International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

SIGN LANGUAGE RECOGNITION USING GENERAL DEEP NEURAL NETWORK BASED ON LARGE SCALE DATASET

MAHENDRA BABU D[1], KRISHNA VAMSI G[2], SRINIVASULU I[3], VIKHIL M[4], Dr.G.KAVITHA[5]

[1]COMPUTER SCIENCE AND ENGINEERING MUTHAYAMMAL ENGINEERING COLLEGE RASIPURAM,TAMILNADU
devinenimahendra@gmail.com
[3]COMPUTER SCIENCE AND ENGINEERING MUTHAYAMMAL ENGINEERING COLLEGE RASIPURAM,TAMILNADU
srinivaschowdaryinturi9199@gmail.com
[2]COMPUTER SCIENCE AND ENGINEERING MUTHAYAMMAL ENGINEERING COLLEGE RASIPURAM,TAMILNADU
krishnavamsigunturu819@gmail.com
[4]COMPUTER SCIENCE AND ENGINEERING MUTHAYAMMAL ENGINEERING COLLEGE RASIPURAM,TAMILNADU
mekalavikhilm@gmail.com
[5]M.S.,Ph.D., *PROFESSOR & HEAD MUTHAYAMMAL ENGINEERING COLLEGE* RASIPURAM,TAMILNADU
Hod.cse@mec.edu.in

ABSTRACT—

People with speech or hearing impairments frequently utilize sign language, which is a system of visual gestures and signs. This research proposes a method for identifying alphabetic gestures used in sign language. The goal of this project is to create a real-time sign language recognition system for deaf and mute individuals utilizing Python, OpenCV, and deep learning techniques like YOLOv5 and Convolutional Neural Networks (CNN).

Index Terms—Sign Language, Deep Learning, Convolutional Neural Networks(CNN),YOLOv5, OpenCV,Gesture Recognition.

Introduction :

Sign language is an essential mode of for individuals with speech and hearing impairments, allowing them to interact with others using visual gestures. However, many people who do not use sign language find it difficult to understand and communicate with the deaf and mute community. This communication gap often limits accessibility and inclusivity, making it challenging for hearing- impaired individuals to fully integrate into society. The development of automated sign language recognition systems can help bridge this gap by providing real-time interpretation of sign language gestures into text or speech.

Recent advancements in deep learning and computer vision have paved the way for more efficient and accurate sign language recognition systems. Traditional sign language interpretation methods require human interpreters, which are not always available and can be expensive. Automated systems utilizing deep neural networks (DNN) and real-time image processing techniques offer a viable solution to this challenge.

This research focuses on implementing an AI-driven system that can recognize hand gestures corresponding to sign language symbols in real time. The system leverages a combination of deep learning models, including YOLOv5 for object detection and Convolutional Neural Networks (CNN) for classification, to accurately interpret gestures.

The primary goal is to develop an automated sign language recognition system capable of real-time gesture detection using a webcam. By leveraging YOLOv5 (You Only Look Once) and CNN models, the system can accurately identify hand gestures and convert them into readable text. This approach enhances accessibility for deaf and hard-of-hearing individuals, facilitating seamless communication with the broader community.

To train the model, a custom dataset was created using OpenCV, capturing 800 images per gesture for training and

200 images per gesture for testing. Preprocessing steps included grayscale conversion, Gaussian blur, adaptive thresholding, and image resizing. The classification model was trained using SoftMax activation and the TensorFlow Object Detection API.

To explore future improvements, such as multi-language sign recognition and integration with speech synthesis technologies.

With the increasing need for inclusive communication solutions, the development of real-time sign language recognition systems can significantly enhance accessibility for the hearing-impaired community.

PROBLEM STATEMENT :

To extract features from the required image Deep Learning SSD ML algorithm is used. For the detection, TensorFlow Object Detection API is used where the extracted features from the pictures taken are passed onto the TensorFlow module which goes to create comparisons with the real time video present within the frame. On detection of any of those features it's visiting, generate a bounding box round the gesture and make the prediction. The prediction goes to be identical because of the label of the image. we are going to be ready to detect linguistic communication in real time using OpenCV.
The project aims to develop a real-time sign language recognition system using deep learning techniques such as YOLOv5 and Convolutional Neural Networks (CNN).The system captures real- time video input via a webcam, processes the data using OpenCV, and recognizes hand gestures.The SSD ML algorithm extracts image features and detects gestures using TensorFlow Object Detection API.
One major challenge is the lack of larger and more semantically rich datasets, limiting translation accuracy.

