Real Time Sign Recognition using YOLOv8 Object Detection Algorithm for Malayalam Sign Language

Esther Daniel[1], V. Kathiresan[2], Priyadarshini .C[1], Golden Nancy .R [3,*], P. Sindhu[4]

[1]Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India

[2]GITAM University (Deemed to be University), Karnataka, India

[3]Department of Artificial Intelligence and Machine Learning, Karunya Institute of Technology and Sciences, Coimbatore, India

[4]Department of Computer Science and Engineering, Rajiv Gandhi College of Engineering, Anna University, Chennai, 602105 , India

Emails: estherdaniell@gmail.com; xyzkathir@gmail.com; priyadharsini@karunya.edu; goldennancy@karunya.edu; sindhupandi26@gmail.com

Abstract

Sign language recognition is important for enhancing message and user-friendliness for the community of deaf and hearing-impaired people. This paper proposes a Malayalam Sign Language (MSL) method using sign language that emerged from the state of Kerala. The main factor contributing to this emergence of such regional sign language is the absence of a standardized and consistent approach to the use of Indian Sign Language (ISL) in various states. This is due to the variations in signs, grammar, and syntax used in different regions. The system uses the You Only Look Once v8 (YOLOv8) algorithm-based object detection method which is based on Convolution Neural Network (CNN), a widely accepted deep learning neural network design employed mainly in computer vision. As the dataset for MSL is not publicly available, we used an MSL video from YouTube provided by the National Institute of Speech and Hearing for training a custom model. We pre-processed the video to extract the frames and annotate them with sign labels. Then, we trained the YOLOv8 algorithm on the annotated frames to detect the hand region and recognize signs in real time. The proposed approach achieved an accuracy of 97.21% calculated from the mean Average Precision value on the MSL dataset. The result achieved outperformed other existing approaches even while using less dataset count compared to others.

Keywords: Malayalam sign language; Yolo, Computer vision; Deep learning; Machine learning; Convolution neural network

1. Introduction

Sign language is essential for communication among hearing-impaired people as it helps them to communicate their needs and interact with each other. Despite its significance, it continues to be a challenge due to the differences in signs shown by people worldwide. Compared with American Sign Language (ASL) or Indian Sign Language (ISL), Malayalam Sign Language (MSL) has been given more research attention. As per the record of the World Federation of the Deaf [8], over 70 million people who are hearing-impaired use sign language as their prime approach to interact. However, these individuals face significant challenges and barriers in many countries, including limited access to education and resources, inadequate instruction, and lack of recognition for sign languages as official languages. In India, for instance, approximately 18 million hearing-impaired persons depend on sign language as their prime approach to communication [9] [10].

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

Despite its crucial role for the community of deaf people, Indian Sign Language (ISL) poses extensive blockades to acceptance and acknowledgment in several states of India due to a lack of standardization and homogeneity in its application process. The main factor contributing to this issue is the absence of a standardized and consistent approach to using ISL in different states. This is due to the noticeable variations in signs, grammar, and syntax used in different regions, creating communication barriers that exist between several sign languages across different states. This leads to many regional sign languages emerging in various Indian states, such as Malayalam Sign Language (MSL). This has led to the emergence of regional sign languages, such as MSL, which is mainly used in Indian states like Kerala.

To address these issues, it is crucial to develop effective sign language recognition systems and enhance the community of the hearing impaired to have access to education and communication. Such advancement can solve the difficulty of non-sign language users have to communicate with sign language users and give the hearing-impaired access to modern technology. While there have been several studies on the recognition of sign language for other languages, there is limited research on the use of MSL systems for recognition. This significant knowledge gap motivated the study of sign language recognition for Malayalam Sign Language, to offer an effective communication method for the hearing-impaired and those who use sign language as their focal means of communication within the Malayalam-speaking community. Creating an MSL recognition system can help bridge this gap and preserve the rich linguistic and cultural diversity it's known for in the deaf community in Kerala, India.

The implementation of CNN with YOLOv8 architecture was added to identify MSL through the video feed. The YOLOv8 pre-trained model is reinstructed using machine learning through modifications to several channels and classifications by the Malayalam Sign Language requirement. A dataset made up of a few photos and 51 classes of static sign language was collected by taking screenshots of each sign from a video provided by the National Institute of Speech and Hearing. In general, the dataset was split into training dataset, validation process, and testing dataset sets, with a ratio of 80:10:10. Performance was measured with the help of the following processes recall, precision, mAP, accuracy, and F1 score metrics.

