REAL-TIME SIGN LANGUAGE RECOGNITION SYSTEM

By

Shahana Akter
Roll: 1907029

&

Shrabanti Debnath
Roll: 1907094

Supervisor:

Md. Repon Islam
Lecturer
Dept. of Computer Science and Engineering
Khulna University of Engineering & Technology

Signature

Department of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna 9203, Bangladesh
November 2023

Acknowledgement

We would like to express our deepest gratitude to our supervisor, Md. Repon Islam, for his continuous guidance, support, and invaluable insights throughout the development of this project. His expertise and encouragement were instrumental in shaping our ideas and refining the sign language recognition system. We are also thankful for the constructive feedback and timely advice that significantly contributed to the success of this project.

Last, but by no means least, we express our deepest gratitude to God for the talents and abilities we were given that made it possible to undertake this project. We acknowledge the blessings and guidance from God, recognizing his role as the ultimate source of strength and inspiration throughout this journey.

Authors

Shahana Akter

Roll: 1907029

&

Shrabanti Debnath

Roll: 1907094

Abstract

In addition to filling a crucial gap in communication accessibility for the Bengali-speaking deaf community, our Real Time Bengali Sign Language Recognition System is a pioneering effort in a relatively unexplored domain. The project's significance lies in the novelty it brings to the field, considering the limited prior work on Bengali sign language recognition. With the help of the advanced capabilities of MediaPipe and LSTM architecture, our system ensures accurate and almost real-time interpretation of complex hand gestures, enriching communication for the hearing-impaired. Furthermore, the inclusion of instructional videos within the website underscores our commitment to empowering users with the ability to learn and understand Bengali sign language. This educational aspect aligns with our broader objective of creating an inclusive platform that not only recognizes but also facilitates the learning and adoption of sign language. By embracing innovation and addressing the unique linguistic characteristics of Bengali sign language, our project contributes to the global pursuit of fostering inclusivity and communication equity for diverse communities.

Contents

List of Figures

1 Introduction

1.1 Background

The field of sign language recognition has made great steps in the area of communication accessibility, but little is known about the unique peculiarities of Bengali sign language. Acknowledging the lack of complete solutions available for this language setting, our initiative is taking a bold step forward. We purposefully exclude direct acquisition from existing sources, in contrast to traditional sign language recognition systems, to ensure uniqueness in resolving the particular difficulties presented by Bengali sign language.

The goal of this project is to reduce the communication gap that often exists between the deaf people who speak Bengali and traditional language processing tools. By focusing on the specific linguistic complexities of Bengali sign language, our project not only acknowledges the underexplored territory but also takes a significant step towards empowering the hearing-impaired population with a robust and customized communication tool. Our innovative project was motivated by the lack of specific Bengali sign language identification tools.

1.2 Objectives

Bengali Sign Language, a mode of communication crucial for the hearing-impaired in Bengali-speaking regions, lacks comprehensive technological support. Existing systems fall short in addressing the linguistic intricacies of Bengali sign language, necessitating a specialized approach. By employing advanced technologies, specifically MediaPipe and LSTM architecture, our initiative aims to bridge this communication gap effectively. We imagine a future where the hearing-impaired in Bengal can seamlessly express themselves using their native sign language, facilitated by an innovative recognition system. This venture is not just about technology; it is a commitment to inclusivity, ensuring that every individual, irrespective of hearing abilities, can engage fully in the rich linguistic field of Bengali sign language. The project aims to:

- Develop a robust real-time Bengali sign language recognition System with the help of MediaPipe and LSTM architecture.

- Embed the system within a user-friendly website, ensuring widespread accessibility.
- Enhance real-time gesture analysis accuracy, providing an effective communication tool.
- Address the scarcity of Bengali sign language recognition solutions by offering a novel and innovative system.

1.3 Scope

Our project's broad focus is on developing a powerful system for recognizing Bengali sign language that can be easily integrated into online platforms. Beyond recognition, we also build an intuitive user interface with state-of-the-art technologies such as MediaPipe and LSTM architecture. In order to enable successful communication for the deaf community in Bengali-speaking regions, this technology promises to accurately translate complex motions in Bengali sign language in real-time. Our project aims to enable people with hearing impairments to express their ideas and thoughts using their native sign language by creating an inclusive platform. The impact extends beyond technology to include the promotion of a more unified and empathetic society in which language variety is valued and accessibility is seen as a basic human right.

1.4 Unfamiliarity of the problem

The identification of Bengali sign language is yet mainly unexplored, which poses a special problem for the field of communication accessibility. It is remarkable how few specialized works there are in this particular linguistic environment. In order to maintain creativeness, our approach purposefully departs from traditional sign language recognition algorithms by avoiding direct acquisition from pre-existing sources or courses. This divergence highlights the necessity of a novel strategy to tackle the underdeveloped state of Bengali sign language technology. The problem is novel because there isn't any prior experience with dealing with the complexities of Bengali sign language. Our initiative breaks new ground by avoiding tried-and-true approaches and making use of the innovative linguistic environment in conjunction with the cutting-edge technologies of MediaPipe and LSTM architecture. This purposeful divergence from the norm opens the door for a revolutionary advancement in the field of sign language recognition, especially for Bengali language.

