Language model for specific professional fields
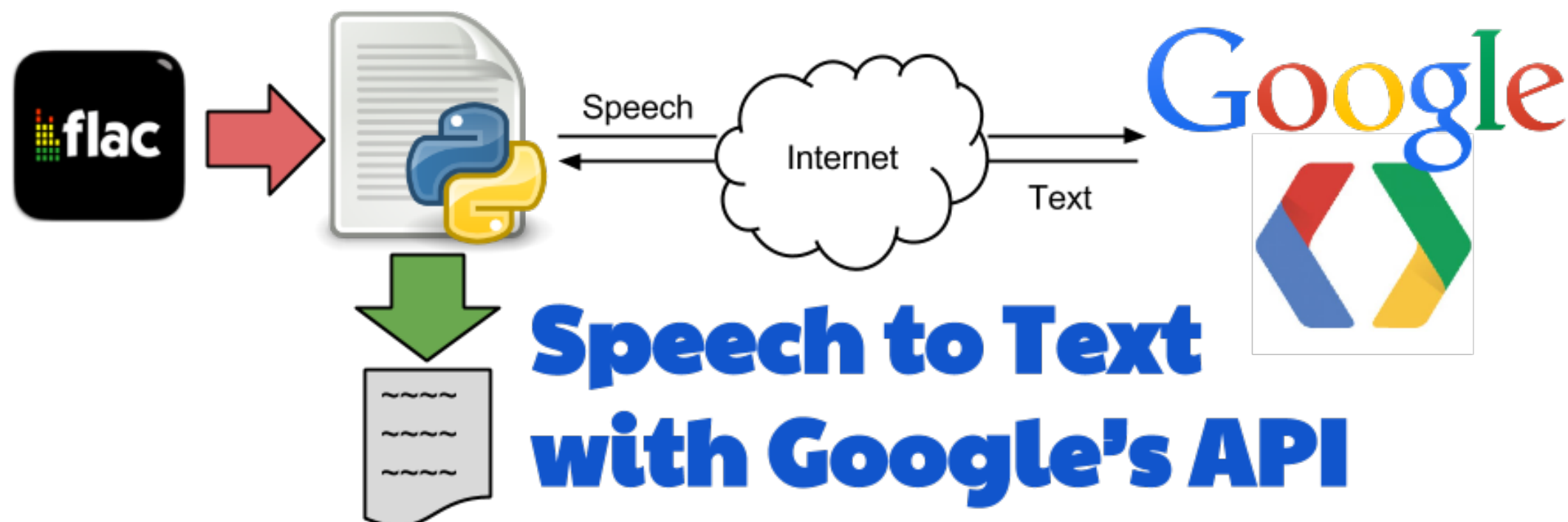
Sahba Yahya