OBJECTIVES :

To develop a deep learning-based model for recognizing hand gestures in sign language with high accuracy.
To create and preprocess a dataset of sign language gestures train and test the model effectively.
To implement real-time recognition of sign language gestures using OpenCV and YOLOv5.
To evaluate system performance using metrics such as F1-score, precision, recall, and Mean Average Precision (mAP).
To optimize computational efficiency for smooth real-time processing on various hardware platforms, including mobile devices and embedded systems.
To explore future improvements, such as multi-language sign recognition and integration with speech synthesis technologies.more effectively by delivering timely, tailored support.

A. ABBREVIATIONS AND ACRONYMS

The project “Sign language recogination” employsvarious technologies and techniques, including American Sign Language(ASL), Convolution Neural Network(CNN), Deep Learning(DL), Deep Neural Network(DNN), Graph Convolution Neural Networks(GCNN),Indian Sign Language(ISL),K-nearest Neighbors Algorithm(KNN), Machine Learning(ML) Recurrent Neural Network(RNN), Right of Person With Disabilities(RPWD), You Only Look Once(YOLO).

SYSTEM DESIGN AND BLOCK DIAGRAM :

A. PROPOSED SYSTEM

A proposed system for sign language detection would involve multiple components working together to recognize and interpret sign language gestures. Here is a possible overview of the components: Video capture: A camera or webcam captures video footage of the signer's hand gestures and facial gestures and facial expressions. Hand detection: Computer vision techniques are used to identify the signer's hand(s) in the video footage, even if they are partially obscured or moving quickly. and tracking: Once the hand(s) have been detected, computer vision algorithms are used to track their movements over time, allowing the system to recognize when a sign has started and ended.

Advantages:

Sign Language detection system shows what the position of hands in viewfinder of camera module means with good accuracy. It can then be used to help people who are just beginning to learn Sign Language or those who don't know sign language but have a close one who is deaf..

B. BLOCK DIAGRAM

A block diagram is a graphical representation of a system or process that shows relationships between components or steps. It uses blocks/boxes and arrows to illustrate complex systems, identify relationships, and analyze/optimize performance.

DATA SET COLLECTION :

Dataset Collection & Preprocessing:

To gather high-quality sign language images and preproces them for deep learning model training.

Data Collection:

A dataset of 800 images per gesture is collected for training. 200 images per gesture are used for testing.
The images are captured using OpenCV and a webcam in different lighting conditions.

Preprocessing:

Grayscale Conversion: Converts images to black and white to reduce computational complexity.
Gaussian Blur: Removes noise from the image.
Adaptive Thresholding: Enhances contrast to highlight the hands.
Image Resizing: Standardizes image dimensions to 128x128 pixels for consistent processing.
Dataset Augmentation (Optional)
Rotation, Scaling, and Flipping are applied to increase dataset variability.
This improves the model's ability to recognize gestures in differentorientations.

IMPLEMENTION OF THE DETECTION MODEL :

To detect hand gestures in real-time using deep learning models. Video Capture:

A camera/webcam continuously records real-time video. Frames are extracted at a defined frame rate (e.g., 30 FPS).

Hand Detection:

The YOLOv5 object detection model is used to identify hands video frame.

A bounding box is drawn around detected hands. Gesture Recognition:

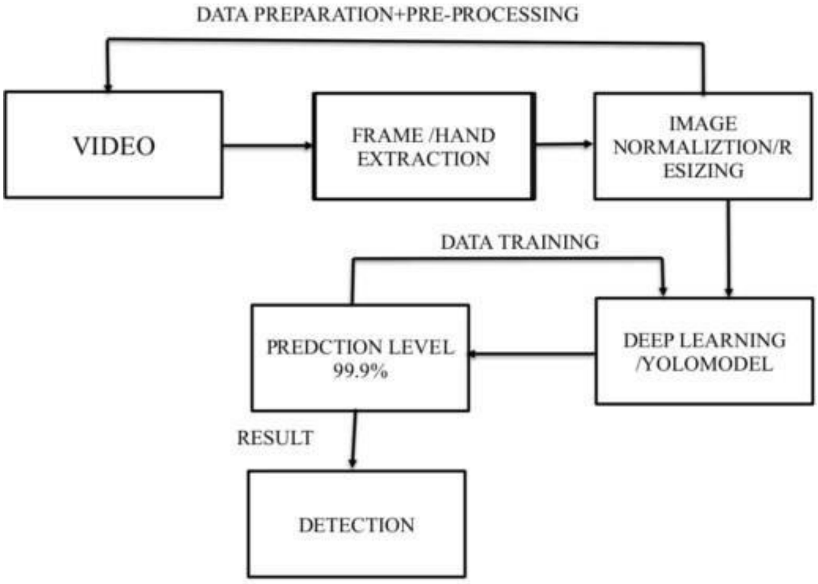


Figure 1.Deep Learning Block Diagram

Gesture Recognition:

The extracted hand image is passed to a CNN model.

The model classifies the gesture based on learned features.

Prediction Thresholding:

If a gesture is detected for 50 consecutive frames, it is considered valid.

