Literature Review

1. Existing implementations

The techniques which can be used in the process of hand gesture recognition can be used in any sign language recognition regardless of the language they originate from. Due to this reason, the background research was not constrained only to the Arabic sign language recognition researches. During this phase it was noted that both machine learning and deep learning techniques have been used for hand gesture recognition. The difference lied in the preprocessing of the raw data. For deep learning based researches they used minimal level of preprocessing but for the machine learning algorithm based research considerable amount of preprocessing had been done. In some instances, a combination of different ML techniques has been used. Some built all the models from scratch while some researchers used pre trained publically available models along with additional techniques.

https://www.doi.org/10.1155/2021/2995851 (KNN) focused on using K-nearest neighbor algorithm along with classical image processing techniques to identify 28 Arabic letters using 9240 images and managed to achieve a 99.5% accuracy which was the highest accuracy found during the background research of Arabic sign language recognition system.

Mohandes M and his team[Mohandes M., Automatic translation of Arabic text to Arabic sign language, *AIML Journal*. (2006) **6**, no. 4, 15–19.], have used support vector machines for the process of sign language detection along with MLP, naïve bayes classifier and special hardware to capture hand movement and details and achieved a 98% accuracy.

https://www.doi.org/10.1007/S41870-020-00518-5 uses KNN and SVM together in their system to identify hand gestures. Additionally, these techniques were fused with Dynamic Time Wrapping (DTW) to detect hand gestures across time. Their system managed to achieve 92.3% of accuracy for a single hand and 93% of accuracy when identifying both hands.

[https://www.doi.org/10.1109/ICCSRE.2019.8807586] developed their system based on a convolutional neural network and managed to achieve better performance that the systems which were using k-nearest neighbor and Support Vector machines

[https://doi.org/10.1109/icenco.2010.5720438] introduced an Arabic alphabet signs recognition system that converts recognized signs to speech. The system is unable to perform real-time recognition. It focuses on static and simple moving gestures. To convert the images to feature vectors, Principal Component Analysis (PCA) is applied. KNN is used in the classification stage.

[https://doi.org/10.1109/cisim.2010.5643519] developed an Arabic alphabet signs translator with an accuracy of up to 91.3%. The inputs of their system are features extracted from a video of signs and the output is text representing the letter. To get the final output, a MLP and a Minimum Distance Classifier(MDC) was used.

[https://www.doi.org/10.1109/tensymp54529.2022.9864539] focuses on American Sign Language Identification. They use a deep learning approach based on Long-Short Term Memory(LSTM) model which is a type of Recurrent Neural Network(RNN). The system is developed using Tensorflow and Keras.

In addition to approaches where they build model from scratch, some research groups focused on transfer learning where they use pre trained models as base models. [https://www.doi.org/10.3390/app132111625] focused on identifying 32 Arabic alphabet characters and achieved a 98% accuracy by using ResNet and InceptionResNet. In addition to this [https://www.doi.org/10.1155/2023/5195007],[ https://www.doi.org/10.3390/s23167156],[ https://www.doi.org/10.1109/ITIKD56332.2023.10099710] are some other researches which utilized transfer learning.

According to the information gathered during the background research, it was noted that in order to achieve mentioned accuracy in most of the papers, the systems should run in controlled environments.

b.)Arabic Sign Language

<Intro and images from the dataset>

c.)Media Pipe Library

<Content can be extracted from the blog>

d.)MLP

<Some content can be extracted from the blog>

e.) Text to speech implementation

f.) ROS framework and its use in this project