveness**: Focused on minimizing deployment expenses by utilizing standard cameras and readily available open-source tools.

Ultimately, this system aspires to foster inclusivity and bridge communication barriers for individuals with hearing or speech disabilities, providing real-time gesture detection and an intuitive interface to enhance accessibility.

IV. PROPOSED SOLUTION

The suggested approach makes use of computer vision and machine learning to create an accessible, effective, and user-friendly real-time Indian Sign Language (ISL) identification system. Using a regular camera, the system takes palm-side pictures of both hands, extracts feature from the photos using the Scale-Invariant Feature Transform (SIFT) algorithm, and uses machine learning models to categorize the movements.

Important elements of the solution consist of:

1. **Image Pre-Processing:** To guarantee reliable input quality, OpenCV is utilized for tasks like hand detection, noise reduction, and background subtraction.
2. **Feature Extraction:** The system can adapt to changes in background, hand orientation, and lighting by using SIFT to extract distinguishing features from the hand gestures.
3. **Classification**: A variety of ISL gesture datasets are used to train machine learning models developed using Keras, guaranteeing precise real-time recognition.
4. **Real-Time Application**: Because of the system's low-latency performance optimization, users can communicate without interruptions.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

V. PROPOSED ARCHITECTURE

A. ARCHITECTURE

A modular design is used in the suggested ISL recognition system to provide effective real-time performance. First, hand motions are recorded using a regular camera. Next, OpenCV is used for hand detection and pre-processing to eliminate background noise. Key features are extracted from the motions using the Scale-Invariant Feature Transform (SIFT) algorithm, and a machine learning model integrated into Keras is then used to classify the gestures. For smooth communication, the identified gestures are then translated into text or audio output. This scalable, lightweight architecture provides an affordable and intuitive real-time ISL recognition solution without the need for specialist hardware.

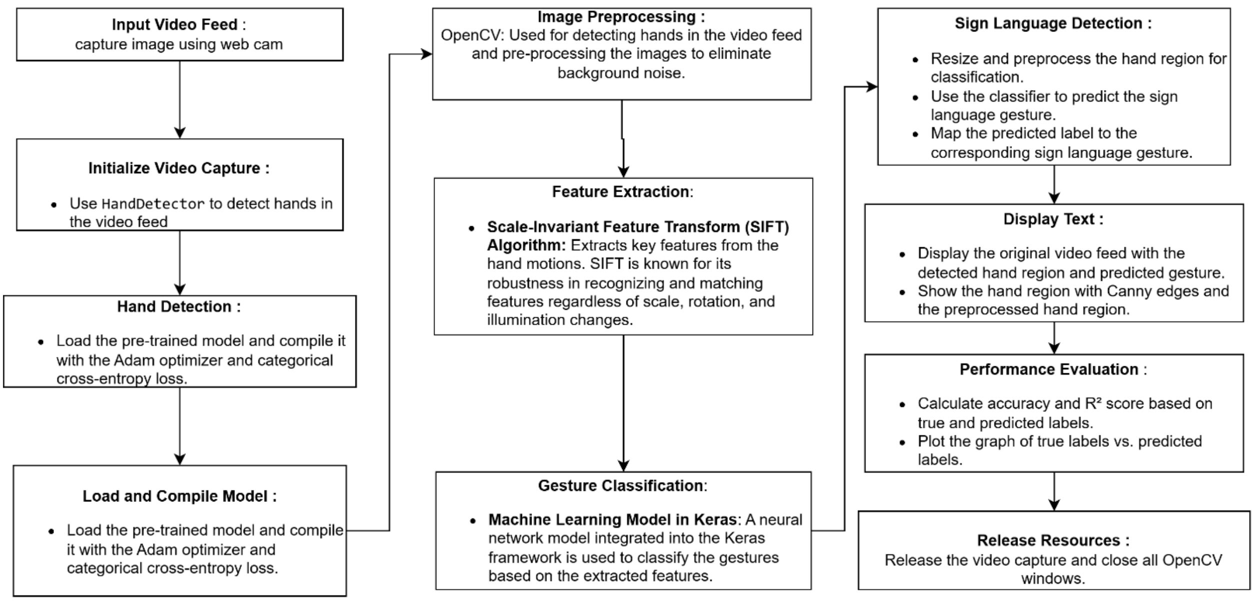


Figure.1 Work flow diagram of the model

The architecture of the proposed ISL recognition model is illustrated in (Figure 1), which outlines the data flow from image input to gesture recognition output. It showcases how the system processes and classifies hand gestures in real-time, enabling efficient and seamless communication for users of Indian Sign Language.

