Project Stage II Report on

Empowering Deaf with Indian Sign Language Interpreter

Submitted in partial fulfillment of the requirements for the Degree of

Bachelor of Engineering

in

Artificial Intelligence and Data Science

by

Mayuri Gade (B190352015) Pranali Hagare (B190352021)

Shivanjali Nimbalkar (B190352047) Soham Vaidya (B190352066)

Under the guidance of

Mr. Pradip Shendage

Vidya Pratishthan's

Kamalnayan Bajaj Institute of Engineering and Technology

Baramati-413133, Dist-Pune (M.S.) India

Department of Artificial Intelligence and Data Science

Academic Year 2023-24

VPKBIET, Baramati

Certificate

This is to certify that the Project Stage II Report on

Empowering Deaf with Indian Sign Language Interpreter

SUBMITTED BY

Mayuri Gade (B190352015) **Pranali Hagare (B190352021)**

Shivanjali Nimbalkar (B190352047) **Soham Vaidya (B190352066)**

in the partial fulfillment of the requirement for the award of Degree of Bachelor of Engineering in Artificial Intelligence and Data Science at Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati, under the Savitribai Phule Pune University, Pune. This work is done during year 2023-24 Semester-II, under our guidance.

Mr. Prandip Shendage	Dr. P. M. Paithane	Dr. R. S. Bichkar
Project Guide	Head of Dept.	Principal

Examiner 1:- - - - - - -

Examiner 2: - - - - - - -

Acknowledgements

It gives immense pleasure in submitting this project report on **"Empowering Deaf with Indian Sign Language Interpreter"** The success of this project showcases both our hard work and the valuable guidance provided by our guide.

We would like to express our heartiest thanks to our project guide **Mr.Pradip Shendage** for his undying support which made us possible to make this project knowledgeable. He not only provided us the literature and guidance to study but also the platform which required for us to prepare best for this project. We also equally indebted to our project co-ordinator **Mrs.Rohini Naik** who has been a constant source of inspiration and guiding star in achieving our goal. We extend our sincere gratitude to **Bharat Forge Lab** for generously providing access to their high-performance server infrastructure, equipped with GPU capabilities, and facilitating seamless internet connectivity, which significantly contributed to the success of our project. We thank **Dr.P.M.Paithane**, Head of Department of AI&DS Engineering and respected Principal **Dr.R.S.Bichkar** for supporting and providing all facilities to complete the work. We are thankful to and fortunate enough to get constant encouragement, support and guidance from entire teaching staff of Department of AI&DS engineering which helped us in successfully completing our project work.

Mayuri Mahesh Gade (B190352015)
Pranali Sandeep Hagare (B190352021)
Shivanjali Vijay Nimbalkar(B190352047)
Soham Nilesh Vaidya(B190352066)

Abstract

There are over 140 different sign languages used globally. Even the same word or concept can have different signs in different countries/regions. For example, the sign for "I" may involve a different hand shape or movement in American Sign Language and Indian Sign Language. So a recognition model trained on signs from one region may not generalize well to other regions. Sign languages evolve over decades just like spoken languages. New signs emerge and old signs can change or become obsolete. For instance, with new technology words coming into use, their signs also get introduced into the vocabulary. The deaf and hard-of-hearing community in India is greatly impacted by studies in the critical subject of Indian Sign Language (ISL) identification. The recognition of ISL poses several unique challenges compared to spoken language recognition. This research endeavors to develop a real-time Indian Sign Language (ISL) recognition system to bridge communication gaps for deaf children between 1^{st} and 5^{th} grade. We have prepared a curated dataset comprising 50 emergency words. The objective is to convert ISL gestures into textual format, enabling effective communication for deaf children. We employ state-of-the-art techniques such as deep learning, computer vision, and neural networks for robust sign language recognition, ensuring accessibility and inclusivity for the target demographic. For building dataset, the system uses the mediapipe library of computer vision. A system is proposed for dynamic hand gesture recognition using multiple deep learning architectures for hand segmentation. Mainly LSTM is used for training the dataset. And secondly, LSTM models are used for Classification of signs.

Keywords : Indian Sign Language ,Mediapipe,Computer Vision,Long Short-Term Memory

List of Figures

List of Tables

Notations and Abbreviations

1. CV : Computer Vision
2. Mp : MediaPipe
3. ML: Machine Learning
4. AI: Artificial Intelligence
5. LSTM: Long-Short Term Memory
6. ISL: Indian Sign Language
7. DL: Deep Learning

Contents

Chapter 1

Introduction

1.1 Introduction

Around 466 million people worldwide experience hearing loss, with 34 million being chil dren. People who are deaf often rely on sign language for communication[6]. Different parts of the world use various sign languages, which are fewer in number compared to spoken languages. India, for instance, has its own sign language called Indian Sign Language (ISL).In many developing countries, there are limited schools for deaf students, contributing to a high unemployment rate among adults with hearing loss. Ethnologue data indicates that in India, where 1% of the population is deaf, the literacy rate and the number of deaf children attending school are notably low[3]. This underscores the need for increased awareness, support, and educational opportunities for the deaf community.

Deaf people often find it hard to communicate effectively with those who can hear[5]. The current technology isn't very good at understanding and translating sign language accurately on its own. This makes it tough for deaf individuals to express themselves and be understood by others.The problem lies in the complexity of sign language the gestures, expressions, and movements can be tricky for technology to grasp fully[6]. Existing tools might not catch all the details, leading to misunderstandings and communication problems.To make things better, we need smarter technology[4]. This could mean improving computer programs using things like artificial intelligence and computer vision.These enhancements would specifically focus on understanding the unique aspects of sign language.In short, we need to work on making technology better at understanding sign language. This improvement can make communication easier for deaf individuals and promote better understanding between people with different hearing abilities.

1.2 Motivation

Deaf people face communication barriers when trying to convey their thoughts to hearing individuals. Existing technology falls short in accurately and autonomously interpreting the sign language phrases used by these people, hindering effective communication. The project aims to bridge the communication gap between deaf and hearing individuals by developing and implementing advanced technologies for sign language recognition and interpretation. By creating innovative solutions, we strive to facilitate seamless communication, fostering understanding and inclusivity.The overarching goal is to empower the deaf community by providing them with the tools and resources necessary for effective communication and integration into various aspects of society. Through the implementation of technology, education, and community engagement, we seek to enhance the overall well-being and opportunities for individuals with hearing impairments.

Chapter 2

Literature Survey

Table 2.1: Literature Survey on Comparison of Techniques for Sign Language classification

Paper	**Technique Used**	**Advantages**	**Gaps**
[1] Mohammed Nadeem et.al.2022 .	Classification VGG-16 and LSTM model and detection model YOLO v5	Accuracy-98 %,best classification	Less dataset, only 8 output signs, under fitting
[2] Nikolas Adaloglou et.al. 2022	Reinforcement Learning, Bidirectional LSTM, Temporal Pooling	Novel approach focusing on video representation and semantic boundary.	Cannot handle complex sign language sequences optimally
[3] Jian Zhao et.al. 2022	Deep Learning, CNN	The method achieves promising performance on two public SLT datasets.	Order of words in a the ground truth sentence often differs from sign language video.

[4] Pengpai Wang et.al. 2021	L1 regularization and classification	Potential in human computer interaction.	Limited decoding instruction set.
[5] Wengang Zhou et.al. 2020	Reinforcement Learning , Deep Learning	By combining weakly supervised learning with reinforcement learning it enhances performance in SLR.	Obtaining large-scale dataset is challenging.
[6]JESTIN JOY et.al. 2019	Nasnet and Inception V3, Tensor flow, Sign-Quiz interface	If model training data will change then this model will support other languages, first work on ISL	Only alphabets, find out behavior of Sign-Quiz, less accuracy.

In proposed system the classification VGG-16 and LSTM model and detection model YOLO v5 techniques are used. The advantage of these models use is its accuracy is 98 % and it best classifies the results. But the limitation of the technique is less dataset as it contains only 8 output signs.It may cause underfitting of the model[1] In the given system Reinforcement Learning, Bidirecctional LSTM and Temporal Pooling technique is used. The main benefit of these techniques is it focuses on video capturing. But in case of complex sign language these techniques cannot handle the complexity. Authors proposed the techniques used are deep learning and CNN[2]. These methods achieves promising performance on two public SLT datasets. The limitation of the techniques is order of words in the ground truth sentence often differes from sign language video. The project explores the application of L1 regularization and classification techniques in the context of human-computer interaction[3]. The limitation mentioned in the research paper is presence of limited decoding instruction set. This could imply a constraint on the variety or complexity of instructions or gestures that the system can accurately decode or classify. Attention-Based 3D-CNNs for Large-Vocabulary Sign Language Recognition" The research paper focuses on Sign Language Recognition(ISL) and implies techniques of weakly supervised learning and reinforcement learning[4]. These techniques are designed to enhance the performance of Sign Language Recognition systems. The challenging task in the research is obtaining a large-scale dataset for Sign Language Recognition. The techniques involved use application of Nasnet and Inception V3, utilizing the TensorFlow framework, in the context of a SignQuiz interface[5]. The primary objective of the study appears to be the development of a model that, when trained with different data could potemtially support sign languages other than initial focus on Indian Sign Language(ISL)[6].

2.1 Gaps

Several gaps have been identified in existing Sign Language System only recognizes a limited set of 8 signs. The dataset used for training the models is relatively small. There may be issues with model generalization as it could be underfitting the data. Inability to handle complex sign language sequences, existing system struggles with complex sign language sequences, indicating limitations in its capacity to accurately interpret intricate gestures. There is a discrepancy between the order of words in the ground truth sentence and the sign language video, suggesting challenges in aligning textual descriptions with sign language gestures accurately. The existing system has a restricted set of decoding instructions, potentially limiting its ability to interpret a wide range of sign language gestures effectively. Acquiring large-scale datasets for training the models is challenging, which can hinder the performance and generalization of the system.

Chapter 3

Proposed System

3.1 Problem Definition

To develop a Indian Sign Language recognition system using deep learning and OpenCV to identify sign language gestures in real-time, offering diverse suggestions based on detected gestures to enhance communication experiences for deaf individuals.

3.2 Objectives

1. To Collect and curate a diverse dataset of images featuring various hand gestures used in sign language.
2. To Develop a real-time sign language translation system for seamless communication between deaf and hearing individuals.
3. To Enhance the user interface for inclusivity, ensuring a user-friendly design that accommodates both deaf and hearing individuals, fostering ease of communication.
4. To Enable context understanding in sign language technology and improve interpretation accuracy and prevent miscommunication.

3.3 Scope of Project

1. **Real-Time Speech-to-Text Transcription:**
   The proposed system employs advanced speech-to-text technology to provide real-time transcription of spoken language during lectures and discussions. This ensures

that deaf students have immediate access to verbal information, enabling them to follow along with course content just as their hearing peers do.

2. **User-Centric Personalization:**
   The system is designed with user-centric personalization features, adapting to the individual needs and preferences of each student. This includes customizable settings for text display, such as font size and color, to enhance readability and comfort.

3. **Seamless Integration with Educational Platforms:**
   The system is designed to seamlessly integrate with existing educational platforms, such as Learning Management Systems (LMS) and virtual classroom software. This ensures a smooth and uninterrupted educational experience for both students and educators.

3.4 Project Constraints

1. **Cultural and Linguistic Diversity:**
   India is culturally and linguistically diverse, with different regions having their own variations of sign languages. Adapting the system to accommodate these variations while maintaining a standardized approach can be challenging.

2. **Limited Standardization of ISL:**
   Unlike some spoken languages, sign languages may not have standardized forms across all regions. The lack of a universally accepted standard for ISL can pose challenges in creating a system that caters to the diversity of sign language users.

3. **Accessibility in Rural Areas:**
   Access to technology and internet connectivity in rural areas may be limited. Ensuring that the system is accessible to users in both urban and rural settings is crucial for inclusively.

