

Comparative analysis of deep learning architectures for Sri Lankan sign language recognition using MediaPipe landmarks

Wijesinghe W.H.M.M.J.^{1*}, Wijerathna R.M.K.K.¹

¹Department of Computing and Information Systems,
Faculty of Computing, Sabaragamuwa University of Sri Lanka

*maheshikajeewanthi99@gmail.com

In Sri Lanka, Sri Lankan Sign Language (SLSL) is the main way of communicating among the deaf community; however, the present recognition systems do not have the culturally adaptive methods that come with high accuracy. This study is going to fill this gap through a comparison of the different deep learning approaches for SLSL recognition based on hand landmark sequences. The dataset used is the SSL400 from Kaggle (50 SLSL signs, 3,092 training samples, 546 test samples), where MediaPipe has tracked 132-dimensional hand landmarks per frame to cover 50-frame sequences. There were four architectures created and examined: TCN-SE, BiLSTM-Attention, 1D CNN-GRU Hybrid, and Lightweight Transformer. Out of the four, TCN-SE was the one that got the highest accuracy of 90.48%, which was way above the accuracy of BiLSTM-Attention (52.20%), CNN-GRU (49.45%), and Lightweight Transformer (50.37%). The analysis of feature importance showed that the combination of multiscale dilated convolutions and adaptive attention mechanisms truly capture both the short-term and long-term temporal patterns in sign language movements. TCN-SE with a training time of just 30 minutes can be called a suitable candidate for practical real-time SLSL recognition systems. The current study provides the first thorough comparative analysis of the cuttingedge architectures solely for SLSL recognition, thus giving the researchers the practical guidance to design the sign language recognition technologies for the hearing-impaired community that are not only efficient but also lowcost.

Keywords: *Deep Learning; MediaPipe; Sign Language Recognition; Sri Lankan Sign Language; Temporal Convolutional Networks*