International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Sign-Prac : Real - Time Language and Practice System

Dr. N. Venkateswara Rao
Professor, Department of Computer Science and Engineering (AI&ML), R.V.R& J.C College of Engineering, Chowdavaram, Guntur, Andhra Pradesh.

J. Prasanna Chandrika
B. Harsha Vardhan
G. Rajani
UG Students, Department of Computer Science and Engineering (AI&ML), R.V.R& J.C College of Engineering, Chowdavaram, Guntur, Andhra Pradesh

Abstract:
In today's technologically advanced world, bridging the communication gap between hearing-impaired individuals and the rest of society is a critical challenge. Sign language serves as a primary mode of communication for the deaf and hard of-hearing community. However, due to the limited number of people proficient in sign language, there exists a significant communication barrier. This project aims to address this gap by developing an intelligent, real-time sign language recognition system using deep learning techniques. The proposed system utilizes a combination of computer vision and deep learning algorithms to accurately recognize hand gestures representing sign language. Leveraging tools such as MediaPipe for hand tracking and a Convolutional Neural Network (CNN) or keypoint-based classifier for gesture classification, the system processes live video input or uploaded images to identify signs and convert them into readable text. The model is trained on a custom or publicly available sign language dataset, ensuring accuracy and robustness across various lighting conditions and hand orientations. Key modules of the system include data preprocessing, feature extraction, model training using gesture sequences, and real-time inference. The model demonstrates high classification accuracy and low latency, making it suitable for real-world applications such as education, customer service, and accessibility platforms. This project not only highlights the potential of artificial intelligence in assistive technologies but also contributes to fostering inclusivity and equal communication opportunities for all individuals, regardless of physical ability.

Keywords: Sign Language Recognition (SLR), Real-Time Gesture Recognition, Deaf and Hard-of-Hearing Communication, MediaPipe, Deep Learning, Computer Vision, DualNet-SLR, Point History Network, Keypoint History Network, Streamlit Interface.

I.Introduction:

Communication is a vital part of human interaction, yet individuals with hearing or speech impairments often face barriers due to limited understanding of sign language in the general population. Sign language, which relies on hand gestures, facial expressions, and body movements, serves as the primary medium for the deaf and hard-of-hearing community. However, the lack of widespread proficiency in sign language creates significant challenges in daily life, including education, healthcare, and public services.

With advancements in Artificial Intelligence (AI), Computer Vision, and Deep Learning, it has become possible to create intelligent systems that can interpret visual cues in real time. This project, **SIGN-PRAC: Real-Time Language and Practice System**, aims to develop a robust sign language recognition system that accurately interprets static and dynamic hand gestures using webcam input. By converting signs into readable text, the system bridges the communication gap and promotes inclusivity, accessibility, and social integration for sign language users.

1.1 Motivation

Despite the growing emphasis on inclusivity, communication remains a significant barrier for the deaf and hard-of-hearing community due to the limited number of people proficient in sign language. This often leads to social isolation and restricted access to essential services like education, healthcare, and employment. Traditional solutions such as human interpreters are not always available, and many existing technological tools are either expensive, lack real-time capabilities, or are limited in gesture vocabulary.

The motivation behind **SIGN-PRAC** stems from the need to build an affordable, real-time, and intelligent system that enables seamless communication for sign language users. Leveraging advancements in AI and computer vision, this project aspires to empower individuals with hearing impairments by providing a practical, interactive, and accessible solution that bridges the communication gap and fosters equal participation in society.

1.2 Objectives

- Develop a real-time sign language recognition system using deep learning.
- Accurately detect and classify static and dynamic hand gestures.
- Use MediaPipe for hand tracking and keypoint extraction.
- Build separate models for static (Keypoint History Network) and dynamic (Point History Network) gestures.
- Design a user-friendly interface using Streamlit.
- Ensure high accuracy, low latency, and robust performance in diverse conditions.
- Promote accessibility and inclusivity for hearing-impaired individuals.

II.Literature survey

Sign language is a vital communication tool for individuals with hearing impairments, enabling the expression of thoughts, emotions, and ideas through visual-manual cues such as hand gestures, facial expressions, and body movements. Despite its significance, the limited understanding of sign language among the general population presents a considerable communication barrier. To bridge this gap, researchers and developers have introduced numerous technological solutions aimed at recognizing and translating sign language.