Chapter 4

Proposed System Architecture

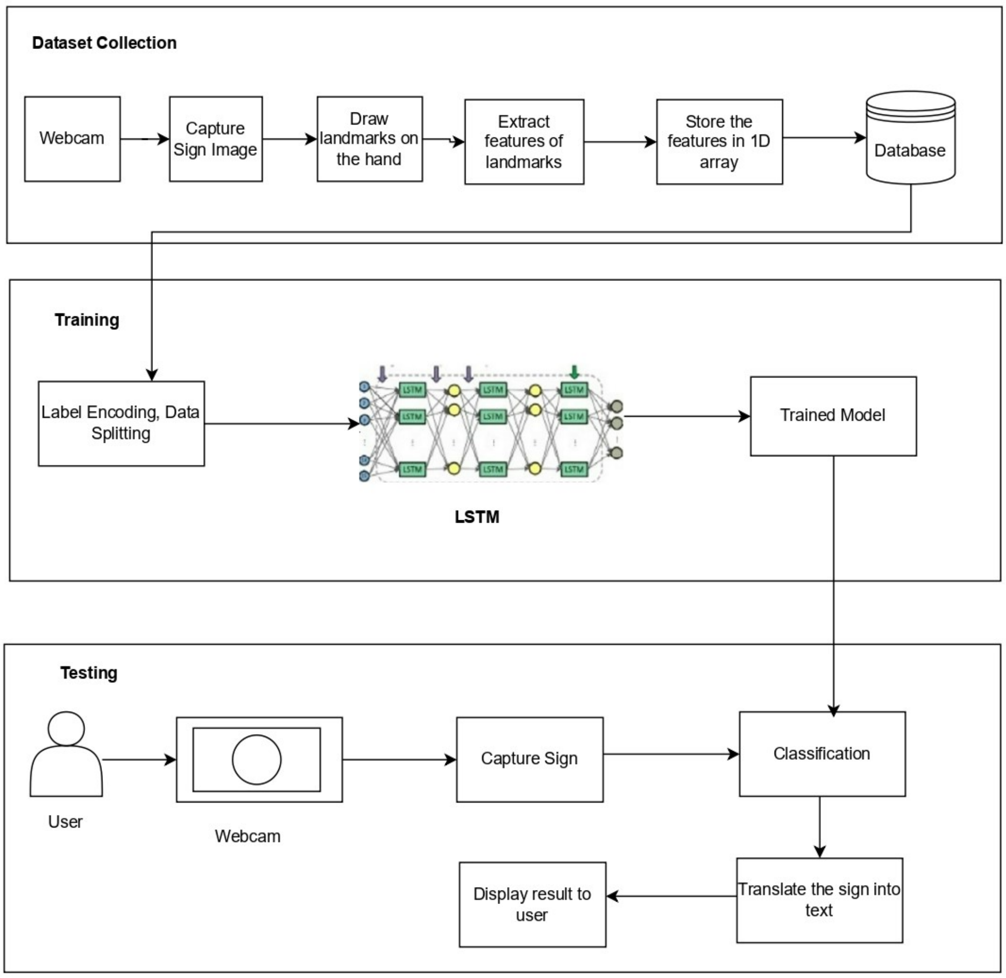


Figure 4.1: System Architecture

4.1 Architecture

4.1.1 The system is mainly divided into 3 Modules:

1. **Dataset collection:**

   The dataset contains 8 words including, "Alone", "Call", "Flower", "Food", "I am Good" "Stop", "There is Gun". Each word contains 50 sequences and each sequence contains 30 frames. Each frame stores features in the form of a 1D array, extracted from the landmarks present on the hand.The landmarks are capture using Mediapipe library. Right hand has 21 landmarks and the left hand also has 21 landmarks. Each landmark has three values (X, Y, visibility), where visibility indicates how visible the landmark is on the hand. The value of each landmark is stored in a 1D array using a .npy file. For each sign, we have 30 frames * 50 sequences, total 1500 features stored for one sign.Dataset contains 8 classes of sign with 500 sequences and 12000 features.

2. **Model Training :**

   Train the model using LSTM neural network. LSTM is a type of recurrent neural network designed to capture long-term dependencies in sequences .Specify the number of LSTM layers, the number of units in each layer, and other hyperparameters such as dropout rates to prevent overfitting. Split the dataset into training and validation sets.Training set is used to train the model and validation set is used to monitor the model's performance. Train the LSTM model using an appropriate optimizer (eg.Adam) and a suitable loss function(eg.categorical crossentropy). Monitor the training process using metrics like training loss and validation loss.

3. **Real time Testing :**

   In real Time testing user open the camera and system capture the sign and feed to the train model. Model classify the sign by comparing its features with the features in dataset.After comparing the features with nearly or exactly same features it gives the output in text format.

4.2 Mathematical Model

Mathematical Model use in LSTM model to build the system is as follow:

Here LSTM layer(i = 1,2,3)

X=input

1.**Input Gate(i_i)**

$$i_t^{(i)} = \sigma(W_i^{(i)} \cdot [h_{(t-1)}^{(i)}, x_t] + b_i^{(i)}) \tag{4.1}$$

- $W_{(i)}$ is weight and $b_{(1)}$is biases for input gate.
- σ is Sigmoid activation.
- h^{t-1} preceding hidden state

2. **Forget Gate(f_t)**

$$f_t^{(i)} = \sigma(W_f^{(i)} \cdot [h_{(t-1)}^{(i)}, x_t] + b_f^{(i)}) \tag{4.2}$$

3. **Cell State(C_t)Update:**

$$C_t^{(i)} = f_t^{(i)} \cdot C_{(t-1)}^{(i)} + i_i^{(i)} \cdot \overline{C} \tag{4.3}$$

$$\overline{C} = ReLU(W_c^{(i)} \cdot [h_{(t-1)}^{(i)}, x_t] + b_c^{(i)}) \tag{4.4}$$

- $\overline{C}$ - Candidate cell state

4. **Output Gate(O_t)**

$$O_t^{(i)} = \sigma(W_O^{(i)} \cdot [h_{(t-1)}^{(i)}, x_t] + b_O^{(i)}) \tag{4.5}$$

5. **Hidden State(h_t):**

$$h_t^{(i)} = O_t^{(i)} \cdot ReLU(C_t) \tag{4.6}$$

Here Dense layer(j = 1,2)

6. **Dense Layer 1 and 2:**

input:($O^{(3)}$) =[$h_1^{(3)}$,$h_1^{(3)}$,......$h_{numtimesteps}^{(3)}$]

Forward Pass: for t = 1,2,....,$numtimesteps$:

$$Z_t^{(j)} = W^{(j)} \cdot O_t^{(j)} + b^{(1)} \tag{4.7}$$

$$A_t^{(j)} = ReLU(Z_t^{(j)}) \tag{4.8}$$

7. **Output layer:**

   input:$(A^{(2)})$ =$[A_1^{(2)}$,$A_1^{(2)}$,......$A_{num_time_steps}^{(3)}]$

   Forward Pass: for t = 1,2,....,num-time-steps:

$$Z_t^{(3)} = W^{(3)} \cdot A_t^{(2)} + b^{(2)} \tag{4.9}$$

$$Y_t^{(3)} = Softmax(Z_t^{(3)}) \tag{4.10}$$

   Output gives the results of sign with text.

4.3 Proposed Algorithm

Algorithm 1: Proposed Algorithm

Input : Array of features extracted from the landmarks present on the hand

Output: Sign Detection

```
Training Phase;
LSTM;
for epoch = 0 to epochs = 150 do
    img_array ← Sequential mode();
    for i = 0 to 2 do
        img_array ←LSTM array with 64 unit and return_sequences=True;
    end
    img_array ←LSTM array with 128 unit and return_sequences=True;
    img_array ←LSTM array with 64 unit and return_sequences=False;
    img_array ←Dense layer with softmax activation;
    Output()←img_array;
end
Testing Phase;
Image processing(image);
output image ← Sign Language Detection(image);
Hand-Landmarks[ ] ← Feature Extraction(Output_image);
for i = 0 to size(Hand-Landmarks) do
    output text += Classify Sign with Trained Model(Hand-Landmarks[i]);
end
System Implementation Phase;
trained model ← LSTM-Based sequential Model(dataset);
Sign Language Detection ← image processing(image);
display(text);
```

Chapter 5

Requirement Specification

5.1 Hardware Requirements

1.Processors: Intel Core i3 processor (Minimum)

2.Laptop camera/webcam - 720 pixels

5.2 Software Requirements

1)Jupyter Notebook - Development Environment

2)Python(3.11.5) – Programming Language

3)OpenCV Libraries-Image Capturing

4)Mediapipe - Detecting hand landmarks

5)Draw.io, Gantt chart – SRS

6)Latex – Documentation

7)GitHub – Version Control

5.3 Performance Requirements

The benefits of using Long-Short Term Memory for classification of sign include exceeding model's accuracy, scalability for large datasets, generalizability across a variety of environments, optimised resource utilisation, minimal training time, robustness to variations, interpretability, energy efficiency.

5.4 Software Quality Attributes/Requirements

For smooth interaction, the system should have high level of real-time responsiveness, which guarantees quick and precise identification of signs indicated by the users.

5.5 Security Requirements

The security requirements for your topic, "Empowering Deaf with Indian Sign Language Interpreter" include ensuring secure handling and storage of data containing arrays representing the signs. Privacy measures should be in place to anonymize or aggregate personal data and access control measures should restrict access to authorized or aggregate personal data.Strong user authentication mechanisms, secure deployment practices, and regular security audits are also essential to ensure the system's security.

5.6 Other Requirements

The system should also meet performance requirements, such as classification of signs efficiently to ensure real-time. It should be able to handle a large volume of data and be scalable to accommodate future growth. The system should also be user-friendly, with an intuitive interface for easy interaction.Reliability is crucial, requiring the system to be stable and available, minimizing downtime or disruptions. Compatibility with different devices and platforms should be ensured for broader usability.

Chapter 6

Project Planning

This plan serves as the foundation for carrying out and monitoring every project activity. It will be used and updated to reflect the project's actual accomplishments and plans over its entire duration.

6.1 Project Estimates

Project estimation is most important thing in software Engineering. Project estimation is used to calculate the duration and cost of the project in prior basis, the process of finding an estimate, or approximation, which is a value that is unable for some purpose even if input data may be incomplete, uncertain, or uncertain, or unstable. corresponding population parameter. The sample provides information that can be pro- Jetted,through various formal or informal processes, to determine a range most likely to describe the missing information. An estimate that turns out to be incorrect will be a overestimate if the estimate exceeded the actual result and an underestimate if the estimate fell short of the actual result. Project Estimation report is size oriented matrix. Here, constructive cost model Le COCOMO-2(1990) is used to estimate effort And time duration by using size of software. Suze in terms of the KLOC(kilo Line Of Code). It include only executable lines not variable, function declaration.

6.1.1 Effort, Duration, and Personnel

The effort, duration, and personnel estimates for the project can be calculated using the following formulas:

$$\text{Effort (E)} = A_b \times (\text{KLOC})^{B_b}$$
$$\text{Duration (D)} = C_b \times (E)^{D_b}$$
$$\text{Personnel (P)} = \frac{E}{D}$$

Where:

$$\text{KLOC} = 1.1 \text{ (estimated)}$$

The project estimate is primarily determined by three factors:

Effort Applied

Effort applied is calculated as the product of coefficients a and b, where a is the effort constraint and b is the estimated KLOC (thousands of lines of code). Typically, estimation involves the value of a statistic.

$$\text{Effort Applied} = A_b \times (\text{KLOC})^{B_b}$$
$$\text{Effort applied} = 2.4 \times (1.1)^{1.05}$$
$$\text{Effort Applied} \approx 2.6 \text{ PM}$$

Development Time

Development time is the time required for development and is estimated to be 4 months.

$$\text{Development Time} = C_b \times (E)^{D_b}$$
$$\text{Development Time} = 2.5 \times (2.6)^{0.38}$$
$$\text{Development Time} = 3.59 \text{ months}$$

Number of Team Members

The number of team members is a positive integer and is estimated to be 4.

