

✓ Introduction

Capstone Project Submission

- Name: Adam Marianacci
- Scheduled project review date/time: 3/19/2024/1:30pm EST

✓ Business Understanding

It is my task to help companies with computer vision for their A.I. tennis ball machines.

✓ Data Understanding

I created all the data used in this project. I used roughly 3,000 images that consisted of a balanced class of forehand and backhand tennis shots. The limitations of the dataset was that it was a fairly small set of data. The images used only consisted of myself hitting forehand and backhand tennis shots. The data used for this project was useful to build a convolutional neural network for image classification because there was a minimal amount of variance in the data used. The neural net was able to learn specific patterns about forehands and backhands with minimal "noise" in the data because of the consistency of the data used.

Dataset: [ForehandsandBackhands](#)

✓ Data Preparation

I started by filming myself in a few different locations hitting approximately 200 forehands and 200 backhands. I then used an open source editing program called "Shotcut" to edit all of my swings from the start of my swing up until the point of contact with the tennis ball. I then converted these edited video shots into an mp4 format so that they could be extracted into frames. One shot clip yielded roughly 15 images of data. I extracted the frames into folders and labeled the frames as either belonging to one of two classes "1" for forehand and "0" for backhand. I constructed an iterator into my pipeline to be able to call on specific "batches" of data. I then scaled my data in the pipeline to make sure that all the images used were the same size. I then set up a train, test, split which was 80% for training, 10% for validation, and 10% for testing. I then set up a way to save and load my data after it has been trained into my pipeline so that more data could be input in the future.

For reproducibility I uploaded my data to Kaggle and tested my notebook in Google Colab to ensure my notebook would run smoothly. If you wish to access the data and run this notebook on another platform such as Google Colab you will have to have a Kaggle account.

- In the cell below you will have to enter your Kaggle username and API Key. This [creating token](#) link will help give you instructions on how to obtain an API key.
- This [link](#) is also helpful for more information on how to use Kaggle

Uncomment the cell below if you wish to work from another notebook

```
1 import json
2 import os
3 from pathlib import Path
4
5 # please enter your Kaggle username and api key here
6 username = 'adammarianacci'
7 key = ''
8
9 # your api key
```

```
10 api_key = {
11 'username': username,
12 'key': key}
13
14 # uses pathlib Path
15 kaggle_path = Path('/root/.kaggle')
16 os.makedirs(kaggle_path, exist_ok=True)
17
18 # opens file and dumps python dict to json object
19 with open (kaggle_path/'kaggle.json', 'w') as handl:
20     json.dump(api_key,handl)
21
22 os.chmod(kaggle_path/'kaggle.json', 600)
23
```

Uncomment the cell below if you would like to download the data.

```
1 # This downloads the images for training and testing
2 !kaggle datasets download -d adammarianacci/forehands-and-backhands
3 !unzip forehands-and-backhands.zip
```

[Show hidden output](#)

```
1 # Importing the necessary libraries
2
3 import pandas as pd
4 import numpy as np
5 import matplotlib.pyplot as plt
6 import seaborn as sns
7 import cv2
8 import imghdr
9 import os
10 import PIL
11 import tensorflow as tf
12 import pickle
13 import hashlib
14
```

```

15 from tensorflow import keras
16 from tensorflow.keras import layers
17 from tensorflow.keras.models import Sequential
18 from tensorflow.keras.layers import Conv2D, MaxPooling2D, Dense, Flatten, Dropout
19 from tensorflow.keras.metrics import Precision, Recall, BinaryAccuracy
20 from tensorflow.keras.models import load_model
21 from tensorflow.keras.optimizers import Adam, Adamax
22 from tensorflow.keras.regularizers import l2
23 from sklearn.model_selection import train_test_split, cross_val_score
24 from sklearn.metrics import accuracy_score, precision_score, confusion_matrix
25 import warnings
26 warnings.filterwarnings('ignore')

```

The cell below enables you to select videos within a specified folder. The function enables you to extract frames from the video file and puts them into a specified folder within the directory. It iterates thru all video files ending with 'mp4' and 'avi' and saves the the extracted frames into a specified folder. If you wish to use this function uncomment the cell below and use your own file paths for the 'video_directory' and 'output directory'

```

1 # # Directory containing the video files
2 # video_directory = '/home/adam/Desktop/video_test'
3
4 # # Output directory for extracted frames
5 # output_directory = '/home/adam/Desktop/frame_test'
6
7 # # Function to extract frames from a video file
8 # def extract_frames(video_file, output_dir):
9 #     # Open the video file
10 #     vid = cv2.VideoCapture(video_file)
11 #     if not vid.isOpened():
12 #         print(f"Error: Couldn't open video file '{video_file}'")
13 #         return
14
15 #     # Create output directory if it doesn't exist
16 #     os.makedirs(output_dir, exist_ok=True)
17
18 #     # Initialize frame counter

