INDIAN SIGN LANGUAGE WITH EMOTION RECOGNITION

Submitted in partial fulfillment of the requirements
of the course **Innovative Product Development (IPD) III**

Year 3, Sem V Computer Engineering

By

Hriday Ranka 60004210203
Aaditya Rajesh 60004210206
Darshit Sarda 60004210208
Haardhik Kunder 60004210250

Guides:

Aniket Kore
Professor

Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

University of Mumbai

2023-2024

CERTIFICATE

This is to certify that the project entitled **"INDIAN SIGN LANGUAGE WITH EMOTION RECOGNITION"** is the bonafide work of **Hriday Ranka (60004210203), Aaditya Rajesh (60004210206), Darshit Sarda (60004210208), Haardhik Kunder (60004210250)** submitted as a project work for the course **Innovative Product Development (IPD) III, Year 3, Semester V, B.Tech Computer Engineering**

Prof. Aniket Kore
Internal Guide

Dr. Meera Narvekar
Head of Department

Dr. Hari Vasudevan
Principal

IPD Project Report Approval for BTech Semester V

This project report entitled **INDIAN SIGN LANGUAGE WITH EMOTION RECOGNITION** by Hriday Ranka, Aaditya Rajesh, Darshit Sarda and Haardhik Kunder is approved for the Innovative Product Development (IPD) IV examination **of Year 3, semester V, B.Tech Computer Engineering**

Examiners

1.________________________________

2.________________________________

Date:

Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that We have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from thesources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Hriday Ranka 60004210203

Aaditya Rajesh 60004210206

Darshit Sarda 60004210208

Haardhik Kunder 60004210250

Date:

Abstract

This research project introduces a groundbreaking real-time sign language recognition system with simultaneous emotion analysis, utilizing advanced machine learning techniques. Aimed at addressing the communication challenges faced by hearing-impaired individuals, the system provides a seamless bridge between sign language users and those who communicate verbally. The project's core innovation lies in the integration of Indian Sign Language (ISL) interpretation and emotional context, enhancing the overall communicative experience.

The need for such a system arises from the existing gaps in communication and the time-consuming nature of learning sign language. The proposed solution eliminates the dependency on physical interpreters and caters to various domains, including medical, educational, legal, and training contexts. By recognizing ISL signs and emotions concurrently, the system enhances accessibility and inclusivity for the hearing-impaired community.

The methodology involves two main modules: Hand Sign Classification and Facial Emotion Recognition. The Hand Sign Classification module utilizes Media Pipe and a customized LSTM model trained on a self-generated Indian Sign Language dataset. Achieving an impressive accuracy in real-time testing, this module effectively converts sign language gestures into text. The Facial Emotion Recognition module employs SVM and CNN models, extracting facial coordinates to classify fundamental emotions. This module contributes to a more comprehensive communication experience by conveying the emotional context of the user.

The research explores novel aspects, including the replacement of existing datasets with a self-generated Indian Sign Language dataset. The integration of sign and emotion detection into a single system is a unique feature, differentiating this project from previous approaches. The use of an LSTM model for hand sign classification contributes to enhanced accuracy in recognizing sign language gestures.

The implementation phase involves the development of a user-friendly Graphical User Interface (GUI), allowing real-time testing with live sign language demonstrations. Modules are implemented with meticulous pseudocode and algorithms, ensuring the system's accuracy and efficiency.

In conclusion, this research project represents a significant contribution to the field of assistive communication technologies. The proposed real-time sign language recognition system with emotion analysis offers a holistic solution, revolutionizing communication for the hearing-impaired.

Contents

List of Figures

List of Tables

List of Abbreviations

Sr. No.	Abbreviation	Expanded form
i	DSS	Decision Support System
ii	GUI	Graphical User Interface
iii	LSTM	Long Short-Term Memory
iv	SVM	Support Vector Machine
v	CNN	Convolutional Neural Network
vi	ASL	American Sign Language

Sign-language to Text with Emotion Detection

1. Introduction

In the field of assistive communication technologies, the combination of machine learning and sign language recognition is seen as an innovative way to improve inclusivity and communication for people with hearing difficulties. The challenges faced by the hearing- impaired community highlight the important role of sign language in enabling meaningful interaction. This research aims to introduce a new system that seamlessly combines real-time Indian Sign Language (ISL) interpretation with emotion recognition. As we explore the changing landscape of assistive technologies, this new solution not only addresses current communication problems but also takes the field into unexplored areas, giving users a comprehensive and instant way to express both language and emotions in their interactions.

The main goal of this research goes beyond the usual approaches, intending to redefine how assistive communication works by blending advanced technology with the subtle expressions of human language. To promote inclusivity and accessibility, the system presented here goes beyond the limitations of typical methods, trying to provide the hearing-impaired community with a tool that understands both Indian Sign Language gestures and the associated emotional context.

As we start on this scientific journey, the meeting point of machine learning algorithms and real-time communication becomes crucial for progress. This research is not just a technological achievement but also shows our commitment to creating a more inclusive and connected world. In this envisioned future, every person, no matter their hearing abilities, can engage in a smooth and meaningful exchange of ideas, emotions, and information. By outlining the details of this transformative system, we contribute significantly to the ongoing conversation in the field, pushing assistive communication technologies towards a future where obstacles are removed, and communication acts as a bridge bringing together diverse communities

2. Need of the Product

2.1. Why is the product needed?

The need for an effective sign language recognition system arises from the challenges faced by hearing-impaired individuals in their social interactions. The proposed system addresses the time-consuming nature of learning sign language and the need for interpretation services in various domains, including medical, educational, legal, and training sessions. By recognizing Indian Sign Language (ISL) signs and emotions simultaneously, the system eliminates the reliance on a physical translator.

2.2. Drawbacks of Existing System

Current technologies often focus on either sign language recognition or emotion detection independently. The proposed system integrates both, providing a comprehensive solution and improving the accessibility of the hearing-impaired community. The imperative for a proficient sign language recognition system is underscored by the pervasive challenges confronted by individuals with hearing impairment in the course of their social interactions. This proposed system represents a significant stride in ameliorating the protracted learning curve associated with attaining sign language proficiency and the indispensable need for interpretation services across diverse domains, including medical, educational, legal, and training contexts. The inherent innovation within the system plays a pivotal role in mitigating the communication barriers that hearing-impaired individuals face daily, transcending mere language comprehension to include the crucial aspect of emotional expression.

2.3. Applications of the Product

The system finds applications in medical, educational, legal, and training settings, offering real-time sign language translation and emotion analysis. It contributes to a more inclusive and efficient communication process. The proposed system's contribution is not confined solely to its capacity to alleviate temporal constraints associated with learning sign language. Instead, it extends to the broader realm of fostering inclusivity across diverse professional and social contexts. As the system enables real-time sign language interpretation and emotion recognition,it facilitates a more nuanced and authentic exchange of ideas, sentiments, and information. This technological advancement aligns with the overarching mission of providing individuals with hearing impairments the means to communicate autonomously and effectively, irrespective of thecontext or domain. In essence, the proposed system emerges as an indispensable tool that not only addresses a critical societal need with technological sophistication but also marks a significant stride towards creating a more inclusive and accessible environment for individuals with hearing impairments. By bridging the communication gap in both linguistic and emotional dimensions, this innovative solution embodies a transformative force in enhancing the communicative agency of individuals within the hearing-impaired community

3. Survey

3.1. Field Survey:

The Components involved are:

1]Preparation:

- **Context Identification:**
  We determine the appropriate settings to conduct the survey—places like deaf communities, sign language centers, educational institutions, or events focused on the deaf and hearing-impaired communities.

- **Technology Setup:**
  Ensure that the survey form is accessible on handheld devices (smartphones, tablets) for ease of use in the field.

- **Team Coordination:**
  Organize the team to administer the survey. That is, include individuals proficient in ISL to assist with communication or clarification for respondents.

2]Execution:

- **Participant Engagement:**
  We approached potential participants within the identified settings. They explain the purpose of the survey and invite them to participate voluntarily.

- **Consent and Explanation:**
  Before respondents start the survey, we explain the nature of the survey, assure confidentiality, and obtaininformed consent. This might involve written consent or verbal affirmation depending on the context.

- **Survey Administration:**
  Participants are provided access to the survey formvia a link or on a tablet. We may offer assistance to those who require support in navigating the survey.

- **Language and Accessibility Consideration:**

  We ensured that the survey is available in formatsaccessible to those proficient in ISL. For example,providing live demonstration on how to fill thesurvey would be highly beneficial.

3]Post-Survey:

- **Data Compilation:**
  Responses collected via the Google Form are storedin the Google Drive associated with the survey creator's account**.**

- **Data Analysis:**
  Once the survey period ends, we analysed the data forpatterns, trends, and insights related to ISL, emotions,and the opinions shared.

- **Insights Generation:**
  We derived conclusions based on the analysed data, noting the prevalence of certain opinions, identifying common challenges, or highlighting areas of interest.

4]Ethical Considerations:

- **Privacy and Confidentiality:**
  It was ensured that participant data is protected,anonymized, and not shared without consent.

- **Respectful Engagement:**
  Throughout the survey process, maintain respect for theparticipants' time, culture, and experiences.

5]Adaptation and Improvement:

- **Feedback Collection:**
  Researchers might gather feedback from participants regarding the survey experience to improve futureiterations or research endeavours.

