|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Academic Year: | 2022/2023 | Term: | Spring 2023 |  |
| Course Code: | ELC 4028 | Course Title: | Artificial Neural Networks and its Applications |

Cairo University

Faculty of Engineering

Electronics and Communications Engineering Department – 4th Year

Neural Networks Applications

- Assignment 2 -

*Submitted to: Dr. Mohsen Rashwan*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | BN | Sec | ID | رقم الجلوس |
| احمد محمود حسيني عطية | 22 | 1 | 9180178 | 34022 |
| علي ماهر عبدالسلام نبيه | 5 | 3 | 9190067 | 34117 |
| محمد احمد طه السيد | 30 | 3 | 9191043 | 34142 |
| محمد حسام عثمان يسن | 35 | 3 | 9191083 | 34147 |
| محمد عاطف ربيع | 43 | 3 | 9190924 | 34155 |

Table of Contents

[1. Multi-layer Perceptron 1](#_Toc132912977)

[Training a Multi-layer Perceptron (MLP) 3](#_Toc132912978)

[Using DCT Features 3](#_Toc132912979)

[Using PCA Features 5](#_Toc132912980)

[Using ICA Features 7](#_Toc132912981)

[2. Convolutional Neural Network (LeNet-5) 10](#_Toc132912982)

[LeNet-5 - No Variations 12](#_Toc132912983)

[Variation #1 - Adding Dropout Regularization 13](#_Toc132912984)

[Variation #2 - Increasing Number of Filters in Conv Layers 14](#_Toc132912985)

[Variation #3 - Adding "Same" Padding to Conv Layers 15](#_Toc132912986)

[Variation #4 - Using "Tanh" Activation 16](#_Toc132912987)

[3. Comparing the Results 17](#_Toc132912988)

[Notes 18](#_Toc132912989)

[4. Digit Spectrograms 19](#_Toc132912990)

[Notes 29](#_Toc132912991)

# 1. Multi-layer Perceptron

import os

import cv2

import numpy as np

import matplotlib.pyplot as plt

import random

from sklearn.utils import shuffle

from scipy.fftpack import dct ,idct

import sklearn

from sklearn.decomposition import PCA, FastICA

from sklearn.metrics import accuracy\_score

from keras.models import Sequential

from keras.layers import Dense, Flatten, InputLayer, Dropout

import time

from tensorflow.keras.callbacks import EarlyStopping

training\_path = "Reduced\_MNIST\_Data\Reduced\_Training\_data"

testing\_path = "Reduced\_MNIST\_Data\Reduced\_Testing\_data"

# Define the list of classes

classes = os.listdir(training\_path)

print(classes)

classes = list(map(int, classes))

print(classes)

# Define an empty list to store the data and labels

X\_train = []

y\_train = []

# Loop over the classes

for class\_name in classes:

    class\_path = os.path.join(training\_path, str(class\_name))

    # Loop over the images in the class folder

    for image\_name in os.listdir(class\_path):

        image\_path = os.path.join(class\_path, image\_name)

        # Load the image and append it to the data list

        image = cv2.imread(image\_path, cv2.IMREAD\_GRAYSCALE)

        X\_train.append(image)

        # Append the label to the labels list

        y\_train.append(class\_name)

# Convert the data and labels lists to NumPy arrays

X\_train = np.array(X\_train)

y\_train = np.array(y\_train)

# Print the shape of the data and labels arrays

print("Training Data shape:", X\_train.shape)

print("Training Labels shape:", y\_train.shape)

X\_test = []

y\_test = []

for class\_name in classes:

    class\_path = os.path.join(testing\_path, str(class\_name))

    for image\_name in os.listdir(class\_path):

        image\_path = os.path.join(class\_path, image\_name)

        image = cv2.imread(image\_path, cv2.IMREAD\_GRAYSCALE)

        X\_test.append(image)

        y\_test.append(class\_name)

# Convert the data and labels lists to NumPy arrays

X\_test = np.array(X\_test)

y\_test = np.array(y\_test)

print("Testing Data shape:", X\_test.shape)

print("Testing Labels shape:", y\_test.shape)

X\_train,y\_train = shuffle(X\_train, y\_train, random\_state=4)

X\_test,y\_test = shuffle(X\_test, y\_test, random\_state=4)

#check if shuffling worked correctly

plt.figure()

plt.subplot(121)

plt.title("Is this {} ?".format(y\_train[1050]))

plt.imshow(X\_train[1050])

plt.subplot(122)

plt.title("Is this {} ?".format(y\_test[1050]))

plt.imshow(X\_test[1050])

plt.show()

# ## DCT Features

# Functions used to extract DCT features

def zigzag(a):

    comp=np.concatenate([np.diagonal(a[::-1,:], i)[::(2\*(i % 2)-1)] for i in range(1-a.shape[0], a.shape[0])])

    return comp[:200]

def dct\_extract(a):

    features=np.zeros((a.shape[0],200))

    for i in range(a.shape[0]):

        z\_features=zigzag(dct(dct(a[i].T, norm='ortho').T, norm='ortho'))

        features[i]=z\_features

    extracted=features.reshape((a.shape[0],-1))

    return extracted

#Extract DCT features for training and testing data

X\_train\_DCT=dct\_extract(X\_train)

X\_test\_DCT=dct\_extract(X\_test)

X\_train\_DCT.shape

# ## PCA Features

pca\_model = PCA(.95) #we want a 95% variance

pca\_model.fit(X\_train.reshape((X\_train.shape[0],28\*28)))

X\_train\_PCA = pca\_model.transform(X\_train.reshape((X\_train.shape[0],28\*28)))

X\_test\_PCA = pca\_model.transform(X\_test.reshape((X\_test.shape[0],28\*28)))

print("For 95% varinace, there are {} components".format(pca\_model.n\_components\_))

X\_train\_PCA.shape

# ## ICA Features

ica\_model = FastICA(n\_components=200)

X\_train\_ICA = ica\_model.fit\_transform(X\_train.reshape((X\_train.shape[0],784)), y\_train)

X\_test\_ICA = ica\_model.transform(X\_test.reshape((X\_test.shape[0],784)))

X\_train\_ICA.shape

## Training a Multi-layer Perceptron (MLP)

### Using DCT Features

# ### 1 Hidden Layer

# Define the model architecture

model\_MLP1\_DCT = Sequential(name='MLP1\_DCT')

model\_MLP1\_DCT.add(Dense(256, activation='relu', input\_shape=(200,)))  # hidden layer

model\_MLP1\_DCT.add(Dropout(0.2)) #dropout regularization

model\_MLP1\_DCT.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP1\_DCT.summary()