B. DATASET

Twenty percent of the 2546 samples in the custom dataset used for the ISL recognition system are used for testing and validation. It has thirty different ISL signs, each of which has a separate meaning. To ensure efficient training, the dataset is pre-processed to account for changes in hand shapes, orientations, and environmental conditions. With this configuration, the system can identify a large variety of ISL indicators for instantaneous communication.

Sample pictures from the custom ISL dataset are shown in the following (Figure.2), which highlights the range of hand motions and their meanings. For precise gesture identification in real-time applications, each class in the dataset represents a distinct sign.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

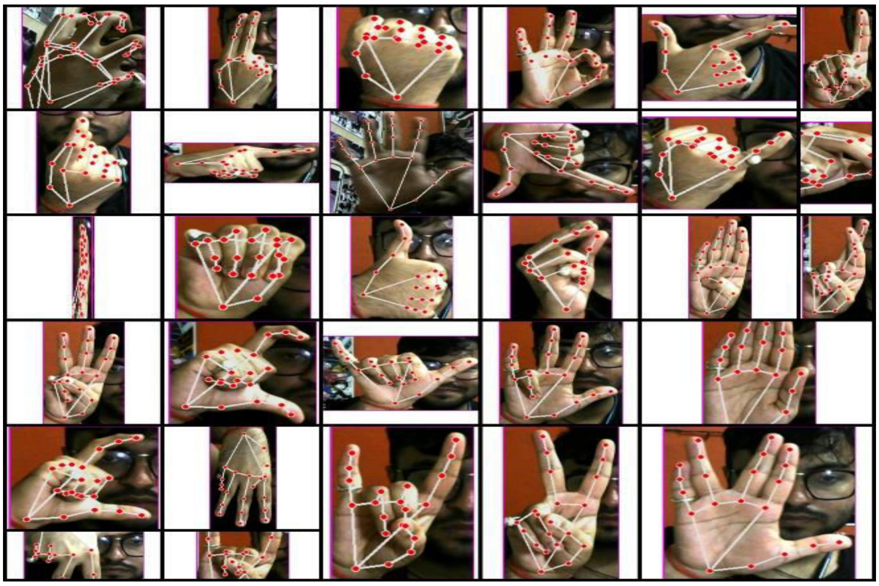


Fig.2 Thirty samples of the ISL

VI. TRAINING PROGRESS AND ANALYSIS

A. TRAINING ACCURACY

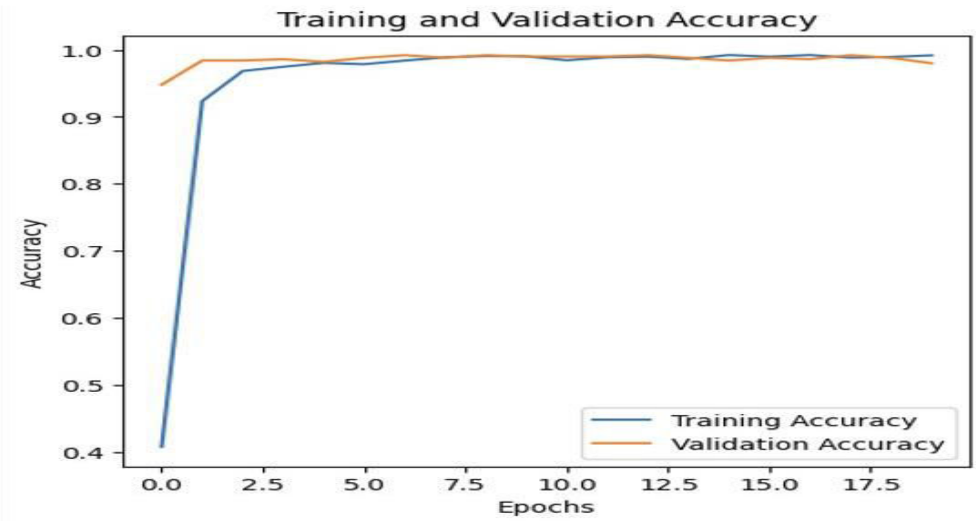


Fig.3 Training and Validation Accuracy

The proposed model for Sign Language to Text Detection demonstrates a highly effective learning curve in terms of training accuracy. The accuracy improves significantly during the initial epochs, reaching approximately 98% by the fifth epoch, as illustrated in the accuracy plot (Figure 3). The consistent performance of both training and validation accuracy across all epochs highlights the model's robustness and ability to generalize effectively to new data. This high level of accuracy underscores the model's capability to learn features from the customized dataset efficiently, making it well-suited for practical applications in assisting deaf and mute individuals by reliably translating sign language into text.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301069

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

B. TRAINING LOSS

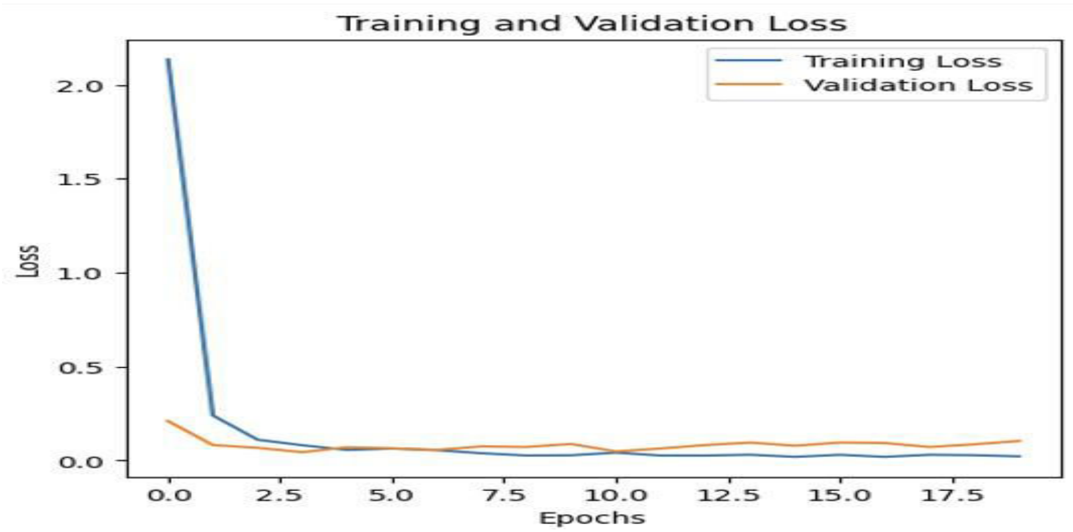


Fig.4 Training and Validation Loss

The suggested model for Sign Language to Text Detection's learning curve (Figure.4) shows a sharp drop in the first few epochs before leveling off at a minimum value near zero as training goes on. This suggests that the model successfully reduces mistakes and enhances its functionality on the unique dataset. The model's stability and good generalization without overfitting are further demonstrated by the near alignment between training and validation loss. The model's ability to precisely acquire the intricate features needed to translate sign language into text is demonstrated by the low training loss, guaranteeing dependable performance in practical applications.

