



Sri Lankan Sign Language Detection System for Service Robots

H. U Rajamanthri
Reg. No.: MS24019422

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

December 2025

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 in IT 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is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources).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:.....

Heshan Rajamanthri

Date:09/11/2025.....

ABSTRACT

Sri Lankan Sign Language Detection System for Service Robots

Heshan Rajamanthri

MSc. in Information Technology

Supervisor: Prof Anuradha Jayakody

December 2025

Sinhala Sign Language (SSL) is the main way many Deaf and hard-of-hearing people in Sri Lanka communicate, but there are very few digital tools to recognize it automatically. This thesis addresses that problem by developing deep learning models that can recognize SSL gestures accurately while also working in real time. A dataset of 5,000 SSL videos covering 50 gestures was created and divided into training, validation, and test sets, with preprocessing steps such as frame resizing, normalization, and data augmentation. Four models were tested: a CNN to learn visual features, an LSTM to capture movement over time, a combined CNN-LSTM model, and YOLOv8 for fast, detection-based recognition. The CNN achieved the best overall performance with 98.32% accuracy, followed closely by the LSTM at 98.18%, while YOLOv8 reached 97.20% accuracy but delivered very high real-time speed of up to 90 frames per second. The CNN-LSTM model performed less well than expected. A working prototype was also built, allowing users to perform signs and see the recognized output in Sinhala or English, with optional speech, showing the system's practical value for inclusive communication. This research establishes a foundational framework for SSL recognition, showing that deep learning models can achieve state of the art accuracy while highlighting the trade offs between accuracy, latency, and deployment feasibility in real world contexts.

ACKNOWLEDGEMENT

This thesis is the culmination of years of dedication, guidance, and support, and I wish to take this opportunity to express my deepest gratitude to all those who made this journey possible.

First and foremost, I am profoundly grateful to my parents and siblings, whose unconditional love, encouragement, and sacrifices have been the foundation of my academic and personal life. Their constant belief in my abilities gave me the strength to persevere through every challenge encountered during this research.

I extend my heartfelt thanks to my supervisor, Professor Anuradha Jayakody, for his invaluable guidance, expertise, and encouragement throughout the course of this study. His constructive feedback, critical insights, and unwavering support were instrumental in shaping this work and ensuring its academic rigor.

I am also indebted to my examiners, whose thoughtful evaluations and recommendations have greatly enhanced the quality of this thesis. Their time, effort, and expertise are sincerely appreciated.

To my friends, I express my gratitude for their encouragement, companionship, and moral support during the long hours of study, experimentation, and writing. Their presence brought balance to this journey and reminded me of the value of friendship in academic pursuits.

A special word of thanks is extended to the Sri Lanka Central Federation of the Deaf, and to all individuals from the Deaf community who contributed to this work. Their willingness to share their knowledge, time, and experiences was vital to the success of this research. Without their collaboration and trust, this project would not have been possible.

Finally, I acknowledge all those who, in various ways, supported me during this journey. Each contribution, no matter how small, has left a lasting impact on me and on this work.

TABLE OF CONTENTS

DECLARATION	ii
ABSTRACT.....	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENTS.....	v
List of Figures	ix
List of Tables	x
Chapter 1 Introduction	1
1.1 Background of the Study.....	1
1.2 Problem Statement	2
1.3 Research Questions / Hypotheses	3
1.4 Objectives of the Study	4
1.5 Scope and Limitations.....	5
1.5.1 Scope	5
1.5.2 Limitations.....	5
1.6 Significance of the Study	5
Chapter 2 Literature Review	7
2.1 Background on Sign Languages and Accessibility	7
2.2 Early Approaches to Sign Language Recognition	7
2.3 Vision Based Static Gesture Recognition	8
2.4 Dynamic and Continuous Sign Language Recognition	9
2.5 Hybrid and Advanced Deep Learning Approaches.....	10
2.6 Regional and Low Resource Sign Language Research.....	11
2.6.1 Sinhala Sign Language (SSL).....	11
2.6.2 Sri Lankan Tamil Sign Language	12
2.6.3 Azerbaijani Sign Language (AzSL)	12
2.6.4 Critical Observations	12
2.7 Real Time Systems and Application Oriented Solutions.....	13
2.7.1 Tutor Systems and Educational Applications.....	13
2.7.2 Translation Systems for Communication	14
2.7.3 Continuous Sentence Level Applications.....	14
2.7.4 Global Inspiration and Comparative Systems	15
2.7.5 Critical Observations on Real Time Systems	15
2.8 Critical Gaps in Literature.....	16
2.8.1 1. Lack of Standardized Datasets	16

