ASL Recognition System:
Bridging Gaps in Communication Accessibility

A Thesis Submitted in partial fulfillment of the Requirements of the
Renée Crown University Honors Program at Syracuse University

Carlo F. Pisacane

Candidate for Bachelor's of Science in Computer Science
and Renée Crown University Honors
May 2025

Honors Thesis in Computer Science

Thesis Advisor: _____
        Dr. Nadeem Ghani, Assistant Teaching Professor

Thesis Reader: _____
        Dr. João Paulo Marum, Assistant Teaching Professor

Honors Director:_____
        Dr. Danielle Smith, Director

Abstract

A concise summary of your research, covering the problem (gaps in existing ASL recognition systems), your proposed solution (ASL recognition tool), methodology, key findings, and conclusions.

Thesis Advisor: Nadeem Ghani
Title: Assistant Teaching Professor, Engineering and Computer Science

Acknowledgments

I am deeply grateful for the support and guidance I have received throughout my academic journey at Syracuse University. I would like to extend my sincere thanks to my advisor, Prof. Nadeem Ghani, whose mentorship was instrumental in shaping my research and academic pursuits. My gratitude also goes to my thesis reader, Prof. João Marum, for his valuable insights and constructive feedback.

I am particularly thankful to Robin Smith, my Honors advisor, for motivating and encouraging me over the past four years during my time in the Honors program and throughout the thesis development process. Their support helped me overcome challenges and remain focused on my goals.

Finally, I am forever indebted to my friends and family, whose encouragement and belief in me have been a constant source of inspiration.

Preface

Why I chose this topic, personal insights, and what I learned, reference: CS example thesis.pdf (Kyle Maiorana)

Contents

Chapter 1

Introduction

Communication is a fundamental human need, and for Deaf and hard-of-hearing individuals, American Sign Language (ASL) serves as a primary mode of expression. ASL is a rich and complex visual language based on hand shape, movement, and nonmanual markers such as facial expressions. Studies emphasizing Deaf-centric design have shown that effective ASL tools must respect the cultural and linguistic practices of the Deaf community (Hibbard et al., 2020 [1]).

Technology has evolved from the past to make communication more accessible, including captioning services, text-based messaging, and video relay services. However, these solutions often rely on real-time interpreters or a shared written language, which can be limiting in spontaneous, in-person conversations.

In recent years, advances in computer vision and machine learning have opened new avenues for automated sign language recognition. Leveraging the power of advanced algorithms and increasingly pervasive hardware (e.g., mobile phone cameras, webcams), engineers and researchers hope to create machines that can recognize ASL hand gestures in real time. While some progress has been made in recognizing static signs or alphabets, challenges remain, especially regarding vocabulary expansion, nonmanual features, and real-world performance. The present work focuses on addressing these challenges by developing a user-centered ASL recognition tool that emphasizes accuracy, speed, and accessibility.

1.1 Research Problem

Current ASL recognition systems have several limitations that make them impractical. Some systems recognize only a limited set of signs, which restricts real-world applicability. Others do not account for nonmanual signs, such as facial expressions or head tilts, which are essential in ASL for the expression of tone, grammatical markers, and emotional context.

Additionally, some tools demand specialized sensors that are expensive or inconvenient, thus deterring widespread use.

There is also a broad gap in designing user interfaces that are attentive to the needs and desires of Deaf individuals. Inaccurate calibration procedures, variability of performance under changing lighting conditions, or excessive latency can lower the reliability of a system. These barriers compound to inhibit the real-world deployment potential of ASL recognition technology in everyday communication (Falvo et al., 2020 [2]).

1.2 Research Objectives

The overall goal of this thesis is to design a real-time ASL recognition system using computer vision and machine learning techniques. Specifically, the system must offer high recognition accuracy. The second aim is to enhance user experience by developing an accessible interface that is easy to install and requires minimal calibration or dedicated hardware. Additionally, the system needs to overcome environmental constraints by performing optimally in various lighting and backgrounds so that real environments can be utilized. Finally, the framework needs to be extensible for future additions of other signs, dynamic gestures, and nonmanual signals. By focusing on these core objectives, the system aims to be a foundation for broader applications in education, assistive technology, and inclusive communication devices.

