

## **Wolof Voice Response Recognition System: A Deep Learning Model for Wolof Audio Classification**

**Authors :** Krishna Mohan Bathula, Fatou Bintou Loucoubar, FNU Kaleemunnisa, Christelle Scharff, Mark Anthony De Castro

**Abstract :** Voice recognition algorithms such as automatic speech recognition and text-to-speech systems with African languages can play an important role in bridging the digital divide of Artificial Intelligence in Africa, contributing to the establishment of a fully inclusive information society. This paper proposes a Deep Learning model that can classify the user responses as inputs for an interactive voice response system. A dataset with Wolof language words 'yes' and 'no' is collected as audio recordings. A two stage Data Augmentation approach is adopted for enhancing the dataset size required by the deep neural network. Data preprocessing and feature engineering with Mel-Frequency Cepstral Coefficients are implemented. Convolutional Neural Networks (CNNs) have proven to be very powerful in image classification and are promising for audio processing when sounds are transformed into spectra. For performing voice response classification, the recordings are transformed into sound frequency feature spectra and then applied image classification methodology using a deep CNN model. The inference model of this trained and reusable Wolof voice response recognition system can be integrated with many applications associated with both web and mobile platforms.

**Keywords :** automatic speech recognition, interactive voice response, voice response recognition, wolof word classification

**Conference Title :** ICCV 2022 : International Conference on Computer Vision

**Conference Location :** Vancouver, Canada

**Conference Dates :** September 22-23, 2022