

AI Powered Sign Language Translator

Group members:

1. Habtamu Tadese
2. Sineshaw Legese
3. Arsema Mandefro
4. Meseret Bolled
5. Abem Tigist
6. Hanna Desalegn

Literature Review

1. Introduction

Sign languages represent a vital mode of communication for a substantial portion of the global population who are deaf or hard of hearing.[1] These languages are characterized by intricate systems of manual gestures, facial expressions, and body movements, forming complete and nuanced linguistic systems distinct from spoken languages.[1] However, individuals relying on sign language frequently encounter considerable obstacles in their daily interactions with the hearing majority, leading to limitations in access to essential aspects of life such as information, education, employment, and social engagement.[4] This communication divide underscores the critical need for effective solutions that can bridge this gap.

Traditionally, human sign language interpreters have played a crucial role in facilitating communication between the signing and non-signing communities. Nevertheless, the reliance on human interpreters presents challenges related to their availability, particularly in remote or emergency situations, as well as the significant costs associated with their services and the logistical complexities of scheduling.[6]

The recent advancements in artificial intelligence (AI) and its subfields, including computer vision, natural language processing, and machine learning, have generated substantial interest in their potential to automate the process of sign language translation.¹ These AI-driven systems aim to provide real-time, cost-effective solutions capable of translating sign language into spoken or written forms and vice versa, thereby promising to enhance accessibility and foster more inclusive communication.[4]

Given the rapid progress and increasing attention in this domain, this literature review seeks to provide a comprehensive overview of the current research landscape in AI-driven sign language translation. It aims to explore the various approaches and AI techniques being employed, identify the challenges that researchers are currently addressing, examine the datasets and resources that are available, discuss the metrics used to evaluate system performance, highlight the diverse applications and potential impact of this technology, and identify any existing open-source projects or publicly accessible tools. This review will serve as a foundational resource for those

embarking on projects in this field. The growing number of publications and projects, exemplified by initiatives like SignGPT.[13] receiving substantial funding, signifies the increasing momentum and importance of this research area.

2. AI-Driven Approaches for Sign Language Translation

Computer Vision-Based Approaches

A dominant paradigm in the field of AI-driven sign language translation involves the utilization of computer vision techniques.¹ These approaches primarily rely on cameras to capture the visual information of a person signing, focusing on the intricate movements of their hands, the expressive nuances of their facial expressions, and the overall dynamics of their body posture.¹ The advancements in deep learning have significantly propelled the capabilities of computer vision in this domain.^[1] Models such as Convolutional Neural Networks (CNNs) are adept at extracting spatial features from video frames, while Recurrent Neural Networks (RNNs), including LSTMs and GRUs, excel at modeling the temporal sequences inherent in sign language.^[1] Increasingly, Transformer networks are being employed for their ability to capture long-range dependencies in sign language.^[2] These systems often concentrate on identifying key visual features such as hand shape, position, movement, and non-manual markers, which are essential for conveying the complete meaning of a sign.^[1] For real-time applications, the integration of object detection algorithms like YOLOv8 and Faster R-CNN with landmark tracking libraries such as MediaPipe enables the accurate and rapid detection and tracking of hand gestures in video streams.^[2] Furthermore, computer vision plays a role in sign language generation through its integration with 3D avatars, which can synthesize signed videos from text or spoken language.^[10]

Sensor-Based Approaches

An alternative to computer vision involves the use of wearable sensors to capture the intricate movements of the hands and body during sign language communication.^[3] These sensor-based methods typically utilize devices such as data gloves equipped with flex sensors, inertial measurement units (IMUs) that track motion and orientation, and surface electromyography (EMG) sensors that detect muscle activity.^[3]

These sensors can directly measure parameters like finger bending, wrist position, and hand motion, which are then processed by AI algorithms for recognition and translation.[3] While sensor-based approaches may offer lower computational demands compared to vision-based methods, they often face limitations in cost, comfort, and the inability to capture non-manual features like facial expressions.[3] Recent advancements include the development of AI-powered rings with micro-sonar technology capable of tracking fingerspelling in American Sign Language.[19]