#Early Stopping to avoid fitting issues

early\_stopping = EarlyStopping(monitor='accuracy', patience=3)

# Compile the model

model\_MLP1\_DCT.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP1\_DCT.fit(X\_train\_DCT, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP1\_DCT.evaluate(X\_test\_DCT, y\_test)

X\_test\_DCT[0].shape

tic=time.time()

model\_MLP1\_DCT.predict(X\_test\_DCT[0].reshape(1,200))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 1 Hidden Layer-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

### 2 Hidden Layers

# Define the model architecture

model\_MLP2\_DCT = Sequential(name='MLP2\_DCT')

model\_MLP2\_DCT.add(Dense(256, activation='relu', input\_shape=(200,)))  # 1st hidden layer

model\_MLP2\_DCT.add(Dropout(0.2))

model\_MLP2\_DCT.add(Dense(128, activation='relu'))  # 2nd hidden layer

model\_MLP2\_DCT.add(Dropout(0.2))

model\_MLP2\_DCT.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP2\_DCT.summary()

# Compile the model

model\_MLP2\_DCT.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP2\_DCT.fit(X\_train\_DCT, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP2\_DCT.evaluate(X\_test\_DCT, y\_test)

tic=time.time()

model\_MLP2\_DCT.predict(X\_test\_DCT[0].reshape(1,200))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 2 Hidden Layers-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# ### 3 Hidden Layers

# Define the model architecture

model\_MLP3\_DCT = Sequential(name='MLP3\_DCT')

model\_MLP3\_DCT.add(Dense(256, activation='relu', input\_shape=(200,)))  # 1st hidden layer

model\_MLP3\_DCT.add(Dropout(0.2))

model\_MLP3\_DCT.add(Dense(128, activation='relu'))  # 2nd hidden layer

model\_MLP3\_DCT.add(Dropout(0.2))

model\_MLP3\_DCT.add(Dense(64, activation='relu'))  # 3rd hidden layer

model\_MLP3\_DCT.add(Dropout(0.2))

model\_MLP3\_DCT.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP3\_DCT.summary()

# Compile the model

model\_MLP3\_DCT.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP3\_DCT.fit(X\_train\_DCT, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP3\_DCT.evaluate(X\_test\_DCT, y\_test)

tic=time.time()

model\_MLP3\_DCT.predict(X\_test\_DCT[0].reshape(1,200))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 3 Hidden Layers-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

### Using PCA Features

# ### 1 Hidden Layer

# Define the model architecture

model\_MLP1\_PCA = Sequential(name='MLP1\_PCA')

model\_MLP1\_PCA.add(Dense(256, activation='relu', input\_shape=(262,)))  # hidden layer

model\_MLP1\_PCA.add(Dropout(0.2)) #dropout regularization

model\_MLP1\_PCA.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP1\_PCA.summary()

# Compile the model

model\_MLP1\_PCA.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP1\_PCA.fit(X\_train\_PCA, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP1\_PCA.evaluate(X\_test\_PCA, y\_test)

tic=time.time()

model\_MLP1\_PCA.predict(X\_test\_PCA[0].reshape(1,262))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 1 Hidden Layer-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# ### 2 Hidden Layers

# Define the model architecture

model\_MLP2\_PCA = Sequential(name='MLP2\_PCA')

model\_MLP2\_PCA.add(Dense(256, activation='relu', input\_shape=(262,)))  # 1st hidden layer

model\_MLP2\_PCA.add(Dropout(0.2))

model\_MLP2\_PCA.add(Dense(128, activation='relu'))  # 2nd hidden layer

model\_MLP2\_PCA.add(Dropout(0.2))

model\_MLP2\_PCA.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP2\_PCA.summary()

# Compile the model

model\_MLP2\_PCA.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP2\_PCA.fit(X\_train\_PCA, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP2\_PCA.evaluate(X\_test\_PCA, y\_test)

tic=time.time()

model\_MLP2\_PCA.predict(X\_test\_PCA[0].reshape(1,262))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 2 Hidden Layers-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# ### 3 Hidden Layers

# Define the model architecture

model\_MLP3\_PCA = Sequential(name='MLP3\_PCA')

model\_MLP3\_PCA.add(Dense(256, activation='relu', input\_shape=(262,)))  # 1st hidden layer

model\_MLP3\_PCA.add(Dropout(0.2))

model\_MLP3\_PCA.add(Dense(128, activation='relu'))  # 2nd hidden layer

model\_MLP3\_PCA.add(Dropout(0.2))

model\_MLP3\_PCA.add(Dense(64, activation='relu'))  # 3rd hidden layer

model\_MLP3\_PCA.add(Dropout(0.2))

model\_MLP3\_PCA.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP3\_PCA.summary()

# Compile the model

model\_MLP3\_PCA.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP3\_PCA.fit(X\_train\_PCA, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP3\_PCA.evaluate(X\_test\_PCA, y\_test)

tic=time.time()

model\_MLP3\_PCA.predict(X\_test\_PCA[0].reshape(1,262))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 3 Hidden Layers-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

### Using ICA Features

# ### 1 Hidden Layer

# Define the model architecture

model\_MLP1\_ICA = Sequential(name='MLP1\_ICA')

model\_MLP1\_ICA.add(Dense(256, activation='relu', input\_shape=(200,)))  # hidden layer

model\_MLP1\_ICA.add(Dropout(0.2)) #dropout regularization

model\_MLP1\_ICA.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP1\_ICA.summary()

# Compile the model

model\_MLP1\_ICA.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP1\_ICA.fit(X\_train\_ICA, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP1\_ICA.evaluate(X\_test\_ICA, y\_test)

tic=time.time()

model\_MLP1\_ICA.predict(X\_test\_ICA[0].reshape(1,200))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 1 Hidden Layer-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# ### 2 Hidden Layers

# Define the model architecture

model\_MLP2\_ICA = Sequential(name='MLP2\_ICA')

model\_MLP2\_ICA.add(Dense(256, activation='relu', input\_shape=(200,)))  # 1st hidden layer

model\_MLP2\_ICA.add(Dropout(0.2))

model\_MLP2\_ICA.add(Dense(128, activation='relu'))  # 2nd hidden layer

model\_MLP2\_ICA.add(Dropout(0.2))

model\_MLP2\_ICA.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP2\_ICA.summary()

