

Title of the Project: Saudi Sign Language Translator: An AI-Powered Solution for Bridging Communication Barriers

Name: Yasir Aljarroof

Institution: Midocean University

Advisor: Ashraf Elsayed

Assignment: Final Report

Abstract:

This research project aims to develop an AI-powered application for translating Saudi Sign Language into text, addressing the communication challenges faced by the deaf community. The project focuses on creating a tool that can recognize and translate static sign language gestures accurately into Arabic text which will help bridge the communication barriers. The methodology involves collecting a dataset of Saudi sign language images and implementing hand detection and gesture recognition to finally translate these into text. The project aims to achieve high accuracy in sign recognition while maintaining real-time processing capabilities. This research specifically addresses the unique challenges of Saudi sign language recognition.

Background and motivation

The deaf community faces significant communication barriers in their daily interactions with hearing individuals. This tool will help the deaf community so that they don't have to rely on sign language interpreters to assist with communication which presents substantial challenges like the availability of sign language interpreters and the lack of automated translation tools. This project addresses this critical gap by using artificial intelligence and computer vision technologies to create an accessible and efficient translation system. The goal is to help deaf people communicate better and feel included while also improving AI for Saudi language and sign language recognition.

Problem statement

The current lack of Saudi Sign Language and automated translation tools creates significant barriers for deaf individuals trying to communicate with hearing people. The challenge lies in developing a system that can accurately recognize and translate a wide range of Saudi signs while being accessible, user-friendly, and practical for everyday use.

Research objectives

1. Develop an AI-powered system for accurate Saudi Sign Language recognition and translation
2. Design and develop a user-friendly interface
3. Achieve high accuracy in sign recognition while maintaining real-time processing capabilities

Literature Review

The advancements in deep learning have made huge progress, which helped in a lot of areas. One of these areas is sign language recognition system, especially for Arabic and Saudi Sign Language, which is aiming to bridge communication gaps between deaf and hearing

communities. Researchers have mostly used convolutional neural networks (CNNs) for static gesture classification, though hybrid models and real-world applications remain areas of active exploration.

Al-Obodi et al. (2020) made a CNN-based system for Saudi Sign Language recognition, achieving high accuracy by training on 40 classes of static gestures, including 28 alphabetical signs, numbers 1-10 signs, and two more gestures for me ”أنا” and ”لا”. Their work demonstrated the viability of CNNs for isolated gestures. However, it had limitations including a restricted vocabulary and reliance on still images rather than continuous sentences. Similarly, El Rwelli et al. (2021) developed a CNN-driven system to translate Arabic Sign Language (ArSL) gestures into vocalized speech, attaining 90% accuracy. However, their model’s training on only 30 sign letters highlighted a recurring challenge in sign language recognition: limited vocabulary scalability.

To address dynamic gestures, hybrid architectures combining CNNs with sequential models have emerged. Noor et al. (2024) proposed a CNN-LSTM framework for real-time ArSL recognition. They have achieved 94.40% accuracy for static gestures, and 82.70% for dynamic ones. They have trained the models on 10 static words and 10 dynamic gestures. The results were great, but the dataset was limited, suggesting the need for larger datasets for both static and dynamic gestures. Al Khuzayem et al. (2024) have developed a CNN-based SSL recognition model and made it into a mobile application, Efhanni, which supports both static and dynamic gestures. Despite its innovation, the system required further testing in varied lighting and background conditions, as the data set is not available to see, and the sample provided had different people doing the signs, but all had same lightning and background (Al-Obodi et al., 2020; Noor et al., 2024).

The efficacy of CNNs for static gestures is further supported by Samonte et al. (2022), whose research and literature review of 20 sign language recognition studies have highlighted CNNs and Connectionist Temporal Classification (CTC) as dominant for static and dynamic recognition.

All together, these studies highlight CNN models robustness in static gesture recognition but reveal some challenges. like limited vocabulary, dynamic gesture recognition less than static performance. Future work must prioritize larger, diverse datasets, and different environments.