A difference threshold (e.g., 20) ensures that incorrect detections are filtered out.

K-Nearest Neighbors (KNN) is a supervised machine learn- ing algorithm that classifies new data points based on their similarity to existing data points.

MODEL TRAINING AND TESTING :

To train a deep learning model to recognize *sign language gestures accurately*.

Feature Extraction Using CNNs:

The system uses Convolutional Neural Networks (CNNs) to extract features such as:

- -Finger position
- Hand orientation
- Motion trajectory

Classification Using SoftMax Function:

The final layer applies SoftMax activation, converting model outputs into probability scores.

The gesture with the highest probability is selected as the final prediction.

Training & Evaluation Metrics:

Accuracy, F1-score, and Confusion Matrix are used to measure performance.

The system achieves an accuracy of 98-99% on a controlled dataset.

Testing & Optimization:

Testing is conducted using real-time video inputs.

Hyperparameters such as learning rate, dropout rate, and batch size are optimized for better performance. can be used to support safety- related interventions by identifying relevant posts in realtime.
KNN is commonly used for classification and regression tasks.

GESTURE CLASSIFICATION AND FORMATION :

To convert recognized gestures into meaningful words and sentences.

Word Formation :

Individual letters detected from gestures are stored in a sequence.
If a letter appears *50 times in a row*, it is added to the output text.

Space Detection:

A predefined "blank" gesture indicates a word break. This allows for continuous sentence formation.

Real-Time Display:

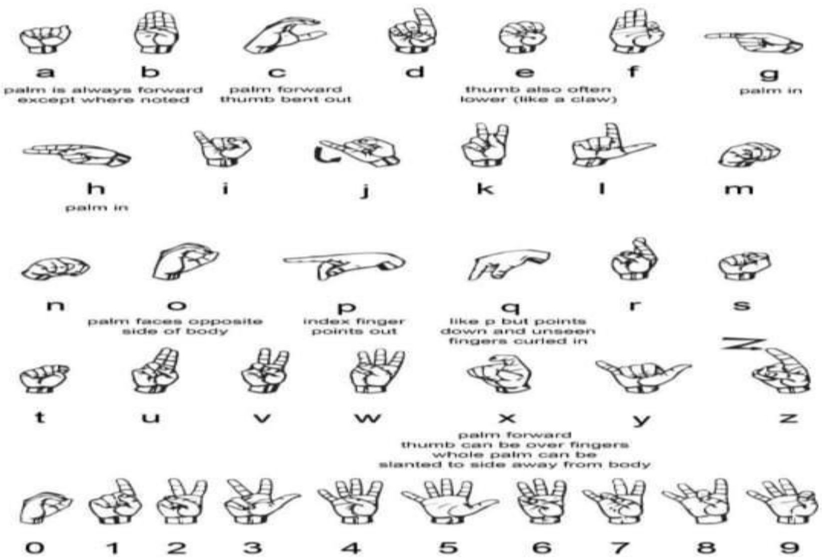
Recognized words are displayed on the *screen*.
The user receives *instant feedback* on detected gestures. Future Enhancements

- -Speech Output Integration: Convert recognized text into spoken words.
- Full Sentence Recognition: Use Recurrent Neural Networks (RNNs) or Transformers to understand context.
- Multi-Language Support: Train the model on multiple sign languages (ASL, BSL, ISL, etc.).

Figure 2 : Image processing FlowChart

Sign language is an essential mode of communication for individuals with speech and hearing impairments, allowing them to interact with others using visual gestures. However, many people who do not use sign language find it difficult to understand and communicate with the deaf and mute community.
Recent advancements in deep learning and computer vision have paved the way for more efficient and accurate sign language recognition systems. Traditional sign language interpretation methods require human interpreters, which are not always available and can be expensive. Automated systems utilizing deep neural networks (DNN) and real-time image processing techniques offer a viable solution to this challenge.
This research focuses on implementing an AI-driven system that can recognize hand gestures corresponding to sign language symbols in real time. The system leverages a combination of deep learning models, including YOLOv5 for object detection and Convolutional Neural Networks (CNN)

Figure 3: Sign Language Alphabets of the ASL



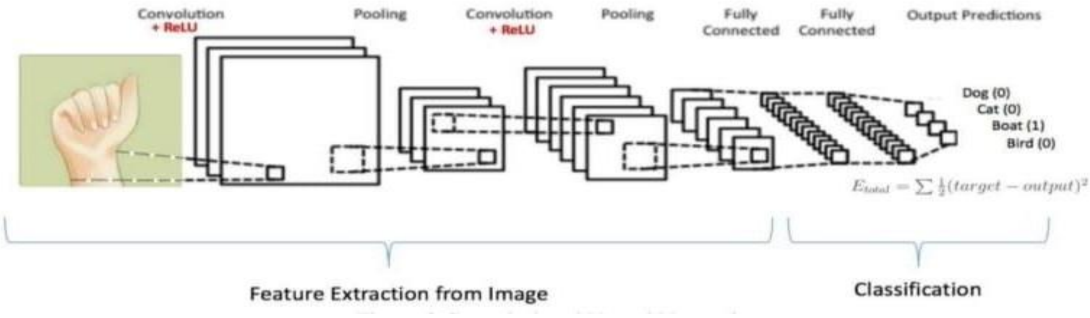


Figure 4 : Convolutional Neural Network

CONCLUSION :

