MINISTRY OF EDUCATION

TECHNICAL UNIVERSITY
OF CLUJ-NAPOCA, ROMANIA

FACULTY OF AUTOMATION AND COMPUTER SCIENCE
COMPUTER SCIENCE DEPARTMENT

SUMMARY
of the License Thesis entitled:

Automated Sign Language Recognition based on Deep Learning

Author: **Denisa-Mariana FILIP**
Advisor: **Prof. dr. ing. Florin ONIGA**

1. Requirements:

This project has as its scope the removal of existing communication barriers between hearing-impaired communities and the rest of society. The main objective is the development of the Sign Language Recognition (SLR) deep learning model that recognizes and classifies sign language gestures from the American Sign Language and integrating it into the Angular frontend of a web application. The other main requirement is the development of a system that converts spoken or written text into its equivalent sequence of sign language letters in the form of images.

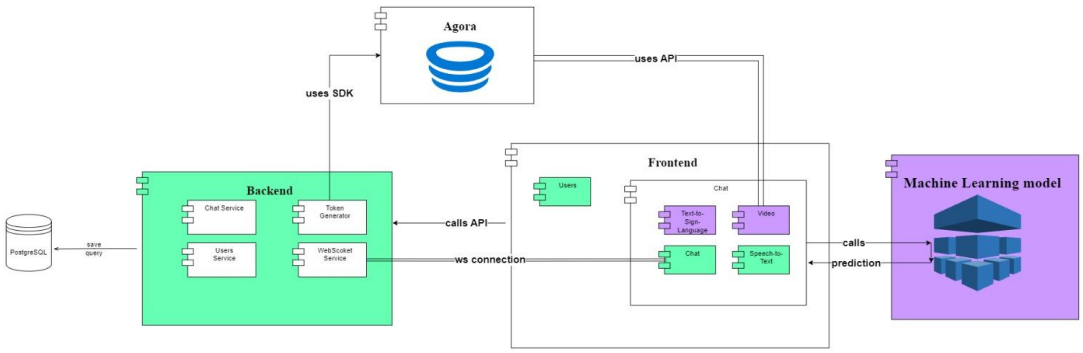

Figure 1. General Application Module Diagram - I developed the purple modules and the associated connections

2. Proposed solutions:

The proposed solution is a web application that captures sign language gestures through video calls and translates them into text (Figure 1). More specifically, my part of the project consisted in a deep learning model for the SLR task, that is a fully connected multi-layer neural network, having one input layer, six hidden layers and one output layer. The model receives as input 21 pairs of x and y coordinates corresponding to

specific points of a hand, along with a label that denotes the gesture that is represented. The architecture of the model is presented in Figure 2.

3. Results obtained:

The obtained result is a SLR model that correctly classifies the 26 letters from the American Sign Language and three extra gestures: *space* (to place a space between the other gestures), *delete* (to delete the most recently signed gesture) and *send message* (to send a message in the chat system).

The model was successfully integrated into the frontend of the surrounding web application, thus empowering the communication between the hearing-impaired community and the rest of the society.

Moreover, all written/spoken messages are converted into a sequence of corresponding sign language images, allowing hearing-impaired users to respond in an efficient manner, since they think in sign language.

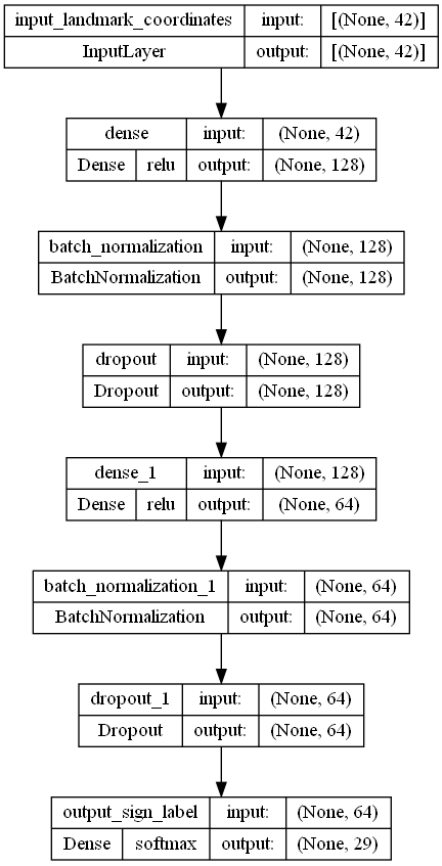

Figure 2. Architecture of Sign Language Recognition Deep Learning Model

4. Tests and verifications:

The accuracy of the model reached a value of 97.35% and a validation loss of 0.09498. The SLR model performed well on each letter and gesture it was trained on,

with the F1-score for each gesture in the range [0.8936, 1.0000], which proves that the Recall and Precision values were high as well.

The SLR model was tested on a dataset that was not used during its training, to avoid the chance of overfitting the model. Moreover, a balanced training dataset was used, each gesture having 3000 samples associated with it.

Accuracy was used to get an overall feel of how the model performs, and the Precision, Recall and F1-score metrics were used to observe where the model had problems in correctly classifying the sign language gestures.

5. Personal contributions:

The personal contributions I had in this project were the **development**, **training**, **evaluation** and **fine-tuning** of the Sign Language Recognition deep learning model, the integration of the model into the frontend of the web application (so that predictions are made on real-time data without the need of round-trips between the frontend and backend of the application) and the translation of messages to sign language images (because sign language users have a difficult time when reading text in English).

6. Documentation sources:

[1] Valentin Bazarevsky and Fan Zhang, "On-Device, Real-Time Hand Tracking with MediaPipe," https://ai.googleblog.com [Accessed 2023.06.13]. [Online]. Available: https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html

[2] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," WSEAS Transactions on Circuits and Systems, vol. 8, pp. 579–588, 2009.

[3] M. S. Abdallah, G. H. Samaan, A. R. Wadie, F. Makhmudov, and Y.-I. Cho, "Lightweight deep learning techniques with advanced processing for real-time hand gesture recognition," Sensors 2023, 23, 2, no. 23, 2022.

Date: 11.07.2023

Author: Denisa-Mariana FILIP

Coordonator: Prof. dr. ing. Florin ONIGA