Google Speech API doesn't recognize words from professional dictionaries

Example
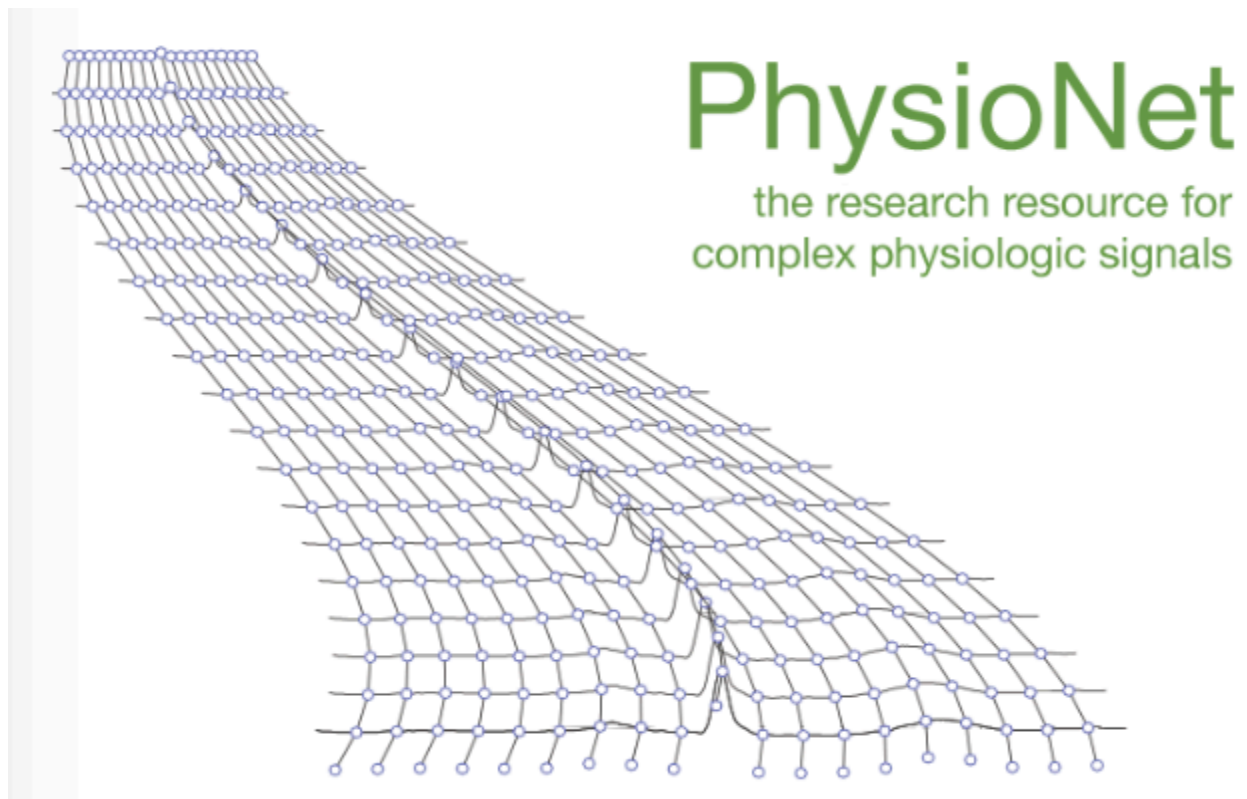

✤ common english language ~5000 words

✤ Doctor have a spoken vocabulary of ~100,000 words

Data Sources (1)

Open speech and language resources

Data Sources (2)

