

**Reconhecimento de Fonemas  
utilizando Redes Neurais  
Convolucionais para Transcrição  
Fonética Automática**

Bauke Alfredo Dijkstra  
Ionildo José Sanches

# O que é o reconhecimento de fonemas

O reconhecimento automático de fonemas (RAF) é a capacidade de extrair características para reconhecer as unidades sonoras das palavras e transcrevê-las.

# Aplicações

Reconhecimento de fala

Identificação de locutores

Identificação de erros de pronúncia

Reconhecimento de emoções

# Modelos de fala do RAF

Palavras isoladas

Palavras concatenadas

Fala contínua

# Como se realiza o RAF

Processamento acústico

Classificação dos fonemas

# Bases de dados

Bases prontas:

TIMIT Acoustic-Phonetic Continuous SpeechCorpus

Bases que necessitam de transformações para realizar o trabalho

VoxForge

Sid

LABSM1.4.

# Processamento acústico

Janela de Hamming

Extração das características do sinal digital:

Mel Frequency Cepstral Coefficients (MFCC)

Perceptual linear prediction (PLP)

Tandem features

# Classificação dos fonemas

Algoritmos de classificação

Hidden Markov Model (HMM)

Multilayer perceptron (MLP)

Deep neural network (DNN)

# Proposta do trabalho

Realizar o reconhecimento de fonemas utilizando deep learning

Framework keras API (Application Programming Interface) de alto nível para o TensorFlow

Obrigado!