The rest of the paper is divided into 5 sections. The context of the work and motivation are presented in the introduction section. Section 1 discusses the numerous studies on sign language recognition that have been done using various methods. The proposed system employed in the development Malayalam language through sign language recognition is described in Section 2. Experimental works are discussed in Section 3. Section 4 summarizes the findings and observations made in this study, and Section 5 wraps up the findings and briefs the system's potential for future development.

2. Related works

Sign language recognition has been an increasingly important area in the field of research as it can help in the process of improving communication and accessibility for hearing-impaired who use sign language and help open new opportunities for them. However, developing sign language recognition systems for specific sign languages can be challenging due to factors such as differences in sign language structure, sign variations, and the limited availability of datasets for training and testing. Although MSL is recognized as a distinct sign language, research on MSL recognition has been limited due to the lack of globally available datasets for MSL. Thus, it is difficult to train and test sign language recognition models specifically for MSL, making the development of such assistive technologies for MSL users a difficult task.

However, with the increasing interest in sign language recognition, there has been some progress in the development of MSL recognition systems. Researchers have attempted to develop MSL recognition systems using various techniques. Despite these efforts, MSL recognition systems are still in the initial stages of development and thus require a need for further research in this area compared to other sign language studies. The need for accurate and reliable MSL recognition systems will create an impact on the lives of users who use MSL as their language by providing them with better access to communication and information.

A robust hand gesture recognition system for MSL was suggested by Ajmi et al. [1] that can efficiently help in tracking static hand gestures using Inception modules employed in CNN for more efficient computation. An application was developed that captures a person's sign gestures for MSL using a webcam and translate them into corresponding text messages and audio in real time. This proposed model was able to achieve a 95% accuracy rate among the models tested in this study.

However, another study by Kunjumon et al. [2] proposed a system that could recognize ISL and then convert it into text and speech in English and Malayalam languages, which could be displayed on a mobile phone with the help of a sensor called Flex, Arduino UNO and Bluetooth HC 05. The system was designed to address the

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

communication difficulties faced by hearing-impaired people who were unable to express their language or thoughts verbally. The system aimed to overcome the barrier by recognizing Indian Sign Language and providing an output in text and speech in two languages with an accuracy rate of approximately 90% with the proposed model. A method for decoding the anticipated labels as text as well as speech and decoding the Indian sign language alphabets and numerals in a live video feed was proposed by Katoch et al. [3]. The technique used the Bag of Visual Words model (BOVW) for segmentation based on the color of the skin and subtraction of the background. Features from SURF were extracted through the images and it is used to map the signs with similar labels through which histograms were generated. The classification was done using CNN and SVM. Additionally, a GUI was developed to provide easy access to the system. The system had all the 36 ISL static alphabets and the digits which was effectively trained and was able to achieve an accuracy of 99%.

Even though the proposed methods by different authors were able to achieve promising results using different algorithms such as Support Vector Machine, CNN, and SURF in computer vision, each of these algorithms has its strengths and limitations. The advancement in recent developments in object detection algorithms has led to the emergence of new algorithms such as You Only Look Once (YOLO) [4] that have shown promising results in sign language recognition.

Areeb et al. [5] proposed ISL hand gesture recognition-based models using deep learning to effectively predict emergency signs. They used a dataset containing video feeds for 8 different emergencies and several frames were taken, which were then deployed to three different models. A 3D Convolutional Neural Network (3D CNN) was the 1st and 2nd model, consisting of a Recurrent Neural Network with Long Short-Term Memory (RNN-LSTM) and a pre-trained model of VGG-16 architecture framework, and the third model was focused on YOLO v5, an effective object recognition algorithm and out of the three models it was found that YOLO based model outperformed the other models, accomplishing a mean average value of 99.6%.The LSTM models are effective for predictions

The study by Daniels et al. [6] developed a sign language-based recognition system using the YOLO algorithm to process input from video-based data in real time. The YOLO algorithm is rooted in a CNN, which is both fast and accurate. The trained model of YOLOv3 was retrained with modifications to the levels and classes available to meet the sign language recognition-based requirements. A dataset based on the Indonesian Sign Language (BISINDO) was collected for the study. In this study, when using the image-based data, the system was successful in achieving 100%. The system performed at a pace of 8 frames per second with a precision value of 77.14%, recall value of 93.1%, accuracy value of 72.97%, and F1 score value of 84.38% when using video based data.

