Project Proposal: Hand Sign Tutor Using Neural Networks

1. Problem Definition

1.1 Overview

Sign language is a crucial means of communication for the deaf and hard-of-hearing community. However, learning sign language can be challenging due to a lack of accessible resources and qualified instructors. Existing solutions often require human intervention or rely on traditional image processing techniques, which struggle with variations in hand positioning, lighting, and background noise.

A **Hand Sign Tutor** powered by neural networks aims to bridge this gap by providing an interactive learning tool that recognizes and evaluates sign language gestures in real time. The system will use deep learning techniques to classify hand gestures and provide immediate feedback to users, making sign language learning more accessible and efficient.

1.2 Target Users

- Individuals learning sign language (students, parents, educators)
- Schools and institutions teaching sign language
- Deaf and hard-of-hearing individuals seeking self-learning tools
- Researchers and developers in AI and accessibility fields

1.3 Challenges in Current Systems

- Difficulty in recognizing hand gestures due to variations in lighting, hand orientation, and individual differences.
- Limited availability of real-time feedback in current sign language learning applications.
- Existing rule-based methods lack adaptability and struggle with complex gestures.

2. Scope of the Project

2.1 What the System Will Cover

- Recognition of a predefined set of hand signs (e.g., American Sign Language (ASL) alphabet and basic words).
- Real-time gesture classification using a neural network model.
- Immediate feedback to users on the correctness of their gestures.
- A user-friendly interface for interactive learning.

2.2 What the System Will Not Cover

- Complex sign language sentences (beyond individual gestures or words).
- Gesture recognition beyond the trained dataset (e.g., different regional sign languages).
- Sign language grammar or sentence structuring.

2.3 Expected Outcomes

- A functional prototype capable of recognizing and evaluating hand gestures in real-time.
- Improved learning experience for users through interactive feedback.
- A scalable system that can be expanded with more sign language datasets.

3. Justification for Using Neural Networks

3.1 Why Neural Networks?

Traditional computer vision techniques struggle with variability in hand shapes, lighting conditions, and camera angles. Neural networks, particularly **Convolutional Neural Networks (CNNs)**, excel at extracting spatial features from images, making them ideal for hand gesture recognition.

3.2 Chosen Neural Network Architecture

- **CNNs**: Effective for recognizing static hand gestures.
- **RNNs or LSTMs** (if needed): Suitable for sequential gesture recognition.
- **Transformer-based Models**: For improved feature extraction and attention mechanisms.

3.3 Expected Performance Metrics

- **Accuracy**: The percentage of correctly classified gestures.
- **Latency**: The time taken for real-time classification.
- **User Satisfaction**: Feedback from users on ease of learning.

4. Literature Review

Several studies have explored the use of neural networks for sign language recognition:

- **Sign Language Recognition using CNNs (2020)**: Achieved 90%+ accuracy on ASL datasets.
- **LSTM-based Gesture Recognition (2021)**: Demonstrated improvements in sequential sign recognition.
- **Transformer Networks for Sign Language (2023)**: Highlighted the potential for better feature extraction and adaptability.

Existing gaps:

- Limited real-time applications for learners.
- High computational costs for complex models.
- Lack of interactive feedback in current solutions.

5. High-Level Design of the Neural Network Model

5.1 Model Architecture

- **Input Layer**: Preprocessed images of hand gestures.
- **Feature Extraction Layer**: CNN layers for spatial feature extraction.
- **Classification Layer**: Fully connected layers with Softmax activation.
- **Output Layer**: Predicted sign class with confidence score.

5.2 Data Collection and Training

- **Dataset**: Publicly available datasets such as RWTH-PHOENIX-Weather, ASL datasets.
- **Data Augmentation**: Rotation, scaling, and brightness adjustments for robustness.
- **Training**: Supervised learning with labeled hand sign images.

5.3 Evaluation Metrics

- **Accuracy and Precision-Recall Scores**
- **Model Latency for Real-Time Performance**
- **User Feedback for Usability Testing**

Conclusion

The proposed Hand Sign Tutor leverages neural networks to create an accessible and effective learning tool for sign language learners. By utilizing CNN-based gesture recognition, the system will provide real-time feedback, making the learning process more interactive and efficient. Future enhancements could include expanding the dataset, supporting additional sign languages, and integrating natural language processing (NLP) for sign language sentence structuring.

Group Members
E/19/174
E/20/099
E/20/371
E/20/416