```

```

18 # # Initialize frame counter
19 # current_frame = 0
20
21 # # Read frames and save them as images
22 # while True:
23 #     success, frame = vid.read()
24
25 #     if not success:
26 #         print(f"End of video reached or couldn't read frame from '{video_file}'")
27 #         break
28
29 #     if frame.shape[0] > 0 and frame.shape[1] > 0:
30 #         # Save the frame as an image
31 #         output_path = os.path.join(output_dir, f'{os.path.splitext(os.path.basename(video_file))[0]}_
32 #         cv2.imwrite(output_path, frame)
33 #         current_frame += 1
34 #     else:
35 #         print(f"Error: Invalid frame size in '{video_file}'")
36
37 # vid.release()
38
39 # # Iterate over video files in the directory
40 # for video_file in os.listdir(video_directory):
41 #     if video_file.endswith('.mp4') or video_file.endswith('.avi'):
42 #         video_path = os.path.join(video_directory, video_file)
43 #         extract_frames(video_path, output_directory)

```

1 # Defining a data directory
2 data_dir = 'data'

The cell below sets up a variable to recognize different images types.

```

1 # Defining image extensions
2 image_exts = ['jpeg', 'jpg', 'bmp', 'png']

```

The cell below shows that my data directory contains my 2 classes.

```

1 # Checking the contents of my data directory
2 os.listdir(data_dir)

['forehands', 'backhands']

```

The cell below sets up a for loop that goes thru the folders and all images in the data directory and removes any images that does not have the appropriate image extensions. Note the skipping output is just skipping directories but there are no issues with the actual files in the directories.

```

1 # Iterate through directories containing images
2 for image_class in os.listdir(data_dir):
3     class_dir = os.path.join(data_dir, image_class)
4     if not os.path.isdir(class_dir):
5         continue # Skip if not a directory
6     for image in os.listdir(class_dir):
7         image_path = os.path.join(class_dir, image)
8         if not os.path.isfile(image_path):
9             print('Skipping non-file:', image_path)
10            continue # Skip if not a file
11        try:
12            tip = imghdr.what(image_path) # Get image type directly
13            if tip is None or tip not in image_exts:
14                print('Image not in ext list or invalid format: {}'.format(image_path))
15                os.remove(image_path)
16        except Exception as e:
17            print('Issue with image {}: {}'.format(image_path, str(e)))
18            # Log the error or handle it appropriately

Skipping non-file: data/forehands/vids
Skipping non-file: data/forehands/frames
Skipping non-file: data/backhands/vids
Skipping non-file: data/backhands/frames

```

Here we start to preprocess the data, we are starting with approx. 6,000 images belonging to two classes that are balanced. This sets up the data pipeline and enables access to "batches" of data using the iterator.

```
1 # loading in tensorflow and keras on our data
2 data = tf.keras.preprocessing.image_dataset_from_directory('data')
    Found 2836 files belonging to 2 classes.
```

```
1 # allow acces to generator from the data pipeline
2 data_iterator = data.as_numpy_iterator()
```

```
1 # allow access to pull a batch (32 images) of data
2 batch = data_iterator.next()
```

batch size of 32, image size of 256x256, 3 (channels of color, RGB)

```
1 # Images represented as numpy arrays
2 batch[0].shape
    (32, 256, 256, 3)
```

```
1 # representing the labels in the batch
2 batch[1]
    array([0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0,
           1, 0, 1, 0, 1, 0, 0, 1, 1, 0], dtype=int32)
```

Below I am plotting some images from a batch to see which classes my images belong to and we can see Forehands = 1 and Backhands = 0.

```
1 # viewing some data to see what class my frames belong to
2 fig, ax = plt.subplots(ncols=4, figsize=(20,20))
3 for idx, img in enumerate(batch[0][:4]):
4     ax[idx].imshow(img.astype(int))
5     ax[idx].title.set_text(batch[1][idx])
```