dictionary.com
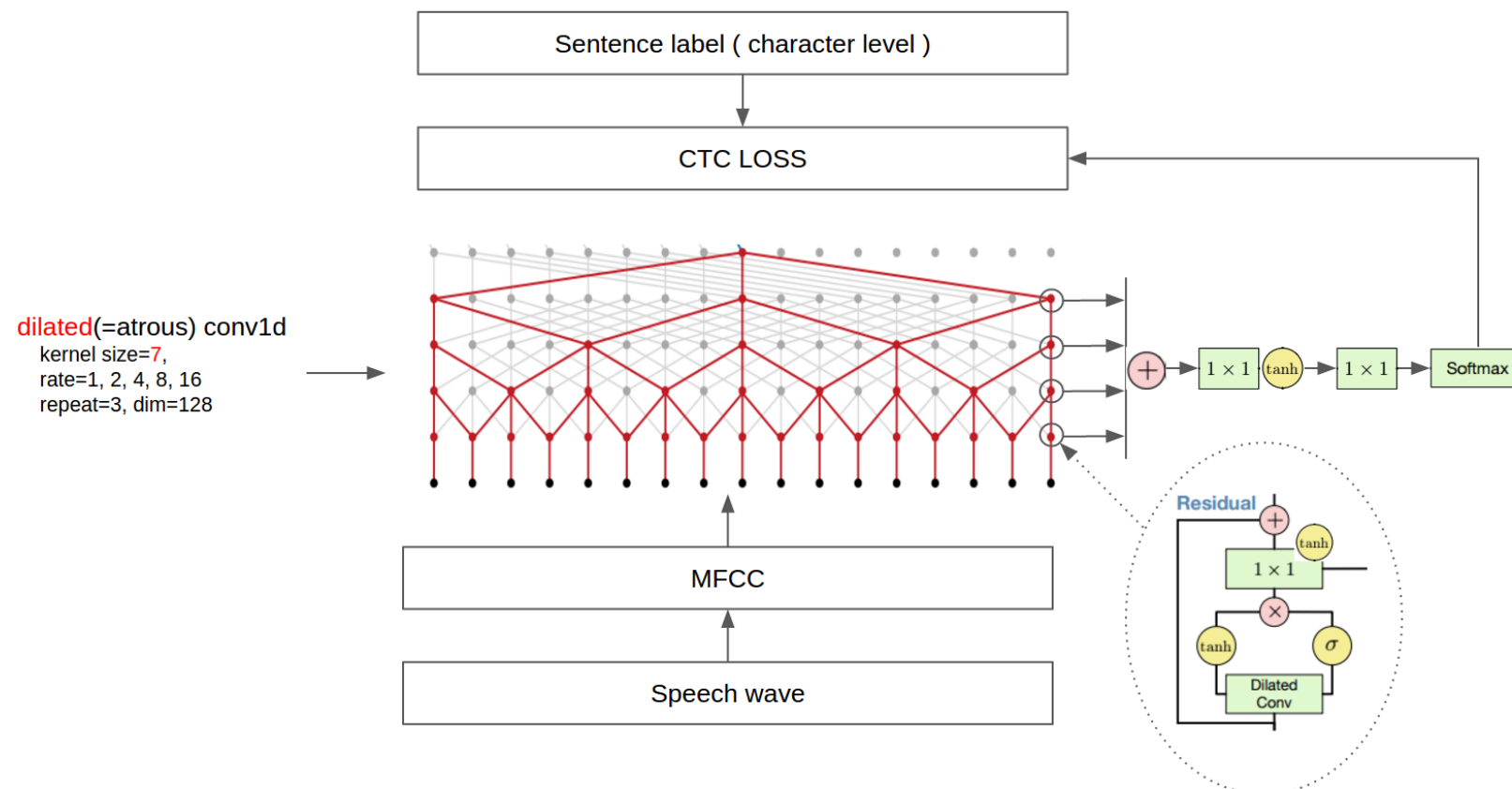
VTCK dataset

Feature extraction

Text + .wav → Feature extraction (FT) → MFCC

Language model (1)
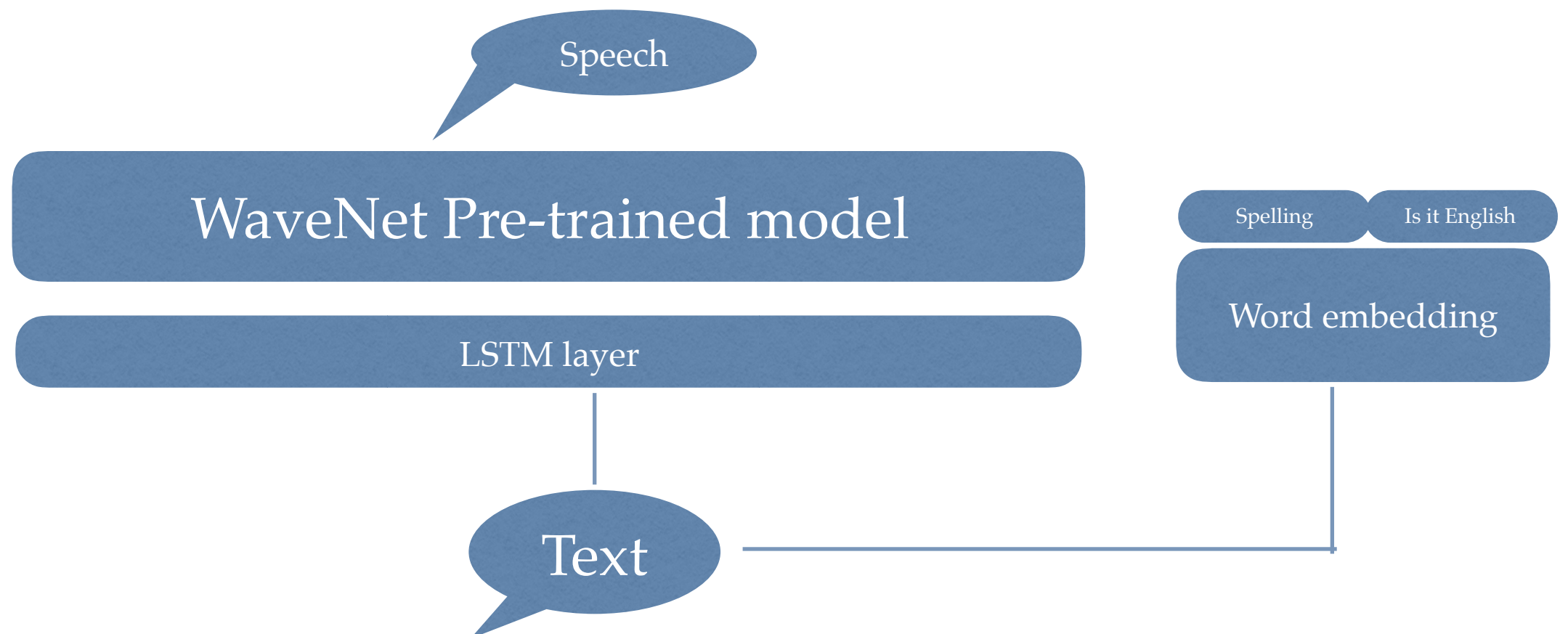
✤ Text to speech WaveNet speech recognition pre-trained model