1.5 Project planning

In order to approach each project stage methodically, we used the Gantt chart, one of the strategic organizational tools that we use in our project planning. Using milestones for important deliveries and a project timeframe, the Gantt chart functions as a visual roadmap. This approach guarantees a methodical flow, from obtaining datasets and training models to developing web interfaces and carrying out exhaustive testing.

By focusing on the Gantt chart, we hope to offer a thorough perspective, encouraging productivity and skillfully resolving any potential roadblocks. Our objective is to successfully finish the project by observing to this well thought-out plan and finding a balance between dependability and innovation. Figure 1.1 illustrates the key milestones and stages in the project's timeline.

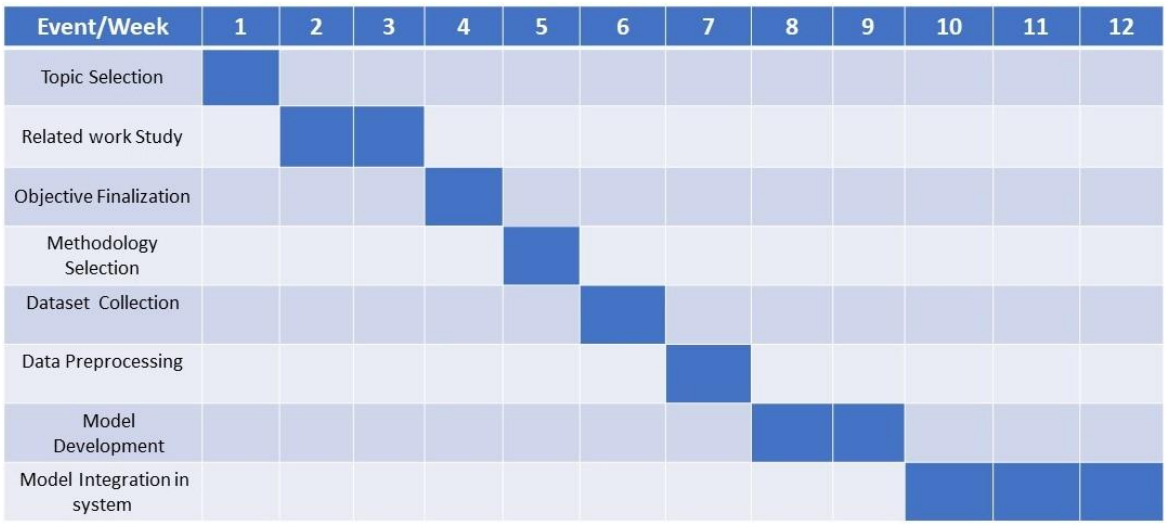


Figure 1.1.1: Project planning of Bengali Sign Language Recognition System

2 Related Works

2.1 Existing Solutions

The current landscape of sign language recognition solutions employs diverse approaches, prominently featuring computer vision and deep learning models such as VGG16, VGG19, InceptionV3, ResNet-50, and AlexNet [1]. Notable examples include [2], which proposes utilizing Mediapipe and RNN to create a Vietnamese Sign Language identification system nevertheless, the model has drawbacks, including imprecise hand movement detection, disorganized labeling, overfitting, and the requirement for better preprocessing to boost performance.

2.2 Limitations in existing solutions

Existing sign language recognition solutions encounter notable limitations such as:

- **Limited Sign Word Datasets:** Existing datasets predominantly concentrate on sign language alphabets, numbers, or characters, neglecting the need for datasets specific to common sign words used in daily communication.
- **Insufficient Variation:** Many datasets lack diversity in terms of hand gestures, background settings, and lighting conditions. This limitation can impact the robustness of sign language recognition systems in real-world scenarios.
- **Communication Barrier:** The Deaf and Dumb (D&D) community faces a significant communication barrier due to the scarcity of datasets focusing on sign words. This hinders the development of effective systems for automatic sign language recognition.
- **Lack of Recognition Systems:** While datasets are crucial, there is a gap in the availability of robust recognition systems for Bengali Sign Language. Existing solutions often fall short in addressing the unique challenges posed by sign words in Bengali Sign Language.

2.3 Discussion of overcoming limitations

Our project's comprehensive solutions to Bengali Sign Language Recognition challenges:

- **Diverse Data Consideration:** Our project incorporated variations in hand gestures, backgrounds, and lighting, enhancing the model's adaptability in real-world scenarios.
- **Recognition Systems:** Recognizing the gap in existing solutions, we developed advanced deep learning models specifically for Bengali Sign Language. Our system address the unique challenges posed by sign words, ensuring robust and accurate recognition.
- **Implementation through Python Script and Website:** To make our solution accessible, we built a user-friendly Python script and a website. The model is seamlessly integrated into these platforms, providing a practical and convenient system for Bengali Sign Language recognition.

By addressing both dataset limitations and the lack of robust recognition systems, and further implementing our solution through practical platforms, our project offers a comprehensive and accessible advancement in Bangladeshi Sign Language recognition.