Hybrid Approaches

Recognizing the complementary strengths of both computer vision and sensor-based techniques, researchers are increasingly exploring hybrid approaches.[2] These methods aim to integrate data from multiple modalities to achieve more robust and accurate sign language translation by leveraging the detailed hand and body movement information from sensors alongside the contextual cues and facial expressions captured by computer vision.[2] Future work includes the potential integration of micro-sonar rings with eyeglasses to capture both fine-grained finger movements and upper body and facial expressions for a more comprehensive ASL translation system.[19]

3. Key AI Techniques in Sign Language Translation

Deep Learning

Deep learning has become the cornerstone of AI-driven sign language translation research.[1] Convolutional Neural Networks (CNNs) are fundamental for extracting spatial features from video frames, enabling the recognition of hand shapes and facial expressions.[1] Recurrent Neural Networks (RNNs), particularly LSTMs and GRUs, are crucial for modeling the temporal sequences of signs in continuous sign language.[1] The adoption of Transformer networks is growing due to their effectiveness in capturing long-range dependencies for end-to-end translation.[2] 3D-CNNs are also explored for extracting spatio-temporal features.[2]

Attention mechanisms enhance these architectures by allowing the model to focus on relevant input parts.[1] Transfer learning, utilizing pre-trained models, helps overcome data scarcity.[2]

Natural Language Processing (NLP)

NLP is crucial for understanding the linguistic structure of sign language, including glosses, grammar, and syntax.[1] NLP techniques are used for text pre-processing and generating spoken language output.[1] Neural Machine Translation (NMT) models translate between sign language glosses and spoken language.[2] Addressing grammatical differences between sign and spoken languages is a key focus.[1] Large Language Models (LLMs) are also being explored for sign language translation and generation.[7]

Machine Learning

Machine learning algorithms are applied for sign language recognition and classification.[3] Techniques like Support Vector Machines (SVMs) and k-Nearest Neighbors (KNN) are used.[1] Machine learning pipelines manage data processing, model training, and evaluation.[5] These techniques are also used for hand pose estimation and feature extraction.[1] Reinforcement learning is explored for training sequence-to-sequence models.[9]

4. Challenges in AI-Driven Sign Language Translation

Sign Language Variations and Linguistic Nuances

The existence of numerous distinct sign languages globally, each with its unique vocabulary, grammar, and syntax, presents a significant hurdle.[1] Variations within the same sign language due to dialects and individual styles further complicate the task. Grammatical differences between sign and spoken languages, and the crucial role of non-manual markers, add layers of complexity.[1] The visual-spatial nature of sign language also poses a challenge for AI models.[8]

Technical Challenges

Technical hurdles include handling hand occlusion, background noise, and lighting variations.[4] Segmenting signs in continuous sign language and achieving real-time processing are also significant challenges.[1] Variability in signing speed and style, and the complexity of representing the multi-dimensional nature of sign language, further complicate the task.[7]

Data Availability and Annotation

A major limitation is the scarcity of large, high-quality annotated datasets for many sign languages.[4] Collecting and annotating sign language data is time-consuming and requires expertise.[8] The lack of standardization in annotation formats also poses a challenge.[17] Efforts are underway to create larger and more diverse datasets through crowdsourcing and other methods.[13] Synthetic data generation and data augmentation techniques are also being explored.[2]

Ethical and Social Considerations

Involving Deaf stakeholders in the development process is crucial.[8] Addressing potential biases in datasets and algorithms is also essential.[12] The technology should meet the actual needs of the Deaf community.[7] The ethical implications of AI potentially replacing human interpreters need consideration.[11] The ultimate goal is to promote inclusivity and equal access for the Deaf community.[4]

5. Datasets and Resources for AI Sign Language Translation

Overview of Existing Datasets

A variety of datasets support research in AI-driven sign language translation, covering languages like ASL, BSL, and German Sign Language.[2] These datasets differ in size, vocabulary, annotation types (glosses, translations), and data modalities (RGB, depth).[2] Examples include RWTH PHOENIX-weather 2014T for German 1, How2Sign for ASL 37, and ASL Fingerspelling Recognition Corpus.[23] WLASL is another significant ASL dataset.[15] Data scarcity, especially for low-resource sign languages, remains a challenge.[2]