6]Reporting:

- **Findings Presentation:**

  The collected data and insights will be summarized intoreport, presentation, or publication. This could involve visual aids, statistics, or qualitative summaries to convey the survey outcomes effectively.

In conclusion, overall, conducting a field survey involves careful planning, ethical considerations, respectful engagement, and a systematic approach to gathering data directly from participants within their natural environments. The insights gained from such surveys can be invaluable for understanding perspectives, challenges, and preferences within a specific community or context.

3.2 Literature Survey

1) A Deep Learning Approach for Real Time Facial Emotion Recognition:

Title	A Deep Learning Approach for Real-Time Facial Emotion Recognition
Authors	Rupali Gill1*, Jaiteg Singh2
Finding	Achieved 93% accuracy in real-time facial emotion recognition using CNN, surpassing existing research.
Algorithm Used	CNN (Convolutional Neural Network)
Datasets Used	LFW, The Extended Yale Face Database B, Google Facial Expression Comparison
Research Gap	Recommends integrating novel facial application technologies for enhanced facial emotion recognition.
Methodology	Applied CNN-based model for recognizing six facial emotions using LFW, The Extended Yale Face Database B, Google Facial Expression Comparison datasets; detailed data collection and emotion identification steps.
Experimentation Results	Demonstrated dataset statistics, CNN model performance, and comparisons with existing facial emotion recognition techniques; achieved 93% accuracy, surpassing other models.
Results and Discussions	Highlighted accuracy, precision, and recall values; compared CNN-based FER with existing architectures, presenting confusion matrix and training/validation results.
Conclusion	Stressed technological advancements in FER; suggested enhancements using new facial application technologies for practical applications.
References	Cited various related works and studies in the domain of facial emotion recognition including surveys, automatic facial expression analysis, and deep learning approaches.
Abstract	Focused on deep learning techniques for understanding human emotions through facial expressions; achieved 93% accuracy in facial emotion recognition using CNN.

2) CNN based Approach for Sign Recognition in the Indian Sign language:

Section	Information
Title	CNN based Approach for Sign Recognition in the Indian Sign language
Authors	Pranav Unkule, Sanchit Agarkar, Chatak Shinde, Usha Verma, Pratik Saurkar
Algorithms used	CNN (Convolutional Neural Network), Canny Edge Filtering
Datasets used	Dataset with more than 400 images of 10 gestures (consisting of alphabets, numbers, and phrases)
Abstract	Vocal communication, difficulties faced by speech-impaired individuals, proposed model bridging communication gap using gestures, converting gestures to text/audio
Methodology	Dataset creation and pre-processing using OpenCV, Canny Edge Filtering, CNN-based classification, Text to Audio conversion
Experimentation Results	Achieved accuracy of 95.31% with a 90:10 training-testing ratio
Results and Discussions	Describes various research papers related to sign language recognition, limitations in existing models, proposed model's enhancements, limitations observed
Conclusion	Proposal for a model using machine learning, image preprocessing, and audio conversion for sign language recognition

3) Real-Time Detection and Classification of Facial Emotion:

Title	Real-Time Detection and Classification of Facial Emotions
Authors	Teerapong Winyangkun, Noparut Vanitchanant, Benjamas Panyangam, Varin Chouvatut
Algorithms Used	Convolutional Neural Networks (CNN), Deep Learning (DL), Facial Emotion Recognition (FER), Haar Cascade, VGG-Face Model, Bayesian Network, Histogram Equalization, Motion Detection using Background Subtraction
Datasets Used	Fer2013 dataset consisting of approximately 30,000 48×48 pixel grayscale images of faces
Abstract	The paper explores the use of DL algorithms, particularly CNN and FER, for facial emotion recognition, achieving 97% accuracy in seven-class recognition. It incorporates techniques like histogram equalization and background subtraction to improve efficiency.
Methodology	The methodology involves the utilization of DL networks, Deepface package, VGG-Face Model, Haar Cascade for facial detection, and techniques like histogram equalization and background subtraction to enhance image quality and emotion recognition.
Experimentation Results	The experimental results show initial accuracy above 90% for all emotions. After applying image-processing techniques, precision and recall measures improved, achieving over 85% F1-score for all classes.
Results and Discussions	Results demonstrate significantly enhanced classification performance after implementing background subtraction and histogram equalization. The model achieved higher accuracy, precision, recall, and F1-score across all emotional categories.
Conclusion	The proposed model successfully detects facial regions in real-time and achieves high accuracy in recognizing seven facial emotions. Additional image-processing techniques further improve the classification performance.
References	A comprehensive list of references used in the research paper for related work, methods, and techniques.

4) An Efficient Real-Time Indian Sign Language (ISL) Detection using Deep Learning :

Title	An Efficient Real-Time Indian Sign Language (ISL) Detection using Deep Learning
Authors	A.Siva SankarReddy, B.Surya, B.V.Prudhvi, N.V. Suresh Krishna, P. Neeraj, V. Hima Deepthi
Algorithms Used	CNN, Media Pipe, OpenCV
Datasets Used	Official data set for Indian Sign Language (ISLRTC)
Abstract	Indian Sign Language detection has become crucially important due to the growing need to bridge the communication gap between the hearing and the non-hearing people in India. The proposed system utilizes deep CNNs, Media Pipe, and OpenCV in real-time to detect and classify ISL gestures, enhancing communication between deaf and hearing users.
Methodology	The methodology involves hand detection and tracking, gesture segmentation, feature extraction using CNN-based techniques, and a user-friendly interface.
Experimentation Results	The experimental results showcase high accuracy and efficiency in detecting and interpreting ISL gestures in real-time video streams or images.
Results and Discussions	The results and discussions section highlights the system's performance, comparing algorithms, and discussing the potential for real-world deployment to facilitate communication between deaf and hearing individuals.
Conclusion	The research concludes that the proposed system using CNNs, Media Pipe, and OpenCV holds promise in enhancing communication and inclusion for deaf individuals.
References	The references section contains numerous sources and studies related to ISL recognition systems, including SVM, CNN, and various techniques applied for gesture recognition.

5) Real Time Conversion of American Sign Language to text with Emotion using Machine Learning

Section	Details
Title	Real Time Conversion of American Sign Language to Text with Emotion using Machine Learning
Authors	Aryan Jamwal, T. Vijayakumar, Vasukidevi G, L. Chandra Sekhar Reddy, Dr TYJ Naga Malleswari, Amara S A L G Gopala Gupta
Algorithms Used	CNN, SVM, HAAR, Kmeans, LSTM
Datasets Used	Self-created dataset, ASL dataset, FER 2013, RAfD[22]
Abstract	Focuses on recognizing different American Sign Language hand signs and associated emotions in real-time, converting them into text; aims to bridge communication gaps between sign and non-sign language users.
Methodology	Includes hand mapping, facial emotion recognition, algorithm and module implementation (SVM, CNN, MediaPipe), various model training and testing methods for real-time detection and conversion.
Experimentation Results	Achieved model accuracy around 80%, real-time conversion success, challenges addressed by different algorithms; incorporates dynamic hand sign and emotion recognition.
Results and Discussions	Discusses model evaluation, precision, recall, and heatmap of hand signs; explores achieved accuracy and existing challenges; emphasizes the achievement of real-time conversion and addressing communication barriers.
Conclusion	Successful achievement of real-time sign language conversion with emotion recognition, closing the communication gap between sign language users and non-signers.
References	Cites various academic papers and resources (references [1] to [24] and more) contributing to the field of sign language recognition and emotion detection through machine learning.

3.3 Outcome of the survey

The results of the survey, based on the survey are as follows:

