Neural Network & Deep Learning

TITLE : Enhancing Communication through AI and deep learning in Sign Language Translation

Submitted by: M V S Satish Reddy

700: 700755869

Author: Hanna ZainEldin, Mohamoud Badawy

Motivation :

Bridging the Communication Gap

- **Accessibility:**
  - Traditional SLT methods have limitations in availability and access, especially in remote areas or for individuals with limited resources.
  - AI-powered solutions offer on-demand and remote accessibility, overcoming geographical barriers.
- **Efficiency and Speed:**
  - AI systems can significantly speed up communication by automating the translation process.
  - Real-time translation capabilities are crucial in dynamic environments like meetings or emergency situations.
- **Scalability :**
  - The demand for SLT services is increasing, and AI offers scalable solutions to address the shortage of qualified interpreters.
  - AI-driven systems can handle large volumes of translation requests efficiently.
- **Personalization:**
  - AI algorithms can adapt to individual user preferences and learning styles, offering personalized feedback and recommendations.
  - This personalization improves user experience and learning outcomes.
- **Innovation:**
  - Investing in AI-powered SLT drives advancements in assistive technology and accessibility research.
  - Integration with AR and wearable devices can further enhance communication experiences.

Objectives:

<u>Building a Comprehensive Translation System</u>

- **Speech-to-Sign Language Translation:**
    - Develop accurate speech recognition algorithms to transcribe spoken language.
    - Utilize NLP techniques to understand the semantic meaning of the transcribed text.
    - Employ computer vision algorithms to generate corresponding sign language gestures.
- **Text-to-Sign Language Translation:**
    - Develop machine translation models to convert written text into sign language glosses or descriptions.
    - Use animation techniques to create lifelike sign language gestures from translated text.
    - Incorporate linguistic rules and cultural considerations for accurate and appropriate translations.
- **Sign Language to Speech/Text Translation:**
    - Develop robust computer vision algorithms for sign language gesture recognition.
    - Integrate natural language generation techniques to produce written text or synthesized speech.
    - Adapt to various sign language dialects and variations for broader applicability.
    - Include accessibility features like text-to-speech for spoken language output.

Problem Statement:

<u>Overcoming Communication Barriers and Accessibility Gaps</u>

- **Limited Accessibility of Traditional SLT Methods:** Highlight the restricted access to qualified interpreters, especially in remote areas or for specific languages.
- **Inefficiency and Time Constraints:** Emphasize the time-consuming nature of traditional translation methods, hindering real-time communication and responsiveness.
- **Scalability Issues:** Address the challenges in meeting the growing demand for sign language translation services due to limited resources and interpreter availability.
- **Lack of Personalization:** Point out the inability of traditional methods to adapt to individual learning styles and preferences, potentially hindering effective communication and learning outcomes.
- **Technological Gaps in Existing AI Solutions:** Mention the limitations of current AI-driven solutions in terms of accuracy, inclusivity, and adaptability to diverse sign languages and signing styles.

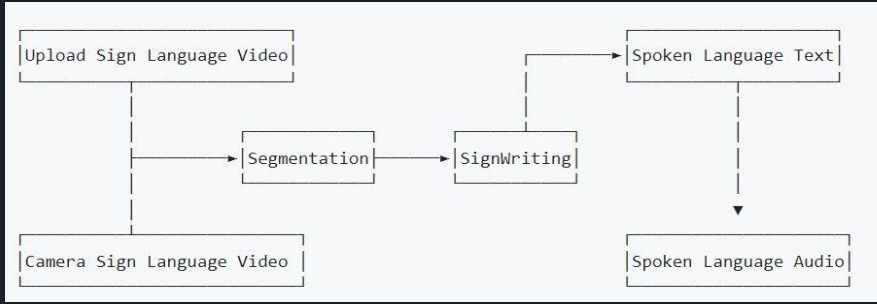
Critical Analysis

- Artificial Intelligence (AI) has revolutionized communication technologies, significantly enhancing accessibility for individuals with hearing and speech impairments. In particular, AI-driven innovations like sign language recognition (SLR) have bridged communication gaps between sign language users and those unfamiliar with sign language. Techniques such as deep learning (DL) and machine learning (ML) effectively interpret sign language gestures, converting them into text or spoken words.
- Despite these advancements, several challenges persist. Variability among sign languages, limited availability of annotated datasets, and the need for real-time processing capabilities are major obstacles affecting the effectiveness of SLR systems. Continued research and development efforts are crucial to improve inclusivity and performance in these technologies. Ethical considerations, including fairness, privacy, and addressing biases in training datasets, are paramount for responsible AI deployment in this domain.
- Moreover, integrating visual and audio data through multimodal approaches holds promise for enhancing communication systems, catering to diverse user needs more effectively. Future advancements should prioritize personalized recognition algorithms that accommodate individual signing styles, thereby fostering greater inclusivity within the deaf and mute community.
- Addressing these challenges is essential for advancing SLR technologies and promoting a more equitable society. While AI offers significant potential in improving communication accessibility, it is imperative to address existing challenges and ethical considerations to ensure these technologies serve all users effectively and fairly.

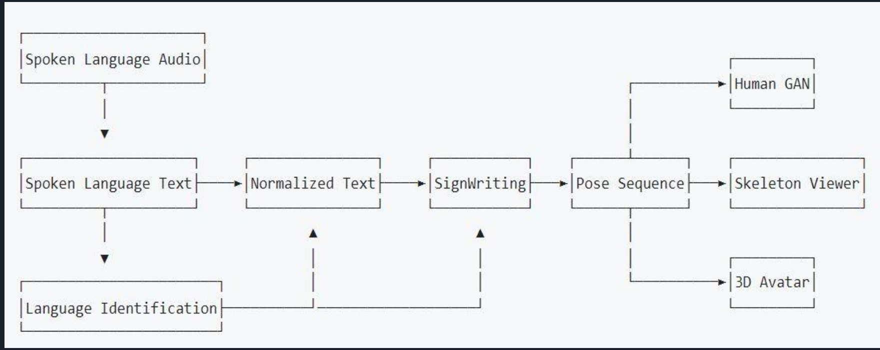
Related work:

Learning from Existing Research

- ## Sign Language Segmentation:
  - Explain current segmentation methods and challenges due to the simultaneous nature of sign language gestures.
  - Stress the importance of understanding sign language structure for improved segmentation techniques.



- ## Sign Language Recognition, Translation, and Production:
  - Discuss previous work on isolated sign recognition and continuous signing sequence recognition.
  - Highlight the advancements and limitations of existing sign language translation and production systems.



Opensourced Datasets :

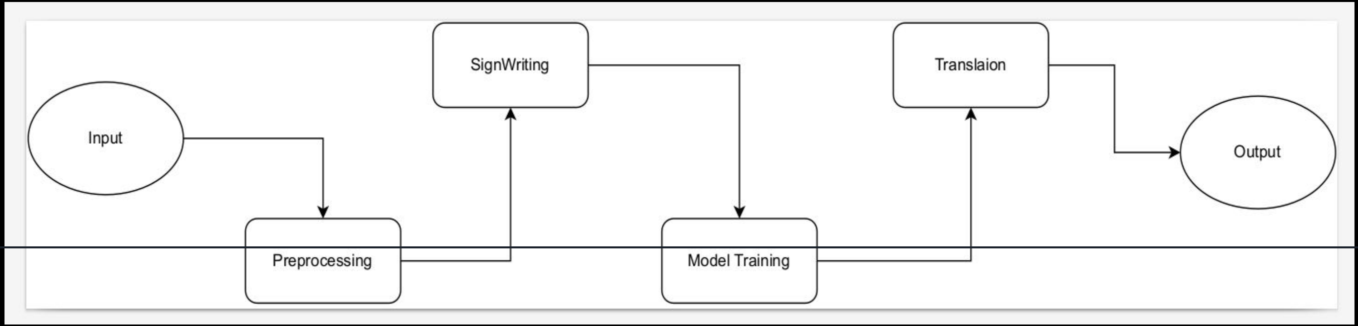
Dataset	Videos	Poses	Versions
aslg_pc12	N/A	N/A	0.0.1
asl-lex	No		2.0.0
rwth_phoenix2014_t	Yes	Holistic	3.0.0
autsl	Yes	OpenPose, Holistic	1.0.0
dgs_corpus	Yes	OpenPose, Holistic	3.0.0
dgs_types	Yes		3.0.0
how2sign	Yes	OpenPose	1.0.0
sign2mint	Yes		1.0.0
signtyp	Links		1.0.0
swojs_glossario	Yes		1.0.0
SignBank	N/A		1.0.0
wlasl	Failed	OpenPose	None
wmtslt	Yes	OpenPose, Holistic	1.2.0
signsuisse	Yes		1.0.0
msasl			None
Video-Based CSL			None
RVL-SLLL ASL			None
ngt_corpus	Yes		3.0.0
bsl_corpus	No	No	3.0.0

Table 1: *Summary of signed language datasets*

Several publicly available datasets can be leveraged for training sign language translation models. Below are datasets selected to evaluate the models used in our project

- Speech-to-Sign Language Translation
- Text-to-Sign Language Translation
- Sign Language-to-Text/Speech Translation

Proposed Solution:



Key Components:

- **Data Acquisition & Preprocessing**
- **SignWriting Representation**
- **Machine Translation Models**
- **API Server**

Experimentation:

- **Bilingual ASL-to-English Translation :**
  Model: "baseline transformer spm factor sign+"
  Dataset: SignBank
  Framework: Joey NMT
  Evaluation Metric: BLEU score
  Results: BLEU score exceeding 30, indicating strong performance in translating ASL to English.

