International Journal of Recent Advances in Multidisciplinary Topics
Volume 5, Issue 1, January 2024
https://www.ijramt.com | ISSN (Online): 2582-7839

Natural Sentence Generation Using Sign Language Gestures

Sushmitha Adiga[1*], S. Jyothi[2], R. Umadevi[3], G. D. Snehalatha[4], C. Surbhi[5]
[1,2,3,4]*Student, Department of Computer Science and Engineering, PES University, Bengaluru, India*
[5]*Assistant Professor, Department of Computer Science and Engineering, PES University, Bengaluru, India*

***Abstract*: Sign language improves communication between hearing-impaired people, and that is the only way to communicate with non-signers. American Sign Language is known for its efficiency in improving communication between hearing-impaired people. This creative application uses technology to overcome the communication gap between non-signers and those with hearing impairments, fostering inclusivity and understanding by creating an application that detects the signs and generates natural sentences using those sequence of words. By utilizing MobileNetV2 for sign language gesture detection and incorporating Natural Language Processing (NLP) techniques for sentence generation, the proposed project not only recognizes the visual expressions of American Sign Language but also converts them into coherent English sentences. This seamless integration of computer vision and language processing technologies holds the potential to expand communication for hearing-impaired and deaf-mute people, providing them with a more effective means to interact with the broader community.**

***Keywords*: WLASL, NLP, MobileNetV2, Deep Learning, CNN, T5, TKinter, Sign Language.**

1. Introduction

Sign language serves as an essential communication tool for individuals facing speech and hearing challenges. Nevertheless, effective communication with these individuals can be challenging for those unfamiliar with sign language. To address this issue, this project aims to develop a system that utilizes deep learning and natural language processing (NLP) techniques to automatically recognize sign language and generate sentences in real-time

The system will use a two-stage approach to achieve its objective. In the first stage, the system will recognize sign language gestures using deep learning models. In the second stage, the system will generate sentences in natural language using NLP techniques, such as natural language processing, and sentiment analysis.

The system's output will be displayed to the user in a user-friendly format, enabling them to communicate with non-sign language users more effectively.

2. Methodology

A. Dataset Description

Our project utilized the WLASL video dataset from Kaggle, containing 21,000 sign videos of 2000 words performed by 100 different signers [10]. Video length varies from 0-10 seconds and each video has a unique video_id and is mapped to the respective gloss using a JSON file.

We focused on a subset featuring the 250 most common daily conversation words. 5 videos of each gloss/word have been used for further processing.

B. Data Pre-processing

Data processing technique is carried out to enhance the features important for further processing and analysis. In general, it involves frame extraction, removing sample noise, normalizing frames, and greyscale conversion of frames to improve the quality of data and facilitate more effective feature extraction.

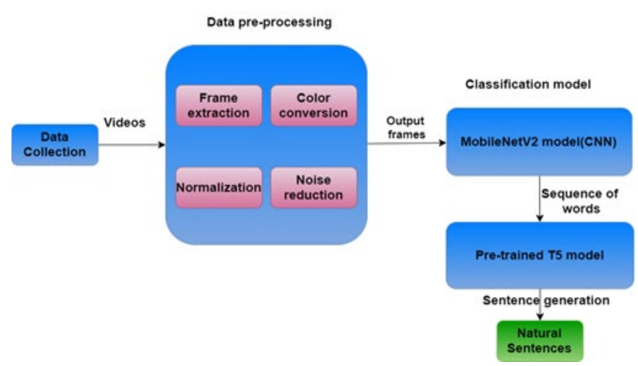

Fig. 1. High-Level design

C. Classification Model

A deep learning pre-trained model MobileNet V2 is trained on a large dataset of sign language videos to detect and recognize various sign language gestures. We are using the Adam optimizer that keeps each parameter's adaptive learning rate constant enabling it to handle sparse gradients and the sparse_categorical_crossentropy loss function that measures the discrepancy between the actual output and the predicted output [7]. The layers of the pre-trained model are frozen so that the weights during training are retained.

D. Sentence Generation

The generation of natural sentences is done by using the

*Corresponding author: adigasushmitha@gmail.com 'pipeline' function from the Hugging Face Transformers library. It creates an NLP pipeline for text generation. The pipeline is initialized with a pre-trained T5 (Text-To-Text Transfer Transformer) model and fine-tuned for common text generation tasks. The sequence of predicted words is filtered in such a way that no words are repeated in a list and this list is fed into the NLP pipeline to get a meaningful sentence. The generated sentence gets printed on a user interface.

3. Results

Sign gesture detection: The trained model successfully detects the signs and can store and print the predicted words or labels.

Capture live: The application can capture the live frames and predict the words.

Upload video files from the system: When a video is uploaded from the local system as an input the application can predict the words and store the sequence of words in a list, which are then given to the NLP model as an input.

Sentence generation: The trained model is capable of generating a sentence using the sequence of keywords that are given as an input.