3 System Design

In order to guarantee smooth operation and user accessibility, the Real-Time Bengali Sign Language Recognition System is methodically constructed, incorporating both hardware and software components.

3.1 Analysis of the system

The Bengali Sign Language Recognition System undergoes a multifaceted analysis to evaluate its performance, usability, and impact. The data flow diagram (DFD) of our system explaining the whole process is shown in Fig. 3.1.1.

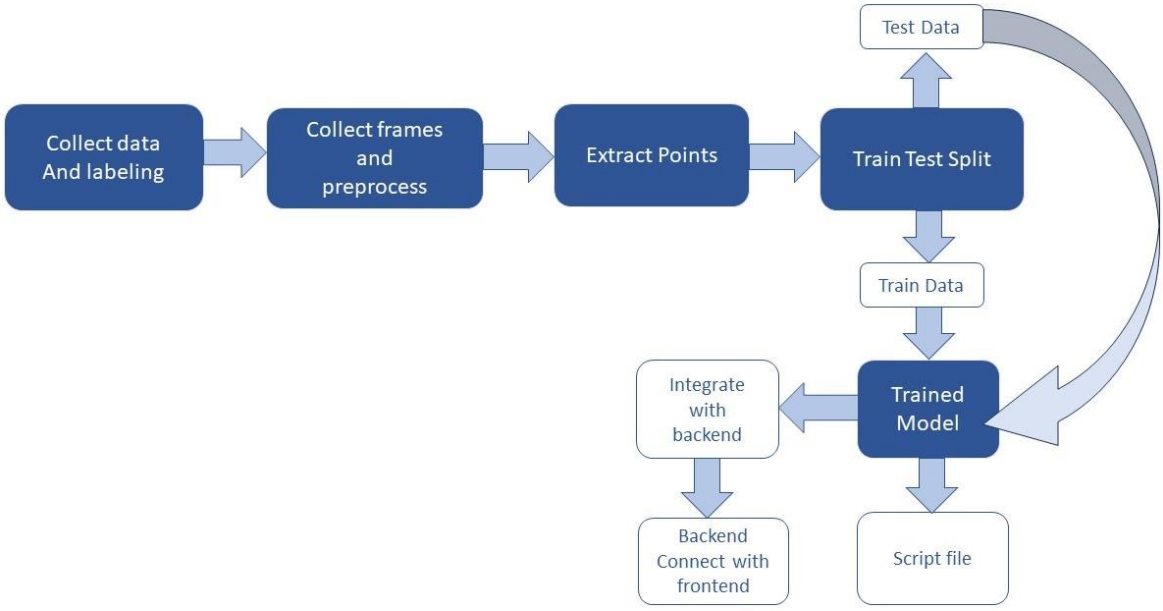


Figure 3.1.1: Data flow diagram (DFD) of the system

In a user scenario employing the Real-Time Bengali Sign Language Recognition System, a person initiates the process by activating their webcam on a designated platform. The system seamlessly captures the live video feed, monitoring the user's sign language gestures in real-time. With the help of advanced technologies such as MediaPipe and LSTM architecture, the system precisely extracts intricate features from the captured video frames. These extracted features are then precisely matched against a pre-trained model, enabling the system to accurately recognize the user's sign language gestures. Subsequently, the recognized gestures are displayed as texts, providing a clear interpretation of the communicated message. This comprehensive process not only showcases the system's capability to capture and analyze dynamic hand movements but also underscores its proficiency in effective communication by translating sign language into easily understandable text. The user, whether hearing-impaired or not, can thereby experience an enhanced and inclusive communication environment through the innovative Real-Time Bengali Sign Language Recognition System.