$$\text{Members} = 4$$

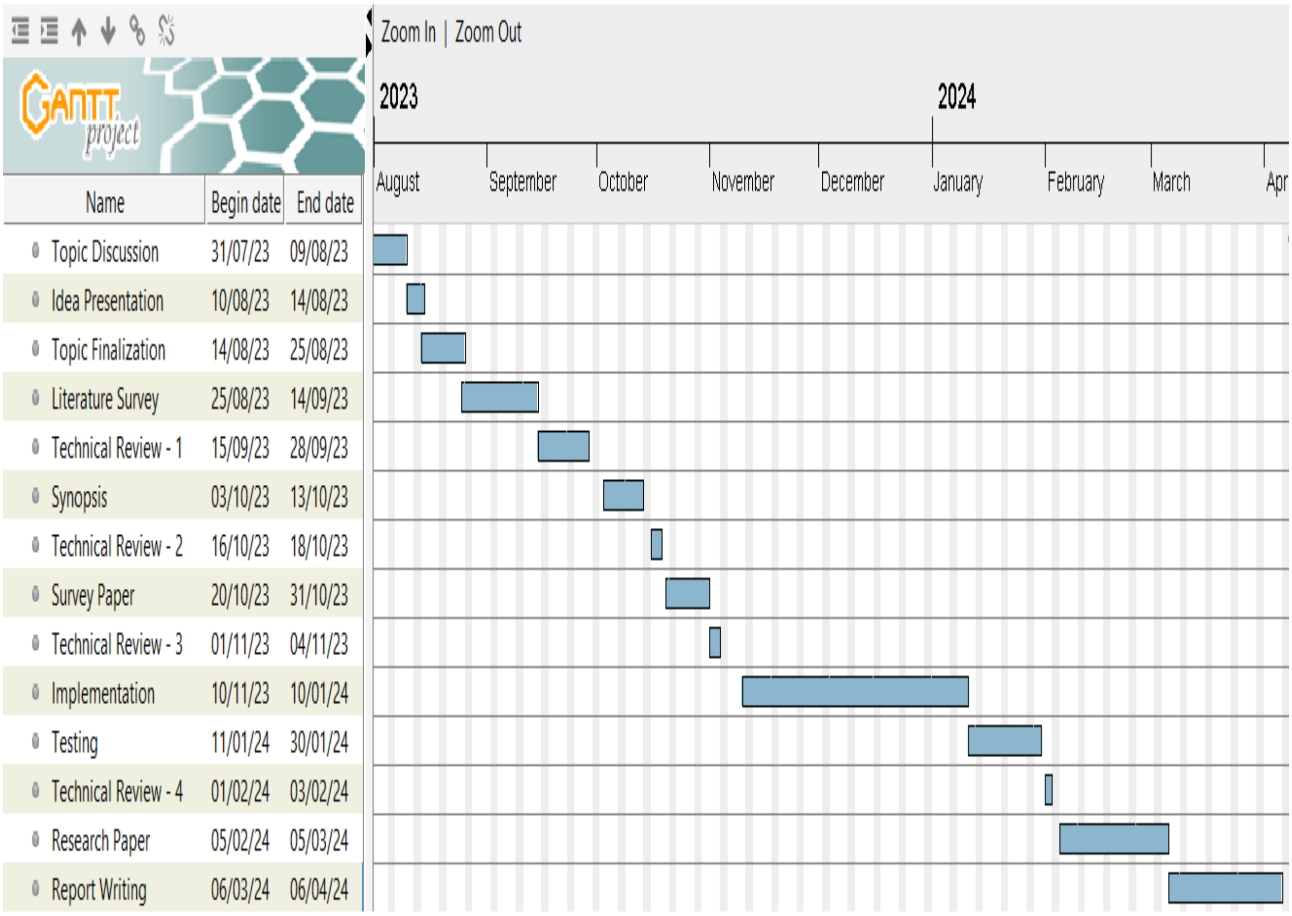
Productivity (P)

$$\begin{aligned}\text{Productivity} &= KLOC/Effort \\ &= 1.1/2.6\text{u}nit \\ &= 0.423 \\ &= 423LOC/Person - Month\end{aligned}$$

6.2 Team Structure

Name	Role
Mayuri Gade	Dataset Collection, Copyright, Implementation
Pranali Hagare	Dataset Collection, Implementation, Research Paper
Shivanjali Nimbalkar	Research Paper, Documentation, Model Deployment
Soham Vaidya	Algorithm Development , Copyright, Research Paper

6.3 Project Breakdown Structure



Name	Begin date	End date
Topic Discussion	31/07/23	09/08/23
Idea Presentation	10/08/23	14/08/23
Topic Finalization	14/08/23	25/08/23
Literature Survey	25/08/23	14/09/23
Technical Review - 1	15/09/23	28/09/23
Synopsis	03/10/23	13/10/23
Technical Review - 2	16/10/23	18/10/23
Survey Paper	20/10/23	31/10/23
Technical Review - 3	01/11/23	04/11/23
Implementation	10/11/23	10/01/24
Testing	11/01/24	30/01/24
Technical Review - 4	01/02/24	03/02/24
Research Paper	05/02/24	05/03/24
Report Writing	06/03/24	06/04/24

Figure 6.1: Time-Line chart

Chapter 7

Project Design

7.1 UML Diagrams

7.1.1 DFD 0:

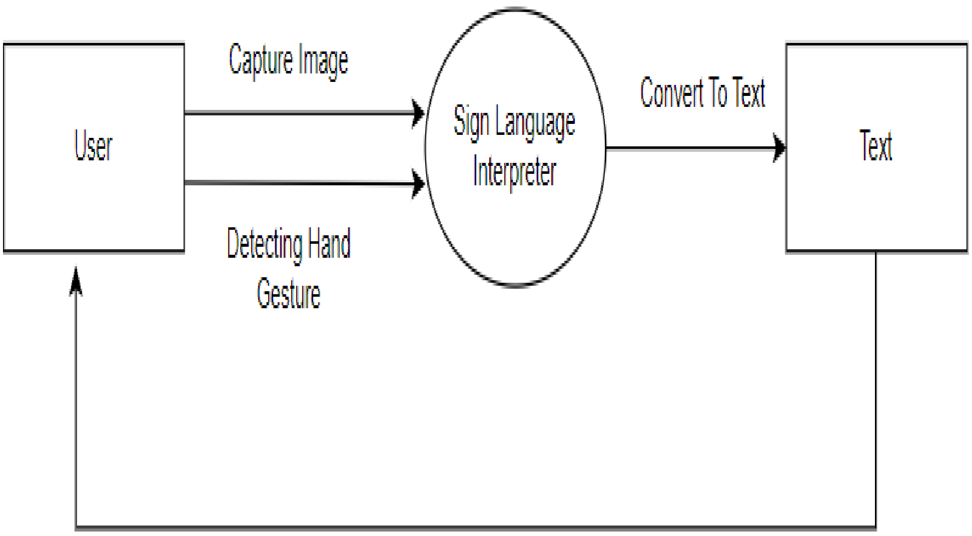


Figure 7.1: DFD 0

7.1.2 DFD 1:

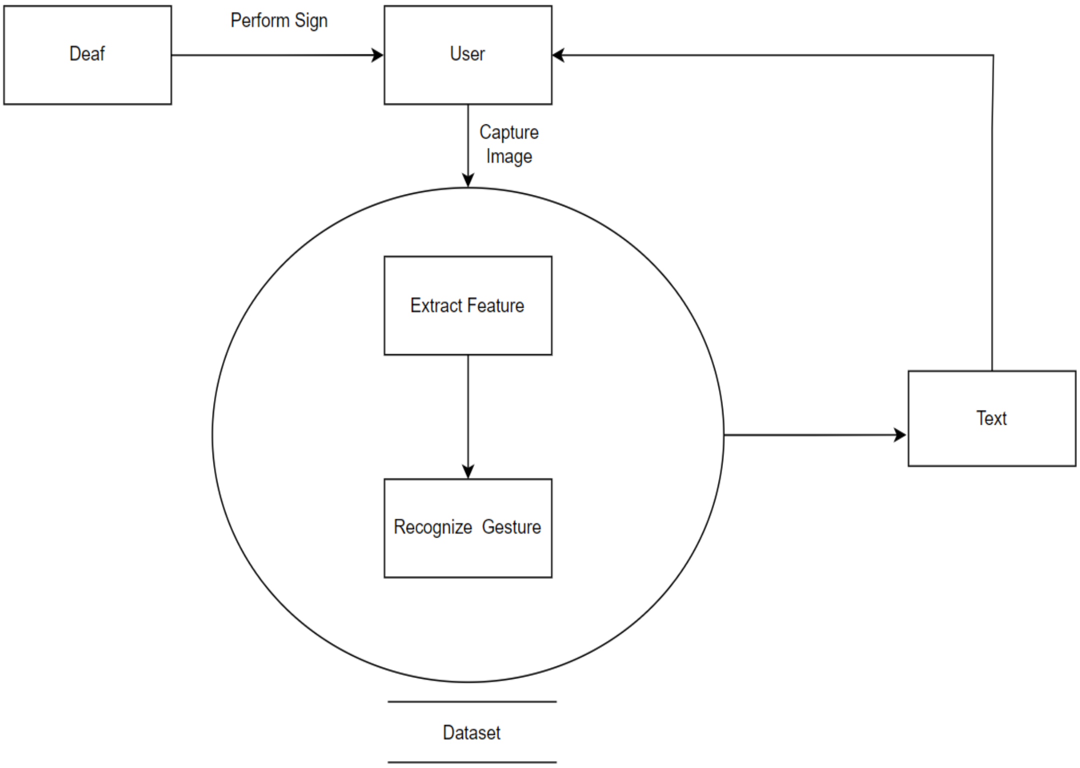


Figure 7.2: DFD 1

7.1.3 Use Case:

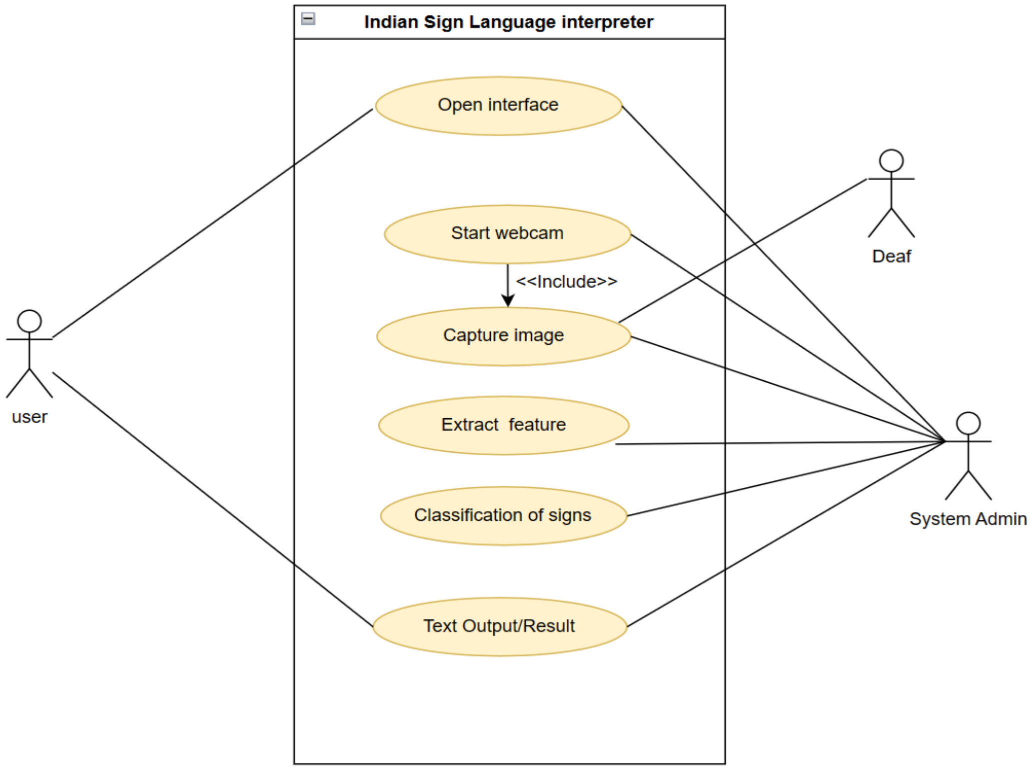


Figure 7.3: Use Cases

7.1.4 Class:

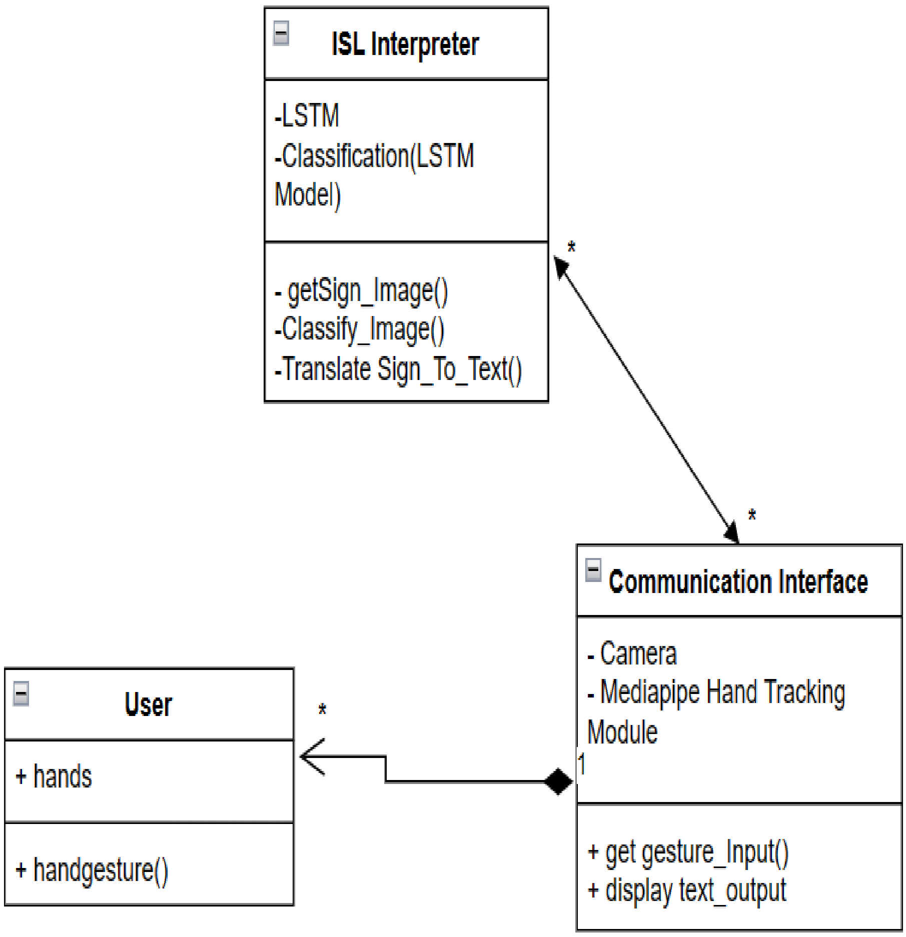


Figure 7.4: Class Diagram

Chapter 8

Results and Experimentation

8.1 Experimental Setup

The SignVaria dataset is curated dataset is a key asset in the study of Indian Sign Language classification. It includes 50 sequences of each sign and 30 frames of each sequence are stored in the array format. It includes the signs of the words Alone, Call, Flower, Food, I am Good, Stop, There is Gun, Hello, Sorry. This dataset is fundamental resource for training and testing classification models, enaabling advancements in the field.