Language model (2)

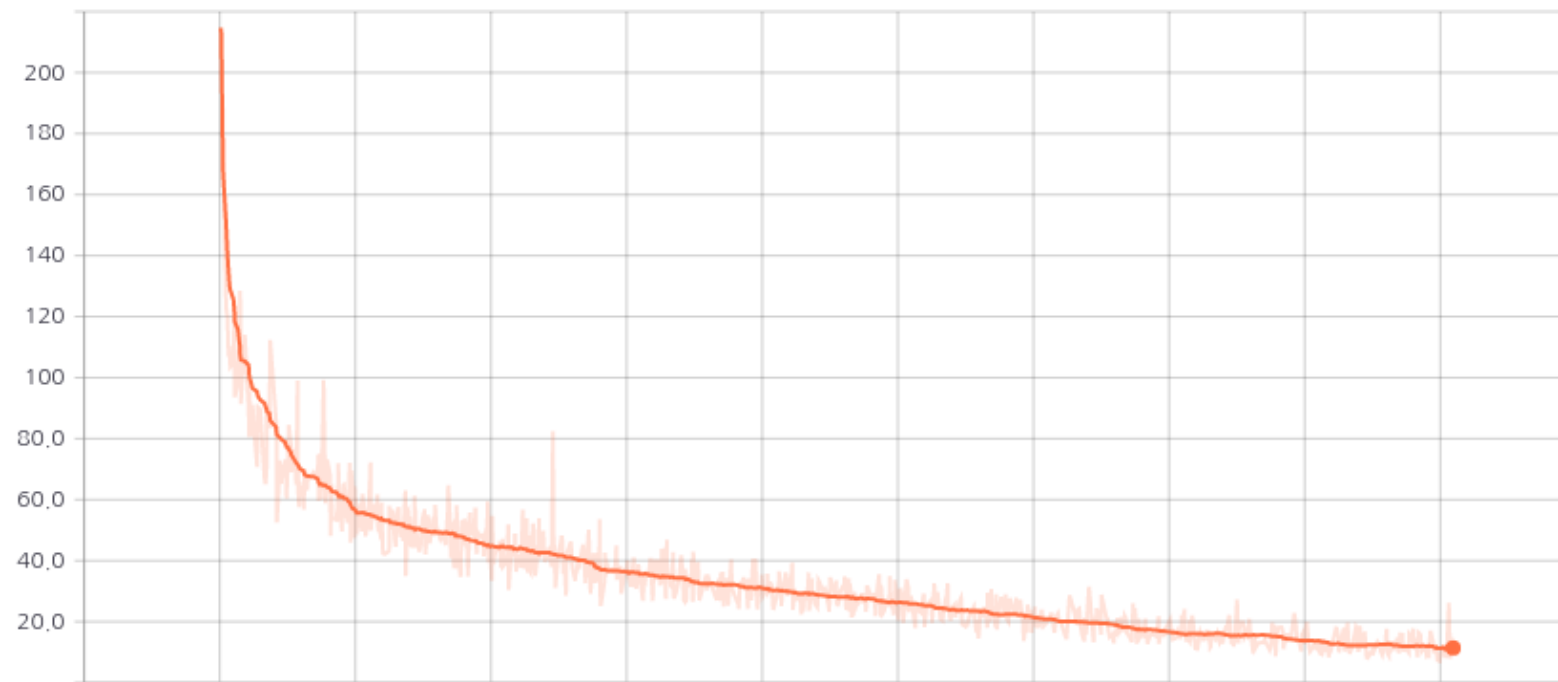✤ The same architecture shows good results when tested on a small speech recognition.

Language Model (3)

Connectionist Temporal classification (CTC)

Output

It is the language of wretchedness — Ground truth

It is the language of Fretchedness — Model output

It is the language of wretchedness — After spelling check

Challenges

- ✤ Need to train the model on larger dataset

- ✤ Out-of-vocabulary (OOV)

- ✤ Takes time to run (speech-to-text)

Improve language model for specific professional (medical terminology)