A system to detect the 26 alphabets of the Indonesian Sign Language System (BISINDO) was proposed by Luthfy et al. [7] using the YOLOv5 method. The system was tested under multiple scenarios, including varying camera distance, background video capture, and lighting levels. The system was found to detect 26 sign language alphabets in real time without being influenced by the background and lighting level. However, it was affected by camera and object distance. The best performance configuration was achieved using a dataset with a distribution of 70:20:10 ratio for data training, data validation, and testing. The batch size of 16 with 300 epochs and 0.01 for the learning rate produced an accuracy value of 99.27%.

Table 1: Test model Specifications and test conditions.

Authors	Algorithm	Dataset	Accuracy
Ajmi et al.[1]	CNN (Inception modules)	MSL	95%
Kunjumon et al. [2]	Hardware-based	ASL & MSL	90%
Katoch et al.[3]	SVM, CNN & SURF	ISL	99%
Areeb et al. [5]	YOLOv5	ISL	99.6%
Daniels et al. [6]	YOLOv3	BISINDO	72.97% (for Video) 100% (for Image)
Luthfy et al. [7]	YOLOv5	BISINDO	99.27%

The studies demonstrated the potential of using CNNs and object-detecting methods such as YOLO to develop real-time sign language systems as shown in Table 1. These types of systems have the strength to greatly improve

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

the communication between people with hearing impairments and those without, by allowing sign language to be translated into text or spoken language in real-time which can hence open new opportunities for the people and reduce the gap they feel between normal people. Deep learning and machine learning algorithms, such as YOLO and LSTM, have shown significant versatility across various domains, including medical imaging and financial forecasting [13-15]. Additionally, the use of deep learning methodologies is expanding to efficiently recognize sign language, enhancing communication for the hearing impaired [16-17]. While extensive research has been conducted on ISL, ASL, and BISINDO, there remains a gap in studies focused on Malayalam Sign Language (MSL).

3. Real time sign recognition using object detection YOLOV8

However, based on previous studies, it was found that the YOLO algorithm outperformed other algorithms because of the way it handles object detection. A single neural-based network is used by the object-based recognition algorithm YOLO to identify items in an image. The images were divided into a grid, and predictions were made for each of the grid cells about the bounding of boxes, object-ness scores, and probabilities of the class [18-19]. YOLOv5 was an improved version of the YOLO algorithm that uses a novel backbone network and a modified detection head to achieve higher accuracy and faster inference. But YOLOv8, the latest version was even further improved. YOLOv8 uses a more efficient backbone network with Cross Stage Partial (CSP) architecture and a decoupled head that distinguishes the branches of classification and detection. It also uses Distance-IoU Loss (DIoU) and Dynamic Feature Loss (DFL) to enhance the reliability of object detection. As YOLOv8 is a cutting-edge object identification algorithm that can precisely recognize hand motions in sign language it is excellent for real-time recognition of sign language. It has a short inference time and can operate on low-power devices [20-21]. Figure 1 represents the dataset converted into the latest YOLOv8 format and the proposed model to recognize sign language.

