FYP-PROPOSAL

Project Title:

SignLink

Project Description:

A Real-Time Pakistan Sign Language Translation and Subtitling into Urdu System for Video Conferencing Platforms

Department:

Department of Computer Science & AI

Project Manager:

Dr. Muhammad Ilyas

Project Advisor:

Dr. Fahad Maqbool (Internal)

Project Team:

1. **Muhammad Hassan Shahbaz (BSCS51F22S007)**
2. **Behlol Abbas (BSCS51F22R038)**

Submission Date:

November 30, 2025

Table of Contents

1. Abstract

The significant communication gap between the Deaf community using Pakistani Sign Language (PSL) and the hearing population in Pakistan presents a critical accessibility challenge. While global research has progressed to continuous sentence-level translation, technological solutions for PSL remain confined to isolated word recognition and alphabet fingerspelling, severely hindering natural interaction in educational, professional, and social contexts. This research proposes SignLink, the first real-time continuous PSL-to-Urdu translation system designed to process natural signing sequences and generate coherent Urdu sentences. The project aims to develop a comprehensive PSL sentence dataset through community collaboration and implement a hybrid deep learning architecture that combines spatial-temporal feature extraction with sequence-to-sequence modeling. The outcome will be a webcam-based application providing real-time translation with simultaneous Urdu text and speech output, validated through user testing with the Deaf community. The expected academic contributions include novel methodologies for continuous PSL recognition and the first publicly available PSL sentence dataset, while the industrial and social benefits encompass enhanced accessibility, reduced dependency on human interpreters, and potential applications in education and public services, ultimately fostering greater digital inclusion in Pakistani society.

2. Background and Justification

Recent years have seen growing interest in making technology more accessible. Major corporations have initiated research and development in this domain, as evidenced by the works we have studied:

Sign language recognition has evolved into a robust research domain globally, with significant progress achieved for major sign languages like American Sign Language (ASL). Modern approaches leverage deep learning architectures including Convolutional Neural Networks (CNNs) for spatial features, Recurrent Neural Networks (RNNs) for temporal dependencies, and Transformers for end-to-end translation [1], [2]. The introduction of pose estimation libraries like MediaPipe has substantially advanced the field by providing robust, real-time human pose and hand landmark data [3]. Notable breakthroughs include Transformer-based frameworks for continuous sign sentence translation [4] and real-time systems using spatio-temporal Graph Convolutional Networks [5].

In stark contrast to global progress, Pakistani Sign Language research remains in its nascent stages, predominantly focused on isolated sign recognition using specialized hardware (Kinect, Leap Motion) [9], [10] or traditional computer vision techniques for limited vocabulary classification [11]-[16]. Comprehensive analyses consistently identify critical limitations, noting the complete absence of continuous PSL datasets and end-to-end translation models capable of processing natural signing flow [17], [18]. While some projects have developed text-to-PSL video playback systems [19], no existing work addresses the fundamental challenge of PSL-to-Urdu sentence translation.

2.1 Justification and Differentiation:

Research Gap

- **Limitation of Existing Systems**: Current PSL recognition technology is restricted to isolated alphabets and words, which is insufficient for meaningful dialogue.
- **Complexity of Continuous Translation:** The transition to sentence-level translation introduces unresolved challenges like segmenting a continuous stream of signs and modeling their non-linear temporal dynamics.
- **Grammatical Discrepancy:** PSL has its own unique grammar that differs from spoken Urdu, a complexity that word-level systems fail to handle.

Project Justification

- **Pioneering Solution:** This project aims to develop the first real-time system for translating continuous PSL sentences into Urdu, directly addressing a critical communication barrier.
- **Technical Synthesis:** We justify our approach by building on existing PSL research while integrating advanced, proven techniques from global sign language recognition projects.
- **Practical Impact**: The primary justification is to create an accessible and practical tool that enables natural communication for the Pakistani Deaf community in social, educational, and professional settings.