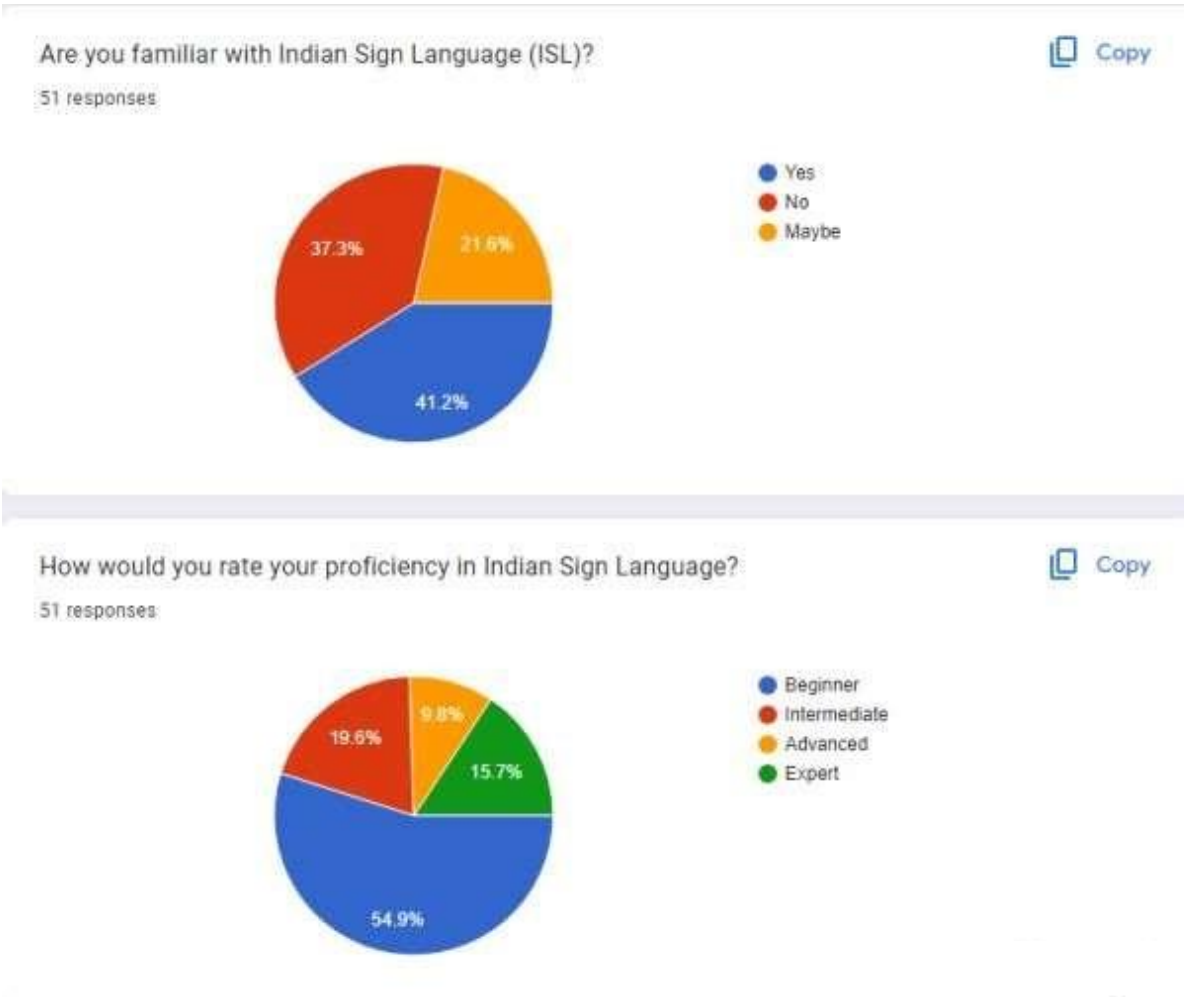


Fig 1. Survey results

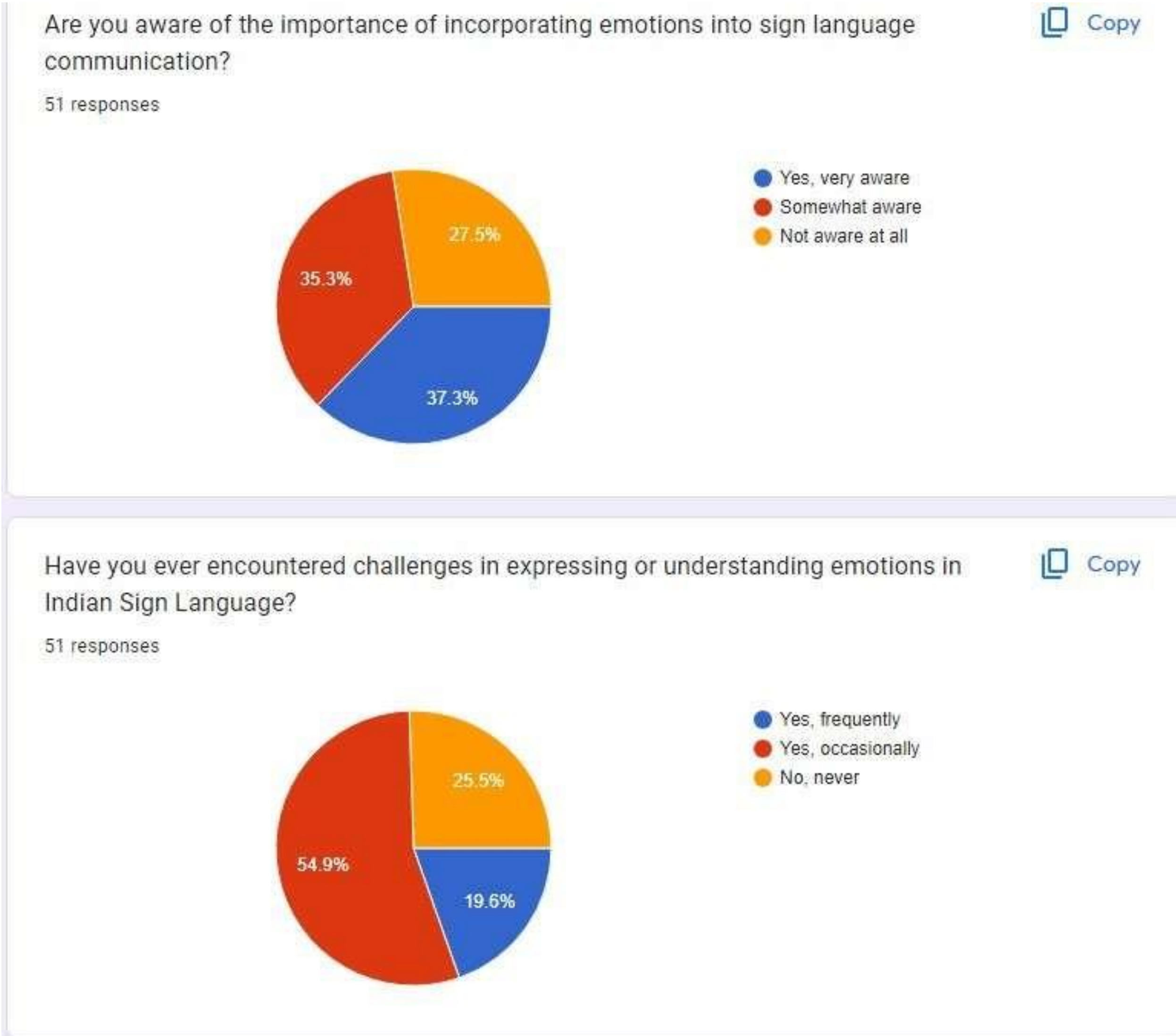


Fig 2. Survey results

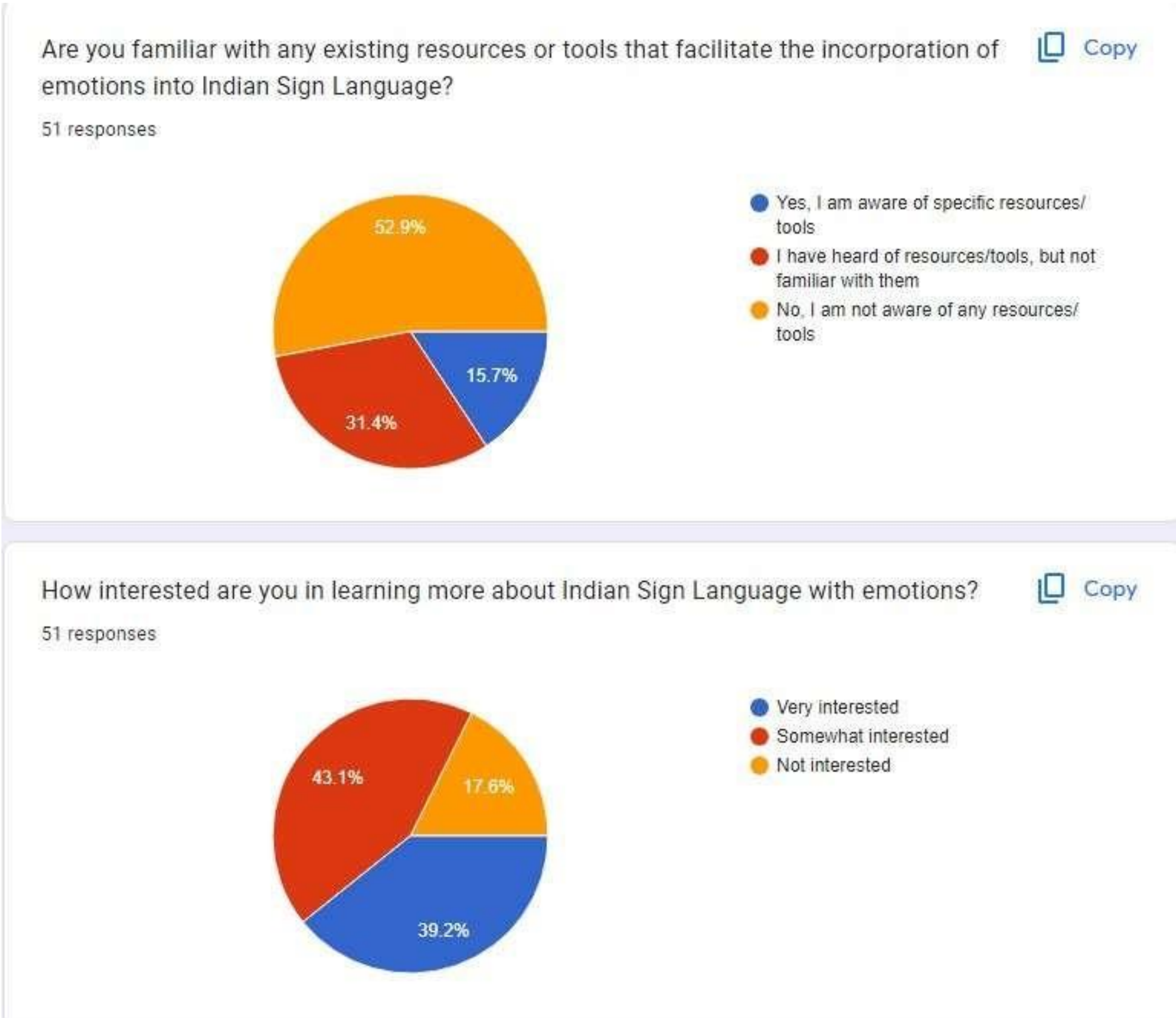


Fig 3. Survey results

Fig 4. Survey results

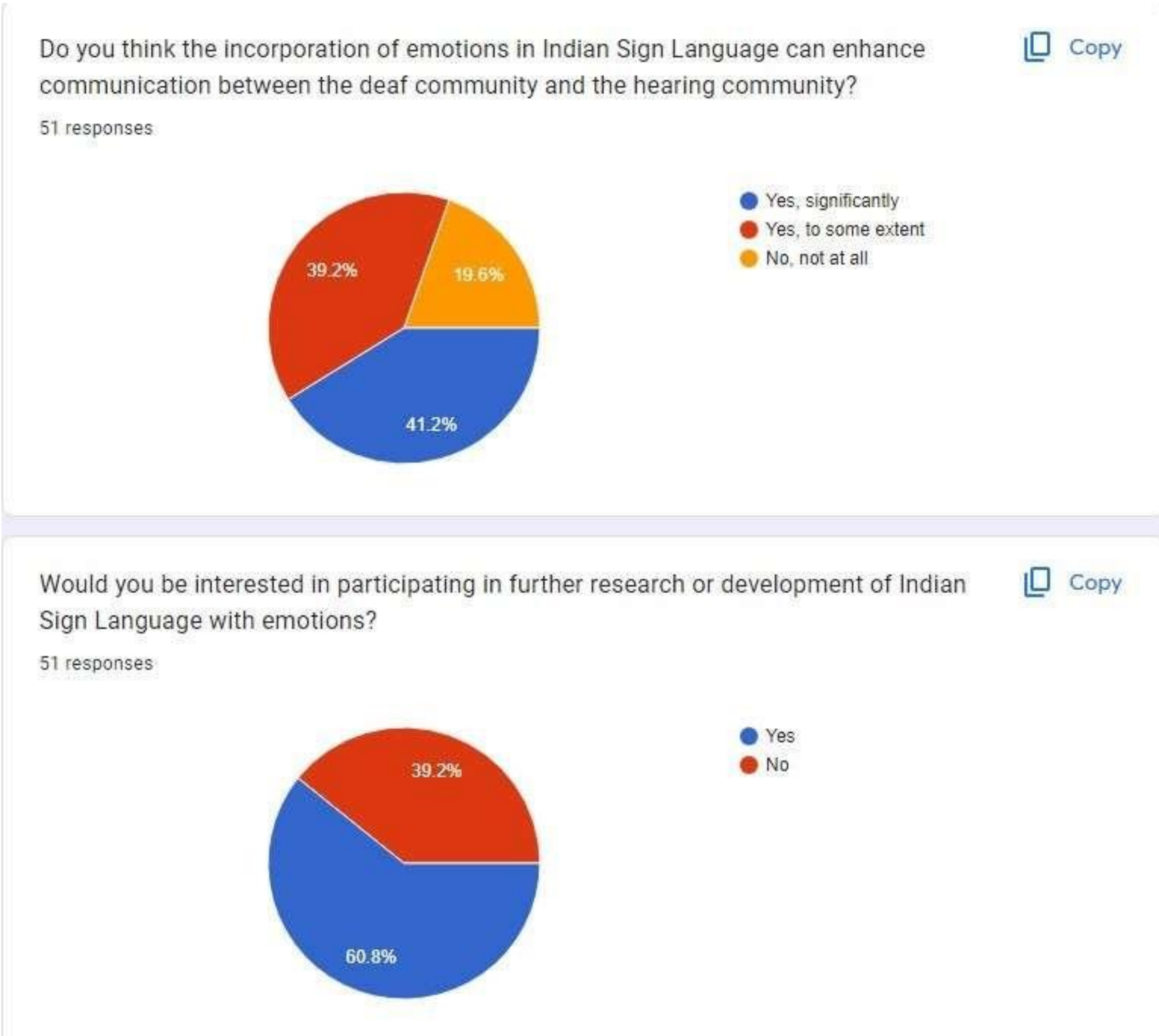


Fig 5. Survey results

The outcomes and conclusions drawn from the survey on Indian Sign Language (ISL) with emotions could be multifaceted based on thecollected data. Here's the conclusions that could be derived from the survey data.

Survey Conclusions:

1. **Awareness and Proficiency:**
   A significant portion of the respondents demonstrated familiarity with Indian Sign Language (ISL), with varying levels of proficiency ranging from beginners to experts. However, a notable percentage indicated they were not aware of the importance of incorporating emotions into ISL communication.

2. **Challenges in Emotional Expression:**
   A substantial number of participants reported encountering challenges in expressing or understanding emotions in ISL. This suggests a need for further resources or tools to facilitate the incorporation of emotions into ISL effectively.

3. **Interest in Learning:**
   A majority of respondents expressed keen interest in learning more about Indian Sign Language with emotions, indicating a potential demand for educational materials or programs focused on this aspect.

4. **Perceived Benefits:**
   Respondents identified several benefits of incorporating emotions into ISL, including improved expressiveness, better emotional understanding, enhanced communication effectiveness, and increased inclusivity.

5. **Potential Applications:**
   The survey highlighted various potential applications or use caseswhere ISL with emotions could be beneficial, such as in education, interpersonal communication, arts, entertainment, and healthcare or therapy.

6. **Community Integration:**
   A significant majority believed that the incorporation of emotions in ISL could significantly enhance communication between the deaf and hearing communities.