8.2 Test Specifications

1. Model Training Performance: Analyze the model's performance for classification of sign using the dataset SignVaria. Analyze metrics on training and validation sets, including loss, F1-score, recall, accuracy.
2. Dataset Compatibility Testing: Examine to verify if the system's classification model can accurately categorize signs according to the signs included in the SignVaria dataset.

8.2.1 Assumptions and Dependencies

8.2.2 Assumption

1. **Quality of Input Data:** The project assumes that the input data, consisting of arrays representing signs captured by the mediapipe hand tracking module, is of

high quality and accurately represents the lable for each sign.

2. **Model Training:** It is assumed that the LSTM model will be trained using a sufficient quantity of labeled data and appropriate hyperparameters to achieve optimal performance.

3. **Software Compatibility:** It is assumed that the software stack, including Python, and deep learning frameworks will be compatible with the chosen hardware environment and operating system.

Dependencies

1. **Data Availability:** The project is dependent on the availability of a sufficient quantity of labeled image data for training and evaluation purposes.

2. **Hardware Resources:** The project relies on access to hardware resources with adequate processing power, memory, and storage capacity to support model training and evaluation.

3. **Software Tools:** Dependencies exist on specific software tools and libraries, such as Jupyter Notebook and deep learning frameworks, which are essential for model development and experimentation.

4. **Training Time:** The project's timeline is dependent on the computational resources available for training the LSTM model, as well as the complexity of the training process, which may impact the duration of experiments and iterations.

8.3 Performance Measures

8.3.1 Metrics Formula

- Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8.1}$$

- Precision

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8.2}$$

- Recall

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8.3}$$

- F1-score

$$F1\text{-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8.4}$$

- Support

$$\text{Support} = \text{Number of true instances for each class} \tag{8.5}$$

8.3.2 Test Loss and Accuracy

The metrics for test loss and accuracy offer valuable information about the LSTM model overall performance on the test set. The test accuracy shows the percentage of correctly classified samples, whereas the test loss shows the average loss computed during the evaluation process.Test Loss and Accuracy of LSTM Model

- Test Loss: The test loss is calculated as 0.3014.
- Test Accuracy: The test accuracy is 0.9625.

8.3.3 Classification Report

For every class in the classification problem, the precision, recall, F1-score, and support are presented in detail in the table shown below. It offers insightful information about the LSTM model effectiveness for certain classes.

Table 8.1: Classification Report

Signs	**Precision**	**Recall**	**F1-score**	**Support**
Alone	1.00	0.92	0.96	13
Call	1.00	1.00	1.00	13
Flower	1.00	1.00	1.00	11
Food	1.00	0.86	0.92	7
I am Good	1.00	0.88	0.93	8

Ok Fine	1.00	1.00	1.00	10
Stop	0.62	1.00	0.77	5
There is Gun	1.00	1.00	1.00	13

8.4 Experimental Results

8.4.1 Result Analysis

The primary approach were examined in the study of Indian sign language classification using SignVaria datset. The Long-Short Term Memory is used for developing the system. The LSTM model was first trained with the dataset SignVaria with 83 % accuracy. After changing some parameters in the model the accuracy increased to 90 %. Again the dataset is recorded using webcam with better labelling the accuracy increased to 96 %.

Output Images:

Figure 8.1: Output Images 1

Output Images:

Figure 8.2: Output Images 1

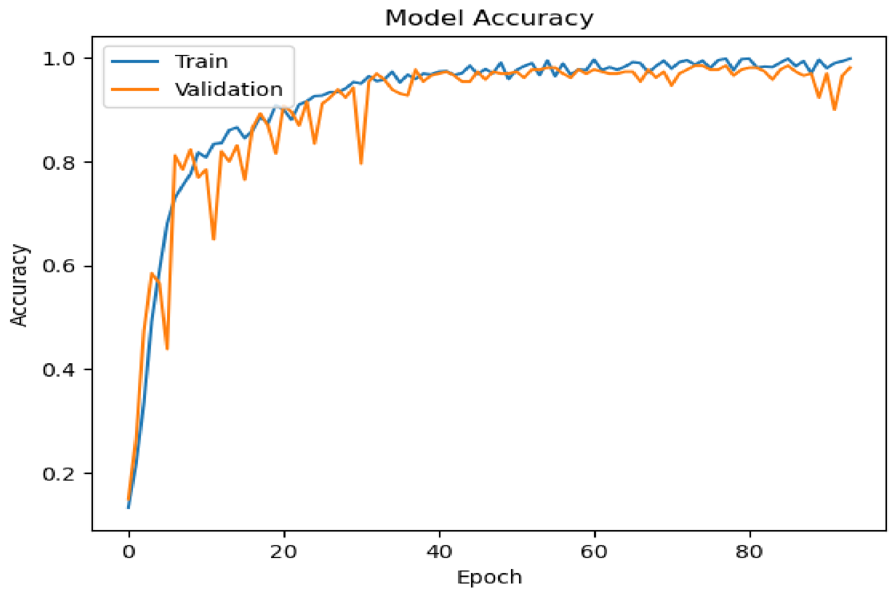


Figure 8.3: Accuracy

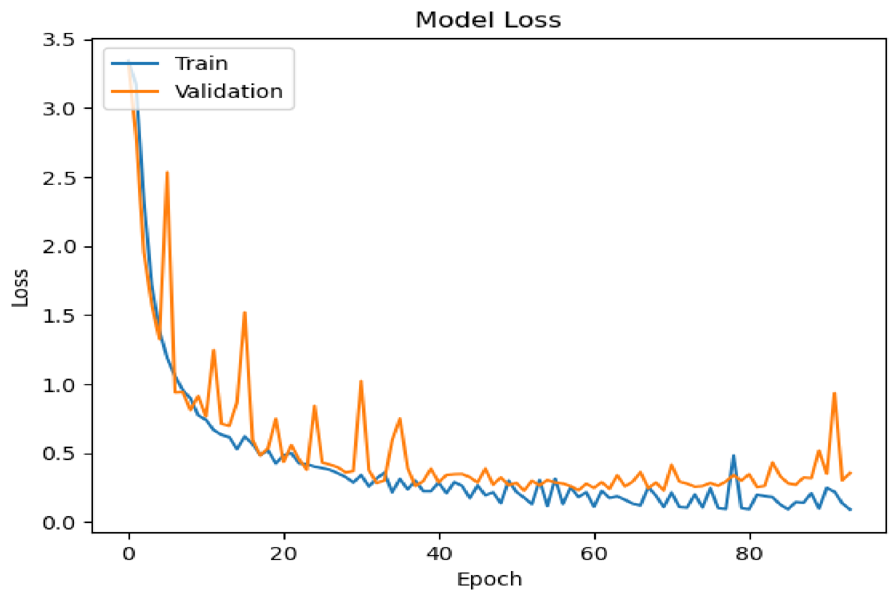


Figure 8.4: Loss

8.4.2 Results Comparison

Table 8.2: Results Comparison

Sr.No.	Authors	Methods	Results
1	**Proposed System**	**LSTM and Mediapipe Handtracking model**	**Accuracy - 96.25 % , Dice Score - 94.82**
2	Mohammed Nadeem et.al.2022	Classification VGG-16 and LSTM model and detection model YOLO v5	Classification VGG-16 improved accuracy - 98 %
3	Nikolas Adaloglou et.al. 2022	Reinforcement Learning, Bidirectional LSTM, Temporal Pooling	Bidirectional LSTM and Reinforcement Learning enhance accuracy to 95.68 %
4	Jian Zhao et.al. 2022	Deep Learning, CNN	CNN model increases accuracy to 80 %
5	Pengpai Wang et.al. 2021	L1 regularization and classification	L1 regularization enhance accuracy to 89.60 %
6	Wengang Zhou et.al. 2020	Reinforcement Learning , Deep Learning	The techniques Reinforcement Learning and Deep Learning improve accuracy to 90.6 %
7	JESTIN JOY et.al. 2019	Nasnet and Inception V3, Tensor flow, Sign-Quiz interface	Nasnet and Inception V3 , Tensor flow improves accuracy to 97 %

8.5 Discussions

- **Sensitivity to Environmental Factors:**
  Environmental elements like background clutter, lighting, and camera quality may have an impact on the sign detection system's performance and affect how accurately signs are recognized.
- **Generalization to Diverse User Groups:**
  The efficiency of the system may differ depending on the age, gender, and cultural background of the user base. Since people express their signs in different ways, it can be difficult to generalize the recommendations to different user groups.
- **Real-Time Processing Requirements:**
  High computational requirements for real-time sign detection and classification could be problematic for devices with constrained processing power or internet. bandwidth

Chapter 9

Conclusion

The dataset consists of 8 classes for different signs. The signs are recorded using mediapipe handtracking module. The deep learning technique such as LSTM is used to implement the project. Project overcomes the objective which was reducing communication barrier between normal people and deaf people. The proposed system gives the accuracy 96.25 %, precision value 0.98, recall value 0.96. An LSTM model is employed for classification tasks, exhibiting encouraging outcomes in converting concise sign language gestures into text.The system works efficiently in real-time for seamless communication between normal individuals and deaf people.

The future scope for the project is that it can improve by adding more words and short phrases to its dataset. Right now, it mainly focuses on unfamiliar signs that aren't in SignVaria. The main challenge is that the system can't recognize new signs. This could be a future goal for the project.

References

[1] Q. M. Areeb, Maryam, M. Nadeem, R. Alroobaea and F. Anwer, ”Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach,” in IEEE Access, vol. 10, pp. 8502-8517, 2022, doi: 10.1109/ACCESS.2022.3142918.

[2] N. Adaloglou et al., ”A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition,” in IEEE Transactions on Multimedia, vol. 24, pp. 1750-1762, 2022, doi: 10.1109/TMM.2021.3070438.

[3] J. Zhao, W. Qi, W. Zhou, N. Duan, M. Zhou and H. Li, ”Conditional Sentence Generation and Cross-Modal Reranking for Sign Language Translation,” in IEEE Transactions on Multimedia, vol. 24, pp. 2662-2672, 2022, doi: 10.1109/TMM.2021.3087006.

[4] P. Wang, Y. Zhou, Z. Li, S. Huang and D. Zhang, ”Neural Decoding of Chinese Sign Language With Machine Learning for Brain–Computer Interfaces,” in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 2721-2732, 2021, doi: 10.1109/TNSRE.2021.3137340.

[5] . Huang, W. Zhou, H. Li and W. Li, ”Attention-Based 3D-CNNs for Large-Vocabulary Sign Language Recognition,” in IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 9, pp. 2822-2832, Sept. 2019, doi: 10.1109/TCSVT.2018.2870740.

[6] J. Joy, K. Balakrishnan and M. Sreeraj, ”SignQuiz: A Quiz Based Tool for Learning Fingerspelled Signs in Indian Sign Language Using ASLR,” in IEEE Access, vol. 7, pp. 28363-28371, 2019, doi: 10.1109/ACCESS.2019.2901863.

