

Link : <http://www.icicel.org/ell/contents/2024/9/el-18-09-12.pdf>

## IMPROVEMENT IN SPEECH RECOGNITION OF INDONESIAN LANGUAGE USING MEL FREQUENCY CEPSTRAL COEFFICIENTS AND LONG SHORT-TERM MEMORY METHOD

Adriani<sup>1</sup>, Ridwang<sup>1,\*</sup>, Agustan Syamsuddin<sup>2</sup>, Muliadi<sup>3</sup> and Usman Umar<sup>4</sup> <sup>1</sup>Department of Electrical Engineering <sup>2</sup>Magister of Elementary Education, Postgraduate Program Universitas Muhammadiyah Makassar Jl. Sultan Alauddin No. 259, Makassar 90221, Indonesia { adriani; agustan }@unismuh.ac.id \*Corresponding author: ridwang@unismuh.ac.id <sup>3</sup>Department of Informatics and Computer Engineering Education Universitas Negeri Makassar Jl. AP. Pettarani Makassar, Sulawesi Selatan 90222, Indonesia muliadi7404@unm.ac.id <sup>4</sup>Department of Medical Electrotechnology Politeknik Muhammadiyah Makassar Jl. DR. Ratulangi No. 101, Makassar, Sulawesi Selatan 90132, Indonesia usmanumar@poltekkesmu.ac.id Received September 2023; accepted December 2023

**Abstract.** Speech recognition technology has witnessed significant advancements in recent years, revolutionizing the way humans interact with machines and devices. The Indonesian language, with its rich phonetic diversity, presents unique challenges for automatic speech recognition systems. This research aims to enhance the accuracy and efficiency of Indonesian speech recognition by employing Mel Frequency Cepstral Coefficients and Long Short-Term Memory neural networks. The study begins by collecting a comprehensive dataset of spoken Indonesian phrases from various speakers, capturing a wide range of dialects and accents. Preprocessing techniques are applied to clean and prepare the audio data, including noise reduction and feature extraction using MFCCs. These MFCCs are used to represent the spectral characteristics of the audio, providing a compact and informative input for subsequent recognition. The core of the research lies in the implementation of LSTM neural networks, a type of recurrent neural network (RNN) known for its ability to capture long-term dependencies in sequential data. The LSTM model is trained on the preprocessed audio data to learn the underlying patterns and relationships in the spoken Indonesian language. The model is fine-tuned through iterations to optimize its performance. Experimental results demonstrate a significant improvement in the accuracy and robustness of the Indonesian speech recognition system when compared to conventional methods. The incorporation of MFCCs and LSTM networks not only enhances the system's ability to handle diverse dialects but also increases its tolerance to background noise and speaker variations. The achieved recognition rates exhibit promising outcomes for practical applications in voice assistants, transcription services, and other voice-controlled technologies. **Keywords:** Speech, Recognition, MFCC, Indonesian language, LSTM