7. **Interest in Further Research:**
   A portion of participants expressed interest in participating in further research or development related to Indian Sign Language with emotions, indicating potential support for future initiatives in this domain.

Recommendations:

1. Educational Initiatives:

The Develop educational resources or programs aimed at incorporating emotions into ISL to address the reported challenges and meet the interest expressed by participants.

2. Tool Development:

Invest in the creation of tools or resources that aid in enhancing emotional expression and understanding within ISL communication.

3. Community Engagement:

Foster initiatives that promote collaboration between the deaf and hearing communities to enhance communication inclusivity and understanding.

4. Further Research:

Encourage and facilitate ongoing research and development efforts in the field of Indian Sign Language with emotions, leveraging the interest expressed by participants.

These conclusions and recommendations would guide future actions, strategies, and initiativesaimed at improving the incorporation of emotions into Indian Sign Language communication based on the insights gained from the survey data.

4. Problem Formulation

4.1. Problem Formulation

The project aims to develop a real-time sign language recognition system with emotion analysis, addressing the challenges faced by hearing-impaired individuals in expressing both signs and emotions simultaneously.

4.2. Product Objectives

The system will recognize ISL signs and emotions in real-time, providing a simultaneous conversion to text and emotion display.

4.3. Novelty

The novelty of this work lies in the simultaneous recognition of sign language and emotions. The paper proposes an improved CNN model using MediaPipe and emphasizes the need for real-time communication.

4.4. Scope of the Project

The project's scope includes applications in medical, educational, legal, and training domains, contributing to the independence of hearing-impaired individuals.