# Compile the model

model\_MLP2\_ICA.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP2\_ICA.fit(X\_train\_ICA, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP2\_ICA.evaluate(X\_test\_ICA, y\_test)

tic=time.time()

model\_MLP2\_ICA.predict(X\_test\_ICA[0].reshape(1,200))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 2 Hidden Layers-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# ### 3 Hidden Layers

# Define the model architecture

model\_MLP3\_ICA = Sequential(name='MLP3\_ICA')

model\_MLP3\_ICA.add(Dense(256, activation='relu', input\_shape=(200,)))  # 1st hidden layer

model\_MLP3\_ICA.add(Dropout(0.2))

model\_MLP3\_ICA.add(Dense(128, activation='relu'))  # 2nd hidden layer

model\_MLP3\_ICA.add(Dropout(0.2))

model\_MLP3\_ICA.add(Dense(64, activation='relu'))  # 3rd hidden layer

model\_MLP3\_ICA.add(Dropout(0.2))

model\_MLP3\_ICA.add(Dense(10, activation='softmax'))  # Output layer

model\_MLP3\_ICA.summary()

# Compile the model

model\_MLP3\_ICA.compile(optimizer='adam', loss='sparse\_categorical\_crossentropy', metrics=['accuracy'])

# Train the model

tic=time.time()

model\_MLP3\_ICA.fit(X\_train\_ICA, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

test\_loss, test\_acc = model\_MLP3\_ICA.evaluate(X\_test\_ICA, y\_test)

tic=time.time()

model\_MLP3\_ICA.predict(X\_test\_ICA[0].reshape(1,200))

toc=time.time()

proc\_time=toc-tic

print("-----MLP With 3 Hidden Layers-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Processing Time for 1 example = {} ms".format(np.round(proc\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# 2. Convolutional Neural Network (LeNet-5)

import os

import cv2

import numpy as np

import matplotlib.pyplot as plt

import random

from sklearn.utils import shuffle

from scipy.fftpack import dct ,idct

import sklearn

from sklearn.decomposition import PCA, FastICA

from sklearn.metrics import accuracy\_score

from keras.models import Sequential

from keras.layers import Dense, Flatten, InputLayer, Dropout, Conv2D, AveragePooling2D

import time

from tensorflow.keras.callbacks import EarlyStopping

training\_path = "Reduced\_MNIST\_Data\Reduced\_Training\_data"

testing\_path = "Reduced\_MNIST\_Data\Reduced\_Testing\_data"

# Define the list of classes

classes = os.listdir(training\_path)

print(classes)

classes = list(map(int, classes))

print(classes)

# Define an empty list to store the data and labels

X\_train = []

y\_train = []

# Loop over the classes

for class\_name in classes:

    class\_path = os.path.join(training\_path, str(class\_name))

    # Loop over the images in the class folder

    for image\_name in os.listdir(class\_path):

        image\_path = os.path.join(class\_path, image\_name)

        # Load the image and append it to the data list

        image = cv2.imread(image\_path, cv2.IMREAD\_GRAYSCALE)

        X\_train.append(image)

        # Append the label to the labels list

        y\_train.append(class\_name)

# Convert the data and labels lists to NumPy arrays

X\_train = np.array(X\_train)

y\_train = np.array(y\_train)

# Print the shape of the data and labels arrays

print("Training Data shape:", X\_train.shape)

print("Training Labels shape:", y\_train.shape)

X\_test = []

y\_test = []

for class\_name in classes:

    class\_path = os.path.join(testing\_path, str(class\_name))

    for image\_name in os.listdir(class\_path):

        image\_path = os.path.join(class\_path, image\_name)

        image = cv2.imread(image\_path, cv2.IMREAD\_GRAYSCALE)

        X\_test.append(image)

        y\_test.append(class\_name)

# Convert the data and labels lists to NumPy arrays

X\_test = np.array(X\_test)

y\_test = np.array(y\_test)

print("Testing Data shape:", X\_test.shape)

print("Testing Labels shape:", y\_test.shape)

X\_train,y\_train = shuffle(X\_train, y\_train, random\_state=4)

X\_test,y\_test = shuffle(X\_test, y\_test, random\_state=4)

#check if shuffling worked correctly

plt.figure()

plt.subplot(121)

plt.title("Is this {} ?".format(y\_train[1050]))

plt.imshow(X\_train[1050])

plt.subplot(122)

plt.title("Is this {} ?".format(y\_test[1050]))

plt.imshow(X\_test[1050])

plt.show()

#reshaping the dataset to fit CNN architectures

X\_train = X\_train.reshape(X\_train.shape[0], 28, 28, 1)

X\_test = X\_test.reshape(X\_test.shape[0], 28, 28, 1)

print(X\_train.shape)

print(X\_test.shape)

## LeNet-5 - No Variations

model = Sequential()

# Convolutional layer 1

model.add(Conv2D(6, (5, 5), activation='relu', input\_shape=(28, 28, 1), padding='valid'))

# Average pooling layer 1

model.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Convolutional layer 2

model.add(Conv2D(16, (5, 5), activation='relu', padding='valid'))

# Average pooling layer 2

model.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Flatten layer

model.add(Flatten())

# Fully connected layer 1

model.add(Dense(120, activation='relu'))

# Fully connected layer 2

model.add(Dense(84, activation='relu'))

# Output layer

model.add(Dense(10, activation='softmax'))

#Early Stopping to avoid fitting issues

early\_stopping = EarlyStopping(monitor='accuracy', patience=3)

model.compile(loss='sparse\_categorical\_crossentropy', optimizer='adam', metrics=['accuracy'])

# Train the model

tic=time.time()

model.fit(X\_train, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

tic=time.time()

test\_loss, test\_acc = model.evaluate(X\_test, y\_test)

toc=time.time()

test\_time=toc-tic

print("-----LeNet-5 - No Variations-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

## Variation #1 - Adding Dropout Regularization

model1 = Sequential()

# Convolutional layer 1

model1.add(Conv2D(6, (5, 5), activation='relu', input\_shape=(28, 28, 1), padding='valid'))

#dropout regularization

model1.add(Dropout(0.2))

# Average pooling layer 1

model1.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Convolutional layer 2

model1.add(Conv2D(16, (5, 5), activation='relu', padding='valid'))

model1.add(Dropout(0.2))

