L46 Project: Decision Making and Planning

Emma Urquhart

January 14, 2024

The primary decisions made, challenges encountered and planning stages of this project are outlined below.

1. *Phase 1: Data Acquisition.* In the initial stages of the project, difficulties were encountered in relation to data acquisition due to the scarcity of publicly-available sign language datasets. The decision was taken to work with the Irish Sign Language Hand Shapes Dataset due to the availability of both static and dynamic sign data, and the easy-to-replicate format of the data, which would make the dataset easy to extend in the future. In addition, as an Irish student, I would like to address the gap in the literature relating to ML methods for ISL translation to increase the level of inclusivity for the deaf community in my home country.

2. *Phase 2: Automated Data Preparation.* In order to fulfil the aim of creating a framework that could facilitate easy extension of the dataset and further training in the future, it was decided to automate the process of data preparation for processing by the neural network. The code is designed to handle data as it appears in the original dataset, with no prior modifications. It can also interact directly with the webcam in the real-time static recognition model and process video clips from the webcam in the case of dynamic recognition. It was also decided to implement mitigating techniques against the challenges posed by a limited dataset were also devised and implemented (use of MediaPipe Hands in pre-processing, data augmentation and an attempt to verify the results through real-world testing). Due to time constraints, the scope of this real-world testing was limited, however, the framework is provided for potential future testing using the same methods.

3. *Phase 3: Static Gesture Recognition.* The next phase involved designing a static gesture recognition model which processes images. Due to the nature of the dataset, there were many more available images than videos, so this was the most promising avenue to pursue. However, the completed model did not require compression in the end due to its lightweight nature. As a result, it was decided to extend the scope of the project to include a more complex model that would benefit more from compression.

4. *Phase 4: Dynamic Gesture Recognition.* It was decided to attempt to build a dynamic gesture recognition model for the three dynamic sign gestures of ISL in order to provide the required scope for exploration of compression techniques. Upon research into the propoerties of different types of neural networks that could enable data patterns to be identified across sequential frames, the LSTM architecture was selected. The presence of this technique in existing research efforts across a range of sign languages in the literature indicated its adaptability and given that ISL is a separate language to the main sign languages, it seemed an appropriate choice.

5. *Phase 5: Model Compression and Evaluation.* The compression techniques learned in class and in the lab practical were re-employed in this application, with the aim of compressing the neural network to a size suitable for deployment to a resource-constrained device. The first approach explored was knowledge distillation, which did not offer the desired benefits as the accuracy degradation was too great for too small of a compression factor. Furthermore, pruning was also explored, including fine-tuning the pruned model. These results were more favourable and lent themselves to further compression in the form of post-pruning quantization. The combined approach resulted in a significant reduction in model size, although causing a degradation in performance to slightly below the desired threshold of 90%.

6. *Phase 6: Analysis of Results, Discussion and Reflection.* This phase of the project involved analysing the results and their implications. The advantages, disadvantages and shortcomings of the approaches implemented were reflected upon and discussed in the final report. Given that one of the initial aims of the project was to provide a baseline for future extension in accordance with the growth of the dataset for the

achievement of more reliable results, it was also important to allocate time to consideration of future directions for extensions of this research.

7. *Phase 7: Preparation of Final Report.* The final stage of the project involved compilation of the final report and the remainder of time was invested in technically and completely describing the study, design decisions made, analysis conducted and results observed.