بسم الله الرحمن الرحيم

Electrical Engineering Department

EE413 Applied Digital Signal Processing

Second Semester 2021/2022

Due Date: May 8, 2022

Term-Project:

Arabic speech recognition

By:

Name	KFUPM ID
Abdullah M. Al-Kathiry	201765290
Abdulrahman A. Al-Arifi	201773310
Majed M. Al-Harthy	201773870
Saleh M. Al-Khalifah	201746210

For:

Dr. Mudassir Masood

Sunday, May 8, 2022

A- Packages needed to be downloaded:

1- ***pydub*** (`conda install -c conda-forge pydub`)

It's used to split the sound file into several files

2- **pysoundfile** (`conda install -c conda-forge pysoundfile`)

It's needed for pytorch to load the sound files.

Other packages (not requiring any installation):

- os
- torch
- tqdm
- random
- IPython.display
- math
- numpy
- from sklearn.metrics import confusion_matrix

B- Dataset:

The dataset is composed of 24 commonly spoken Arabic words. Each word is recorded around 150 times. The sound files contain multiple words where the code will split it into multiple files each of 1 second length. The sampling rate is 8000 Hz. Each sound file contains 8000 samples and if less then zeros will be appended accordingly. The words are separated randomly by the code so that 15% of it are used for testing while the rest are for training. The names of these files are stored in either **"training_list"** or **"testing_list"**. We used dictionary to relate each Arabic word to an English word. Shown below in **table1** each Arabic word with an English label.

Table 1: Arabic words

صدقة	قيام	مياه	مال	لبن	جزر	ذهب	كتاب	عبدالله	رمضان	قرآن	إفطار
sadaga	**qyam**	**myah**	**mal**	**lbn**	**gzr**	**dahab**	**book**	**abdullh**	**Ramadan**	**Quran**	**Eftar**
زكاة	صلاة	ليل	وقود	كهرباء	زمن	ثوب	صياح	سلام	سماء	صالح	سيف
Zakat	**Prayer**	**Night**	**Fuel**	**Electricity**	**zamen**	**thoub**	**syah**	**slam**	**sky**	**saleh**	**saif**

C- Model used:

We used Transfer learning in this project, the model used here is called M5 which is developed in (https://doi.org/10.48550/arXiv.1610.00087) because our dataset is small we use the M5 model pretrained with 35 English word from the speechcommand dataset and use the transfer learning the pretrained model is in pytorch website: https://pytorch.org/tutorials/intermediate/speech_command_classification_with_torchaudio_tutorial.html

C.1 Model for time domain

The class wise accuracy and confusion matrix are presented below. Please refer to the jupyter file to view some predicted words.

1- Class wise Accuracy

```
Accuracy for class: إفطار is 100.0 %
Accuracy for class: كهرباء is 100.0 %
Accuracy for class: وقود  is 100.0 %
Accuracy for class: ليل   is 100.0 %
Accuracy for class: صلاة  is 100.0 %
Accuracy for class: قرآن  is 100.0 %
Accuracy for class: رمضان is 100.0 %
Accuracy for class: زكاة  is 100.0 %
Accuracy for class: عبدالله is 100.0 %
Accuracy for class: كتاب  is 100.0 %
Accuracy for class: ذهب   is 100.0 %
Accuracy for class: جزر   is 100.0 %
Accuracy for class: لبن   is 100.0 %
Accuracy for class: مال   is 100.0 %
Accuracy for class: مياه  is 100.0 %
Accuracy for class: قيام  is 100.0 %
Accuracy for class: صدقة  is 97.0 %
Accuracy for class: سيف   is 100.0 %
Accuracy for class: صالح  is 100.0 %
Accuracy for class: سماء  is 94.1 %
Accuracy for class: سلام  is 95.7 %
Accuracy for class: صياح  is 100.0 %
Accuracy for class: ثوب   is 100.0 %
Accuracy for class: زمن   is 100.0 %
```

2- Confusion Matrix

The list with English words is used as they represent the key of each Arabic word in the dictionary defined

```
['Eftar', 'Electricity', 'Fuel', 'Night', 'Prayer', 'Quran', 'Ramadan', 'Zakat', 'abdullh', 'book', 'dahab', 'gzr', 'lbn',
'mal', 'myah', 'qyam', 'sadaga', 'saif', 'saleh', 'sky', 'slam', 'syah', 'thoub', 'zamen']
[[29  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 25  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 30  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 37  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 26  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 35  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 29  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 23  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 21  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 14  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 28  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0 23  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 21  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0 14  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 18  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 14  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0 32  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 29  0  0  0]
 [ 0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 16  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0 22  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 17]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 18]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 19]]
```

C.2 Model for frequency domain (using DFT)

The class wise accuracy and confusion matrix are presented below. Please refer to the jupyter file to view some predicted words.

1- Class wise Accuracy

```
Accuracy for class: إفطار is 96.6 %
Accuracy for class: كهرباء is 100.0 %
Accuracy for class: وقود  is 100.0 %
Accuracy for class: ليل  is 100.0 %
Accuracy for class: صلاة  is 100.0 %
Accuracy for class: قرآن  is 100.0 %
Accuracy for class: رمضان is 100.0 %
Accuracy for class: زكاة  is 100.0 %
Accuracy for class: عبدالله is 100.0 %
Accuracy for class: كتاب  is 78.6 %
Accuracy for class: ذهب  is 100.0 %
Accuracy for class: جزر  is 100.0 %
Accuracy for class: لبن  is 100.0 %
Accuracy for class: مال  is 100.0 %
Accuracy for class: مياه  is 100.0 %
Accuracy for class: قيام  is 100.0 %
Accuracy for class: صدقة  is 100.0 %
Accuracy for class: سيف  is 100.0 %
Accuracy for class: صالح  is 96.6 %
Accuracy for class: سماء  is 82.4 %
Accuracy for class: سلام  is 95.7 %
Accuracy for class: صياح  is 100.0 %
Accuracy for class: ثوب  is 100.0 %
Accuracy for class: زمن  is 94.7 %
```

2- Confusion Matrix

The list with English words is used as they represent the key of each Arabic word in the dictionary defined

```
['Eftar', 'Electricity', 'Fuel', 'Night', 'Prayer', 'Quran', 'Ramadan', 'Zakat', 'abdullh', 'book', 'dahab', 'gzr', 'lbn',
'mal', 'myah', 'qyam', 'sadaga', 'saif', 'saleh', 'sky', 'slam', 'syah', 'thoub', 'zamen']
[[28  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0]
 [ 0 25  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 30  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 37  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 26  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 35  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 29  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 23  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 21  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  1 11  0  0  0  0  0  0  0  0  0  1  0  1  0]
 [ 0  0  0  0  0  0  0  0  0  0 28  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0 23  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 21  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0 14  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 18  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 14  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 33  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 28  0  0  1  0]
 [ 0  0  1  0  1  0  0  0  0  0  0  0  0  0  0  0  1  0  0 14  0  0  0]
 [ 0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 22  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 17  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 18]
 [ 0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 18]]
```