Proposed Methodology

Data Pre-processing:

- resizing images to a consistent input size.

Data augmentation:

- we would take the existing dataset, and apply transformations such as rotations, flips, shifts, and zoom to balance the dataset.

Model Exploration:

- Since CNN seems to be the most popular choice and has a high accuracy from previous studies, I will choose to develop a custom CNN to learn distinctive features present in the dataset.

Hyperparameter Tuning:

- Experiment with the number of convolutional layers, filter sizes, learning rates, and dropout rates.

Some of the suggested libraries:

- OpenCV and MediaPipe for real time image capture, preprocessing, and hand/body landmark detection.
- PyTorch: for building and training neural networks.

Datasets Description

The dataset I chose for this project is Arabic Sign Language Letters Dataset (ArSL21L). It is available on Kaggle by Munkhjargal Gochoo from United Arab Emirates University and it was Published on 11 February 2022. Its is made of Images of 32 distinct Arabic sign language letter signs, including 28 arabic letters, and the following: لا, لا, لا, لا. This data set has various backgrounds, and was collected from 50 different subjects. The data set has been evaluated using the YOLOv5l model achieving high results.

Timeline of Second Semester Project (10 Weeks)

- **Weeks 1–2: Data Preparation**
 - Checking and verifying the dataset and labels.
 - Pre-process images (resizing, normalization, augmentation).
 - Explore the dataset with visualizations to understand class distribution.
- **Weeks 3–4: Initial Model Development**
 - Design and implement the basic CNN architecture as detailed above.
 - Train on a small subset of the data to validate the architecture's suitability.

- **Weeks 5–7: Model Refinement & Hyperparameter Tuning**
 - Testing additional convolutional layers.
 - Experiment with different hyperparameters (learning rate, batch size, dropout rates).
- **Weeks 8–9: Evaluation**
 - Generate performance metrics (accuracy, precision, recall, F1-score) and confusion matrices.
- **Week 10: Integration into Application Prototype**
 - Integrate the trained CNN into the sign translator application.
- **Weeks 11–12: Documentation**
 - Finalize the project report detailing the design, experiments, results, and any future recommendations.

References

- Al-Obodi, A. H., Al-Hanine, A. M., Al-Harbi, K. N., Al-Dawas, M. S., & Al-Shargabi, A. A. (2020). A Saudi Sign Language Recognition System based on Convolutional Neural Networks. *Building Services Engineering Research and Technology*, 13(11), 3328–3334. <https://doi.org/10.37624/IJERT/13.11.2020.3328-3334>
- Al Khuzayem, L., Shafi, S., Aljahdali, S., Alkhamesie, R., & Alzamzami, O. (2024). Efhamni: A Deep Learning-Based Saudi Sign Language Recognition Application. *Sensors*, 24(3112). <https://doi.org/10.3390/s24103112>
- Noor, T. H., Noor, A., Alharbi, A. F., Faisal, A., Alrashidi, R., Alsaedi, A. S., Alharbi, G., Alsanoosy, T., & Alsaeedi, A. (2024). Real-Time Arabic Sign Language Recognition Using a Hybrid Deep Learning Model. *Sensors*, 24(3683). <https://doi.org/10.3390/s24113683>
- El Rwelli, R., Shahin, O. R., & Taloba, A. I. (2021). Gesture-Based Arabic Sign Language Recognition for Impaired People Using CNN. *International Journal of Advanced Computer Science and Applications*, 12(12), 574–582. <https://doi.org/10.14569/IJACSA.2021.0121273>
- Samonte, M. J. C., Guingab, C. J. M., Relayo, R. A., Sheng, M. J. C., & Tamayo, J. R. D. (2022). Using Deep Learning in Sign Language Translation to Text. *Proceedings of the International Conference on Industrial Engineering and Operations Management, Istanbul, Turkey*, 4036–4044.