Annotation Tools and Methodologies

Glosses are commonly used as intermediate representations.[1] Videos are annotated with bounding boxes and keypoints.[2] Accurate and consistent annotation is challenging.[12] Tools and platforms are being developed for data annotation.[13]

Resources for Different Sign Languages

Dataset and resource availability varies across sign languages, with a focus on ASL and BSL.[2]

Efforts are growing to develop resources for other, including low-resource, sign languages.[2]

Table 1: Key Datasets for AI Sign Language Translation

Dataset Name	Sign Language(s)	Size (Approx.)	Annotation Type(s)	Modality(ies)	Publicly Available
RWTH-PHOENIX-Weather 2014T	German	~53 GB	Glosses, Translations	RGB Video	Yes
How2Sign	ASL	~80 hours	Transcripts, Depth, Keypoints	RGB Video, Depth	Yes (Research Only)
ASL Fingerspelling Recognition Corpus	ASL	N/A	Hand & Facial Landmarks	RGB Video	Yes
WLASL	ASL	N/A	Person Bounding Box, Dialect	YouTube Video	Yes
BSL-1K	BSL	N/A	Glosses	RGB Video	Yes
Auslan-Daily	Auslan	N/A	Translations	RGB Video	Yes
YouTube-ASL	ASL	Large Scale	Parallel Corpus	YouTube Video	Yes

ASL Citizen	ASL	~83K videos	Glosses	Webcam Video	Yes
LSA-T	Argentinian	~45 GB	Joint Positions	RGB Video	Yes
CSL-Daily	Chinese	N/A	Glosses, Translations	RGB Video	Yes
AUTSL	Turkish	~226 hours	Glosses, Translations, BB, KP	RGB Video, Depth	Yes
OpenASL	ASL	~29.5 hours	Glosses	RGB Video	Yes
SignBank+	Multilingual	N/A	SignWriting Tokens	Images	Yes
Spread the Sign (STS)	Multilingual	~50GB	Translations	RGB Video	Yes (with conditions)
MPI-ISL	German	~1000 hours	Glosses, Translations, BB, KP	RGB Video, Depth, IMU	Yes (with conditions)
Dicta-Sign	Greek	~100 hours	Glosses, Translations	RGB Video	Yes (with conditions)

Greek Elementary School Dataset	Greek	~30K pairs	Translations	RGB Video	Yes
INCLUDE	Indian	N/A	Glosses	RGB Video	Yes

6. Performance Metrics for Evaluating Sign Language Translation Systems

Commonly Used Metrics

BLEU score is widely used for evaluating the fluency and adequacy of sign language to spoken language translation.[2] Word Error Rate (WER) measures the accuracy of sign language recognition.[1] Accuracy, precision, recall, and F1 score assess sign recognition and classification.[1] Sign Error Rate (SER) is specific to sign language recognition.[25] Mean Average Precision (mAP) evaluates object detection-based systems.[14]

Suitability and Limitations of Metrics

BLEU score's correlation with human judgment has been discussed.[24] It has limitations in capturing semantic meaning. Factors like latency and user satisfaction are important for real-time systems.[18] New metrics like SignBLEU aim to better capture the multi-channel nature of sign language.[22]

7. Applications and Potential Impact of AI Sign Language Translators

Diverse Applications Across Various Sectors

AI sign language translators have diverse applications in education, healthcare, accessibility services, social communication, and emergency situations.[3]

Impact on the Lives of Deaf and Hard-of-Hearing Individuals and Society

These translators break down communication barriers and promote social inclusion, increasing autonomy and access to information.[4] This can improve educational and employment outcomes, contributing to a more equitable society.[5]

