Microsoft Machine Learning

Final Project

Hand Gesture Recognition for Arabic Sign Language to Text

CAI2_AIS2_S4

Team Members:

- Seif Sherif Assad Ali
- Sama Ahmed ElSayed
- Yusuf Sobhy Sadek Elmeligy
- George Nemr Mellek Poqtor
- Omar Elsayed Elsayed Mousa

Table of Contents

1. Introduction 3
1.1 Project Overview 3
1.2 Motivation 3
2. Project Scope & Features 3
2.1 Dataset Selection & Preprocessing 3
2.2 Model Development & Training 3
2.3 Real-Time Gesture Recognition & Deployment 3
2.4 Future Expansion & Enhancements 3
3. Implementation Plan 4
3.1 Project Timeline & Phases 4
4. Technical Implementation 4
4.1 Tools & Libraries 4
4.2 Model Architecture 4
4.3 Performance Metrics 4
5. Expected Outcomes & Future Work 5
5.1 Expected Outcomes 5
5.2 Future Work 5
6. Conclusion 5

1. Introduction

1.1 Project Overview

This project focuses on developing an AI-based Hand Gesture Recognition System that translates Arabic Sign Language (ArSL) into text using computer vision and deep learning techniques. With the increasing need for accessible communication tools, our goal is to create a real-time recognition system that can bridge the gap between deaf or hard-of-hearing individuals and those who do not understand sign language.

The system will utilize image processing and machine learning models trained on a dataset containing RGB images of Arabic hand gestures representing different letters of the alphabet. The end goal is to accurately classify hand gestures and convert them into written text, which can then be further extended into various applications, such as text-to-speech conversion or integration with conversational AI assistants.

1.2 Motivation

Sign language is the primary mode of communication for millions of individuals around the world. However, Arabic Sign Language (ArSL) lacks robust technological tools that facilitate seamless communication between sign language users and those who do not understand it.

Our motivation for this project is to enhance accessibility and inclusion by developing a highly accurate and scalable Arabic Sign Language recognition system. This system will empower users by enabling real-time, gesture-to-text conversion, making communication easier and removing barriers between Arabic-speaking sign language users and the broader community.

2. Project Scope & Features

2.1 Dataset Selection & Preprocessing

- Utilize the Arabic Sign Language Dataset from Kaggle, consisting of RGB images of different hand gestures.
- Implement data cleaning, augmentation, normalization, and resizing to enhance model performance.

2.2 Model Development & Training

- Utilize Convolutional Neural Networks (CNNs) for image classification and feature extraction.
- Train the model on a preprocessed dataset to recognize Arabic sign language letters.
- Optimize model accuracy, generalization, and real-time inference for practical applications.

2.3 Real-Time Gesture Recognition & Deployment

- Implement real-time gesture detection using a webcam feed.
- Integrate the trained model with a live-streaming pipeline for instant recognition.
- Develop a user-friendly interface to display recognized text.

2.4 Future Expansion & Enhancements

- Extend the project to support full words and sentences, not just individual letters.
- Integrate speech synthesis (text-to-speech) for voice output of recognized text.
- Deploy the system as a mobile application or browser extension to increase accessibility.

3. Implementation Plan

3.1 Project Timeline & Phases

Phase	Tasks
1. Planning & Research	Define objectives, research existing models, finalize dataset
2. Data Collection & Preprocessing	Clean and preprocess data, perform exploratory data analysis
3. Model Development & Training	Implement CNN, train model, optimize performance
4. Real-Time System Integration	Integrate model with webcam feed, create UI
5. Deployment & Testing	Deploy system, conduct real-world testing
6. Final Documentation & Presentation	Prepare final report, conduct live demonstration

4. Technical Implementation

4.1 Tools & Libraries

- **Data Handling & EDA**: pandas, numpy, matplotlib, seaborn, opencv, PIL
- **Machine Learning & Deep Learning**: TensorFlow, Keras, PyTorch, scikit-learn
- **Preprocessing & Augmentation**: OpenCV, torchvision.transforms
- **API & Deployment**: Flask, FastAPI, Docker, Google Cloud/AWS

4.2 Model Architecture

- **Input Layer**: Resized RGB images (128x128 or 224x224 pixels).
- **Feature Extraction**: CNN-based architecture with multiple convolutional layers.
- **Classification Layer**: Fully connected layers with softmax activation.
- **Output**: Predicted Arabic letter as text.

4.3 Performance Metrics

- **Accuracy**: Percentage of correctly classified gestures.
- **Precision & Recall**: Ensuring minimal false positives/negatives.
- **Real-Time Latency**: Optimizing model inference for low-latency predictions.

5. Expected Outcomes & Future Work

5.1 Expected Outcomes

- Develop a high-accuracy gesture recognition model for Arabic Sign Language.
- Ensure real-time performance for practical usability.
- Create an intuitive user-friendly interface to enhance accessibility.
- Promote inclusivity by bridging the communication gap.

5.2 Future Work

- Expand dataset coverage to include regional variations of Arabic Sign Language.
- Integrate Natural Language Processing (NLP) for better text formation.
- Deploy the system as a mobile application for on-the-go accessibility.

6. Conclusion

This project presents an innovative approach to Arabic Sign Language recognition by leveraging deep learning and real-time computer vision techniques. With a structured implementation plan, our system will provide an accessible and practical tool to enhance communication for the deaf and hard-of-hearing community. By continuously improving the model and integrating future enhancements, this project has the potential to significantly impact language accessibility and inclusion in the Arabic-speaking world.