Neural Computing and Applications (2023) 35:13225–13238
https://doi.org/10.1007/s00521-023-08424-0

ORIGINAL ARTICLE

A multi-stack RNN-based neural machine translation model for English to Pakistan sign language translation

Uzma Farooq[1,2] · Mohd Shafry Mohd Rahim[1] · Adnan Abid[2,3]

Received: 16 March 2022 / Accepted: 20 February 2023 / Published online: 11 March 2023

Abstract

Sign languages are gesture-based languages used by the deaf community of the world. Every country has a different sign language and there are more than 200 sign languages in the world. American Sign Language (ASL), British Sign Language (BSL), and German Sign Language (DGS) are well-studied sign languages. Due to their different grammatical structure and word order the deaf people feel difficulty in reading and understanding the written text in natural languages. In order to enhance the cognitive ability of the deaf subjects some translation models have been developed for these languages that translate natural language text into corresponding sign language gestures. Most of the earlier natural to sign language translation models rely on rule-based approaches. Recently, some neural machine translation models have been proposed for ASL, BSL, DGS, and Arabic sign language. However, most of these models have low accuracy scores. This research provides an improved and novel multi-stack RNN-based neural machine translation model for natural to sign language translation. The proposed model is based on encoder–decoder architecture and incorporates attention mechanism and embeddings to improve the quality of results. Rigorous experimentation has been performed to compare the proposed multi-stack RNN-based model with baseline models. The experiments have been conducted using a sizeable translation corpus comprising of nearly 50,000 sentences for Pakistan Sign Language (PSL). The performance of the proposed neural machine translation model for PSL has been evaluated with the help of well-established evaluation measures including Bilingual Evaluation Understudy Score (BLEU), and Word Error Rate (WER). The results show that multi-stacked gated recurrent unit-based RNN model that employs Bahdanau attention mechanism and GloVe embedding performed the best showing the BLEU score of 0.83 and WER 0.17, which outperform the existing translation models. The proposed model has been exposed through a software system that converts the translated sentences into PSL gestures using an avatar. The evaluation of the usability has also been performed to see how effectively the avatar-based output helps compensating the cognitive hearing deficit for the deaf people. The results show that it works well for different granularity levels.

Keywords Pakistan sign language · Deaf people · Neural machine translation · Cognition · Avatar

1 Introduction

More than 5% of the world's population, that is almost 466 million, is affected by hearing disability. The projections show that by 2050 this number may rise to 900 million people [1]. Special languages, generally referred to as sign languages, are used by the deaf community of the world. Just like the natural languages there are different sign languages for different countries [2]. Each gesture comprises of some hand shapes, hand movements around the upper body, and facial expression including shoulder shrugging, lips movement, smile etc [3]. The grammatical structure of the sign language sentences is different from

✉ Adnan Abid
adnanabid7@gmail.com; adnan.abid@vu.edu.pk

Uzma Farooq
uzma.farooq@umt.edu.pk

Mohd Shafry Mohd Rahim
shafry@utm.my

[1] Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

[2] University of Management and Technology, Lahore, Pakistan

[3] Faculty of Computer Science and Information Technology, Virtual University of Pakistan, Lahore 54000, Pakistan