SUMMARY

Intelligent systems for SLR continue to pique the interest of academics and industry practitioners alike as ML and computational intelligence methodologies develop. The SLR systems have evolved from recognizing only static signs and alphabets to a system that can detect dynamic motions that appear in a continuous sequence of images.

Large datasets composed for SLR are still not available for some of the countries that are involved in developing automated SLR tools. We reviewed a couple of models of state-of-art object detection technologies and utilize them in order to perform detections on a specific ESL dataset.

In conclusion, ML based translation application has been developed for the ESL community since such systems have not been studied and researched as extensively. Around 100 images have been captured and trained in order to train the model to detect some most common expressions such as "Tere – Hello", "Ema – Mother", "Laupäev – Saturday", "Meeldima – to like" and letter "Õ". The following conclusions may be derived from the experiments and research carried out in the field of object detection and machine learning:

- The TF Object Identification API includes a comprehensive array of models that can be adapted and exploited for any object detection application. In our case, we utilized Faster R-CNN and SDD models to recognize ESL terms based on a bespoke dataset.

- When compared to SSD-based models, Faster R-CNN models have a greater tendency to attain higher precision in terms of raw mAP. Each detection's accuracy was above 95%.

- As a result, there is no one-size-fits-all model. The method is mainly concerned with picking the most appropriate model based on the individual characteristics and demands of each case situation. The framework diversity, attribute extractor, an input image, and size, software and hardware choices, all have an impact on the object detection outcomes.

- While the experiment conducted on Google's cloud server with high availability VMs may not be capable of real-time detections, it is strongly anticipated that a better

GPU will enable the same model configurations, particularly the SSD ones, to identify the object in real-time time more precisely.

- The dataset used in the preceding tests was acquired using the webcam of the MacBook 2018 Air. Due to the nature of their architecture, SSD models might conduct poorly in recognizing low-quality images. These noisy images are identified mostly in the early convolutional layers of models.