[7] hongfu Ye-Attention-Based Sign Language Recognition Network Utilizing Keyframe Sampling and Skeletal Features 2020

[8] R. Elakkiya-Multi-Semantic Discriminative Feature Learning for Sign Gesture Recognition Using Hybrid Deep Neural Architecture 2023

[9] Kosin Chamnongthai-American Sign Language Words Recognition Using Spatio-Temporal Prosodic and Angle Features: A Sequential Learning Approach 2022

[10] Muhammad Al-Qurishi-Deep Learning for Sign Language Recognition:Current Techniques, Benchmarks,and Open Issues 2021

[11] Kosin Chamnongthai- IDF-Sign: Addressing Inconsistent Depth Features for Dynamic Sign Word Recognition 2023

[12] Akash Deep, Ashitosh Litoria, Chamaria, Akshay Ingole, & Vaibhav Asare, J. ”Realtime Sign Language Detection and Recognition(2022).” In 2022 2nd Asian Conference on Innovation in Technology (ASIANCON). DOI:10.1109/ASIANCON55314.2022.9908995

[13] Sruthi C.J.,Lijiya, J. ”Signet: A Deep Learning based Indian Sign Language Recognition System” In 2019 International Conference on Communication and Signal Processing (ICCSP).DOI:10.1109/ICCSP.2019.8698006

[14] Atharv Ganpatye; Sunil Mane J. ”Motion Based Indian Sign Language Recognition using Deep Learning” In 2022 2nd International Conference on Intelligent Technologies (CONIT). DOI:10.1109/ICCSP.2019.8698006

[15] Aniket Kumar, Mehul Madaan,Shubham Kumar , Aniket Saha, & Suman Yadav ”Indian Sign Language Gesture Recognition in Real-Time using Convolutional Neural Networks” In 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN). DOI:10.1109/SPIN52536.2021.9566005

[16] Pradip N. Shendage, Santosh A. Shinde,Shubham ”A Novel QoS-Based Approach for Web Service Recommendation and Visualization” In International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Volume 4 Issue 5, May 2015 Paper ID: SUB155122

[17] https://indiansignlanguage.org/

Appendix A

Plagiarism Report

turnitin — **Similarity Report ID:** oid:7916:59943237

PAPER NAME
BE Report II.pdf

AUTHOR
Mayuri and Pranali

WORD COUNT
5708 Words

CHARACTER COUNT
33170 Characters

PAGE COUNT
44 Pages

FILE SIZE
4.5MB

SUBMISSION DATE
May 25, 2024 12:40 AM GMT+5:30

REPORT DATE
May 25, 2024 12:40 AM GMT+5:30

9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 6% Internet database
- 4% Publications database
- Crossref database
- Crossref Posted Content database
- 8% Submitted Works database

Excluded from Similarity Report

- Manually excluded text blocks

Appendix B

Base Paper

J. Joy, K. Balakrishnan and M. Sreeraj, "SignQuiz: A Quiz Based Tool for Learning Fingerspelled Signs in Indian Sign Language Using ASLR," in IEEE Access, vol. 7, pp. 28363-28371, 2019, doi: 10.1109/ACCESS.2019.2901863.

paper link:https://ieeexplore.ieee.org/document/8657686

Appendix C

Tools Used

1. Python (3.11.5) :-

   Programming Language: Python is a high-level, readable programming language, with version 3.11.5 incorporating the latest features for various applications.

2. Libraries like OpenCV :-

   Image Capturing: OpenCV is a robust open-source computer vision library widely used for image capturing and processing.

3. Draw.io, Gantt Chart :-

   SRS: Draw.io is a web-based diagramming tool, and Gantt charts provide visual project timelines, both utilized in creating Software Requirements Specifications (SRS).

4. LaTeX :-

   Documentation: LaTeX is a typesetting system ideal for preparing high-quality scientific and technical documents.

Appendix D

Papers Published/Certificates

- **Conference Title:**Institute of Electrical and Electronics Engineers (IEEE)

- **Paper Title:** Empowering Deaf with Indian Sign Language Interpreter using Deep Learning

- **Conference Paper:**

Empowering Deaf with Indian Sign Language Interpreter using Deep Learning

Shivanjali Vijay Nimbalkar
Artificial Intelligence and Data Science
VPKBIET
Baramati, India
shivanjali.nimbalkar.29@gmail.com

Soham Nilesh Vaidya
Artificial Intelligence and Data Science
VPKBIET
Baramati, India
sohamvaidya333@gmail.com

Mayuri Mahesh Gade
Artificial Intelligence and Data Science
VPKBIET
Baramati, India
mayurigade02@gmail.com

Pranali Sandip Hagare
Artificial Intelligence and Data Science
VPKBIET
Baramati, India
pranalihagare6@gmail.com

Pradip N. Shendage
Artificial Intelligence and Data Science
VPKBIET
Baramati, India
pradipshendage27@gmail.com

***Abstract*—In the ever-changing global landscape, the diversity of sign languages, with over 140 distinct variants worldwide, poses a significant challenge in developing universally applicable recognition models. This complexity is compounded by the dynamic nature of sign languages, requiring constant adaptation to incorporate emerging signs tied to technological advancements. This research aims to improve communication accessibility for the deaf community in India through a Real-time Indian Sign Language (ISL) recognition system. Specifically addressing the communication gap experienced by deaf children, our approach leverages advanced technologies such as deep learning, computer vision, and neural networks to convert ISL gestures into a textual format. The system relies on a carefully curated dataset, including emergency words, Devanagari script, and English alphabets in ISL. Using methodologies like the mediapipe library and cvzone hand tracking module, we construct the SignVaria and implement dynamic hand gesture recognition modules with advanced deep learning architectures, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The goal is to overcome regional specificity challenges in sign language recognition, ensuring the integration of cutting-edge technologies for enhanced robustness and accessibility for the target demographic.Our research yielded the most impressive results using the Convolutional Neural Network (CNN), achieving an accuracy of 96.1%.**

***Index Terms*—Indian Sign Language (ISL), Mediapipe library, cvzone hand tracking module, cvzone ,Computer Vision, Hand gesture recognition, LSTM.**

I. Introduction

An estimated 466 million individuals globally confront challenges associated with hearing impairments, with a notable subset of 34 million being children. The preferred communication method within the deaf community is sign language, yet the diversity in sign languages across countries is less extensive than linguistic variations in spoken languages. In the Indian context, Indian Sign Language (ISL) takes precedence. Unfortunately, the scarcity of educational institutions catering to deaf students, particularly in emerging nations, exacerbates the situation. Limited access to sign language education and a lack of understanding contribute to elevated unemployment rates among adults with hearing loss.

Ethnographic data reveals a stark contrast in India, where only 1 percent of the population is deaf, and enrollment of deaf children in educational institutions is notably low. This emphasizes the urgent need for heightened awareness, increased support, and expanded educational opportunities for the deaf community. Effective communication proves challenging for deaf individuals, especially in interactions with those who can hear. Current technology falls short in accurately understanding and translating sign language, making it difficult for the deaf to express themselves in conversations with the hearing.

The crux of the issue lies in the intricate nature of sign language—comprising gestures, expressions, and movements—posing a formidable challenge for technology to fully comprehend. Existing tools often miss subtle nuances, leading to misunderstandings and communication breakdowns. Addressing this challenge involves enhancing computer programs through the integration of advanced technologies such as artificial intelligence and computer vision. This research delves into these complexities, emphasizing a specific focus on understanding the unique intricacies of sign language.

In this research, we make substantial contributions to the field of Indian Sign Language (ISL) interpretation systems. Our work encompasses the creation of a meticulously curated dataset SignVaria representing diverse ISL expressions. Additionally, we introduce a real-time system capable of translating short sign words and hand gesture-based phrases to text, addressing the crucial need for effective communication. Notably, our contributions are underscored by the achievement of significant accuracies, with the CNN model reaching 96 % and the LSTM model attaining 95.8 %. These advancements collectively mark a substantial stride towards enhancing accessibility and inclusivity for the hearing-impaired community.

The subsequent sections of the paper are organized as follows: Section 2 provides a summary of related work.

Section 3 outlines the system architecture. Section 4 details the methodology. Section 5 presents a summary of the results. Finally, Section 6 concludes and summarizes the work.

II. Literature Review

Within the realm of sign recognition systems, a thorough exploration unfolds, encompassing diverse methodologies and technologies with the goal of enhancing communication for individuals facing challenges in hearing and speech. Embarking on this journey, the inaugural paper employs Nasnet and Inception V3 within the TensorFlow framework to construct SignQuiz. It underscores the potential to support various sign languages beyond its initial focus on Indian Sign Language (ISL)[1]. Simultaneously, it delves into human-computer interaction, investigating L1 regularization and classification techniques while acknowledging the limitation of a restricted decoding instruction set. The pursuit of continuous sign language recognition is sustained[2].This endeavor incorporates weakly supervised learning and reinforcement learning techniques to amplify the performance of the Sign Language Recognition (SLR) system[3].Subsequently, a novel framework for continuous identification of sign language emerges, leveraging deep neural networks, specifically utilizing convolutional neural networks for feature extraction and bidirectional recurrent neural networks (RNNs) for sequence learning. This work highlights an iterative optimization process, demonstrating its efficacy in handling limited data and achieving over a 15% improvement on challenging benchmarks. Expanding the horizon, a comprehensive assessment of computer vision-based methods for sign language recognition is conducted, introducing innovative sequence training criteria and pretraining schemes[4]. Further innovations arise as authors introduce the Greel sign language and a dataset containing new RGB+D, signaling the superiority of 2D CNN-based models with an intermediate per-gloss representation in continuous recognition of signs. The scope of Sign Language (SL) extends to real-time applications as well[6].A real-time Indian Sign Language (ISL) recognition system is presented, utilizing fuzzy c-means clustering and achieving a 75% accuracy in gesture labeling[6]. Concurrently, a focus on dynamic signs attains a 70% training accuracy using a CNN to bridge the gap between ISL and English[8]. Another notable vision-based system utilizes CNNs and RNNs for interpretation and translation, achieving a commendable 73.60% accuracy[9].

Further innovations underscore the importance of ISL as a visual language in India, employing a deep CNN with the Inception V3 model to achieve a high accuracy of 93% in gesture recognition[10]. An introduction to real-time gesture prediction through depth+RGB data and semantic segmentation is made, while a method for real-time recognition and text-to-speech conversion is presented, offering practical solutions for communication with speech and hearing disabilities[11].The paper presents a deep learning model for Motion-Based Indian Sign Language (ISL) recognition, employing OpenCV for gesture capture, MediaPipe for key point detection, and a trained LSTM model for sign prediction. With an average accuracy of 92 %, the proposed system offers real-time ISL recognition capabilities, making it suitable for integration with video-conferencing applications, thereby enhancing accessibility for the hearing-impaired in communication scenarios[12].This paper addresses the communication challenges faced by the hearing and speech impaired community by proposing a real-time Convolutional Neural Network (CNN) model for identifying and classifying Indian Sign Language (ISL) gestures. Developed using OpenCV and Keras, the model achieves a high accuracy of 99.91% for classifying 36 ISL gestures, contributing to bridging the communication gap and promoting inclusivity in compliance with the Rights of Persons with Disabilities Act, 2016[13].The survey encompasses the evolution of sign language identification and translation to text, traversing learning tools, continuous recognition frameworks, and vision-based interpretation systems. Each contributes to the advancements in technology aimed at fostering inclusive communication. While perusing the survey, instances of overfitting and underfitting conditions were observed across many papers. Despite the utilization of various models in some instances, the accuracy did not consistently reach 90% and above. Predominantly, authors relied on predefined datasets in their studies.

III. System architecture

A. Data Creation

In pursuit of our goal of creating a meticulously curated dataset-SignVaria, we began by capturing photos of hands using Mediapipe and cvzone's hand tracking module. This allowed us to precisely outline existing landmarks. The obtained data was then stored and labeled appropriately. Information about the entire data creation process is described in the methodology.

B. Training model

The envisaged system comprises two training models designed to train the dataset and extract features from the data. Training constitutes a crucial component of the entire system, with the output of the system dependent on the efficacy of the training models.

C. Testing

The testing phase encompasses the development of an interface that showcases character output. Upon the user accessing the camera, real-time image capture ensues. Subsequently, the classification module systematically categorizes the signs into textual sequences.

D. Hardware Configuration

For our research, we employed PyCharm software alongside hardware featuring an Intel Core i5 12th generation processor, 16 GB of RAM, and a 4 GB graphics card. These specifications represent the minimum requirements, and with further optimization, we can potentially reduce them.