This research presents a deep learning-based real-time sign language recognition system using YOLOv5 and CNN. The developed system demonstrates significant accuracy in recognizing hand gestures and translating them into text. The modular approach enhances flexibility, scalability, and real-time performance, making the system suitable for deployment in assistive communication technologies

Future improvements will focus on:

Expanding the dataset to include more variations in hand gestures and regional sign languages. Enhancing the model's ability to recognize complex sentences rather than individual signs.

Integrating with speech synthesis technology to provide voice- assisted communication. Deploying the model on lightweight mobile and embedded systems to increase accessibility.

The future of sign language detection is poised for significant advancements, particularly through the integration of technologies like MediaPipe and deep learning models. Research is focusing on improving real-time recognition accuracy and accessibility for various sign languages, including American Sign Language (ASL) and Indian Sign Language (ISL). Innovations in gesture recognition and natural language processing (NLP) are expected to enhance the effectiveness of these systems. Development of comprehensive databases that include various sign languages and dynamic symbols.

REFERENCES :

1. M. Rafi, N. Nawal, N. S. N. Bayev, L. Nima, C. Shahnaz, and S. A. Fattah, "Image-based Bengali sign language alphabet recognition for deaf and dumb community," in Proc. IEEE Glob. Humanitarian Technol. Conf., 2019, pp. 1–7.
2. Boukdir, M.Benaddy,O.E.Meslouhi, M.Kardouchi, and M.Akhloufi,''Character-level Arabic text generation from sign language video using encoder–decoder model,'' Displays, vol. 76, Jan. 2023, Art. no. 102340, doi: 10.1016/j.displa.2022.102340.
3. C. K. Sikdar, M. K. Hasan, and M. I. T. Khan, "Bangla Sign Language alphabet recognition using deep convolutional neural networks," in Proc. IEEE Int. Conf. Adv. Comput. (ICAC), 2018, pp. 1–5.
4. F.S.Alamri,A.Rehman,S.B.Abdullahi,andT.Saba,''Intel ligentreal-life key-pixel image detection system for early Arabic sign language learners,'' PeerJ Comput. Sci., vol. 10, p. e2063, Jun. 2024, doi: 10.7717 /peerj cs.2063.
5. H. Luqman, ''ArabSign: A multi-modality dataset and benchmark for continuous Arabic sign language recognition,'' in Proc. IEEE 17th Int. Conf. Autom. Face Gesture Recognit. (FG), Jan. 2023, pp. 1–8, doi: 10.1109/FG57933.2023.10042720.
6. H. R. Jabbari and M. B. Shahraray, "Persian sign language recognition using spatio-temporal features and convolutional neural networks," in Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT), 2019, pp. 1–5.
7. J. Shin et al., "Korean sign language recognition using transformer based deneural network," Appl. Sci., vol. 13, no. 5, 2023, Art. no. 3029.
8. J. Wu, C. Liu, and Y. Zhang, "Chinese sign language recognition with deep learning," in Proc. IEEE Int. Conf. Signal Process. Commun. (ICSPC), 2021.
9. M. M. Balaha, S. El-Kady, H. M. Balaha, M. Salama, E. Emad, M. Hassan,andM. M. Saafan,''Avision- baseddeeplearningapproachfor independent-users Arabic sign language interpretation,'' Multimedia Tools Appl., vol. 82, no. 5, pp. 6807–6826, Feb. 2023, doi: 10.1007/s11042-022 13423-
10. R. Sharma, K. Kumar, and R. Yadav, "Indian sign language recognition using Kmeans clustering and ANN," in Proc. IEEE Int. Conf. Innov. Comput. Commun. (ICICC), 2019, pp. 1–7.
11. S. B. Abdullahi, K. Chamnongthai, V. Bolon-Canedo, and Cancela, ''Spatial–temporal feature-based end-to-end Fourier network for 3D sign language recognition,'' Expert Syst. Appl., vol. 248, Aug. 2024, Art. no. 123258, doi: 10.1016/j.eswa.2024.123258.
12. .S. P. Maity, A. K. Das, and S. K. Ghosh, "Real-time Indian sign language recognition system using skin color segmentation and wavelet transform," in Proc. IEEE Region 10 Conf. (TENCON), 2019, pp. 1–6.