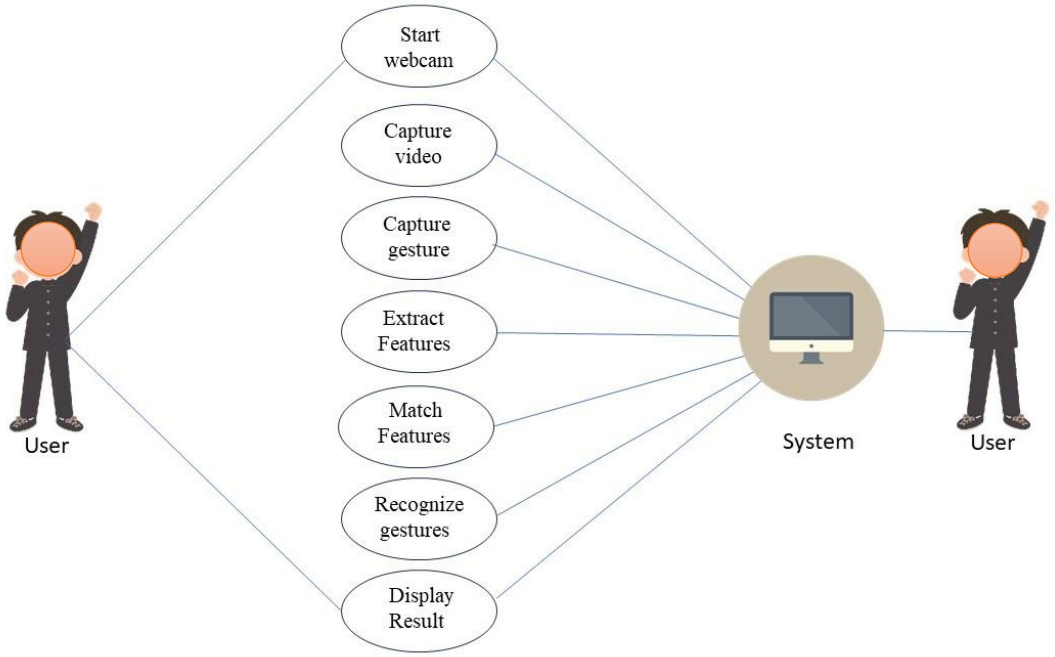


Figure 3.1.2: Use Case Diagram of the System

3.2 System architecture

The system architecture encompasses a comprehensive workflow, seamlessly transforming video input into interpreted Bengali sign language expressions. The step-by-step process involves:

- **Video Stream Conversion:** The system begins by converting the incoming video into a video stream, creating a continuous flow of frames for real-time processing. This ensures a dynamic and responsive interaction with the user.
- **Frame Extraction:** From the video stream, individual frames are extracted to capture the user's hand gestures at each moment. These frames serve as input for subsequent processing, allowing the system to analyze the intricate movements in sign language expressions.
- **Hand Key Points Extraction:** Using the extracted frames, the system employs the MediaPipe framework to identify and extract key points on the user's hands. This

step involves recognizing the positions of various hand landmarks, which are crucial for understanding the sign language gestures accurately.

- **LSTM Neural Network:** The architecture integrates an LSTM (Long Short-Term Memory) neural network, a powerful model for sequential data analysis. Trained on a dataset of Bengali sign language gestures, the LSTM model interprets the sequences of hand key points, capturing the nuanced patterns associated with different signs.
- **Word Recognition:** The final stage involves the interpretation of LSTM outputs to recognize specific words or phrases in Bengali sign language. The system maps the identified patterns to linguistic expressions, providing a meaningful and context-aware representation of the user's intended communication.

This holistic architecture ensures a smooth and coherent pipeline, enabling real-time conversion of hand gestures into linguistically meaningful outputs. The integration of advanced technologies, such as video stream processing, hand key points extraction, and LSTM modeling, contributes to the system's accuracy and responsiveness in facilitating effective communication for the deaf community.

Fig 3.2.1 shows the system architecture of the project:

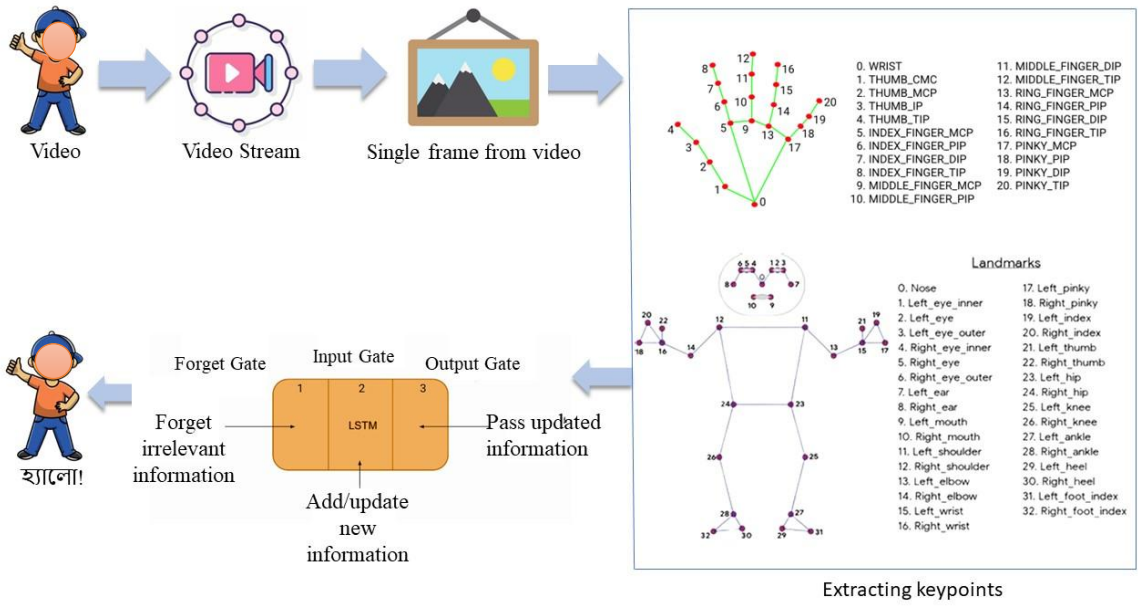


Figure 3.2.1: The System Architecture of the Project

3.3 Tools and Platforms used

Tools and platforms Used in Bengali Sign Language Recognition System are:

- **MediaPipe:** We employed MediaPipe library for hand tracking and holistic gesture recognition.
- **Flask (Backend):** Flask was used to develop the backend of the web application, handle image uploads, processing, and communication with the machine learning model.
- **HTML, CSS, JavaScript (Frontend):** Standard web development technologies were used to create the user interface to allow users to interact with the sign language recognition system.
- **OpenCV:** OpenCV was used for capturing video frames, image processing, and interfacing with the MediaPipe library.
- **Keras (with TensorFlow backend):** Keras was utilized to build, train, and evaluate the LSTM (Long Short-Term Memory) neural network for sign language gesture recognition.
- **Visual Studio Code (VS Code):** VS Code served as the integrated development environment (IDE) for coding, debugging, and version control.
- **Git:** Git was used for version control and collaborative development.