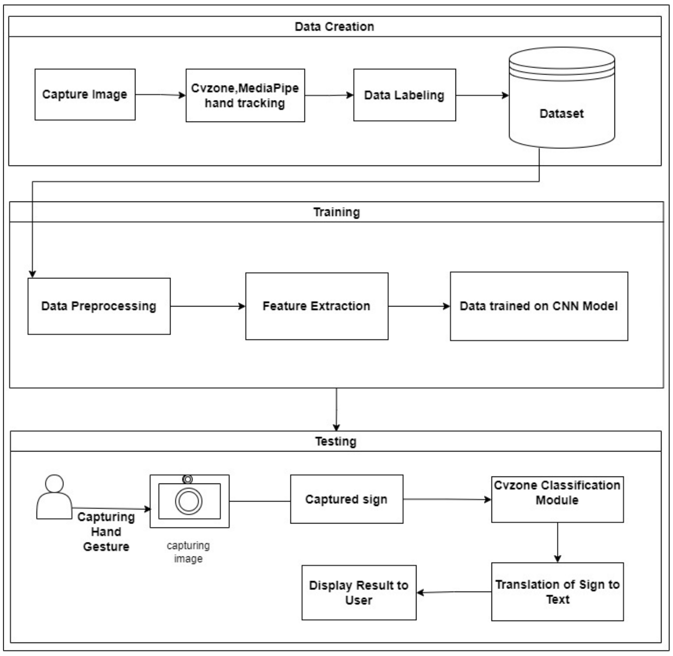


Fig. 1. System Architecture-1

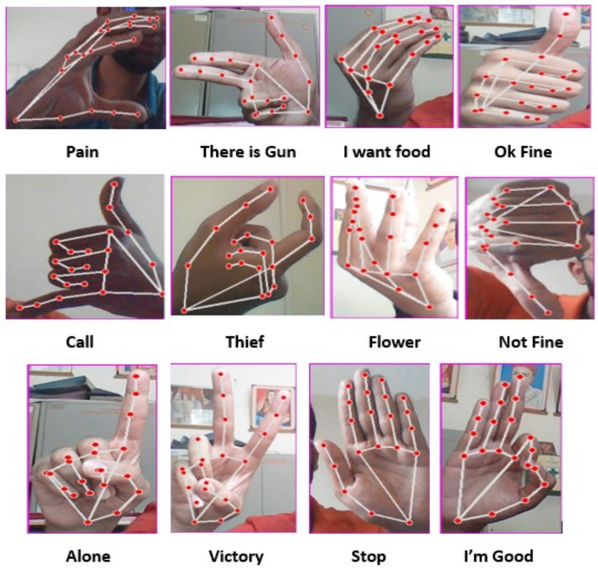


Fig. 2. SignVaria Dataset Signs

IV. Methodology

A. Overview

Through a comprehensive review of the existing literature, we formulated a SignVaria exclusively centered around Indian Sign Language. Our research introduces two distinct methods for generating data. In the proposed system, we undertake a comparative analysis of these two approaches in training data, ultimately shaping the system in a dual manner.

V. Model-1

A. Data Creation

1) Image Capturing: With help of camera, set of images are captured with various hand poses. Eliminating the lighting conditions and by maintaining a plane background played important role to make the model more robust.

2) MediaPipe and Cvzone Hand Tracking: This library is utilized for hand tracking. It involves running the hand tracking model on each captured image to detect and locate the hand(s) within the image.It provides a pre-trained hand tracking model that can be used to identify the landmarks (key points) on the hand. Which provides information about the hand's position and orientation.

3) Data Labeling: We have created separate folder for each label while storing images in the dataset.For each image, we associate the extracted landmark coordinates with the corresponding hand regions in the image.

B. Dataset Features

In First model we collect dataset of 25 different Indian signs. The dataset contain 500 images for each Labelled sign. According to Sign shape image size lies between 17 KB to 24 KB. Description of stored image is as follow:

TABLE I
Image description

Parameters	values
Dimensions	300 X 300
Width	300 pixels
Height	300 pixels
Horizontal resolution	96 dpi
Vertical resolution	96 dpi
Bit depth	24

C. Training approach

1) Convolutional Neural Network: cvzone Classification Module include different machine algorithms such as Support Vector Machines (SVM), Neural Network (CNN), Random forests depending on complexity of images it decides the desirable model for classification tasks. In first model cvzone Classification Module uses CNN to perform tasks such as image classification or object recognition In Indian Sign Language.

In CNN model hidden layer has ReLU and For classification task in output layer softmax activation function is utilised.

TABLE II
Hyperparameters

Parameters
learning rate = 0.2
batch size = 32
epochs = 100

Mathematical Model used in LSTM is as follow:
1.Input Layer: Here input layer contains r,g,b values
Input=X
2.Convolutional Layer:

$$Z_{conv} = W_{conv} \times X + b_{conv} \tag{1}$$

$$A_{conv} = ReLU(Z_{conv}) \tag{2}$$

where:
•W_{conv} is the convolutional filter parameter.
•b_{conv} is bias.
•Z_{conv} Convolutional result before activation.
•A_{conv} output after applying ReLU activation function.
3.Pooling Layer: Input:$A^{[1]}$,output of first convolutional layer.

$$A^{[2]}_{i,j,k} = max(A^{[1]}_{2i2,j,k}, A^{[1]}_{2i,2j+1,k}, A^{[1]}_{2i+1,2j,k}, A^{[1]}_{2i+1,2j+1,k}) \tag{3}$$

4.Flatten Layer:
The Flatten layer simply reshapes the input tensor into a flat vector without any additional parameters. If the output shape of the MaxPooling Layer 1 is (None, 56, 56, 64), then the Flatten layer would reshape it to (None, 56 ×56× 64) = (None, 200, 704).
5.Dense Layer:
When input reaches to Dense Layer 1 it is a flattened to size (None, 200, 704):

$$Z^{(1)} = W^{(1)} \ldots Flattenoutput + b^{(1)} \tag{4}$$

$$A^{(1)} = ReLU(Z^{(1)}) \tag{5}$$

Where:
•$Z_{(1)}$ is the weighted sum.
•$A_{(1)}$ is the output after activation function.
6.Output Layer:
Assuming the input give to the output layer is $A^{(1)}$ and the layer parameters are $W^{(2)}$ (weights) and $b^{(2)}$ (biases):

$$Z^{(2)} = W^{(2)} \ldots A^{(1)} + b^{(2)} \tag{6}$$

$$Y = Softmax(Z^{(2)}) \tag{7}$$

Where:
• Y is the output after applying the Softmax activation function.

VI. Model-2

A. Data Creation

1) Capture Image: For capturing the video cvzone library is used. For capturing one sign image a video gets started in which we have given 50 sequences in which each sequence captures 30 frames.

2) Hand Detection: For detecting hand gestures, landmarks and body poses Mediapipe library is used.

3) Feature extraction: With the help of Mediapipe library landmarks on both hands are detected. There are 33 landmarks on body and 21 landmarks on both right and left hand. Each hand landmark contain 3 features which are X, Y, Z and body landmarks contain 4 features X, Y, Z, visibility.

4) Array Formation: These extracted features are stored in array and this array is saved as .npy file in dataset.

B. Dataset Features

Dataset contain 25 different signs. Each sign is captured by giving 50 sequences and each sequence contains 30 frames . The landmark value for different hand and body poses is different. After extracting the landmark values for different poses that values are stored in one dimensional array in the form of .npy file. One numpy file contain total 258 different values. 258 values are calculated using following formula :
$(33L \times 4F) + (21L \times 3F) + (21L \times 3F) = 258$
Where,
L = Landmarks
F = Features
33L = 33 body landmarks
4F = 4 features of each body landmark
21 L = 21 landmarks for both hands
3F = 3 features of each hand landmark

C. Training Approach

Model 2 uses LSTM which is designed to capture long term dependencies from land marks of hand gestures. Proposed system contain 3 LSTM layers, 3 dense layers. LSTM model contain ReLU where Output layer contain Softmax activation function which help to classify the signs.

Mathematical Model use in LSTM model to build the system is as follow:
Here LSTM layer(i = 1,2,3)

1.Input Gate(i_i)

$$i^{(i)}_t = \sigma(W^{(i)}_i \cdot [h^{(i)}_{(t-1)}, x_t] + b^{(i)}_i) \tag{8}$$

•$W_{(1)}$ is weight and $b_{(1)}$is biases for input gate.
•σ is Sigmoid activation.
•h^{t-1} preceding hidden state
2.Forget Gate(f_t)

$$f^{(i)}_t = \sigma(W^{(i)}_f \cdot [h^{(i)}_{(t-1)}, x_t] + b^{(i)}_f) \tag{9}$$

3.Cell State(C_t)Update:

$$C^{(i)}_t = f^{(i)}_t \cdot C^{(i)}_{(t-1)} + i^{(i)}_i \cdot \overline{C} \tag{10}$$

$$\overline{C} = ReLU(W^{(i)}_c \cdot [h^{(i)}_{(t-1)}, x_t] + b^{(i)}_c) \tag{11}$$

4.Output Gate(O_t)

$$O^{(i)}_t = \sigma(W^{(i)}_O \cdot [h^{(i)}_{(t-1)}, x_t] + b^{(i)}_O) \tag{12}$$

5. Hidden State(h_t):

$$h^{(i)}_t = O^{(i)}_t \cdot ReLU(C_t) \tag{13}$$

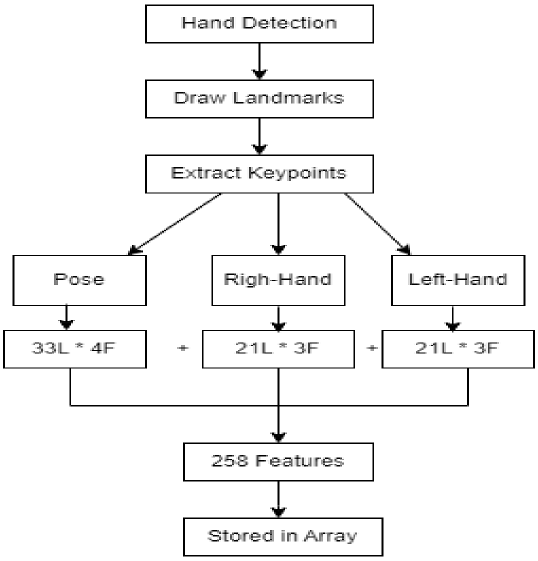


Fig. 3. Data Feature Extraction of Model-2

Here Dense layer(j = 1,2)
6. Dense Layer 1 and 2:
input:$(O^{(3)})$ =[$h_1^{(3)}$,$h_1^{(3)}$,......$h^{(3)}_{numtimesteps}$]
Forward Pass: for t = 1,2,....,$numtimesteps$:

$$Z_t^{(j)} = W^{(j)} \cdot O_t^{(j)} + b^{(1)} \tag{14}$$

$$A_t^{(j)} = ReLU(Z_t^{(j)}) \tag{15}$$

7. Output layer:
input:$(A^{(2)})$ =[$A_1^{(2)}$,$A_1^{(2)}$,......$A^{(3)}_{num_time_steps}$]
Forward Pass: for t = 1,2,....,num-time-steps:

$$Z_t^{(3)} = W^{(3)} \cdot A_t^{(2)} + b^{(2)} \tag{16}$$

$$Y_t^{(3)} = Softmax(Z_t^{(3)}) \tag{17}$$

Output gives the results of sign with text.

VII. Result

Our curated data set SignVaria improves growth of Indian sign language. Proposed system gives output into textual format. When user show sign to system then it gives accurate meaning of that sign in Indian sign language in to text. So far, we have collected 25 words of data and achieved accurate results. For calculating result analysis of sign prediction we use four parameter Accuracy, Precision, recall and F1 score.

1) Accuracy: This indicates Signs with which the model predicts correctly. It's calculated accurate predictions with respect to all predicted result.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{18}$$

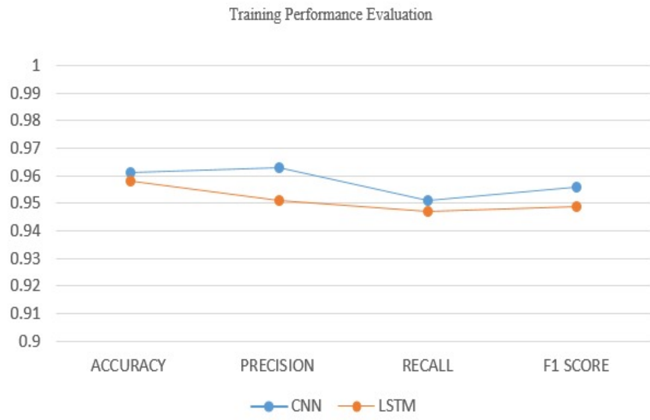


Fig. 4. Performance of models for Indian Sign Language detection

2) Precision: This measures the accuracy of positive predictions. Specifically, it's the fraction of true positive predictions to all positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{19}$$

3) Recall: This metric focuses on the true positive rate. It represents the fraction of actual positive cases that were correctly identified.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{20}$$

4) F1 Score: This is the weighted average of Precision and Recall. It's especially useful when there's an imbalance between classes. An F1 Score close to 1 denotes excellent precision and recall, while a score near 0 shows the opposite.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{21}$$

Our study comprehensively reviews two classification algorithms.we analyzed an Convolutional Neural Network, which achieved an accuracy of 96.1% and LSTM, which achieved an accuracy of 95.8% .Both the models give fine result and correctly translate the sign to it's appropriate text.Given methods of data creation and training help to improve the result. We are able to overcome the over fitting and under fitting problems occurring in other survey papers.