1.3 Significance of the Study

This project holds the potential for shattering communication barriers among the Deaf and hard-of-hearing and for providing meaningful resources for hearing individuals who wish to learn ASL. An effective real-time ASL recognition system can facilitate communication more effectively in public places, schools, and workplaces, particularly where access to interpreters may not be readily available. It can also serve as an interactive learning tool for ASL

learners, offering instant feedback on handshapes to facilitate language learning. The proposed framework can further be extended to encompass the full richness of sign languages, thereby enabling future research studies. By prioritizing usability and involving Deaf/ASL communities in the design, this research underscores the importance of user-centered solutions.

1.4 Thesis Structure This thesis is organized into five main chapters:

1. **Chapter 1: Introduction**

   Provides the background and context of ASL recognition, defines the research problem, outlines objectives, and explains the study's significance.

2. **Chapter 2: Literature Review**

   Examines existing ASL recognition systems, machine learning techniques, and user-centered design principles. Identifies gaps in the current research and sets the stage for the proposed approach.

3. **Chapter 3: Software and Application**

   Details the methodology, including data collection, model architecture, and real-time inference pipeline. Discusses the design choices and rationale behind the system's implementation.

4. **Chapter 4: Results**

   Presents empirical findings, including model performance metrics, real-time testing results, and user feedback. Analyzes both quantitative and qualitative data.

5. **Chapter 5: Conclusion**

   Summarizes key insights, highlights contributions, and suggests avenues for future work. Reflects on the system's potential impact on accessibility and communication technologies.

Chapter 2

Literature Review

2.1 Introduction

This literature review surveys various methods and technologies employed in American Sign Language (ASL) recognition systems. The goal is to establish how our approach, aimed at real-time recognition and user-friendly interaction fits into the existing body of work. By examining both hardware-based and vision-based solutions, we highlight the fundamental challenges and opportunities in creating accessible tools for Deaf and hard-of-hearing communities.

2.2 Overview of Existing ASL Recognition Systems

Recent developments in sign language recognition often revolve around three main technological streams:

- Computer Vision Approaches

- Depth Sensor Approaches

- Wearable Technology Approaches

Each of these streams has strengths and limitations related to cost, accuracy, ease of deployment, and user comfort. This section reviews notable research in these areas and provides the groundwork for our own system's design choices.

2.2.1 Computer Vision Approaches

A substantial portion of ASL recognition research leverages standard RGB cameras and a variety of computer vision techniques. Traditional pipelines often involve skin detection,

feature extraction, and hand segmentation before proceeding to classification. More recently, robust libraries like *OpenCV* and frameworks such as *MediaPipe* have significantly simplified and improved hand detection and tracking.

One notable example is the *MediaPipe Hands* solution, an open-source tool by Google[1]. MediaPipe Hands tracks 21 key landmarks on each hand, as illustrated in Figure 1. This framework provides real-time tracking even under challenging conditions such as varied lighting or partial occlusions, which is crucial for the fluid and rapid gestures of sign language.
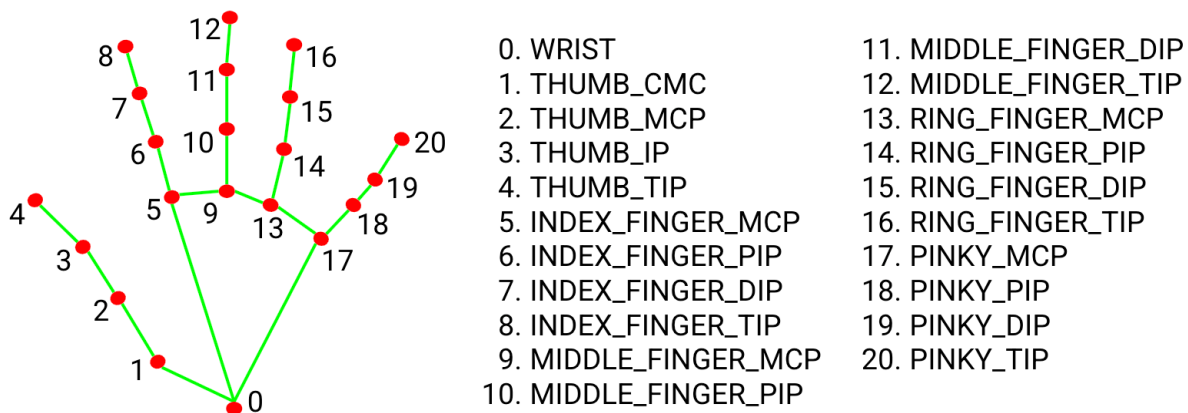

0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Figure 1: MediaPipe Hands landmarks. Each point corresponds to a specific joint or fingertip.