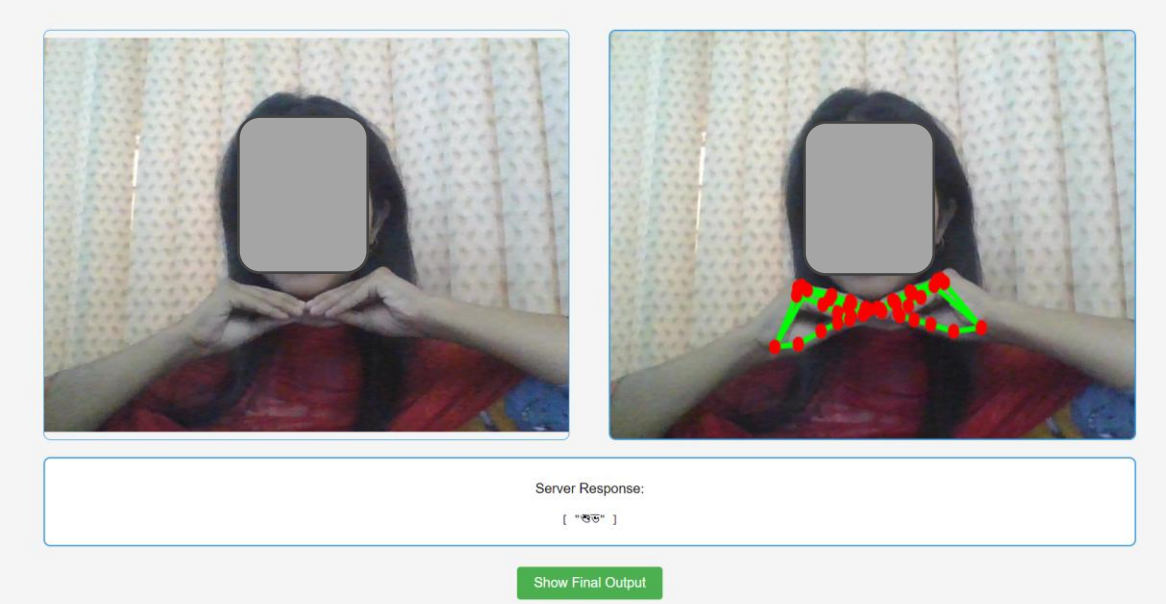
These tools, including Visual Studio Code, collectively form an integrated development environment, combining machine learning, web development, and project management tools to create a robust and accessible Bengali Sign Language Recognition System.