The below cell transforms the images "X" to all be scaled for optimization in the data pipeline. I typed in .map function using lambda in google and found this [solution](#)

```
1 # scaling data in the pipeline
2 data = data.map(lambda X,y: (X/255, y))

1 # The iterator is now scaled from 0 to 1
2 scaled_iterator = data.as_numpy_iterator()

1 # renaming the variable batch so that it will always be scaled
2 batch = scaled_iterator.next()

1 # confirming the min is 0
2 batch[0].min()

0.0
```



```
1 # confirming max is 1
2 batch[0].max()
    1.0
```

Here we are visualizing another batch of data to make sure everything is working on the scaled data.

```
1 # viewing a batch of data to see what class my frames belong to
2 fig, ax = plt.subplots(ncols=4, figsize=(20,20))
3 for idx, img in enumerate(batch[0][:4]):
4     ax[idx].imshow(img)
5     ax[idx].title.set_text(batch[1][idx])
```

Here I am ready to start my train, test, split. I check the length of the data (in batches) and we see we have 89 batches

```
1 # Checking the length of the data in batches
2 len(data)

89
```

I shuffled the data so that when I eventually use it for modeling it is not biased by learning about the images in any particular order.

```
1 # Shuffle the dataset
2 data_shuffled = data.shuffle(buffer_size=len(data))
```

This splits that data for training, testing, and validation by the number by batches. I designated 80% to training which is 158 batches, 10% testing which is 15 batches, and 10% for validation which is 15 batches.

```
1 # Splitting data into 80% training, 10% testing, 10% validation
2 train_size = int(len(data_shuffled)*.8)
3 val_size = int(len(data_shuffled)*.1)+1
4 test_size = int(len(data_shuffled)*.1)+1
```

Confirming the splits equal to 198 batches.

```
1 # 89 batches
2 train_size+val_size+test_size

89
```

The cell below now sets up the variables to either take or skip certain batches.

```
1 train = data_shuffled.take(train_size)
2 val = data_shuffled.skip(train_size).take(val_size)
3 test = data_shuffled.skip(train_size + val_size).take(test_size)
```

✓ Modeling

Modeling

```
1 model = Sequential()
```

The first layer is a convolutional layer with 16 filters, it is 3 pixels x 3pixels, and a stride of 1 so it goes pixel by pixel. The relu activation function takes into account any non linear patterns. The input shape is the size of our images. The next 2 layers are condensing the rows and width. The last 2 layers are dense fully connected layers and a sigmoid activation function is applied for the output to be between 0 and 1. An output of closer to 1 will be part of the forehand class, and an output of 0 will be closer to the Backhand class.

```
1 model.add(Conv2D(16, (3,3), 1, activation='relu', input_shape=(256,256,3)))
2 model.add(MaxPooling2D())
3
4 model.add(Conv2D(32, (3,3), 1, activation='relu'))
5 model.add(MaxPooling2D())
6
7 model.add(Conv2D(16, (3,3), 1, activation='relu'))
8 model.add(MaxPooling2D())
9
10 model.add(Flatten())
11 model.add(Dense(256, activation='relu'))
12 model.add(Dense(1, activation='sigmoid'))
```

The cell below passes through an optimizer, and BinaryCrossentropy to our loss because we are dealing with Binary Classification, and the metrics we are going to be evaluating on will be accuracy.

```
1 model.compile('adam', loss=tf.losses.BinaryCrossentropy(), metrics=['accuracy'])
```

The cell below shows how the model takes in the data.

```
1 model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 16)	448
max_pooling2d (MaxPooling2D)	(None, 127, 127, 16)	0
conv2d_1 (Conv2D)	(None, 125, 125, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 32)	0
conv2d_2 (Conv2D)	(None, 60, 60, 16)	4624
max_pooling2d_2 (MaxPooling2D)	(None, 30, 30, 16)	0
flatten (Flatten)	(None, 14400)	0
dense (Dense)	(None, 256)	3686656
dense_1 (Dense)	(None, 1)	257
Total params: 3696625 (14.10 MB)		
Trainable params: 3696625 (14.10 MB)		
Non-trainable params: 0 (0.00 Byte)		

```
1 # Creating a log directory that is my logs folder
2 logdir = 'logs'
```

The cell below creates a variable 'tensorboard_callback' which allows you to log the model training while it trains.