Researchers have successfully integrated MediaPipe's tracking features into sign language systems to reduce the complexity of manual feature engineering. For instance, Debnath and Joe [3] demonstrate a real-time ASL detection setup where MediaPipe is employed to extract hand pose information, thereby reliably differentiating between common ASL hand shapes and movements.

2.2.2 Depth Sensor Approaches

Beyond standard RGB cameras, *depth sensors* capture three-dimensional information about hand positions. Notably, Avola et al. [4] illustrate how the Leap Motion Controller can track the skeletal structure of the hand with fine-grained accuracy, enabling precise

---

[1] `https://github.com/google-ai-edge/mediapipe/blob/master/docs/solutions/hands.md`

6

measurement of joint angles. This capability is beneficial for distinguishing similar signs that differ only slightly in finger placement or orientation.

2.2.3 Wearable Technology Approaches

Finally, some research efforts rely on *wearable devices*, such as glove-based sensors, to record hand movements. These systems, often equipped with inertial measurement units (IMUs) or bend sensors, can capture motion data in three dimensions. For example, Stefanidis et al. [5] discuss the use of 3D technologies in sign language applications through wearable platforms. While potentially more accurate, such devices can be costly or cumbersome, limiting their broad adoption for everyday use.

2.3 Summary

In summary, contemporary ASL recognition solutions range from purely vision-based approaches (e.g., MediaPipe Hands combined with OpenCV for live detection) to more specialized hardware solutions (e.g., depth sensors and wearable gloves). Among these, live vision-based detection using MediaPipe and OpenCV offers the optimal balance of high accuracy, cost-effectiveness, and ease of integration, making it the best choice for designing robust, real-time ASL recognition systems that can be easily adopted in real-world contexts.

Chapter 3

Software and Application

Include text here

3.1

Include text here

Chapter 4

Results

Include text here

4.1

Include text here

Chapter 5

Conclusion

Include text here

5.1

Include text here

Bibliography

[1] E. Hibbard *et al.*, "Getting a Sign in Edgewise: User-Centered Design Considerations in Creating a Signed Language Mentoring Management System," *Sign Language Studies*, vol. 20, no. 2, pp. 264–300, 2020. Available: `https://www.jstor.org/stable/26983963`

[2] V. Falvo, L. P. Scatalon, and E. F. Barbosa, "The Role of Technology to Teaching and Learning Sign Languages: A Systematic Mapping," in *2020 IEEE Frontiers in Education Conference (FIE)*, Uppsala, Sweden, 2020, pp. 1–9, doi: `10.1109/FIE44824.2020.9274169`. Available: `https://ieeexplore-ieee-org.libezproxy2.syr.edu/document/9274169`

[3] J. Debnath and P. J. I R, "Real-Time Gesture Based Sign Language Recognition System," in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, Chennai, India, 2024, pp. 01–06, doi: `10.1109/ADICS58448.2024.10533518`. Available: `https://ieeexplore-ieee-org.libezproxy2.syr.edu/document/10533518`

[4] D. Avola, M. Bernardi, L. Cinque, G. L. Foresti, and C. Massaroni, "Exploiting Recurrent Neural Networks and Leap Motion Controller for the Recognition of Sign Language and Semaphoric Hand Gestures," *IEEE Transactions on Multimedia*, vol. 21, no. 1, pp. 234–245, Jan. 2019, doi: `10.1109/TMM.2018.2856094`. Available: `https://ieeexplore.ieee.org/document/8410764`

[5] Stefanidis, Kiriakos, Dimitrios Konstantinidis, Thanassis Kalvourtzis, Kosmas Dimitropoulos, and Petros Daras. "3D technologies and applications in sign language." 2020. Available: `https://www.researchgate.net/publication/340966069_3D_technologies_and_applications_in_sign_language`