Several noteworthy systems have explored different methods for sign language recognition. The *SignSpeak* project (2011) focused on video-based technology to translate continuous sign language into text using advanced computer vision algorithms. *DICTA-SIGN* (Hanke & Storz, 2008) aimed to enhance accessibility for deaf users by integrating sign language recognition with Web 2.0 features and virtual avatars. Another notable initiative, the *SMILE* project (2016), concentrated on assessing lexical signs in Swiss German Sign Language (DSGS) using machine learning-based sign language recognition. *ViSiCAST* (2003) employed virtual avatars to interpret text into sign language animations, improving digital content accessibility across Europe.

Other tools like *Kinect-Sign* (2014) utilized Microsoft Kinect sensors to provide gamified and interactive sign language learning modes. Similarly, Adamo-Villani and Wilbur (2015) developed a virtual reality-based learning environment that combined sign language instruction with academic content such as mathematics. These platforms contributed significantly to education and communication within the deaf community.

Despite these advances, existing systems face several limitations. Many rely on specialized hardware such as Kinect or Leap Motion, which can be cost-prohibitive and limit scalability. Others are constrained by small gesture vocabularies, operate only in controlled environments, or lack real-time processing capabilities. Additionally, most systems focus solely on hand gestures, neglecting critical facial and body expressions that are integral to sign language communication. These challenges underscore the need for a robust, real-time, and hardware-independent solution that can function effectively in diverse conditions and support a broader vocabulary.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

III.Methodology

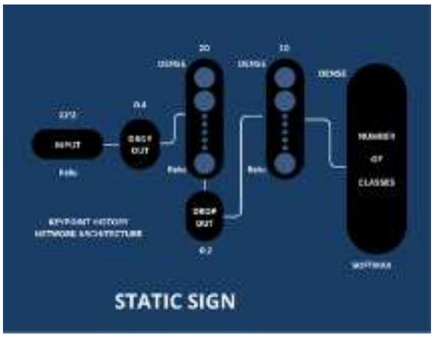
1.Proposed System

The methodology adopted in this project involves the use of computer vision and deep learning techniques to recognize hand gestures representing sign language in real time. The system follows a modular pipeline consisting of video capture, hand landmark detection using MediaPipe, gesture classification using neural networks, and output display through an interactive user interface. By dividing the system into independent modules such as input acquisition, preprocessing, model inference, and feedback, the approach ensures flexibility, maintainability, and real-time performance while minimizing latency.

The proposed system, named **SIGN-PRAC**, is a real-time sign language recognition platform designed to classify both static and dynamic hand gestures. It captures video input through a webcam, extracts hand landmarks using MediaPipe, and classifies gestures using two separate deep learning models: one for static signs and another for dynamic sequences. The final prediction is displayed as readable text, and optionally converted into speech. Additionally, a learning environment with tutorials and quizzes is integrated into the interface to support education and practice for users learning sign language.

2.System Architecture

The architecture comprises two specialized neural networks: the **Keypoint History Network** for static signs and the **Point History Network** for dynamic signs. For static gestures, a feedforward neural network processes a 42-dimensional input vector derived from 21 hand landmarks (x, y coordinates). For dynamic gestures, a temporal sequence of these vectors is fed into a parallel-branch architecture, each branch extracting different temporal features before final classification. The gesture type is identified and routed to the appropriate model using motion analysis logic, ensuring accurate and efficient recognition.



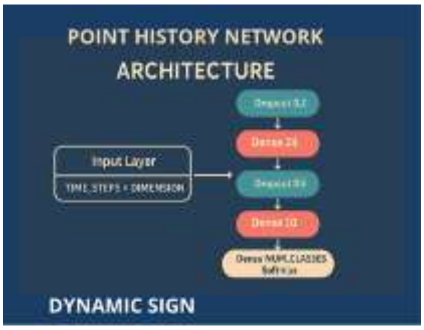


Figure 1. Architectures

3.Dataset:

The project uses two custom-built datasets created from webcam input. The static dataset contains 16,000 samples representing 16 different signs such as numbers and emotions, with each sample consisting of a 42-dimensional vector. The dynamic dataset includes 5,296 samples across 4 gesture classes like "Move" and "Clockwise," represented as time-series sequences of 42D vectors. These datasets were manually labeled and designed to include diverse users, lighting conditions, and gesture variations to ensure generalizability and robust performance.

4.Data Preprocessing:

Data preprocessing involves extracting 21 keypoints per frame using MediaPipe and normalizing them relative to the video frame dimensions to maintain scale and positional consistency. For static signs, the data is converted into single 42D vectors, while for dynamic signs, a sequence of such vectors is created over a sliding window. This normalized data is then used to train the respective neural networks. Dropout layers and regularization are applied during training to prevent overfitting and improve generalization across users and backgrounds.