TABLE III
RESULTS

ALGORITHM	CNN	LSTM
ACCURACY	0.961	0.958
PRECISION	0.963	0.951
RECALL	0.951	0.947
F1 SCORE	0.956	0.949

VIII. Scope of Project

Dedicated to cultivating an inclusive educational setting, we are in the process of crafting a sophisticated system seamlessly

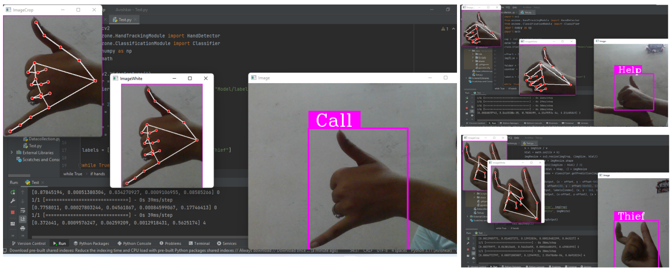


Fig. 5. Output Interface of Model-1

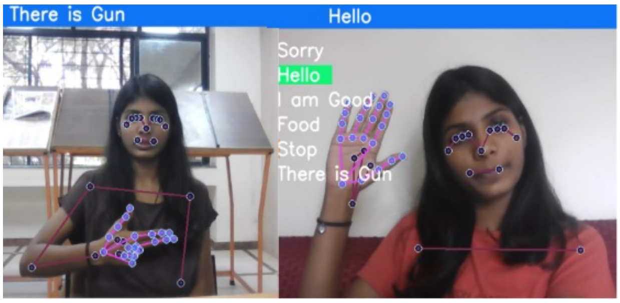


Fig. 6. Output Interface of Model-2

integrated into educational platforms. This integration aims to elevate the learning experience for deaf students.

The proposed system leverages cutting-edge speech-to-text technology, enabling real-time transcription of lectures. This ensures that deaf students gain immediate access to lecture content, equipping them with the same information available to their hearing peers. By facilitating swift access to course materials, this feature goes beyond mere transcription; it empowers deaf students to focus on comprehension, eliminating the struggle to keep pace with spoken information.

IX. Conclusion

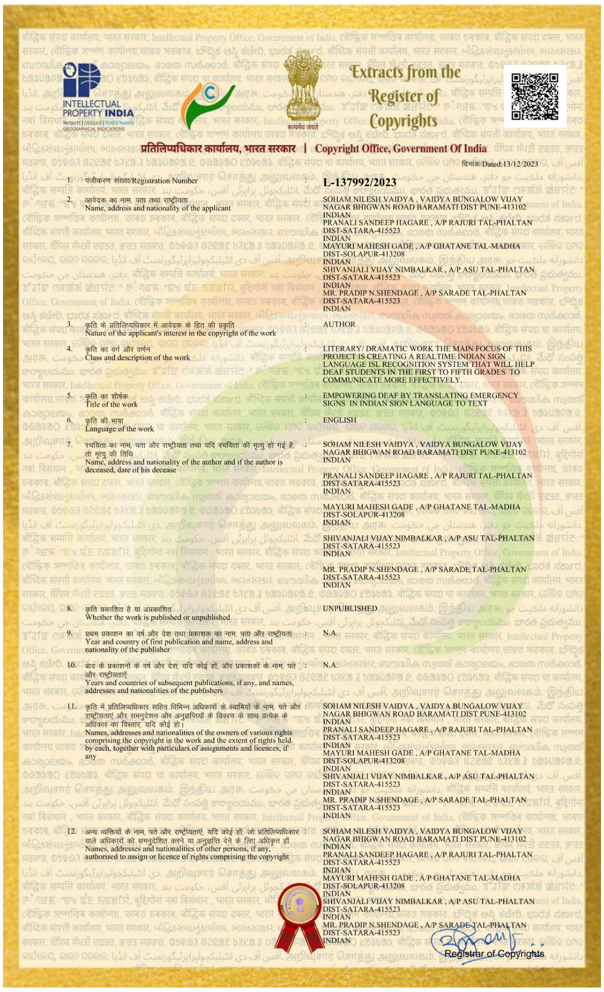
Both approaches yield optimal results in comparison. Employing model 1, we construct the proposed system, offering the advantage of simplified data collection and storage. The primary objective of this system is to alleviate the communication gap between deaf individuals and those without hearing impairments by developing an Indian Sign Language (ISL) interpreter system. We have effectively curated a pivotal dataset SignVaria and implemented three key models. Together, these models exhibit promising outcomes, particularly in the translation of concise sign language expressions into text, with a particular focus on emergency situations. This research marks a significant stride in enhancing real-time communication accessibility for the deaf community.

Looking ahead, our future endeavors aim to enhance the dataset by expanding the vocabulary to encompass additional words and brief sentences. Presently, our system encounters limitations in recognizing novel signs beyond those present in the SignVaria. The inability to detect new signs poses a current challenge. However, we envision overcoming these challenges in the future through the implementation of more robust techniques.

References

[1] Joy, J., Balakrishnan, K., & S. M. (2019). "SignQuiz: A Quiz Based Tool for Learning Fingerspelled Signs in Indian Sign Language Using ASLR." In *IEEE Conference*, Volume 7. DOI: 10.1109/ACCESS.2019.2901863.
[2] Wang, P. "Neural Decoding of Chinese Sign Language with Machine Learning for Brain-Computer." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*,Volume 29. DOI:10.1109/TNSRE.2021.3137340.
[3] Zhou, W., Huang, J., & Zhou, W. (2020). "Attention-Based 3D-CNNs for Large-Vocabulary Sign Language Recognition." *IEEE Transactions on Circuits and Systems for Video Technology*, 29(9), 2822–2832. DOI: 10.1109/TCSVT.2018.2870740.
[4] Cui, R., Liu, H., & Zhang, C. (2021). "A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training." *IEEE Transactions on Multimedia*, 21(7), 1880–1891. DOI: 10.1109/TMM.2018.2889563.
[5] Mariappan, M. H., & Dr. Gomathi, V. (2019). "Real-time Recognition of Indian Sign Language." In *Sixth International Conference on Image Information Processing (ICIIP)*. DOI: 10.1109/ICIIP53038.2021.9702611.
[6] Mistry, P., Jotaniya, V., Patel, P., & Patel, N. (2021). "Indian Sign Language Recognition using Deep Learning." In *International Conference on Artificial Intelligence and Machine Vision (AIMV)*. DOI: 10.1109/AIMV53313.2021.9670933.
[7] Varsha, M., & Nair, C. S. (2021). "Indian Sign Language Gesture Recognition Using Deep Convolutional Neural Network." In *8th International Conference on Smart Computing and Communications (ICSCC)*. DOI: 10.1109/ICSCC51209.2021.9528246.
[8] Likhar, P., Bhagat, N. K., & Rathna, G. N. (2020). "Deep Learning Methods for Indian Sign Language Recognition." In *IEEE 10th International Conference on Consumer Electronics (ICCE-Berlin)*. DOI: 10.1109/ICCE-Berlin50680.2020.9352194.
[9] Gangadia, D., Chamaria, V., Doshi, V., & Gandhi, J. (2020). "Indian Sign Language Interpretation and Sentence Formation." In *2020 IEEE Pune Section International Conference (PuneCon)*. DOI: 10.1109/PuneCon50868.2020.9362383.
[10] Akash Deep, Ashitosh Litoria, Chamaria, Akshay Ingole, & Vaibhav Asare, J. "Realtime Sign Language Detection and Recognition(2022)." In * 2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*. DOI:10.1109/ASIANCON55314.2022.9908995
[11] Sruthi C.J.,Lijiya, J. "Signet: A Deep Learning based Indian Sign Language Recognition System" In * 2019 International Conference on Communication and Signal Processing (ICCSP)*. DOI:10.1109/ICCSP.2019.8698006
[12] Atharv Ganpatye; Sunil Mane J. "Motion Based Indian Sign Language Recognition using Deep Learning" In * 2022 2nd International Conference on Intelligent Technologies (CONIT)*. DOI:10.1109/ICCSP.2019.8698006
[13] Aniket Kumar, Mehul Madaan,Shubham Kumar , Aniket Saha, & Suman Yadav "Indian Sign Language Gesture Recognition in Real-Time using Convolutional Neural Networks" In * 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*. DOI: 10.1109/SPIN52536.2021.9566005
[14] Pradip N. Shendage, Santosh A. Shinde,Shubham "A Novel QoS-Based Approach for Web Service Recommendation and Visualization" In * International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 *Volume 4 Issue 5, May 2015* Paper ID: SUB155122
[15] https://github.com/SohamVaidya12/ISL-Interpretation-system

D.1 Copyright Certificate



Extracts from the Register of Copyrights

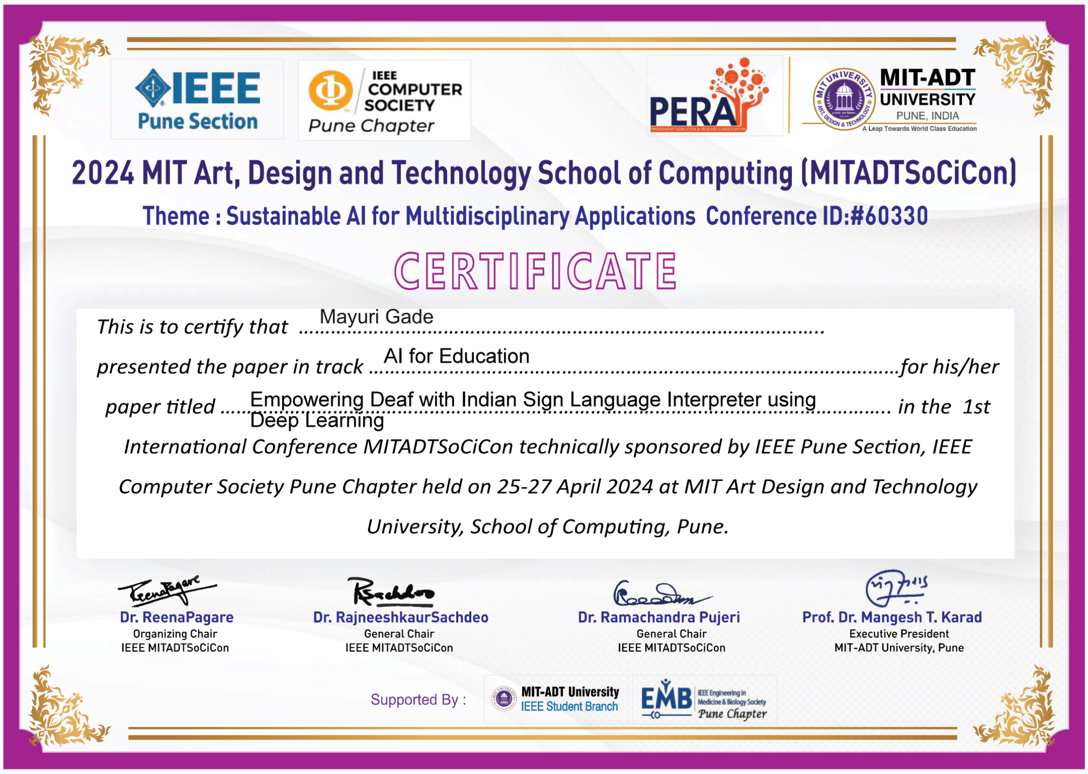
प्रतिलिप्यधिकार कार्यालय, भारत सरकार | Copyright Office, Government Of India