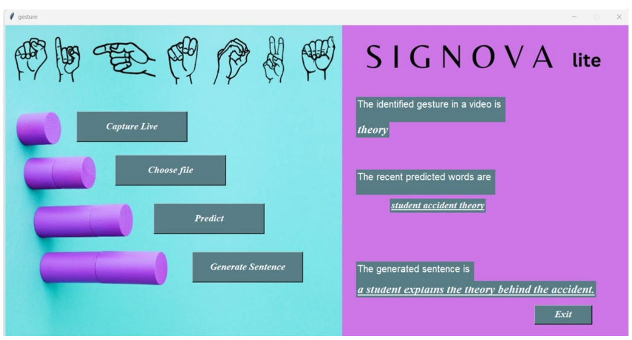

Fig. 2. The user interface with the final output

The proposed model achieved 87% training and validation accuracy for sign gesture detection.

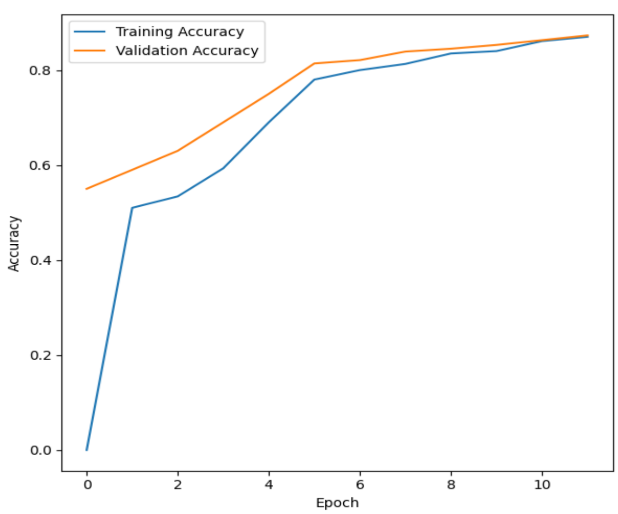

Fig. 3. Training and validation accuracy

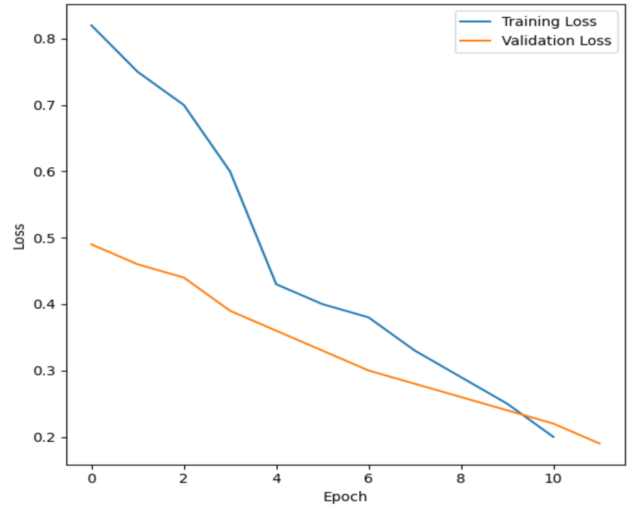

Fig. 4. Training and validation loss

4. Conclusion

The application is created with a user-friendly GUI using the TKinter toolkit. It can detect sign language gestures using live capture in real-time and also using a video uploading method. The repeated words are deleted from a sequence of detected words which are then used in a sentence generation model. The final output of the application is a grammatically correct and meaningful sentence.

References

[1] Ronglai Zuo, Fangyun Wei, Brian Mak, "Natural Language-Assisted Sign Language Recognition," 2023.

[2] S. E. Panneer and M. Sornam, "Recent Advances in Sign Language Recognition using Deep Learning Techniques," 2022 6th International Conference on Trends in Electronics and Informatics

[3] P. Golda Jeyasheeli, N. Indumathi, "Sentence Generation for Indian Sign Language Using NLP," April 2021 Webology 18(Special Issue 01):196-210.

[4] Gajjar, Kevin & Agrawal, Aman & Gonsalves, Arran & Singh, Gargi. (2022). "Sentence Formation Using NLP on the Basis of American Sign Language," International Journal for Research in Applied Science and Engineering Technology.

[5] Saad Hassan, Sooyeon Lee, Dimitris Metaxas, Carol Neidle, and Matt Huenerfauth. 2022, "Understanding ASL Learners' Preferences for a Sign Language Recording and Automatic Feedback System to Support Self-Study," in Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22). Association for Computing Machinery, New York, NY, USA, Article 85, 1–5.

[6] Thakar, Shubham & Shah, Samveg & Shah, Bhavya & Nimkar, Anant. (2022), "Sign Language to Text Conversion in Real Time using Transfer Learning".

[7] Abu-Jamie, Tanseem N. & Abu-Naser, Samy S. (2022), "Classification of Sign-Language Using MobileNet-Deep Learning," International Journal of Academic Information Systems Research.

[8] X. Han, F. Lu and G. Tian, "Sign Language Recognition Based on Lightweight 3D MobileNet-v2 and Knowledge Distillation," ICETIS 2022; 7th International Conference on Electronic Technology and Information Science, Harbin, China, 2022.

[9] Villaplana, Aitana & Martínez-Hinarejos, Carlos-D. (2021), "Generation of Synthetic Sign Language Sentences," 235-239.

[10] D. Li, C. R. Opazo, X. Yu and H. Li, "Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison," 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA, 2020, pp. 1448-1458.