

Trabalhos Práticos – Reconhecimento de Orador

Datasets para Reconhecimento de Orador

Speakers in The Wild (SITW):

- 299 well-known speakers, ≈6000 seconds, with annotated core segments.

VoxCeleb:

- 1251 celebrities, 145000 total utterances, Pre-trained model exists

Ferramentas/Packages para Speech Processing

- Speechpy, python-speech_features: MFCC, filterbank and log-filterbank
- scikits.talkbox: speech features + useful tools (DFT, etc)
- speaker-recognition: recognition system with GUI
- sklearn: general tools for machine learning, with GMM
- bob.bio.spear: VAD preprocessing, audio databases, acoustic features, recognition algorithms
- bob.kaldi: API for Kaldi, a C++ speaker recognition system with feature extraction and models

Ferramentas/Packages para Machine Learning e Deep Learning

- Sklearn: machine learning geral com vários tipos de classificadores
- Tensorflow: ferramenta geral para Machine Learning com ênfase em Deep Learning (CNN, LSTM, RNN)
- Keras: Python API para Deep Learning, backend Tensorflow