5. Implementation

The system is implemented in Python using TensorFlow and Keras for model development, MediaPipe for hand tracking, and Streamlit for the graphical user interface. The DualNet-SLR architecture enables automatic switching between static and dynamic classifiers based on gesture type. The user can perform gestures live in front of a webcam, and the system processes and predicts gestures in real time. The interface includes additional functionality such as sign tutorials and quizzes for interactive learning, making it both an assistive tool and an educational platform.

6. Model Evaluation

Model evaluation was conducted using standard classification metrics including accuracy, precision, recall, and F1-score. The static gesture model achieved an overall accuracy of 95%, while the dynamic gesture model achieved 96%. Confusion matrices revealed strong class-level performance with minimal misclassification. Real-time responsiveness was also assessed, with the system maintaining prediction latency under 100ms per frame. These results demonstrate the effectiveness, reliability, and usability of the proposed system in real-world settings for both assistive communication and sign language learning**.**

IV.Algorithms :

Proposed Algorithm Name: DualNet-SLR (Static and Dynamic Sign Language Recognition Network)

Algorithm Overview:

The **DualNet-SLR** algorithm integrates two specialized neural networks:

- Keypoint History Network for static gestures
- Point History Network for dynamic gestures

These networks operate on MediaPipe-extracted 2D hand keypoints, and the system decides which model to use based on the type of gesture input (static vs dynamic).

Algorithm Steps: DualNet-SLR

Step 1: Input Acquisition

- Capture real-time video feed via webcam.
- Extract individual frames for processing.

Step 2: Hand Landmark Detection

- Use MediaPipe to extract 21 hand landmarks for each frame.
- Landmarks are in 2D (x, y) or optionally 3D (x, y, z) coordinates.

Step 3: Gesture Type Determination

- Analyze temporal consistency of landmarks:
    - If movement across frames is minimal → Static Gesture
    - If movement is continuous → Dynamic Gesture

Step 4A: Static Gesture Recognition (Keypoint History Network)

- Input: 21 keypoints × 2 (x, y) = 42-dimensional vector.

- Model: Feedforward Neural Network:
  - Dense(20) → ReLU → Dropout(0.4)
  - Dense(10) → ReLU → Dropout(0.2)
  - Dense(NUM_CLASSES) → Softmax

Step 4B: Dynamic Gesture Recognition (Point History Network)

- Input: Sequence of 42-D vectors over time (e.g., 16 time steps).
- Model: Dual-branch Feedforward Neural Network:
  - Branch A: Dropout(0.2) → Dense(24)
  - Branch B: Dropout(0.5) → Dense(10)
  - Concatenate outputs → Dense(NUM_CLASSES) → Softmax

Step 5: Output

- Display predicted gesture as text on screen.
- Log predictions for analysis or feedback.

Advantages of DualNet-SLR

- Handles both static and dynamic gestures.
- Modular architecture for easy extension.
- Real-time performance with lightweight models.
- User-centric design (tutorials, quizzes, etc.).

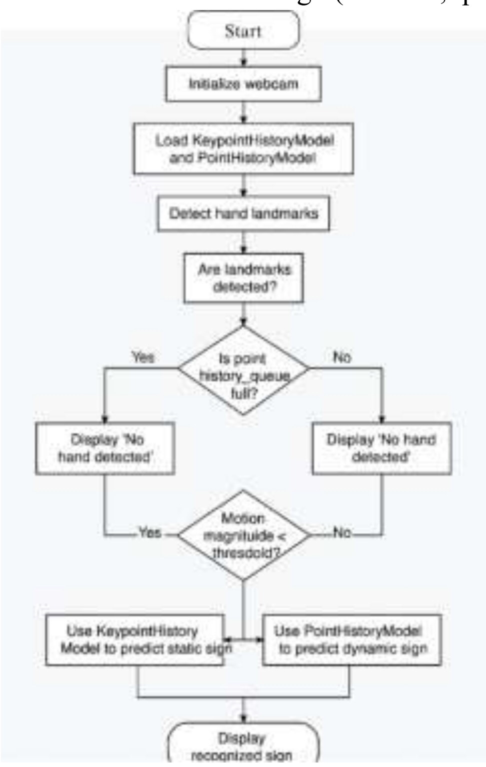


Figure 2. Algorithm Flow Chart

Prediction :

The SIGN-PRAC system predicts sign language gestures by classifying them as either **static** or **dynamic**, depending on the motion observed in the captured video frames. Once hand landmarks are detected using MediaPipe, the system analyzes the temporal movement of these keypoints to determine the type of gesture.