2.8.2 Overemphasis on Static Gestures	16
2.8.3 Limited Adoption of Advanced Architectures	16
2.8.4 Real Time Performance Underexplored	16
2.8.5 Regional Variations and Standardization Challenges	17
2.8.6 Lack of User Centered Evaluation.....	17
2.9 Chapter Conclusion	17
2.10 Dataset Construction, Annotation, and Benchmarking Practices	18
2.11 Landmark Based Pipelines vs. Raw Image CNNs	19
2.12 Preprocessing, Augmentation, and Domain Randomization	19
2.13 Object Detection, Segmentation, and Tracking in Continuous SLR.....	20
2.14 Sequence Modeling: CNN LSTM, Attention, and Transformers	20
2.15 Hybrid and Multimodal Sensing Revisited (Vision + IMU/ EMG).....	21
2.16 Real Time Engineering: Latency, Throughput, and On Device Inference.....	21
2.17 HCI: Feedback, Explainability, and Accessibility	22
2.18 Error Taxonomy and Failure Modes	22
2.19 Language Modeling and Decoding for Continuous SSL	22
2.20 Comparative Analysis of Model Families in Works.....	23
2.21 Ethics, Consent, Privacy, and Community Partnership	23
2.22 Toward a Sinhala Sign Language Research Agenda	24
2.23 Expanded Chapter Conclusion	24
Chapter 3 Research Methodology.....	26
3.1 Introduction	26
3.2 Dataset Development	29
3.2.1 Data Source and Collection (Videos Only)	29
3.2.2 Ethical Considerations.....	32
3.3 Dataset Statistics	32
3.3.1 Demographic and Regional Diversity	33
3.3.2 Dataset Balancing Strategy.....	33
3.3.3 Class Distribution and Partitioning.....	34
3.3.4 Convolutional Neural Network (CNN)	34
3.3.5 Long Short Term Memory Network (LSTM)	38
3.3.6 CNN-LSTM Hybrid Model.....	40
3.3.7 YOLOv8 Model.....	43
3.4 Model Evaluation and Performance Metrics.....	46
3.5 Tools and Frameworks Used.....	50

3.6 Limitations of Methodology	54
Chapter 4 System Design and Implementation.....	58
4.1 Overall System Architecture	58
4.2 Data Flow and Processing Pipeline.....	59
4.3 Feature Extraction Techniques.....	61
4.4 Model Training and Optimization.....	64
4.4.1 CNN Training and Optimization	65
4.4.2 LSTM Training and Optimization	65
4.4.3 CNN-LSTM Hybrid Training and Optimization.....	66
4.4.4 YOLOv8 Training and Optimization	66
4.4.5 Optimization Techniques Across Models.....	67
4.5 Real time Recognition Pipeline.....	67
Chapter 5 Results and Analysis	71
5.1 Experimental Setup	71
5.2 Training Results (Loss/ Accuracy Graphs)	72
5.2.1 CNN Model:	72
5.2.2 LSTM Model:	73
5.2.3 CNN-LSTM Hybrid Model:	74
5.2.4 YOLOv8 Model:.....	75
5.3 Model Performance Evaluation.....	76
5.3.1 CNN Model Performance	77
5.3.2 LSTM Model Performance.....	77
5.3.3 CNN-LSTM Hybrid Model Performance.....	77
5.3.4 YOLOv8 Model Performance	78
5.4 Comparative Analysis of All Models.....	78
5.4.1 Comparative Accuracy and Generalization.....	78
5.4.2 Precision, Recall, and F1 Comparisons	79
5.4.3 Detection Specific Metrics (mAP)	79
5.4.4 Latency and Real Time Performance	80
5.4.5 Comparative Table of Results	80
5.4.6 Practical Trade offs and Insights	80
5.4.7 Conclusion of Comparative Analysis	81
5.5 Error Analysis and Insights	81
5.5.1 Error Sources in CNN Models.....	81
5.5.2 Error Sources in LSTM Models	82

5.5.3 Error Sources in CNN-LSTM Hybrid Models	82
5.5.4 Error Sources in YOLOv8 Models	83
5.5.5 Alphabetic vs. Semantic Gestures	83
5.5.6 Impact of Signer Variability	84
5.5.7 Environmental Influences	84
5.5.8 Methodological Insights from Errors	85
5.5.9 Conclusion of Error Analysis	85
5.5.10 Defining Real Time Performance	86
5.5.11 CNN Real Time Outcomes	86
5.5.12 LSTM Real Time Outcomes	87
5.5.13 CNN-LSTM Hybrid Real Time Outcomes	87
5.5.14 YOLOv8 Real Time Outcomes	87
5.5.15 User Experience in Real Time Testing	88
5.5.16 Methodological Insights from Real Time Outcomes	88
5.5.17 Conclusion of Real Time Outcomes	89
Chapter 6 Discussion	90
6.1 Interpretation of Results	90
6.2 Significance in the Sri Lankan Context	91
6.3 Contributions to Research and Practice	92
6.4 Comparison with Global Sign Language Recognition Studies	92
6.5 Limitations of the Study	93
Chapter 7 Conclusion and Future Work	94
7.1 Summary of Findings	94
7.2 Conclusion	95
7.3 Recommendations	95
7.4 Future Enhancements	96
Chapter 8 References	97
Appendix 1: CNN model	99
Appendix 2: LSTM mode	104
Appendix 3: CNN-LSTM Hybrid model	109
Appendix 4: YOLOv8 model	116

List of Figures

Figure 1. Permisson request letter for SLCFD.....	30
Figure 2. Guesture perfoming in data collection	31
Figure 3. Logos of Jupyter Notebook, Anaconda, Python.....	54
Figure 4. CNN Training vs Validation accuracy	73
Figure 5. CNN Training vs Validation loss	73
Figure 6. LSTM model Training vs validation accuracy	74
Figure 7. LSTM Training vs Validation loss	74
Figure 8. CNN-LSTM Training vs Validation accuracy	75
Figure 9. CNN-LSTM Training vs Validation Loss.....	75
Figure 10. YOLOv8 model accuracy graph.....	76
Figure 11. YOLOv8 model loss graph.....	76

List of Tables

Table 3.1 Train Validation Test Split	34
Table 5.1. Performance across all models	80