# Average pooling layer 2

model1.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Flatten layer

model1.add(Flatten())

# Fully connected layer 1

model1.add(Dense(120, activation='relu'))

model1.add(Dropout(0.2))

# Fully connected layer 2

model1.add(Dense(84, activation='relu'))

model1.add(Dropout(0.2))

# Output layer

model1.add(Dense(10, activation='softmax'))

model1.compile(loss='sparse\_categorical\_crossentropy', optimizer='adam', metrics=['accuracy'])

# Train the model

tic=time.time()

model1.fit(X\_train, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

tic=time.time()

test\_loss, test\_acc = model1.evaluate(X\_test, y\_test)

toc=time.time()

test\_time=toc-tic

print("-----Variation #1 - Adding Dropout-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

## Variation #2 - Increasing Number of Filters in Conv Layers

model2 = Sequential()

# Convolutional layer 1

model2.add(Conv2D(12, (5, 5), activation='relu', input\_shape=(28, 28, 1), padding='valid'))

# Average pooling layer 1

model2.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Convolutional layer 2

model2.add(Conv2D(32, (5, 5), activation='relu', padding='valid'))

# Average pooling layer 2

model2.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Flatten layer

model2.add(Flatten())

# Fully connected layer 1

model2.add(Dense(120, activation='relu'))

# Fully connected layer 2

model2.add(Dense(84, activation='relu'))

# Output layer

model2.add(Dense(10, activation='softmax'))

model2.compile(loss='sparse\_categorical\_crossentropy', optimizer='adam', metrics=['accuracy'])

# Train the model

tic=time.time()

model2.fit(X\_train, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

tic=time.time()

test\_loss, test\_acc = model2.evaluate(X\_test, y\_test)

toc=time.time()

test\_time=toc-tic

print("-----Variation #2 - Increasing no. of Filters-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Testing Time= {} ms".format(np.round(test\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

## Variation #3 - Adding "Same" Padding to Conv Layers

model3 = Sequential()

# Convolutional layer 1

model3.add(Conv2D(12, (5, 5), activation='relu', input\_shape=(28, 28, 1), padding='same'))

# Average pooling layer 1

model3.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Convolutional layer 2

model3.add(Conv2D(32, (5, 5), activation='relu', padding='same'))

# Average pooling layer 2

model3.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Flatten layer

model3.add(Flatten())

# Fully connected layer 1

model3.add(Dense(120, activation='relu'))

# Fully connected layer 2

model3.add(Dense(84, activation='relu'))

# Output layer

model3.add(Dense(10, activation='softmax'))

model3.compile(loss='sparse\_categorical\_crossentropy', optimizer='adam', metrics=['accuracy'])

# Train the model

tic=time.time()

model3.fit(X\_train, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

tic=time.time()

test\_loss, test\_acc = model3.evaluate(X\_test, y\_test)

toc=time.time()

test\_time=toc-tic

print("-----Variation #3 - Adding 'Same' Padding-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

## Variation #4 - Using "Tanh" Activation

model4 = Sequential()

# Convolutional layer 1

model4.add(Conv2D(12, (5, 5), activation='tanh', input\_shape=(28, 28, 1), padding='valid'))

# Average pooling layer 1

model4.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Convolutional layer 2

model4.add(Conv2D(32, (5, 5), activation='tanh', padding='valid'))

# Average pooling layer 2

model4.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

# Flatten layer

model4.add(Flatten())

# Fully connected layer 1

model4.add(Dense(120, activation='tanh'))

# Fully connected layer 2

model4.add(Dense(84, activation='tanh'))

# Output layer

model4.add(Dense(10, activation='softmax'))

model4.compile(loss='sparse\_categorical\_crossentropy', optimizer='adam', metrics=['accuracy'])

# Train the model

tic=time.time()

model4.fit(X\_train, y\_train, epochs=30, batch\_size=32, callbacks=[early\_stopping])

toc=time.time()

training\_time=toc-tic

# Evaluate the model on the test data

tic=time.time()

test\_loss, test\_acc = model4.evaluate(X\_test, y\_test)

toc=time.time()

test\_time=toc-tic

print("-----Variation #4 - Using Tanh Activation-----\n")

print("Training Time = {} s".format(np.round(training\_time, 1)))

print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))

print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

# 3. Comparing the Results

Table : Comparative Analysis for Different Models

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Features** | | | | | |
| **DCT** | | **PCA** | | **ICA** | |
| **Accuracy** | **Training Time** | **Accuracy** | **Training Time** | **Accuracy** | **Training Time** |
| **Classifier** | |
| **K-means Clustering** | 1 | 62.65% | 0.619s | 63.15% | 0.903s | 64.4% | 0.166s |
| 4 | 89% | 1.221s | 88.65% | 1.812s | 81.5% | 0.542s |
| 16 | 93.15% | 3.504s | 93.25% | 4.783s | 89.35% | 1.424s |
| 32 | **95.4%** | 6.798s | 94.75% | 9.291s | 89.1% | 1.872s |
| **SVM** | Linear | 94.35% | 1.808s | 93.85% | 3.814s | 77.8% | 6.240s |
| Non-Linear (RBF) | 97.35% | 2.617s | **97.65%** | 7.158s | 93.8% | 0.783s |
| **Multi-layer Perceptron (MLP)** | | | | | | | |
|  |  | **DCT** | | **PCA** | | **ICA** | |
| **Variations** | **Accuracy** | **Processing Time** | **Accuracy** | **Processing Time** | **Accuracy** | **Processing Time** |
| **MLP** | 1-Hidden | 95.0% | 271.3 ms | **95.20%** | 246.3 ms | 93.20% | 70.8 ms |
| 2-Hidden | 94.30% | 187.5 ms | 93.80% | 413.9 ms | **95.00%** | 88.5 ms |
| 3-Hidden | **95.70%** | 197.5 ms | 94.70% | 218.4 ms | 94.70% | 93.5 ms |
| **CNN – No Features** | | | | | | | |
|  | **Variations** | **Accuracy** | | **Training Time** | | **Testing Time** | |
| **CNN** | No Variations | 97.40% | | 41.9 s | | 515.6 ms | |
| Dropout | **98.60%** | | 76.5 s | | 505.6 ms | |
| Increasing Number of Filters | 98.50% | | 40.9 s | | 493.2 ms | |
| “Same” Padding | 97.80% | | 61.7 s | | 614.4 ms | |
| Tanh Activation | 98.30% | | 50.9 s | | 1004.3 ms | |

## Notes

* The Multi-layer perceptron processing time measurements are based on how much time it takes the model to predict the class of one image.
* The Convolutional Neural Network training time measurements are for different number of epochs, considering Early Stopping was used to avoid fitting issues.
* The time measurements throughout the experiments are heavily dependent on the machine the models are running on and the processes that run on that machine.
* Adding more layers in Fully connected Networks might not always be the best option, as there will be diminishing returns in the accuracy.
* Dropout regularization (and regularization in general) increases the performance of the model, even slightly, as it reduces overfitting, and therefore the model generalizes better.

