Real-Time Sign Language Recognition and Translation Using CNN and BiLSTM

A.Y. 2025-26, III-I

Team: [Your Name], [Your Roll No], [Your Email ID]

Department of Computer Science

July 28, 2025

Problem Statement

- **Main Research Problem**: Developing a real-time system to recognize and translate American Sign Language (ASL) gestures into text to bridge communication gaps for individuals with speech and hearing impairments.
- **Importance**: Enhances accessibility and inclusivity, supporting communication between sign language users and non-users, aligning with Sustainable Development Goals (e.g., SDG 4: Quality Education, SDG 11: Sustainable Cities).

Introduction

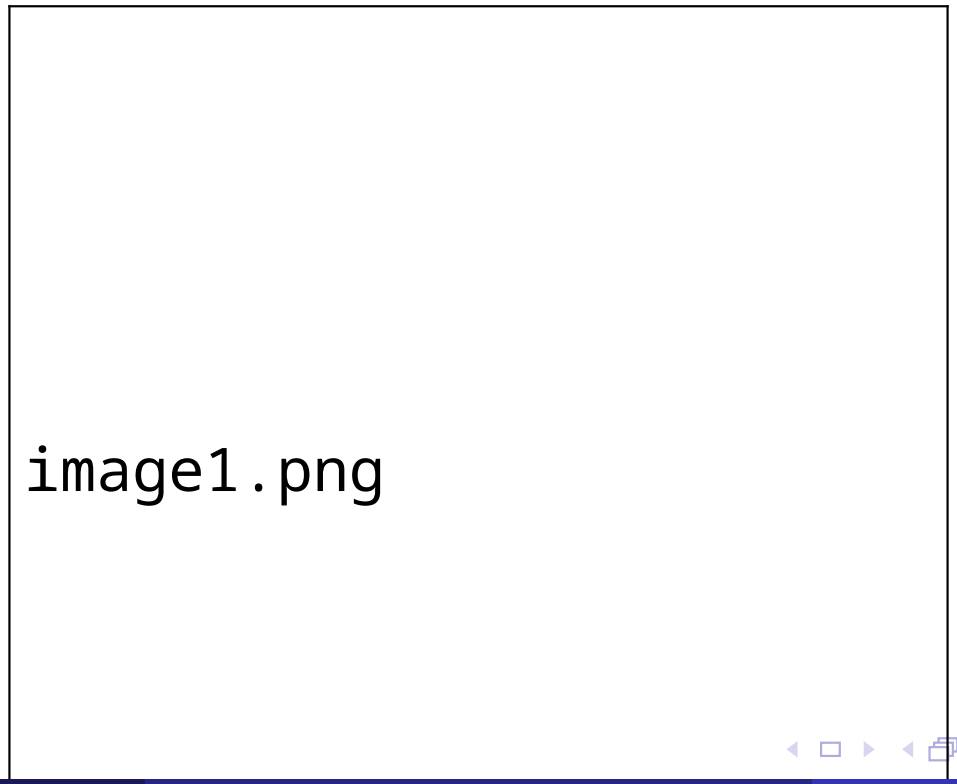
- **Background**: Sign language recognition uses computer vision and machine learning to interpret gestures, critical for individuals with hearing impairments.
- **Domain Importance**: Facilitates seamless communication, reduces barriers, and promotes digital inclusivity.
- **Problem Addressed**: Real-time recognition of dynamic ASL gestures under varying conditions (e.g., lighting, backgrounds).

Objectives of the Paper

- Develop a real-time ASL recognition system using CNN and BiLSTM.
- Achieve high accuracy in recognizing both static and dynamic gestures.
- Translate recognized gestures into text for accessibility.
- Support non-verbal communication with emoji translation.
- Hypothesis: Combining CNN for spatial features and BiLSTM for temporal modeling improves recognition accuracy over traditional methods.

Proposed Methodology

- **Architecture**: Hybrid CNN-BiLSTM model.
- **Steps**:
  1. Capture real-time video streams.
  2. Use CNN to extract spatial features (hand shapes, positions).
  3. Apply BiLSTM to model temporal dependencies (gesture sequences).
  4. Translate recognized gestures into text or emojis.



Tools/Techniques Used

- **Algorithms/Models**: Convolutional Neural Networks (CNN), Bidirectional Long Short-Term Memory (BiLSTM).
- **Datasets**: Likely ASL Alphabet dataset or similar (not specified in excerpt).
- **Libraries/Frameworks**: Python, TensorFlow/Keras, OpenCV for video processing.
- **Hardware**: Standard computing devices with webcam support.

Results & Evaluation

- **Metrics**: Accuracy, Precision, Recall, F1-score (specific values not fully provided in excerpt).
- **Reported Result**: ReLU activation with CNN achieved 92.5% accuracy for two-hand signs (from Table III).
- **Table Example**:

Activation Function	Model	Accuracy
ReLU	CNN	92.5%

Table: Performance Metrics (Partial)

- **Benchmark**: Outperforms traditional RNN-based models (e.g., Sharma et al., Gupta et al.).

Key Findings

- **What Worked Well**:
  - CNN effectively extracted spatial features.
  - BiLSTM captured temporal dependencies, improving dynamic gesture recognition.
  - Real-time performance suitable for practical use.
- **Limitations**: Likely challenges with complex gestures, variable lighting, or dataset diversity (inferred).
- **Suggested Improvements**: Authors may suggest larger datasets or advanced attention mechanisms.

Scope for Implementation/Extension

- **Implementation Plan**: Replicate CNN-BiLSTM model using TensorFlow and test on ASL Alphabet dataset.
- **Proposed Changes**:
  - Add attention mechanism to focus on key frames.
  - Incorporate MediaPipe for improved hand tracking.
  - Extend to other sign languages (e.g., ISL).
- **Goals**: Achieve ¿95% accuracy, deploy on mobile devices, support multi-language signs.

Your Proposed Plan

- **Techniques/Algorithms**: CNN-BiLSTM with attention, MediaPipe for hand detection.
- **Tools/Platforms**: Python, TensorFlow 2.x, OpenCV, MediaPipe.
- **Expected Outcomes**:
  - Real-time ASL recognition with ¿95% accuracy.
  - Mobile-compatible application for accessibility.
  - Support for emoji-based non-verbal translation.

References

- Velula Niresh, “Real-Time Sign Language Recognition and Translation Using CNN and BiLSTM,” 2025 International Conference on Computing Technologies (COCC).
- Sharma et al., “Deep learning-based framework for real-time sign language recognition using CNNs,” [1].
- Gupta et al., “Video-based sign language detection using CNN and LSTM,” [2].
- Wang et al., “Emoji-based translation system for sign language,” [3].
- Kim et al., “Multi-model approach for sign language recognition using RGB-D images and Bi-GRUs,” [5].