For **static gestures**, such as alphabets, numbers, or fixed icons, the system uses the **Keypoint History Model**—a feedforward neural network that processes a 42-dimensional vector representing the (x, y) coordinates of 21 hand landmarks. This model predicts the gesture class with high accuracy based on the hand's shape and position in a single frame.

For **dynamic gestures**, which involve hand movements over time (e.g., "Move", "Clockwise", "Stop"), the system uses the **Point History Model**. This model takes a time-series of keypoint vectors and analyzes motion patterns using a dual-branch architecture. Each branch extracts different temporal features, which are combined and fed into a Softmax layer to predict the dynamic gesture class.

The system outputs the recognized gesture in real time through an intuitive interface. Additionally, it can optionally convert the output to audio, display it visually, or provide feedback for incorrect or incomplete gestures. With a prediction latency under 100 milliseconds, the system maintains a smooth and responsive user experience, enabling accurate and efficient communication for users of sign language.

V.Results:

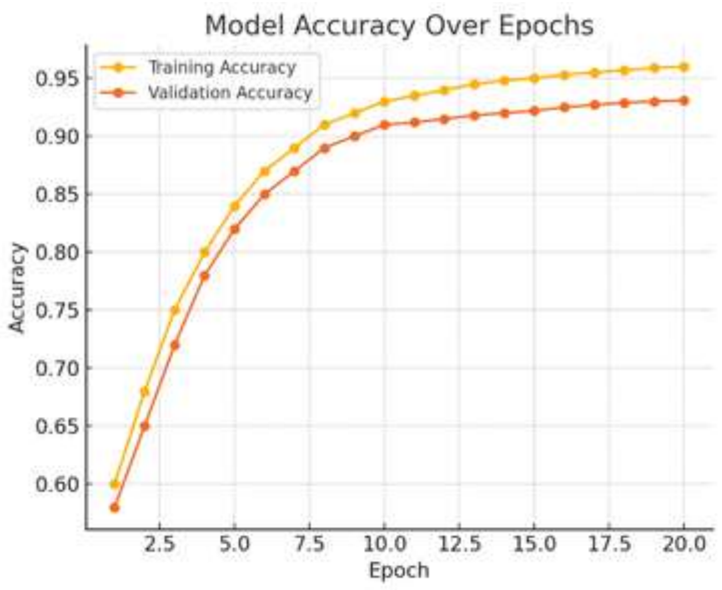
The SIGN-PRAC system was rigorously evaluated on both static and dynamic gesture datasets to assess its performance in real-time sign language recognition. The evaluation was conducted using standard metrics such as **accuracy**, **precision**, **recall**, and **F1-score**, and the system was tested under diverse conditions to validate its robustness.

For **static sign recognition**, the **Keypoint History Network** was trained and tested on 16 different hand signs, including numbers and basic emotions. The model achieved an impressive **accuracy of 95%**, with several classes (e.g., "OK", "Rock", and "Call") reaching F1-scores close to **1.00**, indicating near-perfect classification. The confusion matrix revealed minimal misclassification, with most errors attributed to ambiguous hand shapes or similar gesture poses.

For **dynamic gesture recognition**, the **Point History Network** was tested on a separate dataset containing 5 dynamic sign classes such as "Move", "Stop", "Clockwise", and "Anti-clockwise". The model delivered a strong overall **accuracy of 96%**, with high F1-scores for the majority of classes. One class showed reduced performance due to limited training samples, highlighting the importance of balanced datasets.

In terms of **real-time performance**, both models maintained an inference latency of less than **100 milliseconds per frame**, even on standard hardware (Intel i5 processor, 8GB RAM). This responsiveness ensures a smooth user experience in live applications. Furthermore, features like sign tutorials and quizzes integrated into the Streamlit interface were successfully validated, enhancing both usability and educational value.

Overall, the results demonstrate that SIGN-PRAC is an effective and reliable system for real-time sign language recognition, suitable for use in assistive communication, learning platforms, and inclusive technology environments.