8. Existing Open-Source Projects and Publicly Available Tools

Identification of Open-Source Projects

GitHub hosts numerous open-source projects related to AI for sign language, including lists of papers and code.[21] Specific projects for recognition and translation are available.[16] Open-source libraries and frameworks for building custom translators exist.[20] Community-driven efforts focus on data collection.[16]

Publicly Available Tools and Platforms

Online sign language translation applications and platforms are increasingly available.[18] Web-based demonstrations and mobile applications also exist.[13] Platforms allow contribution to sign language datasets.[23]

9. Conclusion and Future Research Directions

Research in AI-driven sign language translation has made significant strides, utilizing computer vision, sensor-based, and hybrid approaches, powered by deep learning, NLP, and machine learning techniques. While challenges remain in addressing sign language variations, technical limitations, and data scarcity, the field has witnessed the development of numerous datasets, resources, and evaluation metrics. The applications of this technology are vast, promising to revolutionize communication for the deaf and hard-of-hearing community and foster a more inclusive society.

Future research should focus on improving the accuracy and robustness of translation systems, particularly for under-resourced sign languages and in uncontrolled environments. Enhancing the capture and interpretation of non-manual markers and the visual-spatial aspects of sign language is crucial. The development of more comprehensive and standardized datasets with diverse signers and annotations is also needed. Creating user-friendly tools and platforms for real-world applications, addressing ethical considerations through community involvement, and exploring the integration of AI sign language translation with other assistive technologies present exciting avenues for future investigation. Novel AI architectures and learning paradigms hold the potential to further advance the field. Continued interdisciplinary collaboration involving AI

researchers, linguists, and the Deaf community is essential to realize the full potential of AI in breaking down communication barriers.

Reference

1. Automated Sign Language Translation: The Role of Artificial Intelligence Now and in the Future - SciTePress, <https://www.scitepress.org/Papers/2020/101438/101438.pdf>
2. Sign Language Translation: A Survey of Approaches and Techniques, <https://www.mdpi.com/2079-9292/12/12/2678>
3. A Survey of Advancements in Real-Time Sign Language Translators: Integration with IoT Technology - MDPI, <https://www.mdpi.com/2227-7080/11/4/83>
4. Artificial Intelligence Technologies for Sign Language - PMC, <https://pmc.ncbi.nlm.nih.gov/articles/PMC8434597/>
5. AI Powered Sign Language Translator - ResearchGate, https://www.researchgate.net/publication/375601688_AI_Powered_Sign_Language_Translator
6. Engineers Bring Sign Language to 'Life' Using AI - Florida Atlantic University, <https://www.fau.edu/newsdesk/articles/american-sign-language.php>
7. Towards AI-Driven Sign Language Generation with Non-Manual Markers, <https://machinelearning.apple.com/research/ai-sign-language-generation>
8. Why Sign Language Translation Poses a Puzzle for AI - Slator, <https://slator.com/why-sign-language-translation-poses-a-puzzle-for-ai/>
9. Artificial Intelligence for Sign Language Translation -A Design Science Research Study, https://www.researchgate.net/publication/370215194_Artificial_Intelligence_for_Sign_Language_Translation_-_A_Design_Science_Research_Study
10. AUDIO TO SIGN LANGUAGE TRANSLATION USING AI - IJPREMS International Journal of Progressive Research in Engineering Management and Science,

