

Convolutional Neural Network to Recognize Sign Language

Group 6: Ilsa Qadir, Krishma Kapoor, Jasmine Hanjra

Abstract

The purpose of this project is to build a convolutional neural network (CNN) that will accurately recognize sign language gestures. The system will convert the gestures into text, making communication between signers and non-signers efficient and accessible. This deep learning solution aims to bridge communication for people who are deaf or hard of hearing and foster inclusivity in various social settings.

1. Background

- Sign Language is an essential means of communication for millions of people around the world who are deaf or hard of hearing. Oftentimes, language barriers between signers and non-signers can make it difficult to communicate. Using new deep learning technology, it is possible to bridge the gap between signers and non-signers by converting sign language gestures to text.

2. Goals

- The goal of this project is to create a convolutional neural network that will convert sign language gestures into text. The model will be trained to identify hand movements and gestures that correspond to signs and translate them into text making it easier for non-signers to understand what is being said.
 - Develop a CNN model that can classify hand gestures from images or videos.
 - Have high accuracy in recognizing these hand gestures.
 - Integrate the system with text output for real-time communication.

3. Method

- **Data Collection:** This project will be created by using large datasets of sign language gesture images and videos like the ones publicly available from the American Sign Language Dataset.
- **Model Training:** We will train the CNN model based on this data to accurately determine the hand shapes and gestures presented from images or videos.
- **Model Testing:** The model will be tested on accuracy and efficiency.

4. Tools and Technologies

- Programming Language: Python
- Deep Learning Libraries: PyTorch, TensorFlow
- Development Environment: Jupyter Notebook, Google Colab

5. Challenges and Mitigation Strategies

- Data Availability:
 - While large datasets are available, ensuring diversity in hand shapes, sizes, and skin tones can be challenging. This will be mitigated by using data augmentation and collecting additional data from diverse sources.
- Model Accuracy:
 - Achieving high accuracy can be difficult, especially for complex gestures. This will be addressed through model tuning, using ensemble learning techniques, and incorporating additional gesture features like movement tracking.
- Real-Time Processing:
 - Implementing a real-time processing system may introduce latency. Optimizations in code, hardware acceleration (using GPUs), and efficient algorithms will be used to ensure minimal lag.

6. Expected Outcomes

- By the end of the project, we expect to have:
 - A CNN model with high accuracy in recognizing ASL gestures.
 - A functional real-time text conversion system that bridges communication between signers and non-signers.
 - Comprehensive documentation and an easy-to-use interface for future improvements and potential expansions.