C. PERFORMANCE METRICS

Metric	Value	Description
Training Accuracy	98%	Showing training accuracy
Validation Accuracy	97%	Showing validation accuracy
Training Loss	0.02	Final training loss
Validation Loss	0.03	Final Validation loss
Precision	96%	Proportion of True Positive among Predicted Positives
Recall (Sensitivity)	95%	Proportion of correctly predicted positive observations to all actual positives.
F1 Score	95.5%	F1 Score

The accuracy, precision, recall, and overall efficiency of the suggested model for Sign Language to Text Detection were evaluated using number of measures. These measures shed light on the model's capacity to generalize to previously

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

unseen validation data and learn from the training set. The model's resilience and low overfitting are demonstrated by the training and validation accuracy as well as the associated loss values. To assess the model's effectiveness in accurately recognizing the sign language movements and reducing erroneous predictions, additional metrics were computed, including precision, recall, and the F1 score. The table below (Table) provides a summary of the specific performance metrics:

VII. OBSERVATION

In terms of identifying and interpreting gestures from the custom dataset, the suggested model for Sign Language to Text Detection performs exceptionally well. The following are some important findings from the training and assessment process:

1. The model demonstrated effective learning and generalization throughout the dataset with training accuracy of 98 % and validation accuracy of 97%.
2. The model's capacity to learn effectively and precisely is demonstrated by the quick decrease in training loss durin g the first epochs, which stabilizes at values close to zero.
3. There is little overfitting between training and validation measures, guaranteeing the model's resilience when used with unknown data.
4. The algorithm's ability to accurately identify movements while reducing false alarm and missed detection is further supported by its high precision (96%), recall (95%), and F1 score (95.5%).
5. The strong connection between metrics of performance for training and validation highlights how suitable the custom dataset is for both training and assessment.