# 4. Digit Spectrograms

Import Libraries

import os  
from matplotlib import pyplot as plt  
import tensorflow as tf   
!pip install tensorflow\_io  
import tensorflow\_io as tfio  
from tensorflow import keras  
from keras import backend as k  
import cv2  
import numpy as np  
import random  
from sklearn.utils import shuffle  
from scipy.fftpack import dct ,idct  
import sklearn  
from sklearn.decomposition import PCA, FastICA  
from sklearn.metrics import accuracy\_score  
from keras.models import Sequential  
from keras.layers import Dense, Flatten, InputLayer, Dropout, Conv2D, AveragePooling2D,MaxPool2D  
import time  
from tensorflow.keras.callbacks import EarlyStopping

from google.colab import drive  
drive.mount('/content/drive')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force\_remount=True).

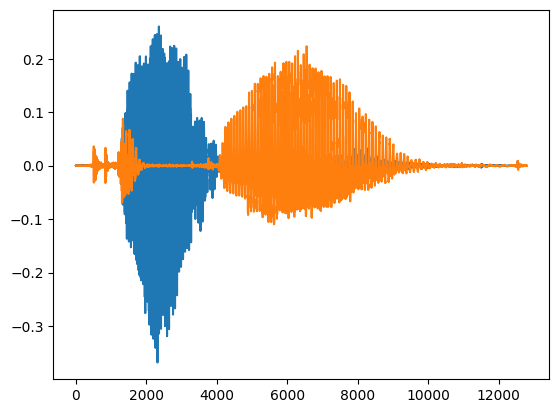
Variables: batch: the process of splitting the training dataset in n batches (mini-batches), classes: number of classifications (labels) of the data, epochs: variations, one epoch is one forward pass + one backward pass on training

#batch\_size = 20  
num\_classes = 10  
epochs = 4

a function that returns audio in numeric representation

def load\_wav\_16k\_mono(filename):  
 # Load encoded wav file  
 file\_contents = tf.io.read\_file(filename)  
 # Decode wav (tensors by channels)   
 wav, sample\_rate = tf.audio.decode\_wav(file\_contents, desired\_channels=1)  
 # Removes trailing axis  
 wav = tf.squeeze(wav, axis=-1)  
 sample\_rate = tf.cast(sample\_rate, dtype=tf.int64)  
 # Goes from 44100Hz to 16000hz - amplitude of the audio signal  
 #wav = tfio.audio.resample(wav, rate\_in=sample\_rate, rate\_out=16000)  
 return wav

TRAIN\_FILE = os.path.join('/content','drive','MyDrive','audio-data','Train','C03n\_0.wav')  
TEST\_FILE = os.path.join('/content','drive','MyDrive','audio-data','Test','C04n\_2.wav')  
  
wave = load\_wav\_16k\_mono(TRAIN\_FILE)  
nwave = load\_wav\_16k\_mono(TEST\_FILE)  
plt.plot(wave)  
plt.plot(nwave)  
plt.show()

Check a sample of audio

TRAIN = os.path.join('/content','drive','MyDrive','audio-data', 'Train')  
TEST = os.path.join('/content','drive','MyDrive','audio-data', 'Test')

Read all audio files and sort

train = tf.data.Dataset.list\_files(TRAIN+'/\*.wav')  
train = sorted(list(train.as\_numpy\_iterator()))  
train = tf.data.Dataset.from\_tensor\_slices(train)  
test = tf.data.Dataset.list\_files(TEST+'/\*.wav')  
test = sorted(list(test.as\_numpy\_iterator()))  
test = tf.data.Dataset.from\_tensor\_slices(test)

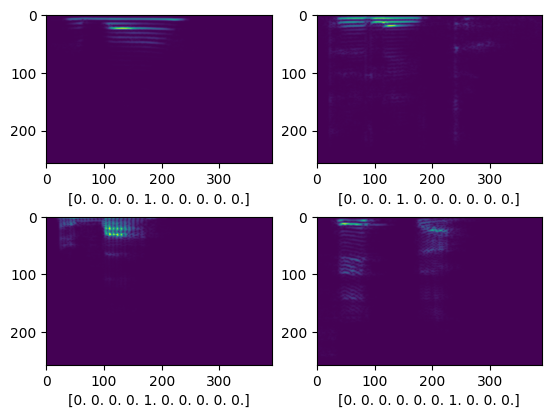
Add Labels

iterations = 0  
i = 0  
train\_label = []  
while iterations!=len(train):  
 iterations +=1  
 train\_label.append(i)  
 i += 1  
 if i == 10 :  
 i = 0  
train\_label=keras.utils.to\_categorical(train\_label,num\_classes)  
trainings = tf.data.Dataset.zip((train, tf.data.Dataset.from\_tensor\_slices(train\_label)))  
#---------------------------------------------------------------#  
iterations = 0  
i = 0  
test\_label=[]  
while iterations!=len(test):  
 iterations +=1  
 test\_label.append(i)  
 i += 1  
 if i == 10 :  
 i = 0  
test\_label=keras.utils.to\_categorical(test\_label,num\_classes)  
testings = tf.data.Dataset.zip((test, tf.data.Dataset.from\_tensor\_slices(test\_label)))

Build Preprocessing Function to get spectogram

def preprocess(file\_path, label):   
 wav = load\_wav\_16k\_mono(file\_path)  
 #wav = wav[:48000]  
 #zero\_padding = tf.zeros([48000] - tf.shape(wav), dtype=tf.float32)  
 #wav = tf.concat([zero\_padding, wav],0)  
 spectrogram = tf.signal.stft(wav, frame\_length=320, frame\_step=32)  
 spectrogram = tf.abs(spectrogram)  
 spectrogram = tf.expand\_dims(spectrogram, axis=2)  
 return spectrogram, label