3. Research Methodology (Still in Progress)

To accomplish our objectives, we will adopt the following methodology:

Phase 1: Literature Review

This is an ongoing foundational activity that informs all subsequent phases. It involves:

- **Systematic Analysis**: Critically reviewing existing research on continuous sign language recognition for global languages (ASL, BSL, CSL) to identify state-of-the-art architectures (Transformers, RNNs, GCNs, CTC loss) [1], [2], [4], [5].
- **Gap Identification:** Analyzing the specific limitations of previous Pakistani Sign Language (PSL) research, which is confined to isolated signs and hardware-dependent systems [9], [10], [11], to solidify the justification for this project.
- **Technology Selection:** Evaluating tools and libraries (e.g., MediaPipe [3], PyTorch, OpenCV) based on their performance, community support, and suitability for real-time application development.
- **Ethical Framework:** Establishing guidelines for ethical data collection, informed consent, and collaboration with the Deaf community, based on best practices in assistive technology research.

Phase 2: Data Collection , Preprocessing and Dataset Development

The research will begin with the creation of the first comprehensive PSL sentence dataset through collaboration with Deaf schools and PSL interpreters. Data collection will focus on common daily

communication sentences (e.g., greetings, basic questions, emergency phrases) rather than isolated signs. The methodology includes:

- Recording 50+ sentence types with multiple signers (target: 10-15 participants).
- Using standard webcams (1080p, 30fps) in consistent lighting conditions.
- Implementing MediaPipe for real-time pose and hand landmark extraction.
- Creating parallel annotations with Urdu text translations.
- Applying data augmentation techniques (rotation, scaling, brightness adjustment) to enhance model robustness.

Phase 3: Model Architecture and Development

A hybrid deep learning architecture will be implemented to handle the spatial-temporal nature of continuous sign language:

- **Spatial Feature Extraction:** MediaPipe will extract 3D coordinates of body, hand, and facial landmarks, reducing input dimensionality from raw video frames to structured pose .
  - **Temporal Sequence Modeling**: A bidirectional LSTM architecture will process landmark sequences to capture temporal dependencies and context.
  - **Sequence-to-Sequence Translation**: Connectionist Temporal Classification (CTC) loss will be employed to handle variable-length input-output sequences without explicit alignment, enabling continuous sentence translation.

Phase 4: System Integration and Application Development

A desktop application will be developed with the following components:

- Real-time Video Processing: OpenCV for webcam feed capture and preprocessing.
- Translation Engine: Integrated model inference pipeline.
- Text-to-Speech: Google gTTS or Pyttsx3 for Urdu speech synthesis.
- User Interface: Tkinter-based GUI with video display, subtitle overlay, and controls.
- Virtual Audio Integration: VB-Cable for injecting synthesized speech into video conferencing applications.

Phase 5: Validation and Evaluation

System performance will be evaluated through:

- **Technical Metrics**: Accuracy, precision, recall, F1-score on test datasets.
- **Temporal Performance:** End-to-end latency measurement (target: <500ms>.
- **User Testing**: Structured evaluation with Deaf community participants using standardized communication tasks.
- **Usability Assessment:** System usability scale (SUS) questionnaires and qualitative feedback collection.

4. Project Scope

Here are some definite scopes and out of scope things are elaborated.

In-Scope

- Integration with at least one platform (e.g., Zoom) as proof of concept.
- Functionality: Real-time sign-to-text translation (Urdu), text-to-speech conversion (Urdu), and live subtitle display.
- Development of a continuous PSL recognition system for 50+ common sentence types.
- Designed for single signers in well lit, clear background environments initially.
- Real-time detection and translation of common sign languages (PSL).
- Plugin/extension for Zoom, Google Meet, etc.
- Output in Urdu audio and onscreen subtitles.
- Support for single signer video call scenarios.

Out-of-Scope

- Support for multiple simultaneous signers.
- Translation of spoken Urdu back to PSL (reverse translation).
- Operation in poor lighting or highly cluttered environments.
- Mobile application development.
- Handling of regional PSL variations across Pakistan.