दिनांक/Dated:13/12/2023

1.	पंजीकरण संख्या/Registration Number	L-137992/2023
2.	आवेदक का नाम, पता तथा राष्ट्रीयता Name, address and nationality of the applicant	SOHAM NILESH VAIDYA , VAIDYA BUNGALOW VIJAY NAGAR BHIGWAN ROAD BARAMATI DIST PUNE-413102 INDIAN PRANALI SANDEEP HAGARE , A/P RAJURI TAL-PHALTAN DIST-SATARA-415523 INDIAN MAYURI MAHESH GADE , A/P GHATANE TAL-MADHA DIST-SOLAPUR-413208 INDIAN SHIVANJALI VIJAY NIMBALKAR , A/P ASU TAL-PHALTAN DIST-SATARA-415523 INDIAN MR. PRADIP N.SHENDAGE , A/P SARADE TAL-PHALTAN DIST-SATARA-415523 INDIAN
3.	कृति के प्रतिलिप्यधिकार में आवेदक के हित की प्रकृति Nature of the applicant's interest in the copyright of the work	AUTHOR
4.	कृति का वर्ग और वर्णन Class and description of the work	LITERARY/ DRAMATIC WORK THE MAIN FOCUS OF THIS PROJECT IS CREATING A REALTIME INDIAN SIGN LANGUAGE ISL RECOGNITION SYSTEM THAT WILL HELP DEAF STUDENTS IN THE FIRST TO FIFTH GRADES TO COMMUNICATE MORE EFFECTIVELY.
5.	कृति का शीर्षक Title of the work	EMPOWERING DEAF BY TRANSLATING EMERGENCY SIGNS IN INDIAN SIGN LANGUAGE TO TEXT
6.	कृति की भाषा Language of the work	ENGLISH
7.	रचयिता का नाम, पता और राष्ट्रीयता तथा यदि रचयिता की मृत्यु हो गई है, तो मृत्यु की तिथि Name, address and nationality of the author and if the author is deceased, date of his decease	SOHAM NILESH VAIDYA , VAIDYA BUNGALOW VIJAY NAGAR BHIGWAN ROAD BARAMATI DIST PUNE-413102 INDIAN PRANALI SANDEEP HAGARE , A/P RAJURI TAL-PHALTAN DIST-SATARA-415523 INDIAN MAYURI MAHESH GADE , A/P GHATANE TAL-MADHA DIST-SOLAPUR-413208 INDIAN SHIVANJALI VIJAY NIMBALKAR , A/P ASU TAL-PHALTAN DIST-SATARA-415523 INDIAN MR. PRADIP N.SHENDAGE , A/P SARADE TAL-PHALTAN DIST-SATARA-415523 INDIAN
8.	कृति प्रकाशित है या अप्रकाशित Whether the work is published or unpublished	UNPUBLISHED
9.	प्रथम प्रकाशन का वर्ष और देश तथा प्रकाशक का नाम, पता और राष्ट्रीयता Year and country of first publication and name, address and nationality of the publisher	N.A.
10.	बाद के प्रकाशनों के वर्ष और देश, यदि कोई हों, और प्रकाशकों के नाम, पते और राष्ट्रीयताएँ Years and countries of subsequent publications, if any, and names, addresses and nationalities of the publishers	N.A.
11.	कृति में प्रतिलिप्यधिकार सहित विभिन्न अधिकारों के स्वामियों के नाम, पते और राष्ट्रीयताएं और समनुदेशन और अनुज्ञप्तियों के विवरण के साथ प्रत्येक के अधिकार का विस्तार, यदि कोई हो। Names, addresses and nationalities of the owners of various rights comprising the copyright in the work and the extent of rights held by each, together with particulars of assignments and licences, if any	SOHAM NILESH VAIDYA , VAIDYA BUNGALOW VIJAY NAGAR BHIGWAN ROAD BARAMATI DIST PUNE-413102 INDIAN PRANALI SANDEEP HAGARE , A/P RAJURI TAL-PHALTAN DIST-SATARA-415523 INDIAN MAYURI MAHESH GADE , A/P GHATANE TAL-MADHA DIST-SOLAPUR-413208 INDIAN SHIVANJALI VIJAY NIMBALKAR , A/P ASU TAL-PHALTAN DIST-SATARA-415523 INDIAN MR. PRADIP N.SHENDAGE , A/P SARADE TAL-PHALTAN DIST-SATARA-415523 INDIAN
12.	अन्य व्यक्तियों के नाम, पते और राष्ट्रीयताएं, यदि कोई हों, जो प्रतिलिप्यधिकार वाले अधिकारों को समनुदेशित करने या अनुज्ञप्ति देने के लिए अधिकृत हों Names, addresses and nationalities of other persons, if any, authorised to assign or licence of rights comprising the copyright	SOHAM NILESH VAIDYA , VAIDYA BUNGALOW VIJAY NAGAR BHIGWAN ROAD BARAMATI DIST PUNE-413102 INDIAN PRANALI SANDEEP HAGARE , A/P RAJURI TAL-PHALTAN DIST-SATARA-415523 INDIAN MAYURI MAHESH GADE , A/P GHATANE TAL-MADHA DIST-SOLAPUR-413208 INDIAN SHIVANJALI VIJAY NIMBALKAR , A/P ASU TAL-PHALTAN DIST-SATARA-415523 INDIAN MR. PRADIP N.SHENDAGE , A/P SARADE TAL-PHALTAN DIST-SATARA-415523 INDIAN

Registrar of Copyrights

D.2 Conference Certificates



IEEE Pune Section
IEEE COMPUTER SOCIETY Pune Chapter
PERA
MIT-ADT UNIVERSITY PUNE, INDIA
A Leap Towards World Class Education

2024 MIT Art, Design and Technology School of Computing (MITADTSoCiCon)

Theme : Sustainable AI for Multidisciplinary Applications Conference ID:#60330

CERTIFICATE

This is to certify that Mayuri Gade presented the paper in track AI for Education for his/her paper titled Empowering Deaf with Indian Sign Language Interpreter using Deep Learning in the 1st International Conference MITADTSoCiCon technically sponsored by IEEE Pune Section, IEEE Computer Society Pune Chapter held on 25-27 April 2024 at MIT Art Design and Technology University, School of Computing, Pune.

Dr. ReenaPagare, Organizing Chair, IEEE MITADTSoCiCon
Dr. RajneeshkaurSachdeo, General Chair, IEEE MITADTSoCiCon
Dr. Ramachandra Pujeri, General Chair, IEEE MITADTSoCiCon
Prof. Dr. Mangesh T. Karad, Executive President, MIT-ADT University, Pune

Supported By : MIT-ADT University IEEE Student Branch; EMB IEEE Engineering in Medicine & Biology Society Pune Chapter

IEEE Pune Section
IEEE COMPUTER SOCIETY Pune Chapter
PERA
MIT-ADT UNIVERSITY PUNE, INDIA
A Leap Towards World Class Education

2024 MIT Art, Design and Technology School of Computing (MITADTSoCiCon)

Theme : Sustainable AI for Multidisciplinary Applications Conference ID:#60330

CERTIFICATE

This is to certify that Pranali Hagare presented the paper in track AI for Education for his/her paper titled Empowering Deaf with Indian Sign Language Interpreter using Deep Learning in the 1st International Conference MITADTSoCiCon technically sponsored by IEEE Pune Section, IEEE Computer Society Pune Chapter held on 25-27 April 2024 at MIT Art Design and Technology University, School of Computing, Pune.

Dr. ReenaPagare, Organizing Chair, IEEE MITADTSoCiCon
Dr. RajneeshkaurSachdeo, General Chair, IEEE MITADTSoCiCon
Dr. Ramachandra Pujeri, General Chair, IEEE MITADTSoCiCon
Prof. Dr. Mangesh T. Karad, Executive President, MIT-ADT University, Pune

Supported By : MIT-ADT University IEEE Student Branch; EMB IEEE Engineering in Medicine & Biology Society Pune Chapter


IEEE Pune Section

IEEE COMPUTER SOCIETY Pune Chapter

PERA

MIT-ADT UNIVERSITY PUNE, INDIA

2024 MIT Art, Design and Technology School of Computing (MITADTSoCiCon)

Theme : Sustainable AI for Multidisciplinary Applications Conference ID:#60330

CERTIFICATE

This is to certify that Shivanjali Nimbalkar presented the paper in track AI for Education for his/her paper titled Empowering Deaf with Indian Sign Language Interpreter using Deep Learning in the 1st International Conference MITADTSoCiCon technically sponsored by IEEE Pune Section, IEEE Computer Society Pune Chapter held on 25-27 April 2024 at MIT Art Design and Technology University, School of Computing, Pune.

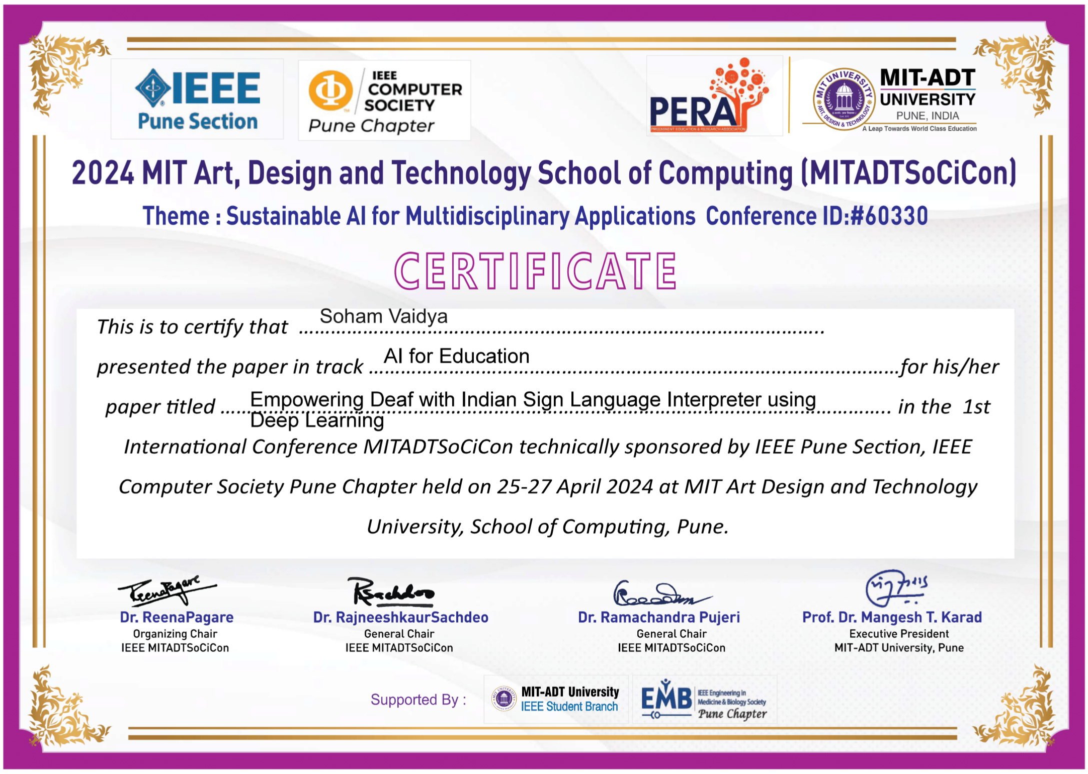
Dr. ReenaPagare, Organizing Chair, IEEE MITADTSoCiCon

Dr. RajneeshkaurSachdeo, General Chair, IEEE MITADTSoCiCon

Dr. Ramachandra Pujeri, General Chair, IEEE MITADTSoCiCon

Prof. Dr. Mangesh T. Karad, Executive President, MIT-ADT University, Pune

Supported By : MIT-ADT University IEEE Student Branch; EMB Pune Chapter


IEEE Pune Section

IEEE COMPUTER SOCIETY Pune Chapter

PERA

MIT-ADT UNIVERSITY PUNE, INDIA

2024 MIT Art, Design and Technology School of Computing (MITADTSoCiCon)

Theme : Sustainable AI for Multidisciplinary Applications Conference ID:#60330

CERTIFICATE

This is to certify that Soham Vaidya presented the paper in track AI for Education for his/her paper titled Empowering Deaf with Indian Sign Language Interpreter using Deep Learning in the 1st International Conference MITADTSoCiCon technically sponsored by IEEE Pune Section, IEEE Computer Society Pune Chapter held on 25-27 April 2024 at MIT Art Design and Technology University, School of Computing, Pune.

Dr. ReenaPagare, Organizing Chair, IEEE MITADTSoCiCon

Dr. RajneeshkaurSachdeo, General Chair, IEEE MITADTSoCiCon

Dr. Ramachandra Pujeri, General Chair, IEEE MITADTSoCiCon

Prof. Dr. Mangesh T. Karad, Executive President, MIT-ADT University, Pune

Supported By : MIT-ADT University IEEE Student Branch; EMB Pune Chapter