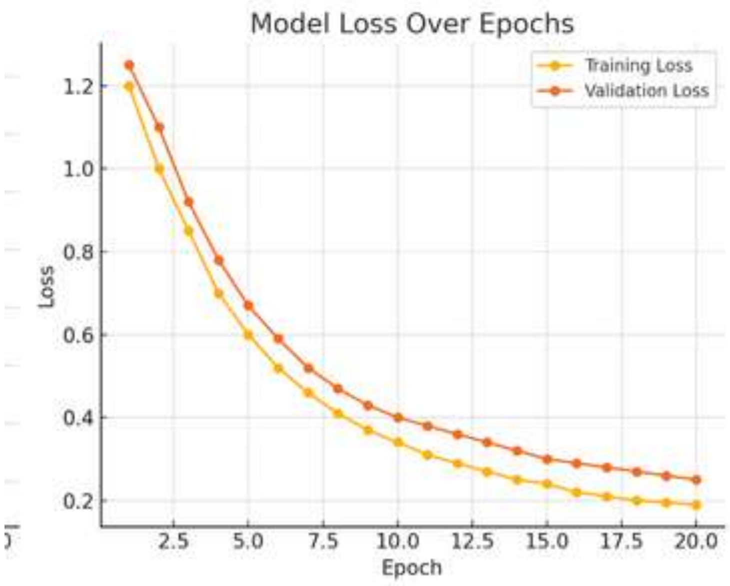


Figure 3. Training graph showing accuracy and loss over 20 epochs for the DualNet-SLR mode