Draw examples of spectogram

for i in range(4):  
 filepath, label = trainings.shuffle(buffer\_size=10000).as\_numpy\_iterator().next()  
 spectrogram, label = preprocess(filepath, label)  
 plt.subplot(2,2,i+1)  
 plt.imshow(tf.transpose(spectrogram)[0])  
 plt.xlabel(label)  
plt.show()

Convert all to Spectogram

# train data  
x\_train = trainings.map(preprocess)  
x\_train = x\_train.cache()  
x\_train = x\_train.shuffle(buffer\_size=1000)  
x\_train = x\_train.batch(16) # 16 at a time  
x\_train = x\_train.prefetch(8)  
# test data  
x\_test = testings.map(preprocess)  
x\_test = x\_test.cache()  
x\_test = x\_test.shuffle(buffer\_size=1000)  
x\_test = x\_test.batch(16) # 16 at a time  
x\_test = x\_test.prefetch(8)

# test one batch  
samples, labels = x\_train.as\_numpy\_iterator().next()  
print(samples.shape)  
print('\n',labels)

(16, 391, 257, 1)  
  
 [[1. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
 [1. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0. 0. 0. 0. 0. 1.]  
 [0. 0. 1. 0. 0. 0. 0. 0. 0. 0.]  
 [0. 1. 0. 0. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 0. 1. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0. 0. 0. 1. 0. 0.]  
 [0. 0. 0. 0. 0. 0. 1. 0. 0. 0.]  
 [0. 1. 0. 0. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 0. 1. 0. 0. 0. 0. 0. 0.]  
 [1. 0. 0. 0. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 0. 1. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 1. 0. 0. 0. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 0. 0. 0. 0. 1. 0.]  
 [0. 0. 0. 0. 0. 1. 0. 0. 0. 0.]  
 [0. 0. 0. 0. 1. 0. 0. 0. 0. 0.]]

Design the CNN architecture

the 1st model

model1 = Sequential()  
# Convolutional layer 1  
model1.add(Conv2D(6, (5, 5), activation='relu', input\_shape=(391, 257, 1), padding='valid'))  
# Average pooling layer 1  
model1.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Convolutional layer 2  
model1.add(Conv2D(16, (5, 5), activation='relu', padding='valid'))  
# Average pooling layer 2  
model1.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Flatten layer  
model1.add(Flatten())  
# Fully connected layer 1  
model1.add(Dense(120, activation='relu'))  
# Fully connected layer 2  
model1.add(Dense(84, activation='relu'))  
# Output layer  
model1.add(Dense(10, activation='softmax'))  
model1.summary()

Model: "sequential\_12"  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
 Layer (type) Output Shape Param #   
=================================================================  
 conv2d\_24 (Conv2D) (None, 387, 253, 6) 156   
   
 average\_pooling2d\_24 (Avera (None, 193, 126, 6) 0   
 gePooling2D)   
   
 conv2d\_25 (Conv2D) (None, 189, 122, 16) 2416   
   
 average\_pooling2d\_25 (Avera (None, 94, 61, 16) 0   
 gePooling2D)   
   
 flatten\_12 (Flatten) (None, 91744) 0   
   
 dense\_36 (Dense) (None, 120) 11009400   
   
 dense\_37 (Dense) (None, 84) 10164   
   
 dense\_38 (Dense) (None, 10) 850   
   
=================================================================  
Total params: 11,022,986  
Trainable params: 11,022,986  
Non-trainable params: 0  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#Early Stopping to avoid fitting issues  
early\_stopping = EarlyStopping(monitor='accuracy', patience=5)  
  
model1.compile(loss = keras.losses.CategoricalCrossentropy(), optimizer='adam', metrics=['accuracy'])  
  
# Train the model  
tic=time.time()  
model1.fit(x\_train, epochs=80, callbacks=[early\_stopping])  
toc=time.time()  
training\_time=toc-tic

Epoch 1/80  
75/75 [==============================] - 18s 27ms/step - loss: 1.3095 - accuracy: 0.6408  
Epoch 2/80  
75/75 [==============================] - 2s 23ms/step - loss: 0.4219 - accuracy: 0.8983  
Epoch 3/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.1560 - accuracy: 0.9658  
Epoch 4/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.1012 - accuracy: 0.9792  
Epoch 5/80  
75/75 [==============================] - 1s 18ms/step - loss: 0.0735 - accuracy: 0.9850  
Epoch 6/80  
75/75 [==============================] - 1s 18ms/step - loss: 0.1282 - accuracy: 0.9808  
Epoch 7/80  
75/75 [==============================] - 1s 18ms/step - loss: 0.1137 - accuracy: 0.9792  
Epoch 8/80  
75/75 [==============================] - 1s 16ms/step - loss: 0.0534 - accuracy: 0.9950  
Epoch 9/80  
75/75 [==============================] - 1s 16ms/step - loss: 0.0187 - accuracy: 0.9967  
Epoch 10/80  
75/75 [==============================] - 1s 16ms/step - loss: 0.0024 - accuracy: 1.0000  
Epoch 11/80  
75/75 [==============================] - 1s 16ms/step - loss: 0.0011 - accuracy: 1.0000  
Epoch 12/80  
75/75 [==============================] - 1s 16ms/step - loss: 7.9277e-04 - accuracy: 1.0000  
Epoch 13/80  
75/75 [==============================] - 1s 16ms/step - loss: 5.9054e-04 - accuracy: 1.0000  
Epoch 14/80  
75/75 [==============================] - 1s 16ms/step - loss: 4.7252e-04 - accuracy: 1.0000  
Epoch 15/80  
75/75 [==============================] - 1s 17ms/step - loss: 3.8069e-04 - accuracy: 1.0000

# Evaluate the model on the test data  
tic=time.time()  
test\_loss, test\_acc = model1.evaluate(x\_test)  
toc=time.time()  
test\_time=toc-tic  
  
  
print("----- #1 - orignal model-----\n")  
print("Training Time = {} s".format(np.round(training\_time, 1)))  
print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))  
print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

19/19 [==============================] - 4s 9ms/step - loss: 2.2276 - accuracy: 0.8567  
----- #1 - orignal model-----  
  