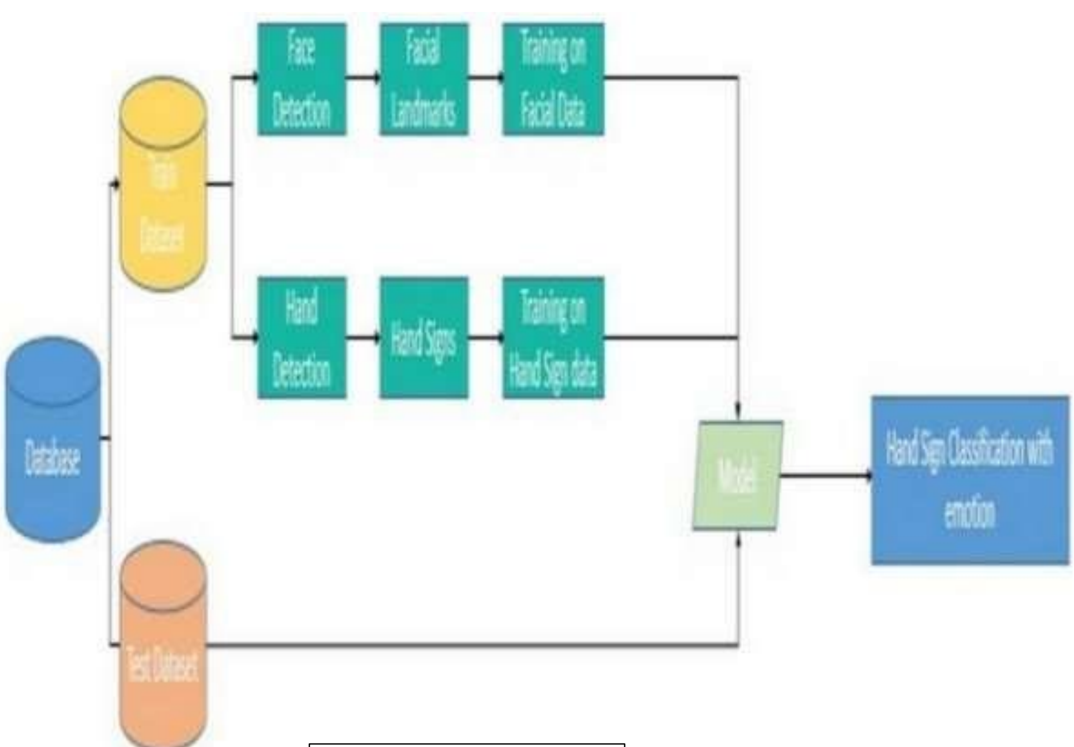


Fig 6. Project Scope

5. Proposed Design

5.1. Proposed Model

Part A) Sign Language Detection

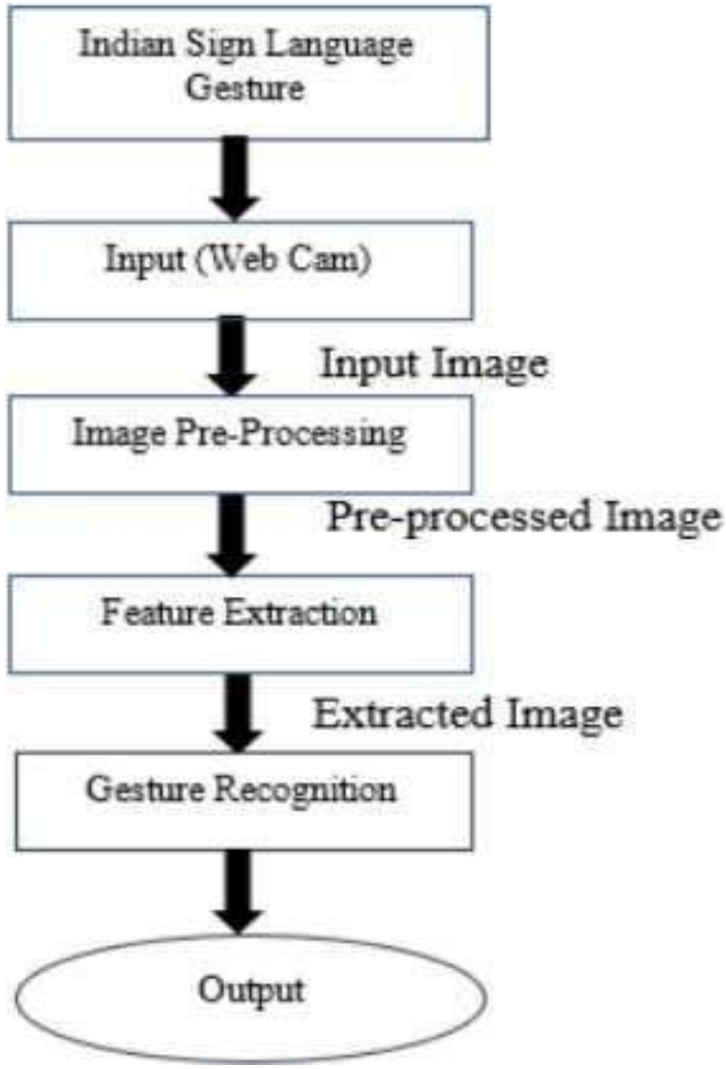


Fig 7. Sign Language Model

Part B) Emotion Detection

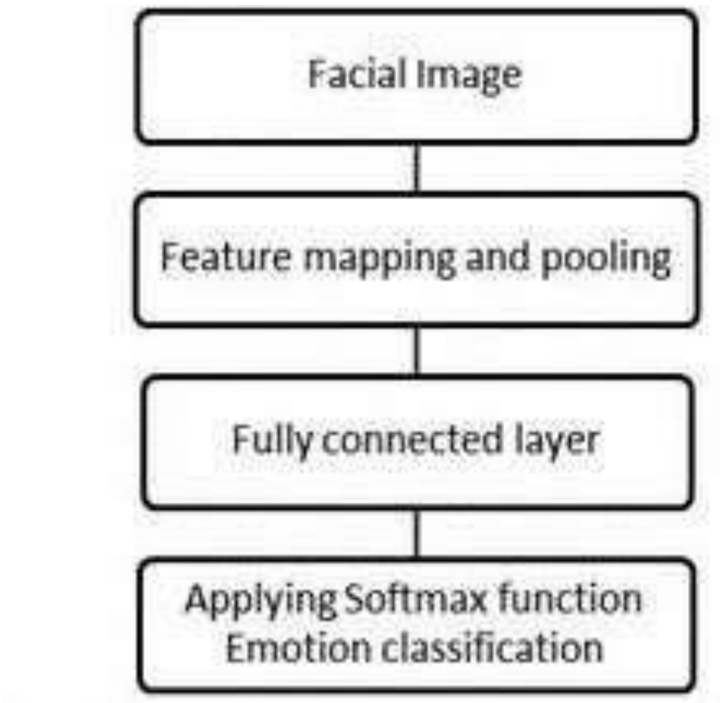


Fig 8. Emotion Detection Model

5.2. Database Design

The system incorporates a self-generated Indian Sign Language dataset and Pre -existing facial emotion datasets for training and testing.

5.3. Use Cases

The system's versatility allows applications in medical consultations, educational settings, legal proceedings, and training sessions.

6. Implementation

6.1. GUI Design

Part A) Sign Language Detection

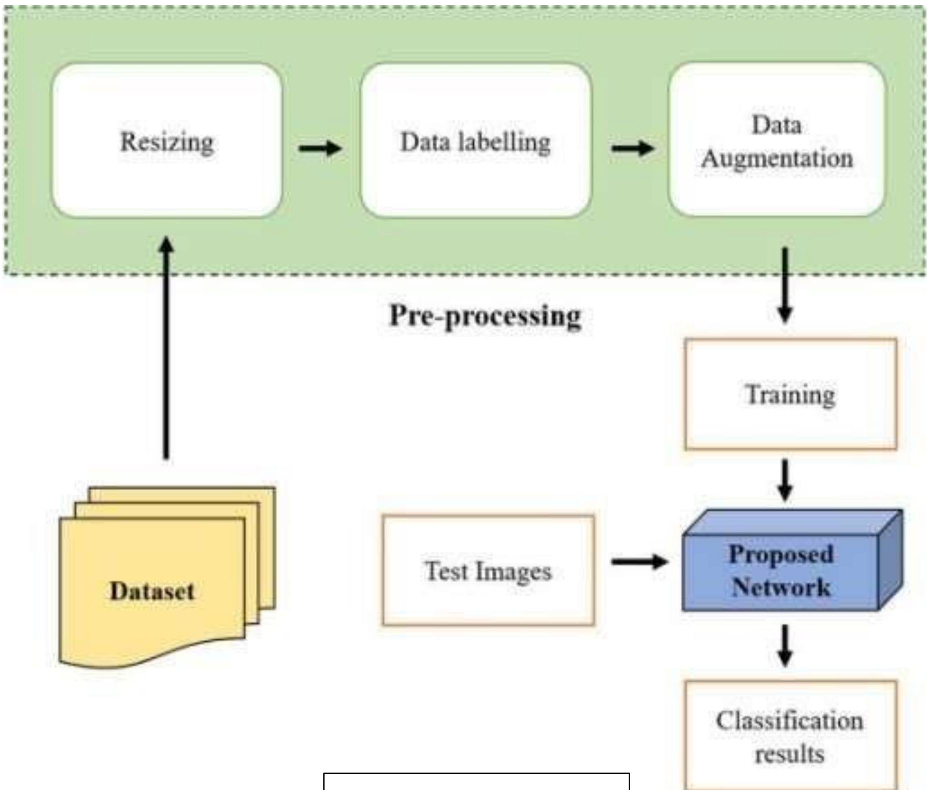


Fig 9. GUI Sign Language

1. Import dependencies: Install and import necessary libraries like MediaPipe, TensorFlow, and Keras.

2. MediaPipe Holistic setup:
    - Initialize MediaPipe Holistic for body, face and hand landmark detection.
    - Define a function to extract keypoints from detected landmarks.
    - Define a function to draw landmarks for visualization (optional).

3. Data collection:
    - Define a function to capture video frames from webcam or file.
    - Pre-process data:
        - Scale keypoints to appropriate range.
        - Convert data to desired format (e.g., time series).
        - Store pre-processed data and labels for training and testing.
4. LSTM model training:
    - Define LSTM model architecture with suitable layers and parameters.
    - Train the model using pre-processed data and labels.
    - Track and monitor training metrics like accuracy and loss.
    - Save the trained model weights.
5. Prediction:
    - Load the saved model weights.
    - Extract keypoints from new video frames.
    - Feed pre-processed keypoints to the trained LSTM model.
    - Predict the sign language action based on the model's output.
6. Evaluation:
    - Test the model performance on a separate testing dataset.
    - Calculate accuracy, precision, recall, and other relevant metrics.
    - Analyze results and identify areas for improvement.
7. Real-time application:
    - Implement continuous video capture and processing loop.
    - Display real-time sign language recognition results.
    - Optionally integrate with other applications or interfaces.

