Localization of AI-Driven Sign Language Recognition

H.T.M.D Samaranayake
MS23002074

A THESIS

SUBMITTED TO

SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

(TECHNOLOGY/MANAGEMENT/SYSTEM)

December 2024

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

_____

Prof. Anuradha Jayakody

Approved for MSc. Research Project:

MSc. Programme Co-ordinator, SLIIT

Approved for MSc:

_____

Head of Graduate Studies, FoC, SLIIT

DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.

Sign: ……H.T.M.D Samaranayake….

H.T.M.D Samaranayake

Date: ……01.31.2025……..

ABSTRACT

Localization of AI-Driven Sign Language Visualization

H. T. M. D Samaranayake

MSc. in Information Technology

Supervisor: Prof. Anuradha Jayakody

December 2024

Sign language recognition (SLR) is a vital research area promoting inclusivity and interaction for the deaf and hard-of-hearing communities. This thesis focuses on automatic SLR using the American Sign Language (ASL) dataset, emphasizing preprocessing, feature extraction, and Long Short-Term Memory (LSTM) networks to enhance accuracy. The process begins with data collection, where video frames from the ASL dataset are resized, normalized, and converted into grayscale to reduce computational load while retaining key features. Data augmentation techniques like rotation, flipping, and scaling are applied to improve the model's generalization.

Feature extraction captures spatial and temporal information critical for SLR. Optical flow is employed to detect hand motion and facial expressions, while Convolutional Neural Networks (CNNs) extract spatial patterns from the video frames. These features are fed into an LSTM network, designed to learn sequential dependencies in the data. LSTMs are effective for understanding dynamic gestures, as they capture both short- and long-term dependencies between frames. The model predicts sign language symbols or words, facilitating real-time recognition.

The thesis further integrates semantic sentence prediction, enabling the system to recognize isolated signs and predict entire sentences. Using Natural Language Processing (NLP), input sentences are mapped to sign language sequences, which are visualized through synthesis models that generate animations. This approach captures handshapes, movements, and expressions essential in ASL.

By combining preprocessing, feature extraction, and deep learning, this research improves SLR accuracy and contributes to communication accessibility. It lays a foundation for advancements in SLR systems for applications in education, healthcare, and human-computer interaction.

ACKNOWLEDGEMENT

While at the Sri Lanka Institute of Information Technology, I had the privilege of being guided by exceptional advisors who often held differing perspectives, which greatly enriched my academic journey. Prof. Anuradha Jayakody was an outstanding mentor who provided invaluable advice and consistently offered constructive criticism on my ideas and writing. His unwavering support and encouragement helped refine my research skills and broaden my intellectual horizons. Additionally, he granted me the freedom to explore and develop my own projects within his research account's time, fostering an environment that nurtured creativity and independent thinking.

TABLE OF CONTENTS

List of Figures