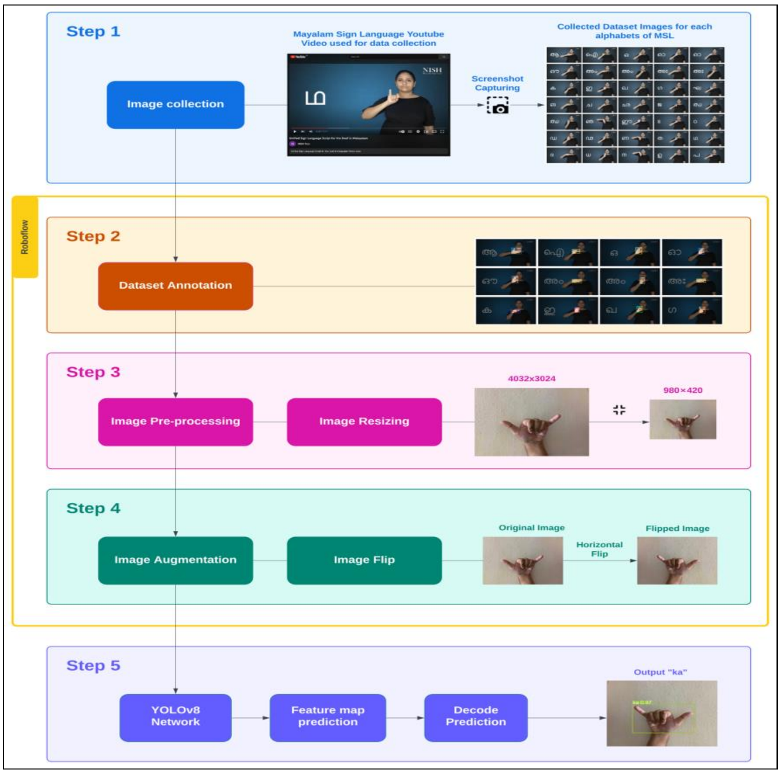


Figure 1. Proposed system for MSL

YOLOv8 algorithm architecture consists of different blocks as displayed in Figure 2. P1, P2 P3, P4, and P5 are feature pyramids, these are the fundamental elements of recognition systems for finding objects at various scales. To increase the model's precision and performance, a new backbone network known as Cross Stage Partial (CSP) was added to the YOLOv8 architecture. The CSP was created to lower the computing burden of the model while preserving accuracy. It operates by dividing the network into two sections and sending input data through each section simultaneously. The input data was subjected to convolutional operations in the first part, while a more

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

involved transformation was carried out in the second. The outcomes from the two sections were then merged and sent to the following layer. As a result, computation resources may be used more effectively, leading to quicker and more precise results. A C3 module, which consists of three Convolution modules and several bottleneck blocks, is used as the backbone of YOLOv5. A new C2f module is utilized in YOLOv8, nevertheless. The bottleneck blocks and 2 ConvModules that make up the C2f module are combined by a Split and Concat operation.

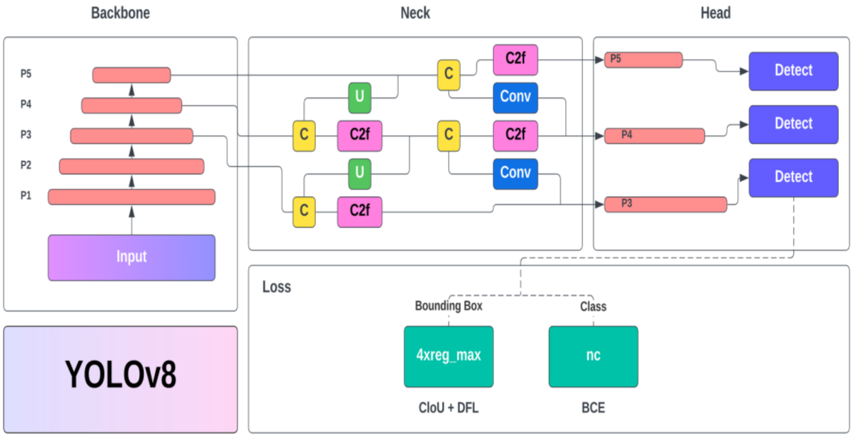


Figure 2. YOLOv8 Model Architecture

The YOLOv8 algorithm also aims to decrease the computational complexity of the model [22-23]. To achieve this, the block size in each stage was reduced from 3, 6, 9, and 3 during the YOLOv5 algorithm to 3, 6, 6, and 3 in YOLOv8 [12]. Furthermore, the SPPF module that is being used in YOLOv5 is adopted during Stage 4 of YOLOv8, as it helps to increase the inference speed of the model compared to the original SPP module. This reduction in computational complexity results in faster and more efficient sign language recognition using the YOLOv8 algorithm.

To overcome the problem of information loss in small objects due to extensive convolution operations, the YOLOv8 architecture's Neck component used the Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) designs. It was used during the Neck section of the architecture where the feature fusion of images was carried out. More network layers allowed the upper features to capture more data, while fewer convolution layers allowed the lower features to retain more location-specific data. When dealing with photos of various sizes, YOLOv5 used the FPN structure to up-sample the bottom feature map and the PAN structure to down-sample the top feature map, producing, respectively, more feature and position information. Following the merging of the two feature outputs, precise results were produced. However, in YOLOv8, FPN and PAN structures are still used, but with a tweak that does away with convolution operations during the up-sampling stage. By making this change, the computational load will be further decreased, and the inference process will be expedited.

In contrast to the coupled head utilized in YOLOv5, the decoupled head in the YOLOv8 design separates the classification and detection heads. Only the classification and regression branches were preserved, with the object branch being eliminated. YOLOv8 uses Anchor-Free to find the object by its center and forecast the distance required from the center till the bounding of the box rather than Anchor-Base, which pre-sets many anchors in the image. With this method, there was no need to set up anchors, and the model was more adaptable to identifying objects of different sizes and forms. YOLOv8's decoupled head offers a method for object detection that is more effective and precise.