- https://www.ijprems.com/uploadedfiles/paper//issue_11_november_2024/37044/final/fin_i_jprems1732604955.pdf
11. Will the profession ASL interpreter die out due to AI? : r/deaf - Reddit,
https://www.reddit.com/r/deaf/comments/lisk1v7/will_the_profession_asl_interpreter_die_out_due/
 12. Systemic Biases in Sign Language AI Research: A Deaf-Led Call to Reevaluate Research Agendas - arXiv, <https://arxiv.org/html/2403.02563v1>
 13. Research Initiative Aims to Revolutionize Sign Language Translation with AI,
<https://hearingreview.com/inside-hearing/research/research-initiative-aims-to-revolutionize-sign-language-translation-with-ai>
 14. Study Uses AI to Interpret American Sign Language in Real-time - Florida Atlantic University, <https://www.fau.edu/newsdesk/articles/artificial-intelligence-sign-language-study.php>
 15. AI-Based ASL Translation: A Study of Deep Learning and 3D Modeling - ResearchGate,
https://www.researchgate.net/publication/379020267_AI-Based_ASL_Translation_A_Study_of_Deep_Learning_and_3D_Modeling
 16. AI Sign Language Translation isn't as far away as you think | by João Pedro Santos,
<https://medium.com/@santosjoaopedro/ai-sign-language-translation-isnt-as-far-away-as-you-think-91d1a29ddff6>
 17. Trends and Challenges for Sign Language Recognition with Machine Learning,
https://www.researchgate.net/publication/373998129_Trends_and_Challenges_for_Sign_Language_Recognition_with_Machine_Learning
 18. Signapse AI | Sign Language Technology, <https://www.signapse.ai/>

19. AI ring tracks spelled words in American Sign Language | Cornell Chronicle,
<https://news.cornell.edu/stories/2025/03/ai-ring-tracks-spelled-words-american-sign-language>
20. sign-language-translator/sign-language-translator: Python library & framework to build custom translators for the hearing-impaired and translate between Sign Language & Text using Artificial Intelligence. - GitHub, <https://github.com/sign-language-translator/sign-language-translator>
21. binbinjiang/SL_Papers: Latest AI Sign Language Papers ... - GitHub,
https://github.com/binbinjiang/SL_Papers
22. SignBLEU: Automatic Evaluation of Multi-channel Sign Language Translation - ACL Anthology, <https://aclanthology.org/2024.lrec-main.1289.pdf>
23. Google - American Sign Language Fingerspelling Recognition - Kaggle,
<https://kaggle.com/competitions/asl-fingerspelling>
24. Understanding MT Quality: BLEU Scores - ModernMT Blog,
<https://blog.modernmt.com/understanding-mt-quality-bleu-scores/>
25. Sign language recognition using modified deep learning network and hybrid optimization,
<https://pmc.ncbi.nlm.nih.gov/articles/PMC11525556/>

Data Research

Introduction

To develop the AI-powered sign language translator with accurate and inclusive results, there needs to be a thorough exploration of data. Ensuring the models are trained properly with diverse and representative datasets is very crucial. This project aims to bridge the communication gaps between the deaf and hard-of-hearing community, also those who do not understand sign language. Delivering this goal requires a deep exploration of datasets. Even though datasets for Amharic sign language are limited, we will explore and study the available datasets that can help us achieve our aim.

Organization

The data research findings will be organized thematically, focusing on the following key areas:

- Data Sources: Overview of datasets used.
- Data Description: Detailed characteristics of the datasets.
- Data Analysis: Insights and trends identified from the data.

Data Description

1. Amharic Sign Language Recognition Dataset

Dataset Name: Amharic Sign Language Recognition Dataset

Data Source: Collected from various sign language teachers, this dataset is aimed at recognizing Amharic Sign Language gestures and translating them into Amharic characters.

Format: The dataset includes video files along with annotated images. The annotations are in XML, TFRecord, and CSV formats, created using the LabelImg tool.

File Types: Videos are in standard formats (e.g., MP4), while images are typically in formats like JPEG or PNG.

Total Size: The dataset is approximately 298 MB, containing numerous video clips and associated image frames.

Number of Classes: It includes 10 classes for recognizing specific characters in Amharic Sign Language, with plans to expand to cover additional signs, words, and sentences.

Data Structure

Organization: The data is organized into videos and images, with corresponding annotations for each gesture. This structure facilitates easy access for training and testing machine learning models.

Annotations: The dataset is annotated to indicate the signs being performed, which is essential for supervised learning tasks.

Preprocessing Steps

Frame Extraction: Video files are processed to extract individual frames, making it easier to analyze and annotate specific gestures.

Resizing: The frames are resized to ensure uniformity for model training.

Feature Extraction: Key features from the frames are extracted to aid in gesture recognition.