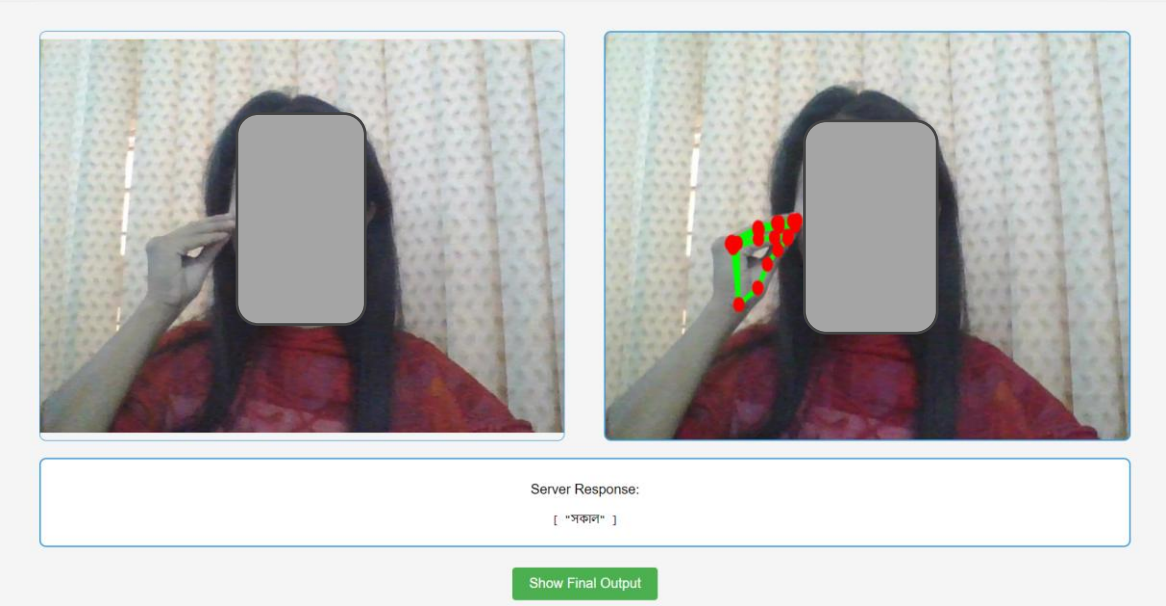
4 Project Implementation

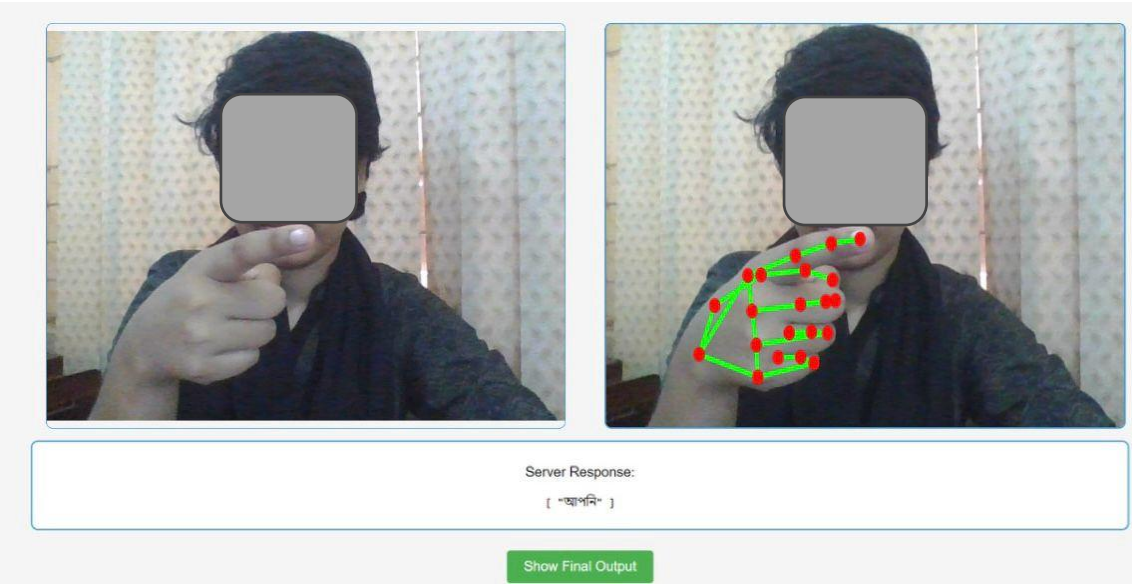
Our Real-Time Sign Language Recognition System was implemented through diverse dataset utilization and advanced technologies like Mediapipe and LSTM architecture. We prioritized user experience with a user-friendly interface. Overcoming challenges like overfitting, our system excelled in capturing intricate hand movements and ensuring robust performance. Strict testing and feedback loops validated its success, highlighting a balance between technical ability and real-world adaptability.