However, in the Loss block, CIoU (Complete IoU) and DFL (Dynamic Focal Loss) two novel components incorporated in YOLOv8 were introduced to increase the localization accuracy for the objects in a detection task. CIoU is a variant of IoU that considers the overlapping areas, aspect ratio, and center point distance between the predicted model and ground model of truth bounding boxes. On the other hand, DFL is a variant of Focal Loss that assigns different weights to the easy and hard examples during training. Finally, the BSE (Box Smooth L1 Error) loss function, which was applied to the regression task, penalizes the variance in the bounding of the box coordinates between the predicted model and the actual ground of truth values.

3.1 YOLOV8 Algorithm

1: The algorithm for the proposed system used for Malayalam Sign Language can be presented as follows,
2: Load the YOLOv8 architecture and the pre-trained weights.

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

3: Load the dataset of Malayalam sign language images and corresponding labels.
4: Split the dataset into two halves as training and validation datasets.
5: Set the criteria for the training dataset, including the rate of learning, batch size, and epoch count.
6: Set up the training loop, iterating over each epoch and mini-batch:
a. Forward pass the input through the model to obtain predictions.
b. Calculate the loss using the DFL, IoU, and classification losses.
c. Backpropagate the loss and update the model parameters.
d. Print and record the training metrics.

$$IoU = \frac{area\ of\ overlap}{area\ of\ union}$$

7: Examine the model based on the performance of the validation set, calculating metrics such as precision, recall, and Mean Average Precision (MAP).

$$AP = \frac{1}{n}\sum_{c=1}^{c=n} AP_c$$

where,

APc = average value of precision for class c

n is the number of classes.

8: Once the model is trained and validated, use it to predict the labels of new sign language images.
9: To predict the labels of a new image:

a. Load the image and convert it to a format that can be input to the YOLOv8 model.
b. Feed the image through the YOLOv8 model and obtain the predicted bounding boxes and corresponding class labels.
c. Apply non-maximum suppression to discard redundant bounding boxes.
d. Use the predicted labels to interpret the sign language gesture and output the corresponding text.

4. Dataset and Experiments

The first step in developing a sign language recognition system was to collect a dataset of sign language gestures. A dataset is a collection of samples that are used to train a machine learning model so that a well-curated and diverse dataset can accurately recognize and classify different signs. Here in this paper, a custom dataset for Malayalam Sign Language (MSL) was created by capturing each frame of a video from YouTube provided by the National Institute of Speech and Hearing shown in Figure 3 as there were no publicly available datasets for MSL. A tailored dataset specifically for MSL in Table 2 was created and ensured it included all the relevant signs and gestures and the dataset was well-balanced, diverse, and suitable for training custom models.

Table 2: Dataset Specification for MSL

Specification	**Value**
Resolution	1280 x 720
Extension	.jpeg
Images	138
Classes	51
Images per class	1-3

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

Figure 3. MSL Dataset

These captured images were labeled with their corresponding English phonetic alphabet instead of the original Malayalam character (for example 'aa' for ആ, 'ee' for ഈ) as labeling each image with Malayalam characters caused the class error. This process of labeling was done manually which involved reviewing each image, marking them with a bounding box, and assigning appropriate labels using a popular image annotation tool called Roboflow [11] as shown in Figure 4. The dataset included 51 sign gestures representing different Malayalam Signs. A few class list details are provided in Table 3.