Model Training

The dataset has been used with two different models: **Faster R-CNN** and **Single-Shot Multibox (SSD) Detector**.

Results: The SSD model demonstrated better accuracy in recognizing Amharic Sign Language gestures compared to Faster R-CNN, which performed well but with slightly lower accuracy.

Rationale for Data Choice

Relevance: This dataset is particularly relevant for the project as it focuses on Amharic Sign Language, addressing a significant gap in existing resources for non-Western sign languages.

Diversity: The data includes gestures from different backgrounds and hand positions, enabling the model to learn from a wide variety of signing styles.

Potential Applications

This dataset can be utilized to develop real-time sign language recognition systems that translate gestures into text, thereby facilitating communication between hearing-impaired individuals and those who do not understand sign language.

2. A Generic Approach towards Amharic Sign Language Recognition

- Source: Collected from signers at Addis Ababa University, the Department of Linguistics, and the Ethiopian Center for Disability and Development (ECDD).
- Sampling Technique: Nonprobabilistic convenience sampling from voluntarily participating deaf individuals.
- Content: Includes all Amharic alphabets, numeric characters (1-7), and two borrowed characters.
- Data Summary:
 - Total classes: 240
 - Total extracted frames: 26,424
 - Selected for a demonstration: 15 words with 2,475 images.
 - Data Split: 80% training, 10% validation, 10% testing.

Dataset Preparation

- Video Processing: Converted videos into sequences of frames.
- Image Preprocessing:
 - Cropping, resizing (from 3264×2448 to 2000×1200 and 1280×720 to 500×400).
 - Color conversion from RGB to BGR using OpenCV.

Feature Extraction

- Method: Utilized a hybrid CNN-LSTM architecture with pretrained models (EfficientNetB0 and ResNet50) for feature extraction.
- Features: Hand shape, motion, and color; extracted using contour-based and region-based techniques.

- Training Strategy: Incremental learning to manage memory usage by training with batches of data (input shapes: $224 \times 224 \times 3$ and $150 \times 150 \times 3$).



Conclusion

In conclusion, this data research for the AI-powered sign language translator has provided valuable insights into how we can improve communication for the deaf and hard-of-hearing community. By examining various datasets, we found that using diverse and high-quality data is crucial for training accurate models that can recognize and translate sign language gestures in real time.

Our review of existing studies showed that advanced machine learning techniques, like CNNs and LSTMs, are effective for this purpose. This research highlights the need for careful data selection to enhance model performance.

Overall, the findings not only contribute to the technical development of the sign language translator but also support the goals of accessibility and inclusivity. By breaking down communication barriers, we hope to promote understanding and equal opportunities for everyone. The insights from this research will guide the next steps in creating a tool that truly meets the needs of its users.

References

<https://onlinelibrary.wiley.com/doi/10.1155/2022/1112169>

<https://data.mendeley.com/datasets/5d3nkyhsrf/1>

https://www.researchgate.net/figure/Ethiopian-sign-language-alphabets_fig1_359650139

Technology Review

1. Introduction

The AI-powered sign language translator project aims to solve communication barriers between the deaf or hard-of-hearing community and those who do not understand sign language. This technology review explores the core tools and frameworks that will be used to develop an accessible, real-time translator using computer vision and deep learning. A review of existing technologies is essential to ensure we use effective, scalable, and inclusive tools that align with our project's goals. The technologies discussed here are selected based on practicality, performance, and alignment with our team's technical strengths.

2. Technology Overview

- **OpenCV:** A powerful open-source library for real-time computer vision tasks. It will be used for video and image preprocessing, hand detection, and frame extraction. OpenCV is highly efficient for low-latency applications.
- **PyTorch:** A flexible deep learning framework used to build and train our model. PyTorch offers dynamic computational graphs, simple syntax, and strong support for research-based development.
- **TensorFlow (Alternative):** An industry-standard framework known for scalable deployment. It can serve as a backup option depending on performance benchmarks and deployment needs.
- **CNN + LSTM Architecture:** Convolutional Neural Networks (CNNs) will handle spatial feature extraction from video frames, while Long Short-Term Memory (LSTM) networks will analyze the temporal flow of gestures.

