Ministry of Higher Education and Scientific Research, Iraq
University of Babylon
information technology collage
Information Security Department
Evening Study

Speaker Verification Using Convolutional Neural Network

Research Presented by The Student

Ruqaya Fadhil Hussein

Supervised by

Prof. Dr. Eman S. Al-Shamery

A Graduate Project Submitted to the Department of Information Security of the College of Information Technology, University of Babylon, in Partial Fulfillment of the Requirements for the Bachelor's degree in Information Security of Information Technology.

2024 **1446 هـ**

ABSTRACT

Speaker verification plays a crucial role in various applications such as voice authentication, speaker identification, and fraud detection. The proposed system aims to accurately classify audio samples into predefined classes, enabling the identification of speakers or distinguishing between different Speaker patterns.

Moving forward, the project outlines the workflow, including steps such as audio visualization, feature extraction, model training using CNN architecture, and evaluation. Audio files are visualized using spectrograms, and features are extracted using Mel Frequency Cepstral Coefficients (MFCCs) to represent audio data effectively. CNN models are then trained and evaluated using a dataset containing labeled audio samples.

The project is implemented using Python libraries such as Librosa for audio processing, TensorFlow and Keras for building and training CNN models. The classification accuracy of the models is assessed using metrics such as accuracy and F1-score, ensuring reliable performance in real-world scenarios, We have thirty-two audio files divided into eight people, four females and four males.

Finally, the trained models are integrated into a user-friendly interface using Gradio, allowing users to upload audio files and obtain predictions for their class labels in real-time.

Our model achieves high performance (0.9615% CNN) in predicting Voices.

When run say Found 32 files belonging to 9 classes, Using 26 files for training. Total params: 6932601 (26.45 MB),Trainable params: 6932601 (26.45 MB) Non-trainable params: 0 (0.00 Byte), Number of training samples = 24, Number of testing samples = 7, loss: 0.1641 - accuracy: 0.9615 - val_loss: 8.5772 - val_accuracy: 0.1667.