Figure 4. MSL Dataset marked with bounding box using Roboflow

Table 3: Dataset class label and its corresponding Malayalam alphabet

Class Label	**Corresponding Malayalam Alphabet**
a	അ
the	ത
ka	ക
ga	ഗ
ja	ജ
er	ഋ

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

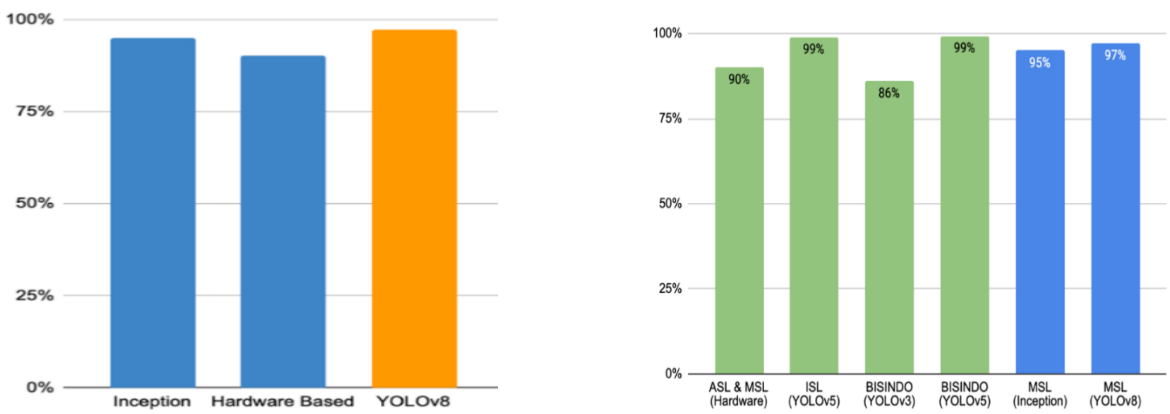
The labelled dataset is split into 3 parts of training, validation, and testing, with a ratio of 80:10:10. The dataset created ensured that each class was balanced with an equal number of sample images. Secondly, the process of pre-processing was applied to the image dataset to decrease the training time and increase the performance using the Roboflow tool. Here, as part of pre-processing, the images were resized to 980x480 from 1280x720, to increase performance. The process of Augmentation was done on the dataset to create new variations of existing images and to increase images in the dataset with the help of the Roboflow tool. This step can significantly make the model more accurate and efficient for sign recognition. Here, a horizontal flip transformation was done, so that the sign-based language system will be able to understand the signs from the left or right hand as the MSL dataset is inadequate. Thus, this allows to improve the reliability of the system.

5. Results and Discussions

The model developed was trained based on an image dataset consisting of images of 51 signs from the MSL vocabulary. The dataset was pre-processed to improve the quality of the images. This was achieved through resizing and augmenting the images to increase the diversity of the training data.

The model was trained and then tested on a separate dataset consisting of images of MSL hand gestures. The results from the model indicate that the proposed model provides promising results in recognizing MSL signs, with an accuracy of 97.21% even using fewer datasets for training purposes. The higher accuracy rate can be compared to the quality of the pre-processed dataset, which included a wide variation of images, and the use of the YOLOv8 architecture, which is highly effective in object detection though the dataset used was less.

The study also compares the deliverance of the proposed model with that of other existing models used for sign language recognition which outperformed the other models with less dataset, including the CNN-based model and hardware as shown in Figure 5.



(a) **(b)**

Figure 5. Comparing the accuracy rate between (a) the Inception module, Hardware, and YOLOv8 proposed in this paper for MSL and (b) with other sign languages

All the detections that were performed produced the right class. However, certain photos have poor confidence, which prevents them from being recognized at a specific threshold.

Figure 4 shows the values of the performance metric used in object detection which measures the quality of the detection by calculating the AP for all the classes from the dataset which is then computed with the mean of these values across all classes. The formula to calculate the Average Precision is,

$$AP = \sum_{n=0}^{l=n-1} [Recall(l) - Recall(l+1)]\ x\ Precision(l)$$

where n is the number of thresholds.

$$Precision = \frac{True\ Positives\ \mathrm{T}(+)}{True\ Positives\ \mathrm{T}(+) + False\ Positives \mathrm{F}(+)}$$

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

$$Recall = \frac{True\ Positives\ \mathrm{T}(+)}{True\ Positives\ \mathrm{T}(+) + FalseNegatives\ F(-)}$$

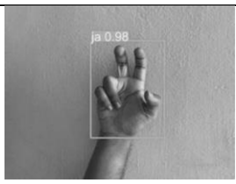
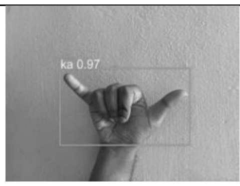
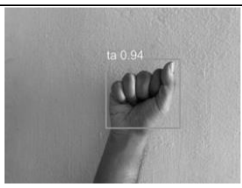
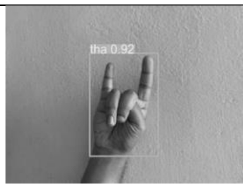
In object detection, T(+) refers to the cases where the predicted bounding of the box overlaps with the ground truth of the bounding box with an IoU (Intersection over Union) greater than a certain threshold, and the class of the predicted box matches the classes with the ground truth box. F(+) refers to the cases where the predicted box will not overlap with any other ground truth box nor overlap with a ground truth box with an IoU below the threshold, and the class of the predicted box does not match the class of the ground truth box. F(-) refers to the cases where there exists a ground truth box but no predicted box overlaps with it with an IoU above the threshold. T(-) refers to the cases where there is no ground truth box and no predicted box.