Part B) Emotion Detection

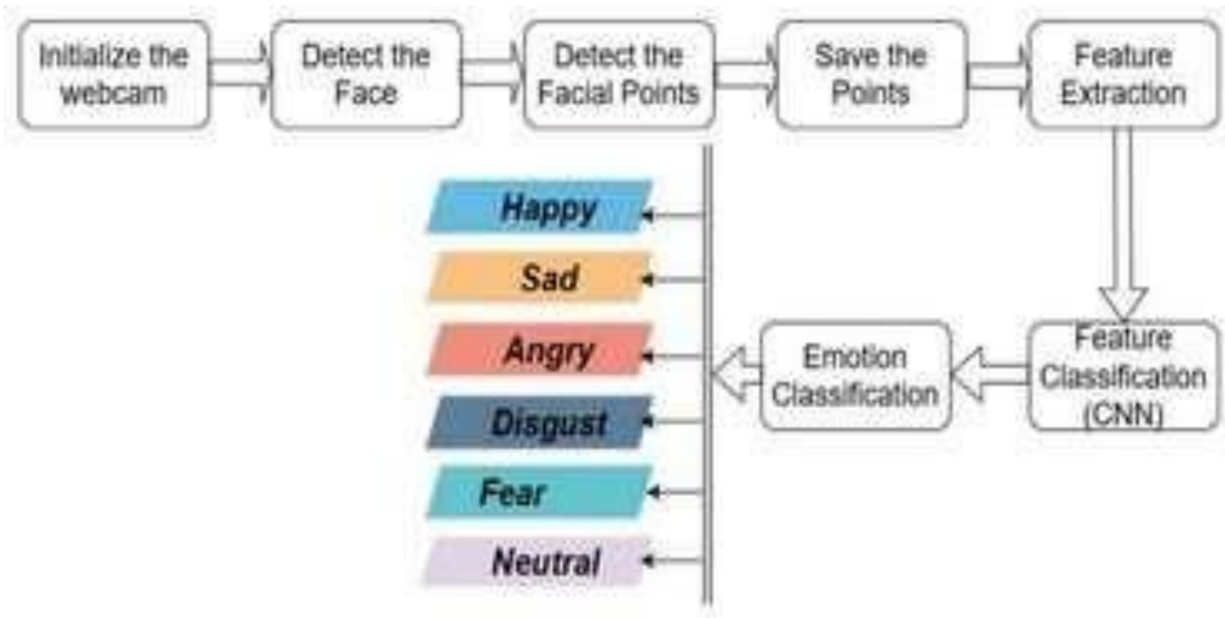


Fig 10. GUI

The implementation of the emotion detection module involves the following key steps:

1. **CNN Model Implementation:**
   - *Dataset Preparation:* Training and testing datasets are carefully curated to include diverse facial expressions, ensuring the model's robustness.
   - *Hyperparameter Tuning:* Experimentation focuses on optimizing CNN model hyperparameters such as filter sizes, convolutional layers, and activation functions to achieve the best possible accuracy.

2. **Real-Time Feasibility:**

- Optimization: Implementing optimizations to ensure that the CNN model operate in real-time, meeting the system's responsiveness requirements.
- Integration with Overall System: The emotion detection module is seamlessly integrated into the broader sign language recognition system, ensuring cohesive functionality.

7. Experimentation & Results

Part A) Sign Language Detection

7.1. Data preprocessing:

```
57]: def draw_styled_landmarks(image, results):
         # Draw face connections
         mp_drawing.draw_landmarks(image, results.face_landmarks, mp_holistic.FACEMESH_CONTOURS,
                                   mp_drawing.DrawingSpec(color=(80,110,10), thickness=1, circle_radius=1),
                                   mp_drawing.DrawingSpec(color=(80,256,121), thickness=1, circle_radius=1)
                                   )
         # Draw pose connections
         mp_drawing.draw_landmarks(image, results.pose_landmarks, mp_holistic.POSE_CONNECTIONS,
                                   mp_drawing.DrawingSpec(color=(80,22,10), thickness=2, circle_radius=4),
                                   mp_drawing.DrawingSpec(color=(80,44,121), thickness=2, circle_radius=2)
                                   )
         # Draw left hand connections
         mp_drawing.draw_landmarks(image, results.left_hand_landmarks, mp_holistic.HAND_CONNECTIONS,
                                   mp_drawing.DrawingSpec(color=(121,22,76), thickness=2, circle_radius=4),
                                   mp_drawing.DrawingSpec(color=(121,44,250), thickness=2, circle_radius=2)
                                   )
         # Draw right hand connections
         mp_drawing.draw_landmarks(image, results.right_hand_landmarks, mp_holistic.HAND_CONNECTIONS,
                                   mp_drawing.DrawingSpec(color=(245,117,66), thickness=2, circle_radius=4),
                                   mp_drawing.DrawingSpec(color=(245,66,230), thickness=2, circle_radius=2)
                                   )
```

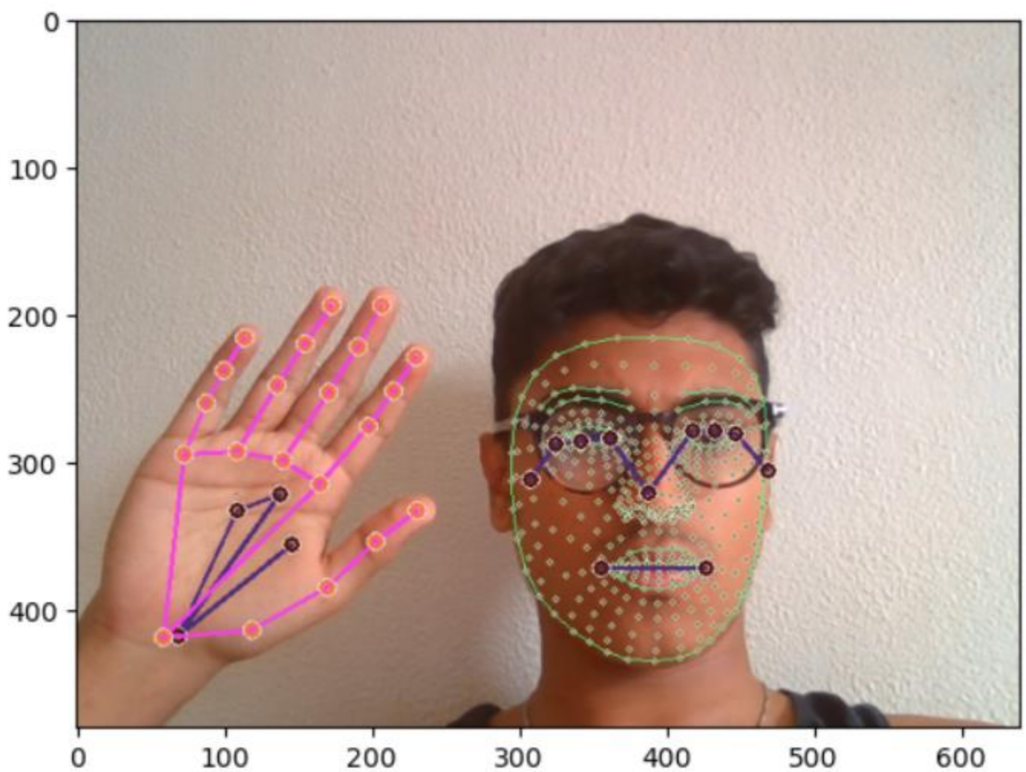


Fig 11. Data Visualization

7.2. Model Fitting:

```
In [101]: actions[np.argmax(res)]
Out[101]: 'tea'

In [70]: model.compile(optimizer='Adam', loss='categorical_crossentropy', metrics=['categorical_accuracy'])

In [71]: model.fit(X_train, y_train, epochs=500, callbacks=[tb_callback])
3/3 [==============================] - 0s 100ms/step - loss: 0.3621 - categorical_accuracy: 0.8353
Epoch 131/500
3/3 [==============================] - 0s 95ms/step - loss: 0.2973 - categorical_accuracy: 0.8588
Epoch 132/500
3/3 [==============================] - 0s 99ms/step - loss: 0.3848 - categorical_accuracy: 0.8588
Epoch 133/500
3/3 [==============================] - 0s 98ms/step - loss: 0.2997 - categorical_accuracy: 0.8706
Epoch 134/500
3/3 [==============================] - 0s 70ms/step - loss: 0.3376 - categorical_accuracy: 0.8118
Epoch 135/500
3/3 [==============================] - 0s 92ms/step - loss: 0.2487 - categorical_accuracy: 0.9176
Epoch 136/500
3/3 [==============================] - 0s 90ms/step - loss: 0.2281 - categorical_accuracy: 0.8941
Epoch 137/500
3/3 [==============================] - 0s 86ms/step - loss: 0.1717 - categorical_accuracy: 0.9176
Epoch 138/500
3/3 [==============================] - 0s 110ms/step - loss: 0.2003 - categorical_accuracy: 0.9176
Epoch 139/500
1/3 [=========>....................] - ETA: 0s - loss: 0.2777 - categorical_accuracy: 0.9062
```

