

Deep Learning and Machine Learning Techniques Applied to Speaker Identification on Small Datasets

Enrico Manfron, João P. Teixeira and Rodrigo Minetto

Research Centre in Digitalization and Intelligent Robotics, Associate Laboratory for Sustainability and Technology, Instituto Politécnico de Bragança, Universidade Tecnológica Federal do Paraná (UTFPR)

In this study, we explore the capabilities of speaker recognition technology for biometric authentication. Our focus was on developing and evaluating Machine Learning and deep learning models, such as the Gaussian Mixture Model, Multilayer Perceptron, Convolutional Neural Network, and Residual Neural Networks for speaker identification. The models were trained and tested on both private and public datasets. We explain our methodology for creating and training models on private and public datasets. The model ResNet50 slightly outperformed the other models on two versions of our private dataset, achieving accuracies of 97.93% and 100% for each. This research explores developing speaker recognition-based access control systems, serving as a resource for future research and improvements in secure and efficient speaker identification solutions.

Keywords: Speaker Identification · Convolutional Neural Network · Deep Learning