Training Time = 38.5 s  
Testing Time = 3582.1 ms  
Test Accuracy = 85.70 %:

the 2nd model

model2 = Sequential()  
# Convolutional layer 1  
model2.add(Conv2D(6, (5, 5), activation='relu', input\_shape=(391, 257, 1), padding='valid'))  
#dropout regularization  
model2.add(Dropout(0.2))  
# Average pooling layer 1  
model2.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Convolutional layer 2  
model2.add(Conv2D(16, (5, 5), activation='relu', padding='valid'))  
model2.add(Dropout(0.2))  
# Average pooling layer 2  
model2.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Flatten layer  
model2.add(Flatten())  
# Fully connected layer 1  
model2.add(Dense(120, activation='relu'))  
model2.add(Dropout(0.2))  
# Fully connected layer 2  
model2.add(Dense(84, activation='relu'))  
model2.add(Dropout(0.2))  
# Output layer  
model2.add(Dense(10, activation='softmax'))  
  
model2.summary()

Model: "sequential\_13"  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
 Layer (type) Output Shape Param #   
=================================================================  
 conv2d\_26 (Conv2D) (None, 387, 253, 6) 156   
   
 dropout\_12 (Dropout) (None, 387, 253, 6) 0   
   
 average\_pooling2d\_26 (Avera (None, 193, 126, 6) 0   
 gePooling2D)   
   
 conv2d\_27 (Conv2D) (None, 189, 122, 16) 2416   
   
 dropout\_13 (Dropout) (None, 189, 122, 16) 0   
   
 average\_pooling2d\_27 (Avera (None, 94, 61, 16) 0   
 gePooling2D)   
   
 flatten\_13 (Flatten) (None, 91744) 0   
   
 dense\_39 (Dense) (None, 120) 11009400   
   
 dropout\_14 (Dropout) (None, 120) 0   
   
 dense\_40 (Dense) (None, 84) 10164   
   
 dropout\_15 (Dropout) (None, 84) 0   
   
 dense\_41 (Dense) (None, 10) 850   
   
=================================================================  
Total params: 11,022,986  
Trainable params: 11,022,986  
Non-trainable params: 0  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#Early Stopping to avoid fitting issues  
early\_stopping = EarlyStopping(monitor='accuracy', patience=3)  
  
model2.compile(loss= keras.losses.CategoricalCrossentropy(), optimizer='adam', metrics=['accuracy'])  
  
# Train the model  
tic=time.time()  
model2.fit(x\_train, epochs=80, batch\_size=32, callbacks=[early\_stopping])  
toc=time.time()  
training\_time=toc-tic

Epoch 1/80  
75/75 [==============================] - 4s 23ms/step - loss: 1.6791 - accuracy: 0.4958  
Epoch 2/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.7131 - accuracy: 0.7817  
Epoch 3/80  
75/75 [==============================] - 2s 22ms/step - loss: 0.5013 - accuracy: 0.8533  
Epoch 4/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.3383 - accuracy: 0.9192  
Epoch 5/80  
75/75 [==============================] - 2s 23ms/step - loss: 0.2278 - accuracy: 0.9350  
Epoch 6/80  
75/75 [==============================] - 2s 23ms/step - loss: 0.2231 - accuracy: 0.9408  
Epoch 7/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.1004 - accuracy: 0.9742  
Epoch 8/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.0944 - accuracy: 0.9733  
Epoch 9/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.1767 - accuracy: 0.9600  
Epoch 10/80  
75/75 [==============================] - 2s 21ms/step - loss: 0.0773 - accuracy: 0.9733

# Evaluate the model on the test data  
tic=time.time()  
test\_loss, test\_acc = model2.evaluate(x\_test)  
toc=time.time()  
test\_time=toc-tic  
  
  
print("---- #2 the 2nd model-----\n")  
print("Training Time = {} s".format(np.round(training\_time, 1)))  
print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))  
print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

19/19 [==============================] - 0s 7ms/step - loss: 0.7661 - accuracy: 0.8633  
---- #2 the 2nd model-----  
  
Training Time = 20.2 s  
Testing Time = 281.0 ms  
Test Accuracy = 86.30 %:

the 3rd model

model3 = Sequential()  
# Convolutional layer 1  
model3.add(Conv2D(12, (5, 5), activation='relu', input\_shape=(391, 257, 1), padding='same'))  
# Average pooling layer 1  
model3.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Convolutional layer 2  
model3.add(Conv2D(32, (5, 5), activation='relu', padding='same'))  
# Average pooling layer 2  
model3.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Flatten layer  
model3.add(Flatten())  
# Fully connected layer 1  
model3.add(Dense(120, activation='relu'))  
# Fully connected layer 2  
model3.add(Dense(84, activation='relu'))  
# Output layer  
model3.add(Dense(10, activation='softmax'))

model3.summary()

Model: "sequential\_14"  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
 Layer (type) Output Shape Param #   
=================================================================  
 conv2d\_28 (Conv2D) (None, 391, 257, 12) 312   
   
 average\_pooling2d\_28 (Avera (None, 195, 128, 12) 0   
 gePooling2D)   
   
 conv2d\_29 (Conv2D) (None, 195, 128, 32) 9632   
   
 average\_pooling2d\_29 (Avera (None, 97, 64, 32) 0   
 gePooling2D)   
   
 flatten\_14 (Flatten) (None, 198656) 0   
   
 dense\_42 (Dense) (None, 120) 23838840   
   
 dense\_43 (Dense) (None, 84) 10164   
   
 dense\_44 (Dense) (None, 10) 850   
   
=================================================================  
Total params: 23,859,798  
Trainable params: 23,859,798  
Non-trainable params: 0  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

#Early Stopping to avoid fitting issues  
early\_stopping = EarlyStopping(monitor='accuracy', patience=5)  
  
model3.compile(loss=keras.losses.CategoricalCrossentropy(), optimizer='adam', metrics=['accuracy'])  
  
  
# Train the model  
tic=time.time()  
model3.fit(x\_train, epochs=80, batch\_size=32, callbacks=[early\_stopping])  
toc=time.time()  
training\_time=toc-tic