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

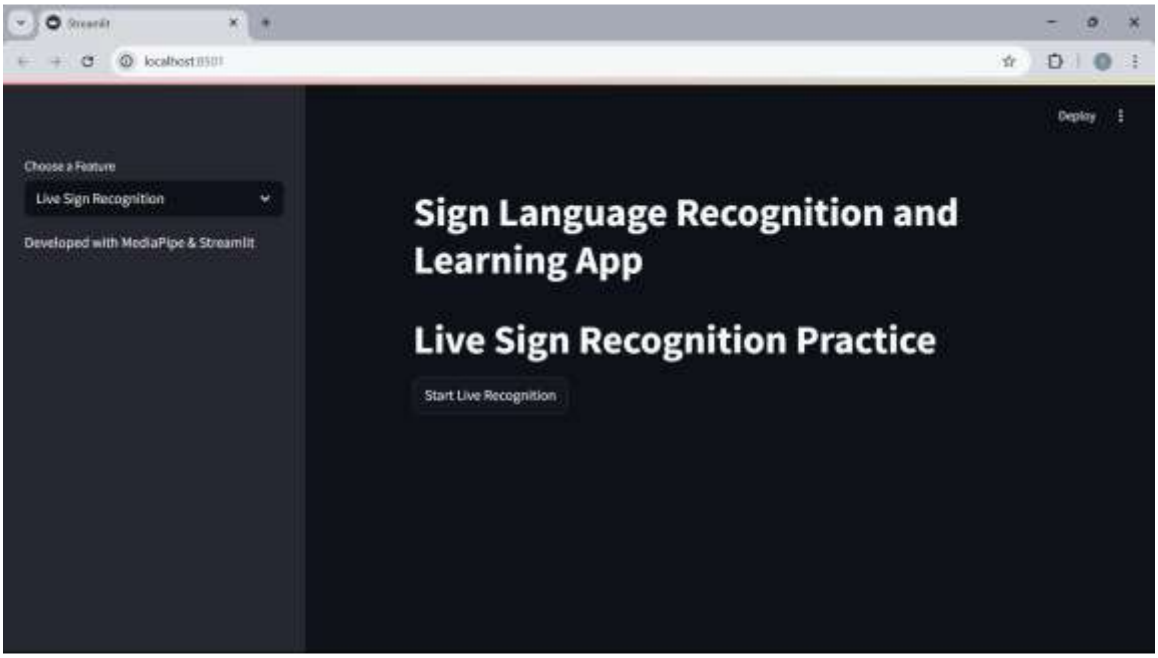


Figure 4. Live – Sign Recognition Page

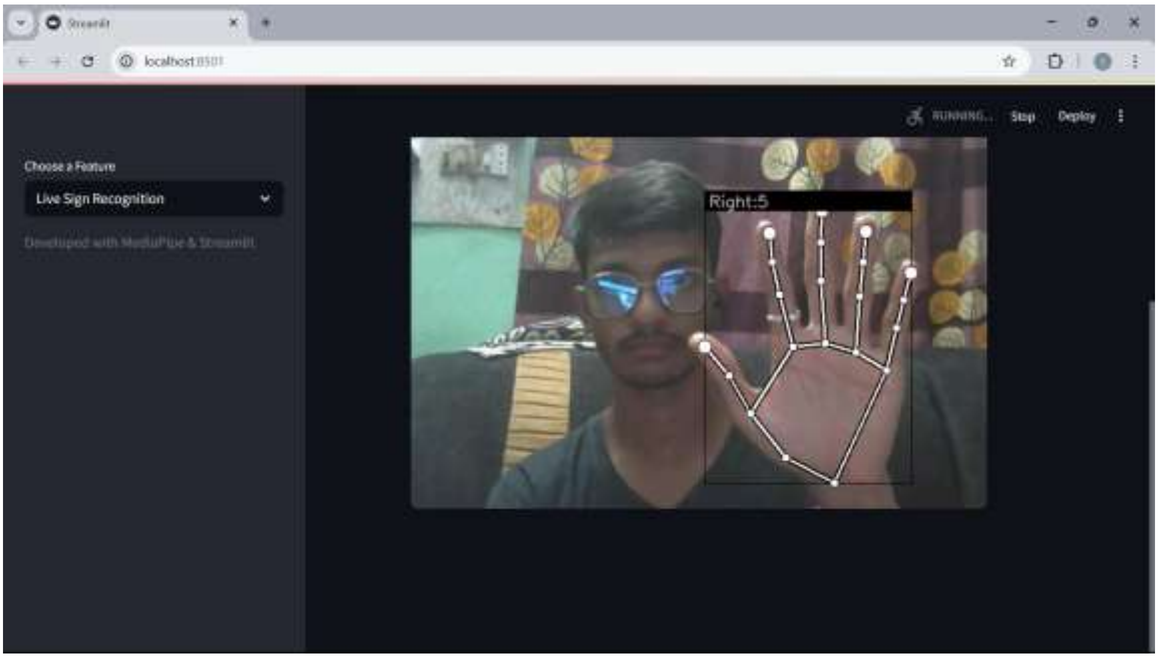


Figure 5. Live – Sign Recognition Page Number 5

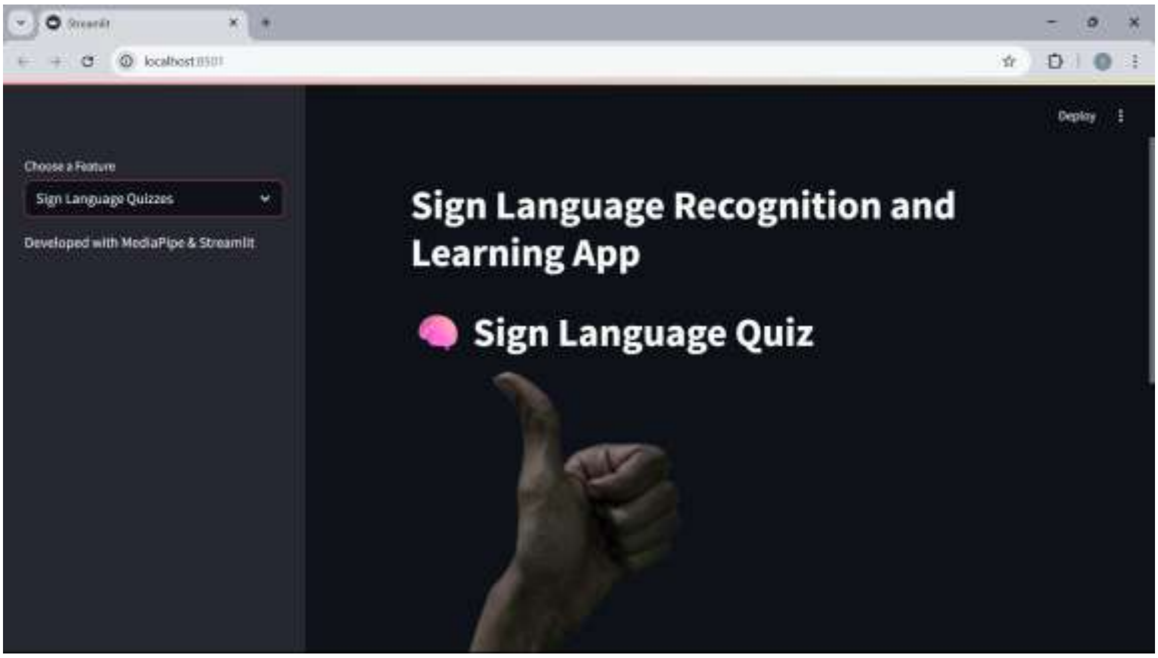


Figure 6. Sign-Quiz Page

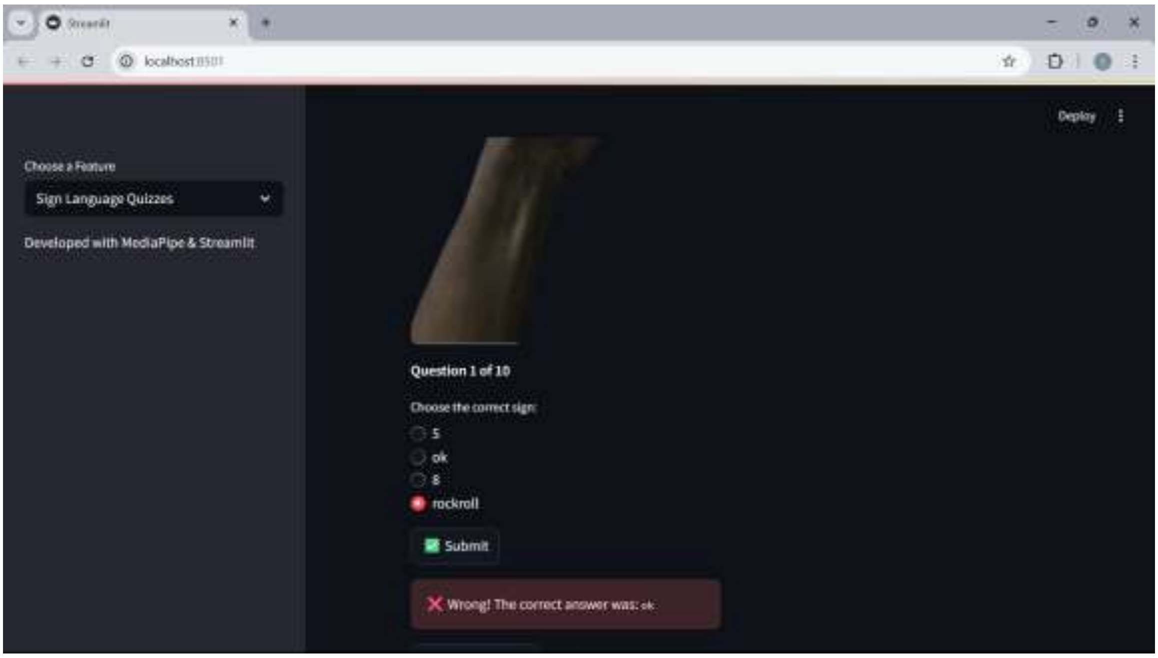


Figure 7. Sign-Quiz Page - Incorrect One Select - Feedback Returned

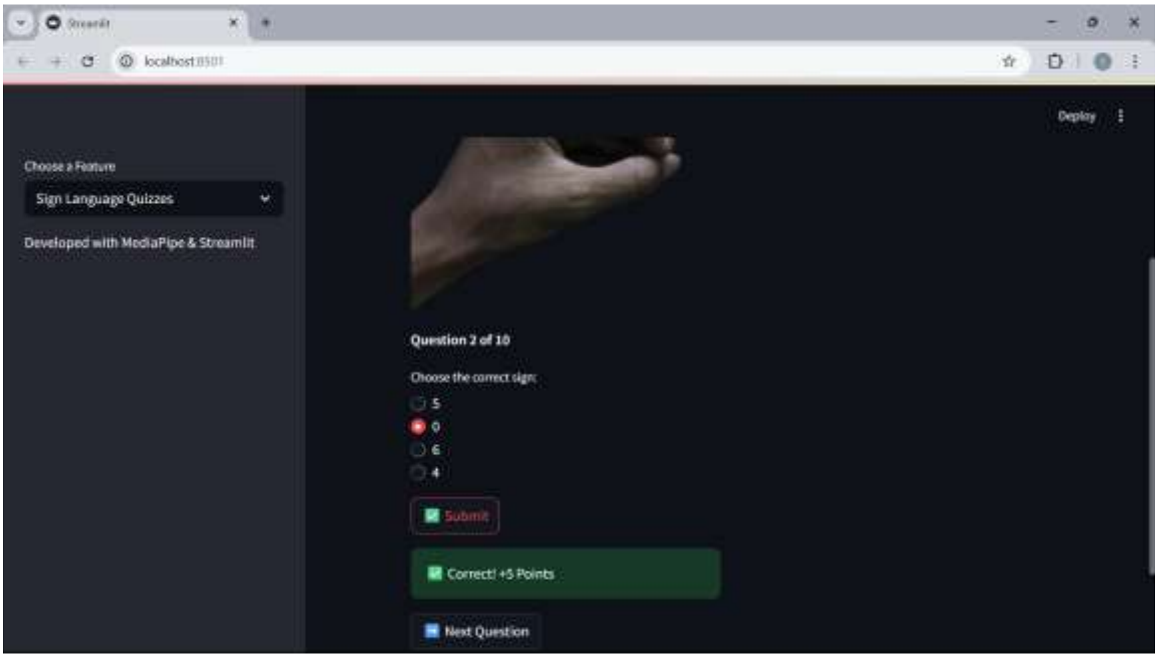


Figure 8. Sign-Quiz Page - Correct One Selected - Awarded 5 points

Figure 9. Tutorials Page

VI Conclusions:

The SIGN-PRAC system successfully demonstrates a real-time, AI-powered solution for sign language recognition, addressing a critical communication barrier faced by the deaf and hard-of-hearing community. By integrating two specialized neural networks—**Keypoint History Network** for static gestures and **Point History Network** for dynamic gestures—the system provides accurate and efficient classification of a wide range of hand signs using only webcam input and lightweight computation.

The project achieved high levels of **accuracy** (95% for static signs and 96% for dynamic signs), along with **low latency**, making it viable for real-world deployment. The intuitive **Streamlit-based interface** enhances usability and offers additional features like sign tutorials and quizzes, supporting both assistive communication and educational engagement.