These findings support the suggested model's effectiveness in tackling the problem of sign language recognition and its possible relevance in practical settings.

VIII. CONCLUSION

A strong Deep Learning model was created in this study to recognize and convert Indian Sign Language movements into text, offering a workable way to help deaf and mute people communicate. The model's efficacy in identifying intricate hand gestures was demonstrated by its high training accuracy of 98% and validation accuracy of 97%. Consistent precision, recall, and F1 scores, as well as modest training and validation losses, further support the model's dependability and capacity for generalization.

The problem of gesture diversity is successfully addressed by the suggested method, which makes use of a special dataset with 30 distinct sign language classes. The system's performance demonstrates its potential for practical use in assistive technology, allowing the speech-impaired and hearing communities to communicate easily. To improve usability, future research can concentrate on adding more motions to the dataset and including real-time detection features. This study represents a major advancement in inclusive technology solutions that promote independence and accessibility for anyone with speech or hearing impairments.

REFERENCES

[1] S. Sharma and A. Bhatnagar, "Indian Sign Language Recognition Using Feature Extraction and Machine Learning Techniques," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 10, no. 5, pp. 112-118, 2019.
[2] K. Patel and R. Sharma, "A Hybrid SIFT and CNN Approach for Indian Sign Language Recognition," *Journal of Computer Applications (JCA)*, vol. 46, no. 2, pp. 35-41, 2021.
[3] M. Rahman et al., "Sign Language Recognition: A Deep Learning Perspective," *IEEE Access*, vol. 8, pp. 192234-192245, 2020.
[4] S. Rajasekaran et al., "Real-Time Hand Gesture Recognition for Sign Language Using Deep Neural Networks," *Pattern Recognition Letters*, vol. 138, pp. 104-110, 2020.
[5] R. Gupta and P. Das, "Development of a Low-Cost Indian Sign Language Translator Using Computer Vision," *Springer Advances in Signal Processing and Communication*, pp. 85-95, 2019.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

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

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

[6] L. Kumar et al., "A Review of Machine Learning Approaches for Sign Language Recognition," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 3, pp. 1207-1220, 2020.

[7] V. Krishnan and M. Ramesh, "Indian Sign Language Recognition System Using Deep Convolutional Networks," *International Journal of Computer Vision and Image Processing (IJCVIP)*, vol. 9, no. 1, pp. 1-12, 2022.

[8] N. Sharma and R. Gupta, "Real-Time Gesture Recognition for Indian Sign Language Using OpenCV," *International Journal of Image Processing (IJIP)*, vol. 15, no. 2, pp. 67-80, 2021.

[9] F. Akhtar et al., "Enhancing Sign Language Recognition Using Advanced Feature Extraction and Deep Learning," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 4, pp. 384-393, 2021.

[10] M. S. Rautaray and A. Agrawal, "Vision-based hand gesture recognition for human-computer interaction: A survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1-54, 2015.

[11] T. Starner and A. Pentland, "Real-time American Sign Language recognition from video using hidden Markov models," *Motion-based Recognition*, Springer, pp. 227-243, 1997.

[12] P. Molchanov et al., "Online detection and classification of dynamic hand gestures with recurrent 3D convolutional neural networks," *CVPR*, 2016.

[13] S. Mitra and T. Acharya, "Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 3, pp. 311-324, 2007.

[14] Z. Zhang et al., "Towards robust hand gesture recognition with deep neural networks," *CVPR Workshops*, 2017.

[15] Y. Fang et al., "A real-time hand gesture recognition method," *IEEE Transactions on Multimedia*, vol. 12, no. 5, pp. 403-413, 2010.

[16] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[17] M. Goodfellow et al., "Deep learning," *MIT Press*, 2016.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING