1. An Application for Translating Sign Language

2. Overview

This project will use various datasets and object detection to build an application which will be used for translating American Sign Language (ASL).

3. Background

Navigating the Signified Landscape: Challenges and Opportunities in ASL Translation

The quest to connect spoken and sign languages takes an exciting turn with American Sign Language (ASL). While progress in translating BSL and LSF is noteworthy, ASL presents unique challenges. Its intricate tapestry of grammar, encompassing spatialization, non-manual features, and dynamic handshape modifications, demands translation models that go beyond simple gesture recognition (Kita & Liddell, 2013).

Deciphering the Puzzle:

Data Scarcity and Diversity: Compared to other sign languages, large-scale ASL video datasets are limited and often lack representation of diverse signers, environments, and handshapes, hindering generalizability (Koller et al., 2016).

Object Detection and Grammatical Nuance: While object detection excels at pinpointing individual handshapes, capturing the subtle syntax and temporal dependencies within ASL sequences remains a hurdle. Recent work by Martinez et al. (2020) utilizing recurrent neural networks shows promise in bridging this gap.

Building upon Breakthroughs:

This project draws inspiration from pioneering research, including:

Abiyev et al.'s (2023) work on Sign Language Translation Using Deep Convolutional Neural Networks: This research delves into DCNNs for hand gesture recognition, laying a foundation for accurate identification within ASL sequences.

Kumar et al.'s (2019) exploration of 3D hand pose estimation using CNNs: This provides valuable insight into precise handshape localization within ASL sequences.

Feng et al.'s (2017) investigation of recurrent neural networks for sign language phrase recognition: This offers valuable lessons in capturing the temporal dynamics of ASL grammar.

Pushing the Boundaries:

We aim to significantly advance ASL translation by:

Building a comprehensive and diverse ASL video dataset: Encompassing a wider range of signers, contexts, and handshapes to address data scarcity and enhance generalizability.

Leveraging cutting-edge object detection and deep learning techniques: Not only identifying individual signs but also analyzing their temporal and spatial relationships to capture ASL's intricate grammar with greater accuracy.

By navigating the complexities of ASL and harnessing the power of advanced technology, this project aspires to break down communication barriers and empower ASL speakers to fully participate in a language-inclusive world.

This project aims to develop an innovative application that aims to facilitate communication for the hearing impaired or individuals who want to learn sign language. Nowadays, the hearing impaired often faces difficulties when communicating, which negatively impacts their daily lives. To overcome these difficulties, the application will offer a feature to understand the sign language that users communicate through images from the camera, using a specially created data set focused on sign language. At the same time, the app will recognize the user's gestures using sign language and translate those gestures into a predictive sentence or word, making communication more effective for the hearing impaired.

References:

Abiyev, R. H., Arslan, M., & Idoko, J. B. (2023). Sign Language Translation Using Deep Convolutional Neural Networks. (KSII)

Feng, W., Zhang, C., & Wu, Y. (2017). Sign language phrase recognition using recurrent neural networks. In International Conference on Pattern Recognition (ICPR) (pp. 2557-2561). IEEE.

Kita, S., & Liddell, S. K. (2013). Embodied cognition in sign languages. Language and cognitive processes, 28(3-4), 345-388.

Koller, O., Kurapati, M., & Bowden, V. (2016). Sign language recognition using 3D convolutional neural networks. In European Conference on Computer Vision (ECCV) (pp. 573-587). Springer, Cham.

Kumar, K., Garg, M., & Khapra, V. (2019). Efficient 3D hand pose estimation using convolutional neural networks. Neurocomputing, 347, 84-97.

Martinez, A. M., Ko, T., & Bowden, R. (2020). Learning ASL sentence structure from sign video using sequential attention mechanisms. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 8402-8411).

4. Key Objectives/ Business Objectives

4.1 Research Questions:

What is the impact of applications for learning sign language used to improve the communication of the hearing impaired?

What difficulties are encountered in the process of understanding the sign language communicated by users through camera images and what methods can be effective to overcome these difficulties?

What is the impact of machine learning models used in the process of recognizing gestures communicated using sign language and converting them into a predictive sentence or word?

How does the performance of crowdsourced datasets and remote sensing imagery compare in improving communication for the hearing impaired?

What is the social impact of the application developed to overcome the communication difficulties faced by the hearing impaired in their daily lives and how is it accepted among users?

Can it be determined which features are important for the app to provide a user-friendly experience and increase the social inclusion of the hearing impaired?

4.2 Key Steps:

Problem Identification and Scope Definition:

Conduct a detailed literature review on the daily communication challenges of the hearing-impaired.

Define the scope of the project and design the targeted solution strategy.

Data Set Creation and Collection:

Develop a specialized data set containing sign language.

Collect images of sign language from users through crowdsourced methods.

Development of Machine Learning Models:

Design and train machine learning models to work on the created data set.

Enhance the models' ability to understand and predict sign language.

Application Development:

Develop an application suitable for use by hearing-impaired individuals and those wishing to learn sign language.

The application will recognize sign language from the user's camera and visually present the predicted sentence or word.

Improvement of Performance Metrics:

Work towards improving performance metrics on the application compared to previous studies.

Evaluate performance objectively, particularly using metrics like R values and similar measurements.

Assessment of Societal Impact and Feedback:

Evaluate feedback from application users and assess the societal impact of the application.

Identify necessary improvements and updates, continuously enhancing the application based on user needs.

5. Datasets and Sources

This project will leverage the following datasets and sources:

Publicly Available Datasets:

ASLLVD (American Sign Language Lexicon Video Dataset):

Over 3,000 ASL signs were recorded in a studio setting. Includes RGB videos, depth videos, and motion capture data.

https://www.researchgate.net/figure/Inconsistent-WLASL-gloss-labels-examples-the-American-Sign-Language-Lexicon-Video_fig1_363883283

MS-ASL (Microsoft ASL dataset): Over 25,000 annotated videos of ASL signs collected from YouTube. Covers a wide range of vocabulary and signer diversity.
https://www.microsoft.com/en-us/research/project/ms-asl/overview/

Custom Data Collection: If necessary, collect additional videos of ASL signs to augment existing datasets.

Ensure diversity in terms of signers, lighting conditions, backgrounds, and sign variations.

Obtain informed consent from participants and adhere to ethical data collection practices.

Data Augmentation: Employ techniques like rotation, flipping, cropping, and color jittering to increase dataset size and variability.

Enhance model robustness and generalization to real-world settings.

Explore other potential datasets or resources, such as The National Center for Sign Language and Gesture Resources (NCSLGR)

The Linguistic Data Consortium (LDC)

Deaf communities and organizations

https://github.com/loicmarie/sign-language-alphabet-recognizer/tree/master/dataset

https://public.roboflow.com/object-detection/american-sign-language-letters