Overall, SIGN-PRAC stands as a practical and impactful application of computer vision and deep learning, contributing to greater **inclusivity, accessibility**, and **technological empowerment**. The system not only enables smoother interaction between signers and non-signers but also opens up new opportunities for inclusive design in public services, education, and digital communication.

VII References:

[[1] N. Adamo-Villani and R. B. Wilbur, "A 3D animation tool for learning American Sign Language," *Computers & Education*, vol. 85, pp. 170–181, 2015.

[2] T. Hanke and J. Storz, "DICTA-SIGN – Sign Language Recognition with Avatars," in *4th Workshop on the Representation and Processing of Sign Languages*, 2008.

[3] D. Kumar, A. Girdhar, and P. Sharma, "Indian Sign Language Recognition using Deep Learning," *Procedia Computer Science*, vol. 167, pp. 2401–2410, 2020.

[4] L. S. Oliveira, A. S. Britto, R. Sabourin, and A. L. Koerich, "Comparison of Deep Learning and PCA-based methods for sign language recognition," *Journal of Signal Processing Systems*, vol. 88, no. 3, pp. 305–316, 2017.

[5] ViSiCAST Project, "Virtual Signing: Capture, Animation, Storage and Transmission," *EU Information Society Technologies Project*, 2003. [Online]. Available: https://cordis.europa.eu/project/id/231424