Fig 12. Model Fitting

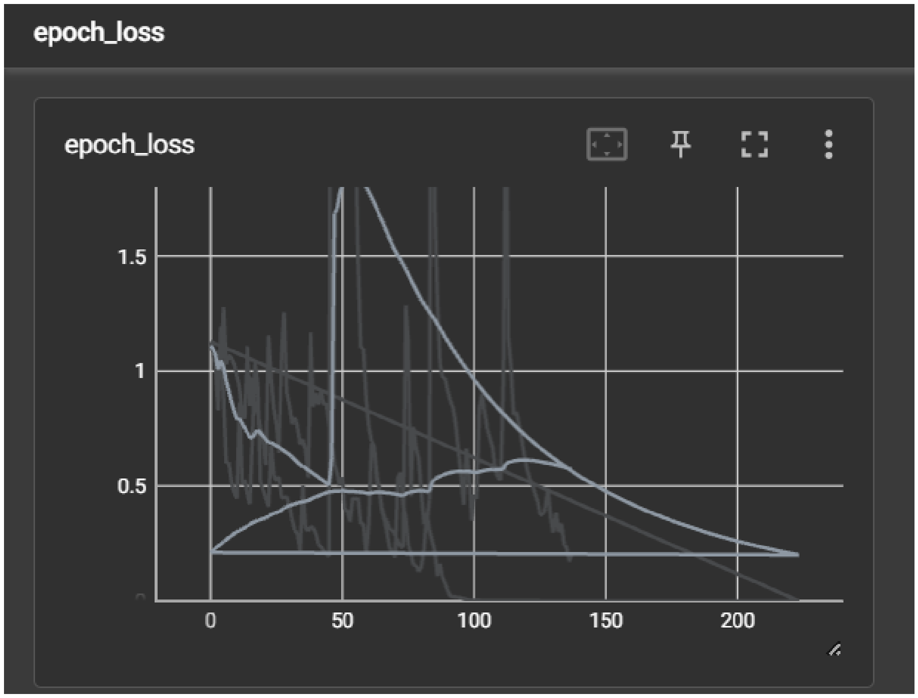
7.3. Predictions:

Evaluation using confusion matrix

```
In [77]: from sklearn.metrics import multilabel_confusion_matrix, accuracy_score

In [78]: yhat = model.predict(X_train)
3/3 [==============================] - 0s 35ms/step

In [79]: ytrue = np.argmax(y_train, axis=1).tolist()
         yhat = np.argmax(yhat, axis=1).tolist()

In [80]: multilabel_confusion_matrix(ytrue, yhat)
Out[80]: array([[[53,  3],
                 [ 2, 27]],

                [[54,  2],
                 [ 3, 26]],

                [[58,  0],
                 [ 0, 27]]], dtype=int64)

In [81]: accuracy_score(ytrue, yhat)
Out[81]: 0.9411764705882353
```

Fig 13. Accuracy Prediction

7.4. Model evaluation and results:



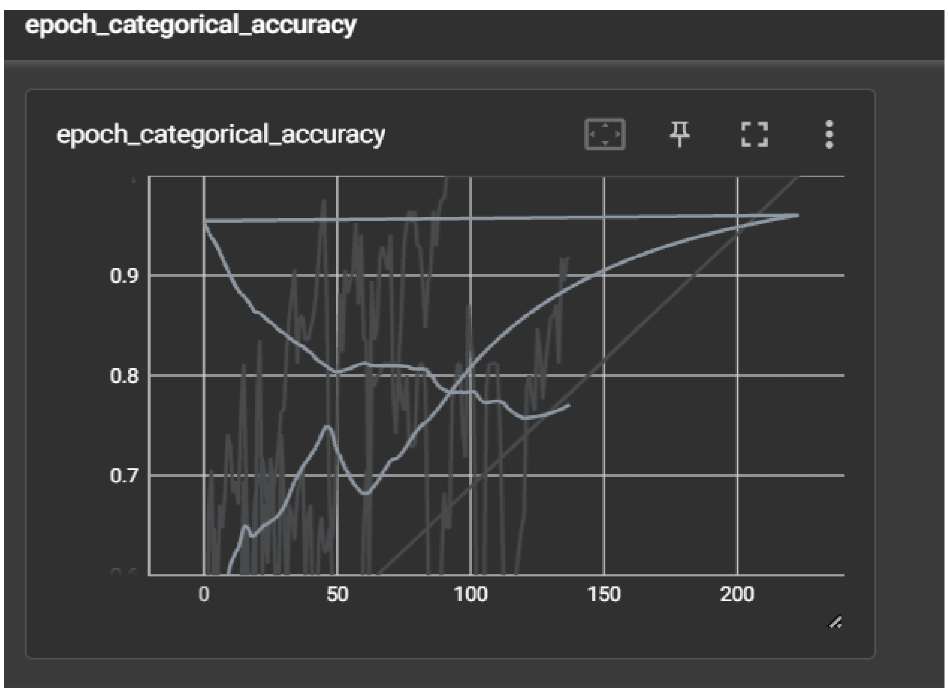


Fig 14. Result Evaluation

Real-time working:

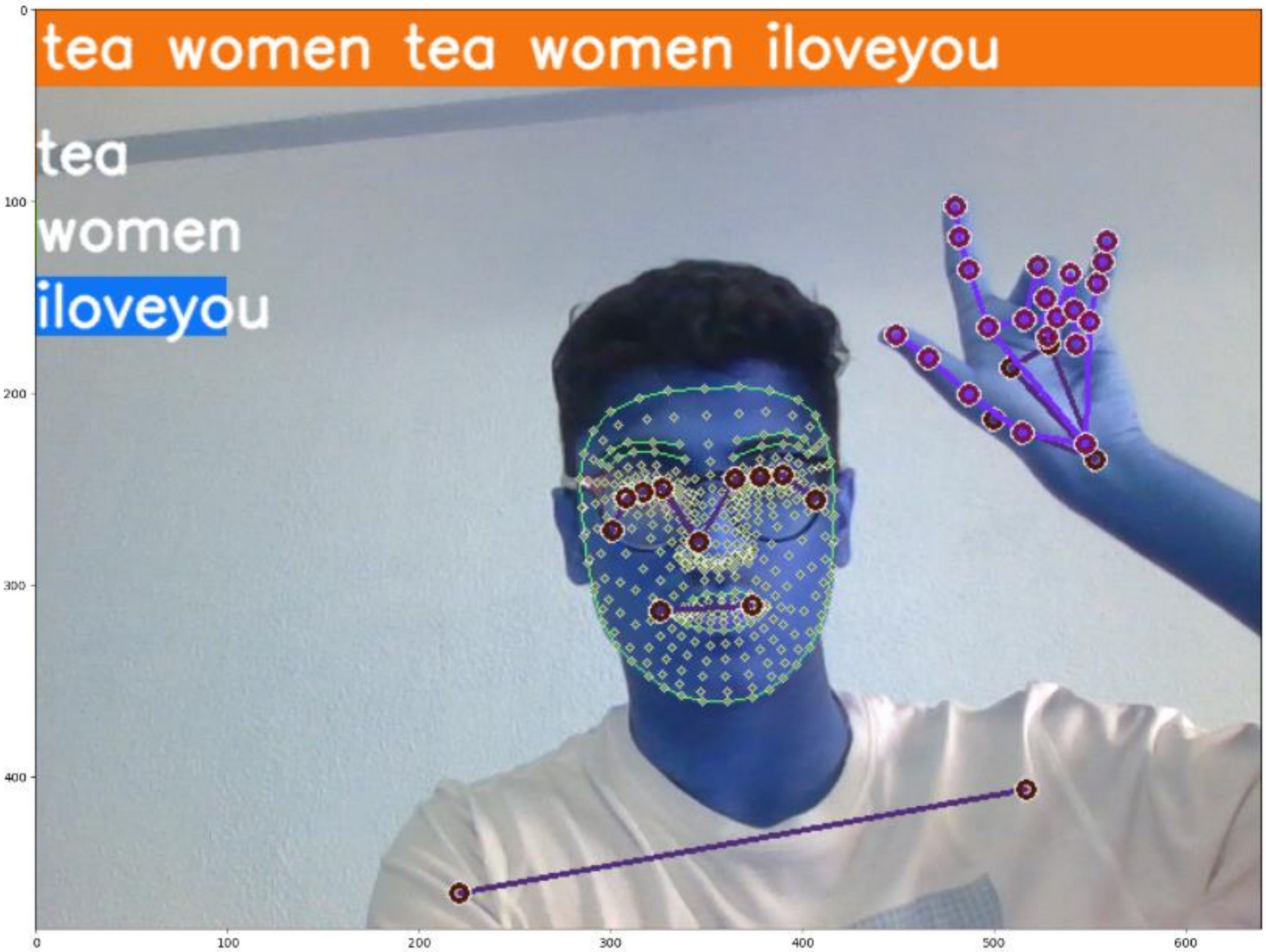


Fig 15. Real Time Output

Part B) Emotion Detection

7.1. Datasets / Tables

FER -2013 Dataset

The data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The training set consists of 28,709 examples and the public test set consists of 3,589 examples.