Epoch 1/80  
75/75 [==============================] - 4s 28ms/step - loss: 1.3406 - accuracy: 0.6283  
Epoch 2/80  
75/75 [==============================] - 2s 25ms/step - loss: 0.2879 - accuracy: 0.9225  
Epoch 3/80  
75/75 [==============================] - 2s 24ms/step - loss: 0.2405 - accuracy: 0.9417  
Epoch 4/80  
75/75 [==============================] - 2s 24ms/step - loss: 0.1124 - accuracy: 0.9717  
Epoch 5/80  
75/75 [==============================] - 2s 24ms/step - loss: 0.1616 - accuracy: 0.9700  
Epoch 6/80  
75/75 [==============================] - 2s 25ms/step - loss: 0.0184 - accuracy: 0.9967  
Epoch 7/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0040 - accuracy: 1.0000  
Epoch 8/80  
75/75 [==============================] - 2s 30ms/step - loss: 0.0016 - accuracy: 1.0000  
Epoch 9/80  
75/75 [==============================] - 2s 28ms/step - loss: 9.5883e-04 - accuracy: 1.0000  
Epoch 10/80  
75/75 [==============================] - 2s 24ms/step - loss: 6.4324e-04 - accuracy: 1.0000  
Epoch 11/80  
75/75 [==============================] - 2s 24ms/step - loss: 4.5650e-04 - accuracy: 1.0000  
Epoch 12/80  
75/75 [==============================] - 2s 24ms/step - loss: 3.5766e-04 - accuracy: 1.0000

# Evaluate the model on the test data  
tic=time.time()  
test\_loss, test\_acc = model3.evaluate(x\_test)  
toc=time.time()  
test\_time=toc-tic  
  
  
print("----- #3 - the 3rd model-----\n")  
print("Training Time = {} s".format(np.round(training\_time, 1)))  
print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))  
print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

19/19 [==============================] - 0s 8ms/step - loss: 1.0015 - accuracy: 0.9100  
----- #3 - the 3rd model-----  
  
Training Time = 27.5 s  
Testing Time = 439.6 ms  
Test Accuracy = 91.00 %:

the 4th model

model4 = Sequential()  
# Convolutional layer 1  
model4.add(Conv2D(12, (5, 5), activation='tanh', input\_shape=(391, 257, 1), padding='valid'))  
# Average pooling layer 1  
model4.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Convolutional layer 2  
model4.add(Conv2D(32, (5, 5), activation='tanh', padding='valid'))  
# Average pooling layer 2  
model4.add(AveragePooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))  
# Flatten layer  
model4.add(Flatten())  
# Fully connected layer 1  
model4.add(Dense(120, activation='tanh'))  
# Fully connected layer 2  
model4.add(Dense(84, activation='tanh'))  
# Output layer  
model4.add(Dense(10, activation='softmax'))  
model4.summary()

Model: "sequential\_15"  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
 Layer (type) Output Shape Param #   
=================================================================  
 conv2d\_30 (Conv2D) (None, 387, 253, 12) 312   
   
 average\_pooling2d\_30 (Avera (None, 193, 126, 12) 0   
 gePooling2D)   
   
 conv2d\_31 (Conv2D) (None, 189, 122, 32) 9632   
   
 average\_pooling2d\_31 (Avera (None, 94, 61, 32) 0   
 gePooling2D)   
   
 flatten\_15 (Flatten) (None, 183488) 0   
   
 dense\_45 (Dense) (None, 120) 22018680   
   
 dense\_46 (Dense) (None, 84) 10164   
   
 dense\_47 (Dense) (None, 10) 850   
   
=================================================================  
Total params: 22,039,638  
Trainable params: 22,039,638  
Non-trainable params: 0  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

early\_stopping = EarlyStopping(monitor='accuracy', patience=5)  
  
model4.compile(loss = keras.losses.CategoricalCrossentropy(), optimizer='adam', metrics=['accuracy'])  
  
# Train the model  
tic=time.time()  
model4.fit(x\_train, epochs=80, batch\_size=32, callbacks=[early\_stopping])  
toc=time.time()  
training\_time=toc-tic

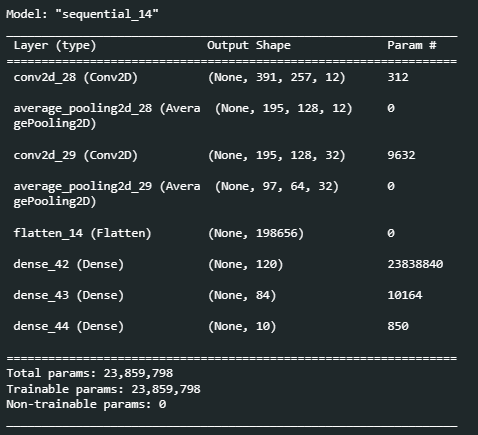
Epoch 1/80  
75/75 [==============================] - 5s 27ms/step - loss: 1.2450 - accuracy: 0.6425  
Epoch 2/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.2714 - accuracy: 0.9442  
Epoch 3/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0994 - accuracy: 0.9850  
Epoch 4/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0397 - accuracy: 0.9983  
Epoch 5/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0149 - accuracy: 1.0000  
Epoch 6/80  
75/75 [==============================] - 2s 27ms/step - loss: 0.0085 - accuracy: 1.0000  
Epoch 7/80  
75/75 [==============================] - 2s 28ms/step - loss: 0.0063 - accuracy: 1.0000  
Epoch 8/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0050 - accuracy: 1.0000  
Epoch 9/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0042 - accuracy: 1.0000  
Epoch 10/80  
75/75 [==============================] - 2s 26ms/step - loss: 0.0035 - accuracy: 1.0000

# Evaluate the model on the test data  
tic=time.time()  
test\_loss, test\_acc = model4.evaluate(x\_test)  
toc=time.time()  
test\_time=toc-tic  
  
print("----- #4 - the 4th model -----\n")  
print("Training Time = {} s".format(np.round(training\_time, 1)))  
print("Testing Time = {} ms".format(np.round(test\_time\*1000, 1)))  
print('Test Accuracy = {:.2f} %:'.format(np.round(test\_acc, 3)\*100))

19/19 [==============================] - 0s 11ms/step - loss: 0.3012 - accuracy: 0.9067  
----- #4 - the 4th model -----  
  
Training Time = 23.6 s  
Testing Time = 440.5 ms  
Test Accuracy = 90.70 %:

## Notes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Part 4** | | | | |
|  | **Variations** | **Accuracy** | **Training Time** | **Testing Time** |
| **CNN** | 1st model | 85.7 % | 38.4 s | 3582.1 ms |
| 2nd model | 86.30 % | 20.2 s | 281 ms |
| 3rd model | 91 % | 27.5 s | 439.6 ms |
| 4th model | 90.7 % | 23.6 s | 440.5 ms |

For this problem we choose the following architecture that have the most accuracy: **91 %** using the dataset ‘digits\_audio\_from0to10 ’which have **120 speakers for training** set and **30 speakers for test set**.