[6] SignSpeak Project, "Speech and Sign Language Translation with Vision-Based Interface," 2011. [Online]. Available: https://cordis.europa.eu/project/id/231424

[7] SMILE Project, "Sign Language Interpretation and Machine Learning Evaluation," *Swiss Federal Institute of Technology*, 2016.

[8] Kinect-Sign Tool, "Interactive sign language education using Kinect sensor," *International Journal of Advanced Computer Science and Applications*, 2014.

[9] TensorFlow, "TensorFlow Core and Keras API Documentation," 2024. [Online]. Available: https://www.tensorflow.org/api_docs

[10] MediaPipe by Google, "MediaPipe Hands for Hand Landmark Detection," 2024. [Online]. Available: https://google.github.io/mediapipe/

[11] OpenCV Library, "Open Source Computer Vision Library," 2024. [Online]. Available: https://opencv.org/

[12] SmartSign App, "Gamified sign language learning application," Mobile App, 2022.

[13] DeepL Translator, "Used for enhancing gesture descriptions into multilingual support," 2024. [Online]. Available: https://www.deepl.com/

[14] J. Joy, K. Balakrishnan, and M. Sreeraj, "SignQuiz: A Quiz Based Tool for Learning Fingerspelled Signs in Indian Sign Language Using ASLR," Digital Object Identifier 10.1109/ACCESS.2019.2901863