5. High level Project Plan

The project plan developed using MS Project, spans 6 months (October 2025 to March 2026) with key milestones. Below is a summarized Gantt-style overview:

Activity	Duration	Resources Assigned	Milestone/Submission Date
Literature Review & Requirements Gathering	2 weeks (Oct 1-14)	Project Manager, Team Lead	Requirements Document - Oct 14, 2025
Data Collection & Preprocessing	4 weeks (Oct 15-Nov 11)	All Group Members (4 people)	Dataset Ready - Nov 11, 2025
Model Training & Development	6 weeks (Nov 12-Dec 23)	Team Lead + 2 Members (ML expertise)	Prototype Model - Dec 23, 2025
Integration & API Development	4 weeks (Jan 1-28)	Project Manager + 2 Members (Software dev)	Integrated Module - Jan 28, 2026
Testing, Iteration & Documentation	4 weeks (Jan 29-Feb 25)	All Group Members	Final Testing Report - Feb 25, 2026
Final Presentation & Submission	2 weeks (Feb 26-Mar 11)	Project Manager	Complete Project - Mar 11, 2026

6. References

[1] O. Koller, "Advances in sign language translation and generation," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV) Workshops, 2021.

[2] O. K. et al., "Sign Language Recognition and Translation with Heterogeneous Modality," Pattern Anal. Mach. Intell., vol. 43, no. 5, 2021.

[3] F. Zhang et al., "MediaPipe Hands: On-device Real-time Hand Tracking," arXiv preprint arXiv:2006.10214, 2020.

[4] N. C. Camgoz, O. Koller, S. Hadfield, and R. Bowden, "Sign language transformers: Joint end-to-end sign language recognition and translation," in Proc.CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2020.

[5] T. Zhu, W. Li, P. Wei, and L. Wang, "Real-time continuous sign language recognition and translation with a spatial-temporal graph convolutional network," in Proc. Conf. Multimedia Expo (ICME), 2021.

[6] H. D. P. W. T. H. K. K. D. B. D. R., "American sign language recognition and translation using deep learning," in Proc. IEEE UEMCON, 2018.

[7] Y. Huang, W. Li, and L. Wang, "A method for continuous Chinese sign language recognition and translation based on key frames and attention mechanism."

[8] N. Adaloglou et al., "A comprehensive study on deep learning-based methods for sign language recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 9, 2022.

[9] S. M. Hassan, "Pakistan Sign Language Recognition Using Kinect Sensor," in Proc. Int. Conf. Frontiers Inf. Technol. (FIT), 2017.

[10] A. R. et al., "Static Sign Word Recognition for Pakistani Sign Language Using Leap Motion Controller," in Proc. Int. Conf. Image Vis. Comput., 2019.

[11] T. Mahmood and A. I. Umar, "Urdu Sign Language Recognition Using Principal Component Analysis," in Proc. Int. Conf. Emerg. Technol. (ICET), 2018.

[12] S. Ahmed, H. Shafiq, Y. Raheel, N. Chishti, and S. M. Asad, "Pakistan sign language to Urdu translator using Kinect," Comput. Sci. Inf. Technol., vol. 3.

[13] M. Naseem, S. Sarfraz, A. Abbas, and A. Haider, "Developing a prototype to translate Pakistan sign language into text and speech while using convolutional neural networking," J. Educ. Pract., vol. 10, no. 15.

[14] H. Ahmed, S. O. Gilani, M. Jamil, Y. Ayaz, and S. I. A. Shah, "Monocular vision-based signer-independent Pakistani sign language recognition system using supervised learning," Indian J. Sci. Technol., vol. 9, no. 25, 2016.

[15] A. A. Mujeeb et al., "A neural-network based web application on real-time recognition of Pakistani sign language," Eng. Appl. Artif. Intell., vol. 135, 2024.

[16] H. Zahid et al., "A computer vision-based system for recognition and classification of Urdu sign language dataset," PeerJ Comput. Sci., vol. 8, 2022.

[17] B. Hassan, M. S. Farooq, A. Abid, and N. S. Khan, "Pakistan sign language: Computer vision analysis & recommendations," VFAST Trans. Softw. Eng., vol. 4, no. 1, 2016.

[18] H. Zahid et al., "Recognition of Urdu sign language: a systematic review of the machine learning classification," PeerJ Comput. Sci., vol. 8, 2022.

[19] A. Abbas and S. Sarfraz, "Developing a prototype to translate text and speech to Pakistan sign language with bilingual subtitles: A framework," J. Educ. Technol. Syst., vol. 47, no. 2, 2018.

Supervisor's Signature

_____