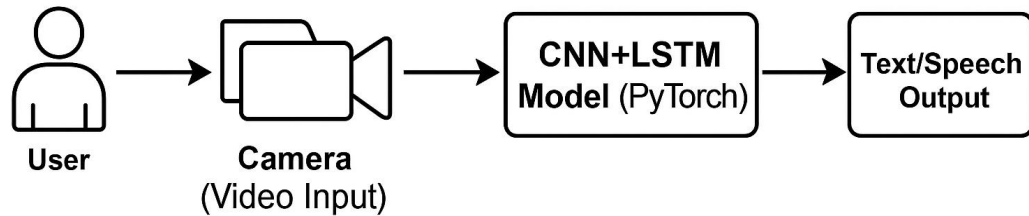


Figure 1: System architecture for AI-powered sign language translator

- Google Colab: Development environment that provides free GPU access (e.g., Tesla T4, K80). Ideal for training deep learning models in a collaborative and resource-efficient way.
- Flask: A lightweight Python web framework used to deploy the model as an API endpoint for real-time translation. It supports fast and easy integration with the front-end UI.
- Amharic Sign Language Dataset: This project will use a publicly available or custom-created Amharic Sign Language dataset, which includes labeled video frames or sequences of gestures. It is essential for local relevance and inclusivity.

3. Relevance to the Project Each technology selected directly addresses a challenge within the sign language translator pipeline:

- OpenCV handles the complexity of preprocessing raw video input.
- CNN + LSTM models are specifically suited for visual-to-sequence tasks like sign recognition.

- PyTorch facilitates fast prototyping and experimentation with different model architectures.
- Google Colab ensures development can proceed without heavy local hardware.
- Flask makes deployment easy and scalable. These technologies are all lightweight, cost-effective, and support real-time performance, which is crucial for accessibility tools.

4. Comparison and Evaluation

Technology	Strengths	Weaknesses
OpenCV	Fast, well-documented, real-time processing	Lacks deep learning integration
PyTorch	Flexible, easy to debug, research-friendly	Slightly less optimized for mobile deployment
TensorFlow	Production-ready, scalable	More complex syntax for beginners
Flask	Lightweight, easy to use, Python-based	Not ideal for high-traffic APIs

5. Cases and Examples

- Zhang et al. (2021): Used CNN + LSTM for ASL recognition with high accuracy and real-time feedback.
- Sharma et al. (2020): Developed a translator using CNN + RNN architectures that outperformed traditional image processing techniques.
- Google's Teachable Machine: Offers visual classification models through transfer learning, demonstrating the feasibility of gesture recognition with minimal training. These examples show that combining deep learning and computer vision for gesture translation is both effective and scalable.

6. Gaps and Research Opportunities

Most existing models focus on American Sign Language (ASL), which makes localization for Amharic Sign Language both a challenge and an opportunity. Challenges include limited labeled datasets and lack of research on regional gesture variations. This project contributes to filling that gap by training models specifically on Amharic gestures, creating new datasets if necessary, and optimizing performance for low-resource environments.

7. Conclusion

This technology review confirms that the selected stack—OpenCV, PyTorch, CNN+LSTM, Google Colab, Flask, and Amharic Sign Language dataset—is well suited for the development of a real-time sign language translator. These tools are chosen not just for their popularity, but for their specific alignment with the technical and social goals of this project. By integrating modern AI frameworks and local language support, this solution will make a meaningful impact on accessibility, inclusivity, and communication equity.

8. References

- Zhang et al. (2021). Real-time Sign Language Recognition Using CNN-LSTM.
- Sharma et al. (2020). Deep Learning-Based Sign Language to Text Translation.
- OpenCV Documentation. <https://opencv.org>
- PyTorch Documentation. <https://pytorch.org>
- TensorFlow Documentation. <https://www.tensorflow.org>
- Flask Documentation. <https://flask.palletsprojects.com>
- Google Colab. <https://colab.research.google.com>