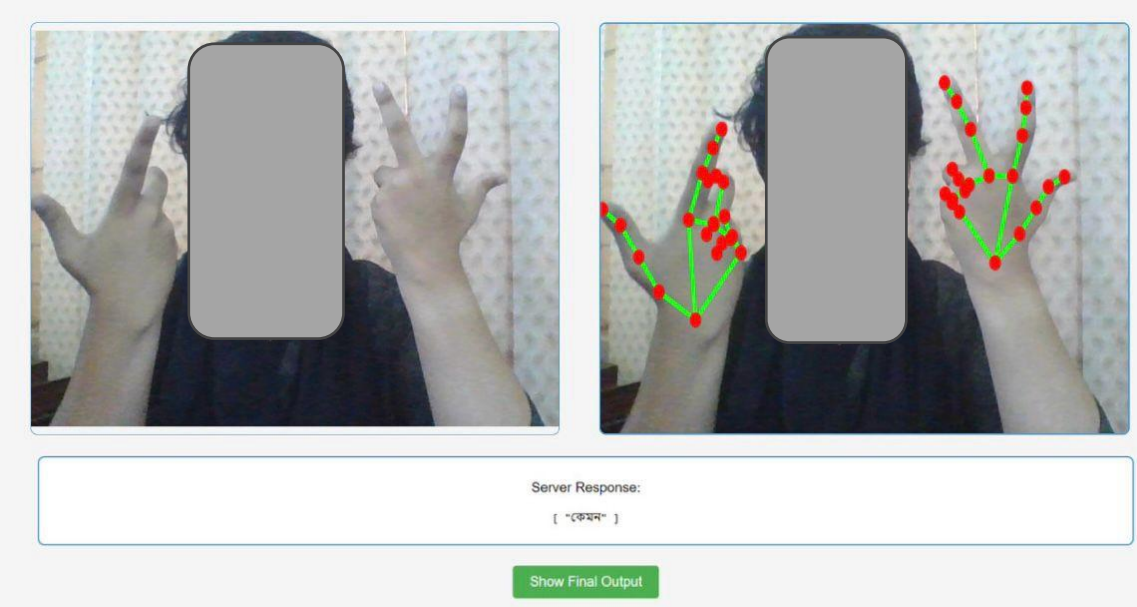
4.1 System implementation

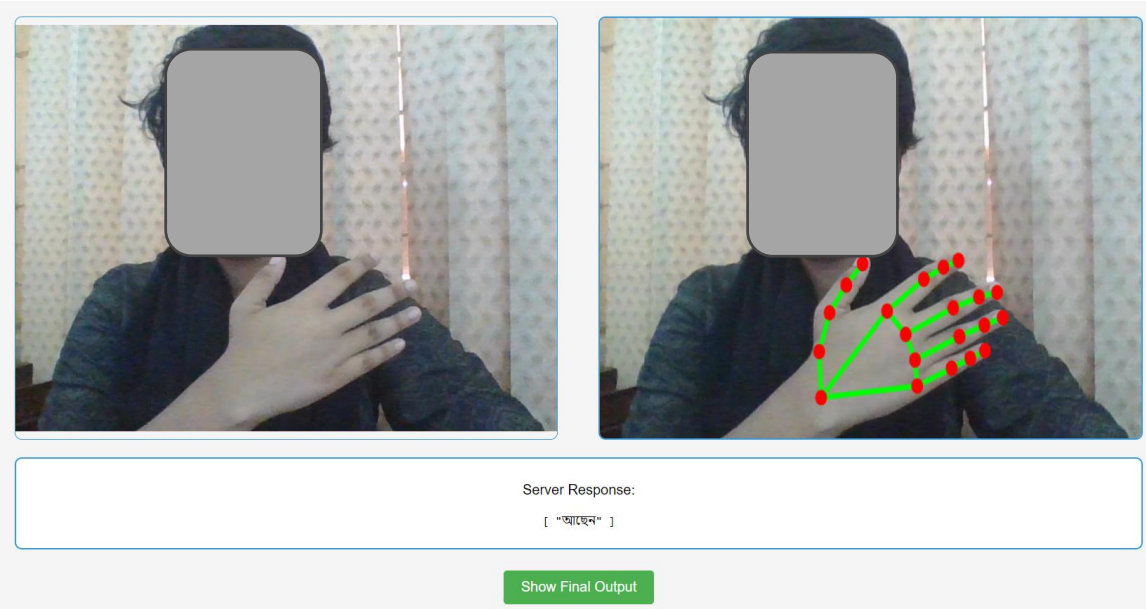
Our Real-Time Sign Language Recognition System (Website Front End):






Server Response:
["আপনি"]
Show Final Output


Server Response:
["কেমন"]
Show Final Output



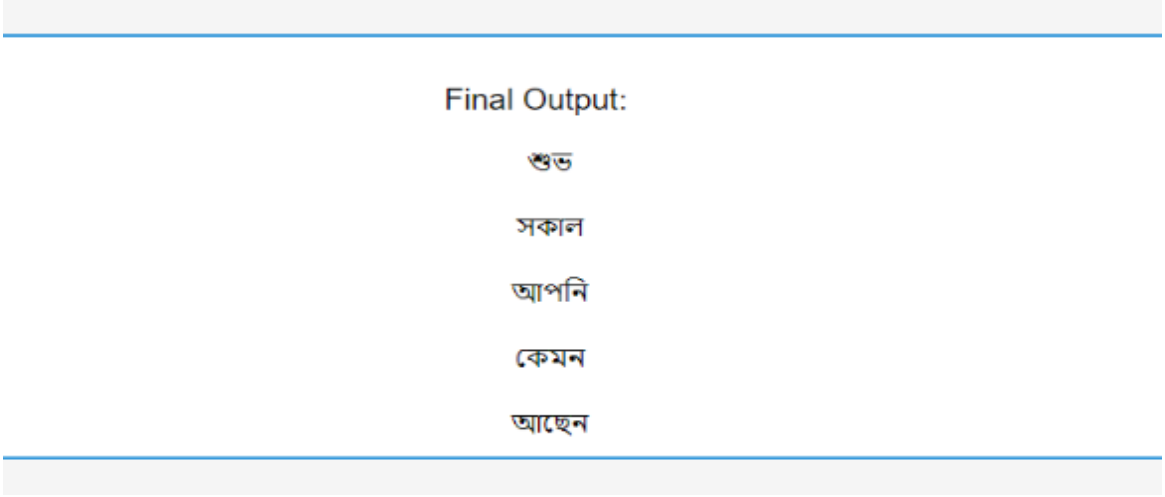


Figure 4.1.1: System Output

4.2 Morality or ethical issues

Our project prioritizes ethical considerations by steering clear of direct acquisition from existing sources or courses. This departure from conventional sign language recognition systems ensures the originality and integrity of our work. Additionally, we acknowledge and respect intellectual property rights and proper citation practices, fostering a responsible and ethical approach to research and development.

4.3 Socio-economic impact and sustainability

The socio-economic impact of our Bengali Sign Language Recognition System extends to societal inclusivity, addressing communication barriers for the hearing-impaired in Bengali-speaking regions. This technological advancement contributes to cultural inclusivity and empowers the deaf community. Furthermore, the project's sustainability is evident in its potential to evolve with expanding datasets and improved models, ensuring long-term relevance and impact.