- **Comparison with Alternative Approaches:**
  1. Direct comparisons were challenging due to variations in data, evaluation methods, and methodologies.
  2. The SignWriting-based approach appears competitive, potentially offering more accurate and nuanced translations compared to systems relying on glosses or other intermediate representations.

Referenc
es

1 R. E. Johnson and S. K. Liddell, “Toward a phonetic representation of signs: Sequentiality and contrast,” Sign Language Studies, vol. 11, no. 2, pp. 241–274, 2011.
2 W. Sandler, “The phonological organization of sign languages,” Language and linguistics compass, vol. 6, no. 3, pp. 162–182, 2012.
3 D. Brentari and C. Padden, “A language with multiple origins: Native and foreign vocabulary in american sign language,” Foreign vocabulary in sign language: A cross-linguistic investigation of word formation, pp. 87–119, 2001.
4 A. Moryoussef, “sign.mt: Effortless real-time sign language translation,” https://sign.mt/, 2023.
5 M. Borg and K. P. Camilleri, “Sign language detection “in the wild” with recurrent neural networks,” in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019, pp. 1637–1641.
6 A. Moryoussef, I. Tsochantaridis, R. Aharoni, S. Ebling, and S. Narayanan, “Real-time sign language detection using human pose estimation,” in Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16. Springer, 2020, pp. 237–248.
7 A. Pal, S. Huber, C. Chaabani, A. Manzotti, and O. Koller, “On the importance of signer overlap for sign language detection,” arXiv preprint arXiv:2303.10782, 2023.
8 B. G. Gebre, P. Wittenburg, and T. Heskes, “Automatic sign language identification,” in 2013 IEEE International Conference on Image Processing. IEEE, 2013, pp. 2626–2630.
9 C. D. Monteiro, C. M. Mathew, R. Gutierrez-Osuna, and F. Shipman, “Detecting and identifying sign languages through visual features,” in 2016 IEEE International Symposium on Multimedia (ISM). IEEE, 2016, pp. 287–290.
10 D. McKee and G. Kennedy, “Lexical comparison of signs from american, australian, british and new zealand sign languages,” The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima, vol. 20, no. 0, p. 0, 2000.
11 P. Santemiz, O. Aran, M. Saraclar, and L. Akarun, “Automatic sign segmentation from continuous signing via multiple sequence alignment,” in 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops. IEEE, 2009, pp. 2001–2008.

Referenc
es

12 I. Farag and H. Brock, "Learning motion disfluencies for automatic sign language segmentation," in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019, pp. 7360–7364.

13 W. Sandler, "Prosody and syntax in sign languages," Transactions of the philological society, vol. 108, no. 3, pp. 298–328, 2010.

14 M. De Sisto, D. Shterionov, I. Murtagh, M. Vermeerbergen, and L. Leeson, "Defining meaningful units. challenges in sign segmentation and segment-meaning mapping," https://aclanthology. org/volumes/2021. mtsummit-at4ssl/, pp. 98–103, 2021.

15 A. Imashev, M. Mukushev, V. Kimmelman, and A. Sandygulova, "A dataset for linguistic understanding, visual evaluation, and recognition of sign languages: The k-rsl," in Proceedings of the 24th Conference on Computational Natural Language Learning, 2020, pp. 631–640.

16 O. M. Sincan and H. Y. Keles, "Autsl: A large scale multi-modal turkish sign language dataset and baseline methods," IEEE access, vol. 8, pp. 181 340–181 355, 2020.

17 R. Cui, H. Liu, and C. Zhang, "Recurrent convolutional neural networks for continuous sign language recognition by staged optimization," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7361–7369.

18 N. C. Camgoz, S. Hadfield, O. Koller, H. Ney, and R. Bowden, "Neural sign language translation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7784–7793.

19 M. De Coster, D. Shterionov, M. Van Herreweghe, and J. Dambre, "Machine translation from signed to spoken languages: State of the art and challenges," Universal Access in the Information Society, pp. 1–27, 2023.

20 M. Muller, S. Ebling, E. Avramidis, A. Battisti, M. Berger, ¨ R. Bowden, A. Braffort, N. C. Camgoz, C. Espa ¨ na-Bonet, ˜ R. Grundkiewicz et al., "Findings of the first wmt shared task on sign language translation (wmt-slt22)," in Proceedings of the Seventh Conference on Machine Translation (WMT), 2022, pp. 744– 772.

21 A. Moryoussef and Z. Jiang, "Signbank: Multilingual sign language translation dataset," arXiv preprint arXiv:2309.11566, 2023.

Thank you :)