After calculating the Average Precision (AP) for each class in the dataset, the following mean Average Precision is calculated using the formula below,

$$mAP = \frac{1}{n}\sum_{c=1}^{c=n} AP_c$$

where AP_c = Average Precision value of class c and n is the number of classes.

Table 4: Sign detection using the proposed YOLOv8 for all the classes.

Class Labels	**Output**	**Class Labels**	**Output**
JA		**KA**	
TA		**THE**	

Here, the metrics/mAP50(B) was calculated as the mean Average Precision (mAP) with 50% (0.5) Intersection over Union (IoU) threshold for bounding boxes, while the metrics/mAP50-95(B) was calculated as the mAP at a range of IoU thresholds from 50% to 95% (0.5 to 0.95). Here, the metrics/mAP50 (B) value is 0.90, indicates the model detected the signs in the images with an IoU threshold of 0.5. Additionally, the metrics/mAP50-95(B) value was 0.70, which clearly states that the model was better at accurately identifying the signs at a range of IoU thresholds from 0.5 to 0.95. Table 4. Shows that the model developed is performing well in detecting signs in the Malayalam Sign Language dataset.

6. Conclusion and Future Work

The developed model using YOLO performed well in recognizing Malayalam Sign Language (MSL) with a model that was trained on a dataset that included images of 51 MSL signs. The model's evaluation also revealed that it performs well in detecting signs in the MSL dataset, despite having some constraints, it was able to achieve significant results. However, there is still scope for improvement in the accuracy of the model. One possible future direction for this research could be to collect a larger and more diverse dataset for MSL, which could aid in more effectively training the model hence, improving the accuracy of the recognition that can be used by the public. Furthermore, a look into using transfer learning to train the model on other sign languages and then test its performance on MSL can also be done to improve the model. Smartphone cameras provide input and initial processing, and edge cloud computing uses advanced machine learning models built on Malayalam datasets to detect sign language in real time. This direction represents a future for real-time sign language detection [24-25]. Overall, this research contributes to the growing area of sign-based language recognition models, and their potential applications in bridging communication barriers for individuals who use sign language.

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

References

[1] N, Ajmi., Akhila Raj. B. R, Basina. B, Varsha. S, and Vishnu S Kumar. “Malayalam Sign Language Recognizer – IJERT.” Malayalam Sign Language Recognizer – IJERT, May 17, 2022. https://doi.org/10.17577/IJERTV11IS050083.

[2] Kunjumon, Jinsu, and Rajesh Kannan Megalingam. “Hand Gesture Recognition System For Translating Indian Sign Language Into Text And Speech.” 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), November 2019. https://doi.org/10.1109/icssit46314.2019.8987762.

[3] Katoch, Shagun, Varsha Singh, and Uma Shanker Tiwary. “Indian Sign Language Recognition System Using SURF with SVM and CNN.” Array 14 (July 2022): 100141. https://doi.org/10.1016/j.array.2022.100141.

[4] Ultralytics. “Ultralytics | Revolutionizing the World of Vision AI,” n.d. https://ultralytics.com.

[5] Areeb, Qazi Mohammad, Maryam, Mohammad Nadeem, RoobaeaAlroobaea, and Faisal Anwer. “Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach.” IEEE Access 10 (2022): 8502–17. https://doi.org/10.1109/access.2022.3142918.

[6] Daniels, Steve, NanikSuciati, and ChastineFathichah. “Indonesian Sign Language Recognition Using YOLO Method.” IOP Conference Series: Materials Science and Engineering 1077, no. 1 (February 1, 2021): 012029. https://doi.org/10.1088/1757-899x/1077/1/012029.

[7] Luthfy, Dicky, CasiSetianingsih, and Marisa W. Paryasto. "Indonesian Sign Language Classification Using You Only Look Once." eProceedings of Engineering 10, no. 1 (2023).

[8] Our Work - WFD World Federation of the Deaf, "Our Work - WFD," n.d. https://wfdeaf.org/our-work/.

[9] Stewart Kirubakaran S, V.P. Arunachalam, S. Karthik and S. Kannan (2022) Towards Developing Privacy-Preserved Data Security Approach (PP-DSA) in Cloud Computing Environment Computer Systems Science and Engineering Tech Science Press. DOI: 10.32604/csse.2023.02669044:1882-1895.

[10] National Association of the Deaf, n.d. http://nadindia.org/.