The system's sustainability lies in its scalability with expanding datasets, continuous model improvement, and long-term relevance. It adapts to growing datasets, undergoes iterative improvements, and engages community feedback, ensuring its effectiveness over time. This commitment to adaptability and refinement positions the technology as a valuable, lasting resource for the Bengali-speaking deaf community.

4.4 Financial analyses and budget

While the project emphasizes research and development, it is essential to consider potential financial aspects. A comprehensive financial analysis includes budget planning for components such as dataset creation, model training, web interface development, and ongoing maintenance. This financial planning ensures effective resource allocation, contributing to the project's success and sustainability.

5 Conclusion

Filling a void in this field, especially for Bengali sign language, our web-based Bengali Sign Language Recognition system enables real-time gesture-to-text translation. The creation of a custom Bengali Sign Language video dataset ensures authenticity. Integrated into a user-friendly website with instructional videos, our platform not only recognizes gestures but also educates, promoting inclusivity. Detailed project planning, including Gantt charts, ensured efficiency. This innovative system aims to bridge communication gaps for the Bengali-speaking deaf community, marking a significant step in sign language technology.

5.1 Challenges Faced

Developing the Bengali Sign Language Recognition System posed several challenges, each requiring thoughtful solutions:

- **Limited Bengali Sign Language Resources:** Scarcity of available datasets and references for Bengali sign language presented a significant obstacle. To overcome this, we created our own dataset through careful collaboration with the deaf community.
- **Real-time Gesture Analysis:** Achieving real-time analysis of hand gestures without compromising accuracy was a complex challenge. Optimizing the system's speed while maintaining precision demanded the integration of advanced algorithms and model architectures.
- **User Interface Design for Accessibility:** Designing a user-friendly web interface that caters to the needs of the hearing-impaired community presented challenges in terms of accessibility. The interface had to be intuitive, inclusive, and visually clear to facilitate ease of use.
- **Integration of MediaPipe and LSTM Architecture:** Merging the capabilities of MediaPipe for gesture detection and LSTM architecture for sequence learning required seamless integration. Compatibility issues and synchronization intricacies were addressed to ensure a cohesive system.

5.2 Future Work

The current Bengali Sign Language Recognition System lays the foundation for future advancements and enhancements:

- **Expand Gesture Vocabulary:** Enrich the system by incorporating a more extensive set of Bengali sign language gestures, covering a broader vocabulary. This expansion will enhance communication capabilities.
- **Continuous Learning and Adaptation:** Implement deep learning models capable of continuous learning to adapt to variations in sign language over time. This dynamic adaptation ensures improved accuracy and responsiveness.
- **User Feedback and Iterative Development:** Establish a feedback mechanism involving users from the Bengali-speaking deaf community. Gather insights, preferences, and challenges to refine the system iteratively for better user acceptance and satisfaction.
- **Mobile Application Development:** Extend the accessibility of the system by developing a mobile application. This allows users to carry the sign language recognition tool on their smartphones, promoting on-the-go communication.
- **Real-time Translation to Text and Speech:** Integrate real-time translation of recognized sign language gestures into both textual and spoken Bengali. This feature extends the system's utility for various communication scenarios.

By pursuing these future avenues, the Bengali Sign Language Recognition System can evolve into a more robust, inclusive, and user-centric tool, significantly contributing to the empowerment of the Bengali-speaking deaf community.

References

[1] B. Duy Khuat, D. Thai Phung, H. Thi Thu Pham, A. Ngoc Bui, and S. Tung Ngo, "Vietnamese sign language detection using Mediapipe," 2021 10th International Conference on Software and Computer Applications, Feb. 2021, doi: https://doi.org/10.1145/3457784.3457810.

[2] Md. M. Islam, Md. R. Uddin, M. J. Ferdous, S. Akter, and Md. Nasim Akhtar, "BdSLW-11: Dataset of Bangladeshi sign language words for recognizing 11 daily useful BdSL words," Data in Brief, vol. 45, p. 108747, Dec. 2022, doi: https://doi.org/10.1016/j.dib.2022.108747.

[3] "Mediapipe Gesture Recognition - Gesture Control in Zoom Call," learnopencv.com, Jun. 07, 2022. https://learnopencv.com/gesture-control-in-zoom-call-using-mediapipe/ (accessed Nov. 30, 2023).

[4] S. saxena, "LSTM | Introduction to LSTM | Long Short Term Memor," Analytics Vidhya, Mar. 16, 2021. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/

[5] "The Python Tutorial — Python 3.7.4 documentation," Python.org, 2019. https://docs.python.org/3/tutorial/index.html

[6] "Welcome to Flask — Flask Documentation (3.0.x)," flask.palletsprojects.com. https://flask.palletsprojects.com/en/3.0.x/

[7] OpenCV, "OpenCV library," Opencv.org, 2019. https://opencv.org/

[8] "Keras | TensorFlow Core | TensorFlow," TensorFlow, 2019. https://www.tensorflow.org/guide/keras