[6] I. Papastratis *et al.*, "Artificial Intelligence Technologies for Sign Language," *Sensors (Basel, Switzerland)*, vol. 21, no. 17, p. 5843, Aug. 30, 2021, doi: `10.3390/s21175843`. Available: `https://pmc.ncbi.nlm.nih.gov/articles/PMC8434597/`

[7] A. Gulli and S. Pal, *Deep Learning with Keras*. Packt Publishing, 2018. Available: `https://books.google.com/books?id=20EwDwAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false`

[8] T. Liu, W. Zhou, and H. Li, "Sign language recognition with long short-term memory," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix,

AZ, USA, 2016, pp. 2871–2875, doi: `10.1109/ICIP.2016.7532884`. Available: `https://ieeexplore-ieee-org.libezproxy2.syr.edu/document/7532884`

[9] E. Alpaydin, *Introduction to Machine Learning*, The MIT Press, 4th ed., 2020. Available: `https://books.google.com/books?id=tZnSDwAAQBAJ&printsec=frontcover#v=onepage&q&f=false`

[10] K. Takahashi, "hand-gesture-recognition-mediapipe," GitHub repository, Available: `https://github.com/kinivi/hand-gesture-recognition-mediapipe` (Accessed: Feb. 26, 2025)

[11] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 15, 1997, doi: `10.1162/neco.1997.9.8.1735`. Available: `https://ieeexplore.ieee.org/document/6795963`

[12] S. Z. Gurbuz *et al.*, "Multi-Frequency RF Sensor Fusion for Word-Level Fluent ASL Recognition," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11373–11381, Jun. 15, 2022, doi: `10.1109/JSEN.2021.3078339`. [Online]. Available: `https://ieeexplore-ieee-org.libezproxy2.syr.edu/document/9425571`

[13] J. T. S. Ru and P. Sebastian, "Real-Time American Sign Language (ASL) Interpretation," in *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, Vellore, India, 2023, pp. 1–6, doi: `10.1109/ViTECoN58111.2023.10157157`. Available: `https://ieeexplore-ieee-org.libezproxy2.syr.edu/document/10157157`

[14] C.-H. Chuan, E. Regina, and C. Guardino, "American Sign Language Recognition Using Leap Motion Sensor," in *2014 13th International Conference on Machine Learning and Applications*, Detroit, MI, USA, 2014, pp. 541–544, doi: `10.1109/ICMLA.2014.110`. Available: `https://ieeexplore.ieee.org/document/7033173?arnumber=7033173`

[15] E. A. Borges-Galindo *et al.*, "Sign Language Interpreting System Using Recursive Neural Networks," *Applied Sciences*, vol. 14, no. 18, Art. no. 8560, Sep. 23, 2024, Available: `https://www.mdpi.com/2076-3417/14/18/8560`

[16] M. Soundarya, M. Yazhini, N. Thirumala Sree, N. Sornamalaya, and C. Vinitha, "Sign Language Recognition Using Machine Learning," in *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, Chennai, India, 2024, pp. 1–7, doi: `10.1109/ACCAI61061.2024.10602025`. Available: `https://ieeexplore.ieee.org/document/10602025`

[17] T. Liu, W. Zhou, and H. Li, "Sign language recognition with long short-term memory," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, 2016, pp. 2871–2875, doi: `10.1109/ICIP.2016.7532884`. Available: `https://ieeexplore-ieee-org.libezproxy2.syr.edu/document/7532884`

[18] C. K. M. Lee, K. K. H. Ng, C.-H. Chen, H. C. W. Lau, S. Y. Chung, and T. Tsoi, "American sign language recognition and training method with recurrent neural network," *Expert Systems with Applications*, vol. 167, pp. 114403, 2021, ISSN 0957-4174. Available: `https://doi.org/10.1016/j.eswa.2020.114403`

[19] C. Olah, "Understanding LSTM Networks," Understanding LSTM Networks, Aug. 17, 2015. [Online]. Available: `https://colah.github.io/posts/2015-08-Understanding-LSTMs/`