7.2. Test Cases

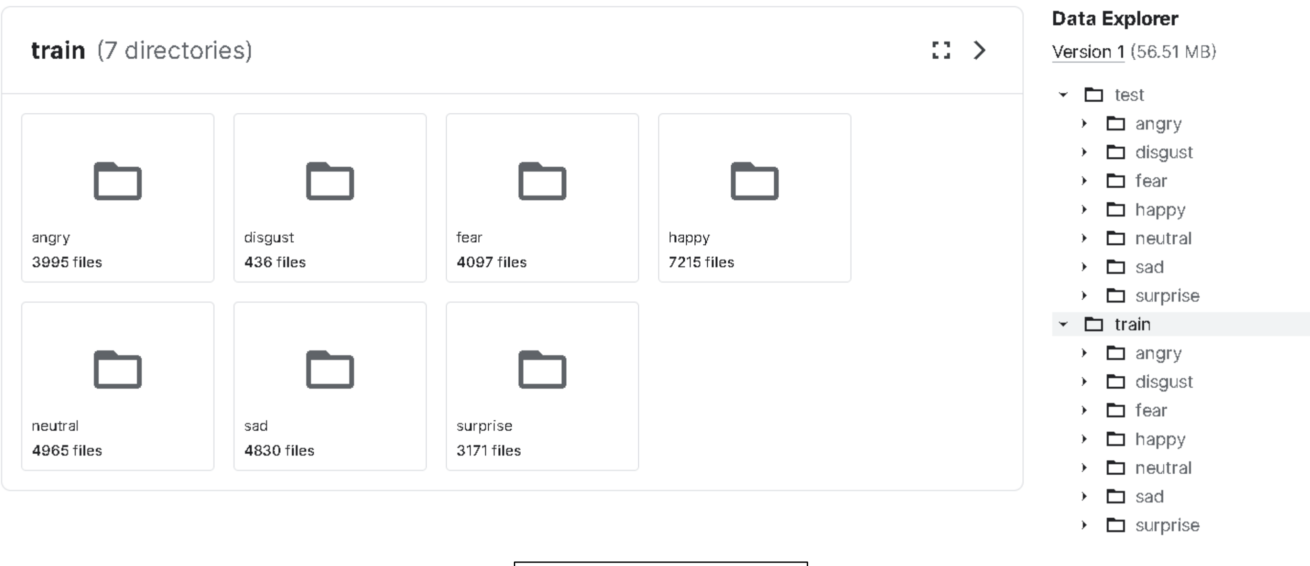


Fig 16. Test Cases

7.3. Parameter Tuning Experiments

```
reduce_learningrate = ReduceLROnPlateau(monitor='val_loss',
                                        factor=0.2,
                                        patience=3,
                                        verbose=1,
                                        min_delta=0.0001)

callbacks_list = [checkpoint,reduce_learningrate]
```

Fig 17. Parameter Tuning

7.4. Results

Accuracy -79.22% on Test

-65.17% on Validation

```
Epoch 46: ReduceLROnPlateau reducing learning rate to 2.0480002416167767e-11.
28/28 [==============================] - 14s 502ms/step - loss: 0.5664 - accuracy: 0.7879 - val_loss: 1.0363 - val_accuracy: 0.6519 - lr: 1.0240e-10
Epoch 47/48
28/28 [==============================] - ETA: 0s - loss: 0.5735 - accuracy: 0.7875
Epoch 47: val_accuracy did not improve from 0.65214
28/28 [==============================] - 15s 521ms/step - loss: 0.5735 - accuracy: 0.7875 - val_loss: 1.0362 - val_accuracy: 0.6519 - lr: 2.0480e-11
Epoch 48/48
28/28 [==============================] - ETA: 0s - loss: 0.5675 - accuracy: 0.7922
Epoch 48: val_accuracy did not improve from 0.65214
28/28 [==============================] - 14s 500ms/step - loss: 0.5675 - accuracy: 0.7922 - val_loss: 1.0362 - val_accuracy: 0.6517 - lr: 2.0480e-11
```

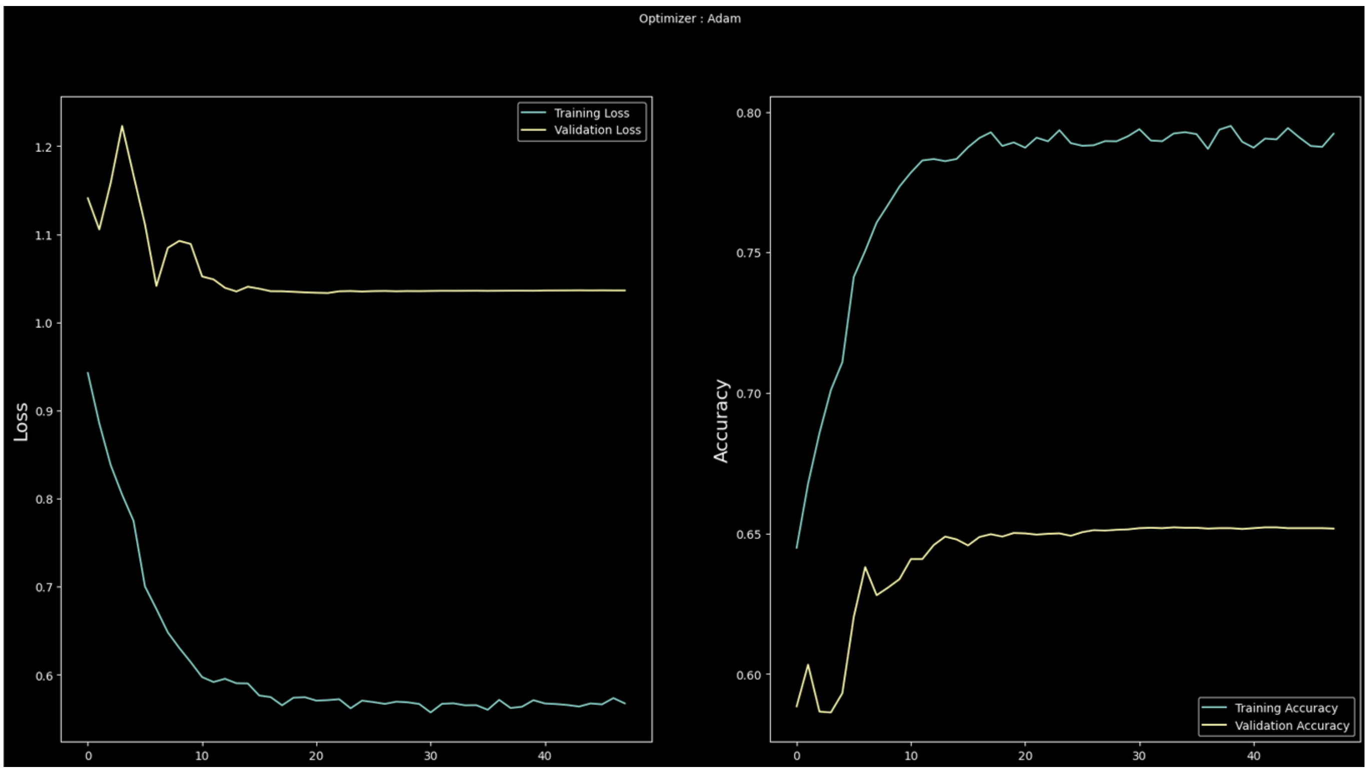


Fig 18. Results

8. Conclusion

The comprehensive real-time sign language recognition system, coupled with emotion analysis, brings forth a revolutionary solution for inclusive communication. By integrating LSTM for sign language and CNN for emotion detection, the system caters to the varied needs of the hearing-impaired community, addressing the limitations of existing systems.

9. References / Bibliography

[1] L. Ku, W. Su, P. Yu and S. Wei, ”A real-time portable sign language translation system,” 2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS), Fort Collins, CO, 2015, pp. 1-4, doi: 10.1109/MWSCAS.2015.7282137.

[2] Mahesh Kumar NB. ”Conversion of Sign Language into Text”. International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9, (2018).

[3] Tiku, Kohsheen & Maloo, Jayshree & Ramesh, Aishwarya & R, Indra. (2020). Real-time Conversion of Sign Language to Text and Speech. 346-351. 10.1109/ICIRCA48905.2020.9182877.

[4] Riyantoko, P.A., Sugiarto, & Hindrayani, K.M. (2021). Facial Emotion Detection Using Haar-Cascade Classifier and Convolutional Neural Networks. Journal of Physics: Conference Series, 1844.

[5] Hand sign recognition from depth images with multi-scale densityfeatures for deaf mute persons -Tniya Sahana , Soumi Paul ,Subhadip Basu, Ayatullah FarukMollah, ‘Procedia ComputerScience”, Elsevier, Volume 167, 2020, Pages 2043-2050.

[6] Wafa Mellouk, Wahida Handouzi, Facial emotion recognition using deep learning: review and insights, Procedia Computer Science, Volume 175, 2020,Pages 689-694.

[7] E. Abraham, A. Nayak and A. Iqbal, ”Real-Time Translation of Indian Sign Language using LSTM,” 2019 Global Conference for Advancementin Technology (GCAT), BANGALURU, India, 2019, pp. 1-5, doi: 10.1109/GCAT47503.2019.8978343.

[8] Bheda, Vivek and Dianna Radpour. Using Deep Convolutional Networks for Gesture Recognition in American Sign Language. ArXiv abs/1710.06836 (2017): n. pag.

[9] Tolentino, Lean Karlo S. et al. Static Sign Language Recognition Using Deep Learning. International Journal of Machine Learning and Computing 9 (2019): 821-827.

[10] Bantupalli, K, Xie, Y. (2018). “American Sign Language Recognition using Deep Learning and Computer Vision”. IEEE International Conference on Big Data (Big Data), (2018).

[11] S. Shahriar et al., ”Real-Time American Sign Language Recognition Using Skin Segmentation and Image Category Classification with Convolutional Neural Network and Deep Learning,” TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), 2018, pp. 1168-1171, doi: 10.1109/TENCON.2018.8650524

[12] Umang Patel and Aarti G. Ambedkar, “Moment Based Sign Language Recognition for Indian Language” 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA).

[13] M.V, Beena. “Automatic Sign Language Finger Spelling Using Convolution Neural Network : Analysis.” (2017).

[14] Hu Peng, “Application Research on Face Detection Technology based on Open CV in Mobile Augmented Reality”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 8, No. 2 (2015)