[11] Hearing Like Me. "Understanding India's Deaf Community | Hearing Like Me," August 6, 2021. https://www.hearinglikeme.com/notes-from-indias-deaf-community/.

[12] Roboflow Annotate. “Roboflow Annotate,” n.d. https://roboflow.com/annotate.

[13] Ju, Rui-Yang, and Weiming Cai. "Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm." arXiv preprint arXiv:2304.05071 (2023).

[14] Stewart Kirubakaran et.al A Brief Assessment on Stock Market Forecast using Long Short Term Memory (LSTM) Algorithm published in 2023 7th IEEE International Conference on Trends in Electronics and Informatics (ICOEI). DOI: 10.1109/ICOEI56765.2023.10125607.

[15] Selvakumar, S. K., Ashok, V., Selvadurai, J. J., Ebenezer, V., Nivin, A. J., & Aron, D. A. (2023). Divination of stock market exploration using long short-term memory (LSTM). Multidisciplinary Science Journal, 6(1), 2024001. https://doi.org/10.31893/multiscience.2024001.

[16] Issac, K., Daniel, E., & Varghese, B. G. (2023, May). A Deep Learning Approach for Recognizing Malayalam Sign Language Gestures. In 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 93-97). IEEE.

[17] Seetha, S., ChristlinShanuja, C., Daniel, E., Chandra, S., & Raj, S. (2022, April). Sign Language to Sentence Interpreter Using Convolutional Neural Network in Real Time. In International Conference on Computational Intelligence in Pattern Recognition (pp. 387-400). Singapore: Springer Nature Singapore

[18] Gunvantray, T. D., & Ananthan, T. (2024, March). Sign Language to Text Translation Using Convolutional Neural Network. In 2024 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 1-5). IEEE.

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024

[19] Konoor, A. M., & Padmavathi, S. (2022, June). A Hybrid Approach on Lexical Indian Sign Language Recognition. In International Conference on Soft Computing and Signal Processing (pp. 171-180). Singapore: Springer Nature Singapore.

[20] Gokulraj, V. A., Chirranjeavi, M., Aaruran, S., & Harikumar, M. E. (2023, October). Transhumeral Sensory System to Control Robotic Arm. In 2023 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER) (pp. 185-190). IEEE.

[21] Padmini, P., Paramasivam, C., Lal, G. J., Alharbi, S., & Bhowmick, K. (2022). A Real-Time Oral Cavity Gesture Based Words Synthesizer Using Sensors. Computers, Materials & Continua, 71(3).

[22] Jia, W., & Li, C. (2024). SLR-YOLO: An improved YOLOv8 network for real-time sign language recognition. Journal of Intelligent & Fuzzy Systems, 46(1), 1663-1680.

[23] Zhang, L. J., Fang, J. J., Liu, Y. X., Le, H. F., Rao, Z. Q., & Zhao, J. X. (2023). CR-YOLOv8: Multiscale object detection in traffic sign images. IEEE Access, 12, 219-228.

[24] Durga, S., Surya, S., & Daniel, E. (2018, May). SmartMobiCam: Towards a new paradigm for leveraging smartphone cameras and IaaS cloud for smart city video surveillance. In 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1035-1038). IEEE.

[25] Lim, K. Y., Priyadarshi, A., Nadiah, N. F., Lee, J. H. J., Lau, J. X., Chew, C. K. L., ... & Wong, D. (2024, March). Towards Seamless Communication for Sign Language Support: Architecture, Algorithms, and Optimization. In International Conference on Science and Technology Education (pp. 401-410). Cham: Springer Nature Switzerland.

[26] S. Hemamalini ,V. D. Ambeth Kumar ,R. Venkatesan,S. Malathi. (2023). Relevance Mapping based CNN model with OSR-FCA Technique for Multi-label DR Classification. Journal of Fusion: Practice and Applications, 11 (2), 90-110.

[27] C. S. Manigandaa,V. D. Ambeth Kumar,G. Ragunath,R. Venkatesan,N. Senthil Kumar. (2023). De-Noising and Segmentation of Medical Images using Neutrophilic Sets. Journal of Fusion: Practice and Applications, 11 (2), 111-123.

[28] V. Sathya Preiya,V. D. Ambeth Kumar,R. Vijay,Vijay K.,N. Kirubakaran. "Blockchain-Based E-Voting System with Face Recognition." Fusion: Practice and Applications, Vol. 12, No. 1, 2023 ,PP. 53-63

DOI: https://doi.org/10.54216/FPA.170110
Received: November 24, 2023 Revised: March 15, 2024 Accepted: July 11, 2024