

Deep Learning-Driven Sign Language Recognition: A Multimodal Approach for Gesture-to-Text Translation Using CNN-RNN Architectures

Krishna Jitendra Jaiswal¹, Amaan Khan², Hitanshu Budhdev³, Mr. Sonali Gandhi⁴

Students, Department of Computer Engineering^{1,2,3}

Assistant Professor, Department of Computer Engineering⁴

Thakur College of Engineering and Technology, Mumbai, India

1032221316@tcetmumbai.in, 1032221318@tcetmumbai.in,

1032210401@tcetmumbai.in, sonali.gandhi@tcetmumbai.in

Abstract: Sign language serves as a vital communication tool for individuals with hearing and speech impairments. However, a significant barrier exists when others do not possess a strong understanding of sign language often requiring interpreters to facilitate communication. To reduce the reliance on human interpreters, this research aims to develop an intelligent system capable of recognizing and translating sign language gestures into meaningful grammatically correct sentences. The proposed system processes both images and videos to interpret gestures, converting them into text that can be translated into any desired language. The system leverages state-of-the-art deep learning techniques, such as Convolutional Neural Networks (CNNs) for identifying key features in images and Recurrent Neural Networks (RNNs) for understanding temporal sequences in video data. By employing these advanced neural networks, the model is able to comprehend hand movements, facial expressions and other non-verbal cues to construct coherent sentences. Additionally, the system integrates natural language processing (NLP) to refine the output, ensuring the resulting sentences are grammatically correct. Our approach addresses common challenges such as differentiating between subtle hand gestures and reducing the impact of environmental noise in images. This solution holds the potential to significantly enhance communication for the hearing and speech impaired offering an efficient interpreter-free method of translating sign language into widely spoken languages.

Keywords: Sign language recognition, deep learning, CNN, RNN, gesture detection, video processing, image processing, natural language processing, interpreter-free communication, neural networks