```
1 tensorboard_callback = tf.keras.callbacks.TensorBoard(log_dir=logdir)
```

The cell below sets up the training process at 20 epochs for the baseline model. It will also be evaluating on our validation data as well. The below cell is commented out to save time as it takes around 45 minutes to run. The model was quickly overfitting after a few epochs so to save time I adjusted the model to 3 epochs which is currently our best performing model. This also saves the model as an h5 file in the specified folder 'models'

```
1 # This line of code will initiate training, it will take approx 45 minutes
2 #hist = model.fit(train, epochs=20, validation_data=val, callbacks=[tensorboard_callback])
3 # model.save(os.path.join('models', 'forehandbackhandmodel_20.h5'))

1 # This line of code will initiate training, it will take approx 15 minutes
2 hist = model.fit(train, epochs=3, validation_data=val, callbacks=[tensorboard_callback])
3 model.save(os.path.join('models', 'forehandbackhandmodel_3.h5'))

Epoch 1/3
71/71 [=====] - 197s 2s/step - loss: 0.3908 - accuracy: 0.8615 - val_loss: 0.0308
Epoch 2/3
71/71 [=====] - 188s 2s/step - loss: 0.0090 - accuracy: 0.9978 - val_loss: 0.0033
Epoch 3/3
71/71 [=====] - 187s 2s/step - loss: 0.0012 - accuracy: 0.9996 - val_loss: 1.0379
```

Plotting the performance which led to 100% accuracy and a decreasing loss close that dropped close to 0.

```
1 # Visualizing our loss
2
3 fig = plt.figure()
4 plt.plot(hist.history['loss'], color='teal', label='loss')
5 plt.plot(hist.history['val_loss'], color='orange', label='val_loss')
6 fig.suptitle('Loss', fontsize=20)
7 plt.legend(loc="upper left")
8 plt.show()
```

```
1 # Visualizing our accuracy
2
3 fig = plt.figure()
4 plt.plot(hist.history['accuracy'], color='teal', label='accuracy')
5 plt.plot(hist.history['val_accuracy'], color='orange', label='val_accuracy')
6 fig.suptitle('Accuracy', fontsize=20)
7 plt.legend(loc="upper left")
8 plt.show()
```

After we have trained on our model on training and validation data and have visualized the loss and metrics of accuracy we can see how it performs on testing images.

```
1 # loading in the trained model to evaluate on testing images
2 #model = load_model(os.path.join('models', 'forehandbackhandmodel_3.h5'))
```

Evaluating performance on our testing data

```
1 precision = Precision()
2 recall = Recall()
3 accuracy = BinaryAccuracy()

1 for batch in test.as_numpy_iterator():
2     X, y = batch
3     y_hat = model.predict(X)
4     precision.update_state(y, y_hat)
5     recall.update_state(y, y_hat)
6     accuracy.update_state(y, y_hat)
```

```

6     accuracy.update_state(y, y_hat)
1/1 [=====] - 1s 547ms/step
1/1 [=====] - 0s 415ms/step
1/1 [=====] - 0s 459ms/step
1/1 [=====] - 1s 769ms/step
1/1 [=====] - 1s 665ms/step
1/1 [=====] - 1s 614ms/step
1/1 [=====] - 0s 470ms/step
1/1 [=====] - 0s 430ms/step
1/1 [=====] - 0s 432ms/step

```

Our testing data is performing at 100% accuracy.

```

1 print(f'Precision:{precision.result().numpy()}, Recall:{recall.result().numpy()}, Accuracy:{accuracy.result().numpy()}')
    Precision:1.0, Recall:1.0, Accuracy:1.0

```

```

1 # Performing a test on an image not in our batch
2 img = cv2.imread('data/backhands/frames/backhandvids_1004.jpg')
3 plt.imshow(cv2.cvtColor(img, cv2.COLOR_BGR2RGB))
4 plt.show()

```



```
1 resize = tf.image.resize(img, (256,256))  
2 plt.imshow(resize.numpy().astype(int))  
3 plt.show()
```

Predicting on a random image the model has never seen before. This preprocesses an image an image array and feeds it into the NN 'model' and returns a prediction probability score.

```
1 y_hat = model.predict(np.expand_dims(resize/255, 0))  
1/1 [=====] - 0s 127ms/step
```

```
1 # It is predicting what our image is
2 y_hat
   array([[0.00126342]], dtype=float32)
```

The cell below is a conditional statement that prints out what the predicted class is. We are taking the probability score of 'y_hat' which is our target class prediction and setting a threshold of 0.5. If the probability score is greater than 0.5 the predicted class will be a forehand and if it is less than 0.5 the predicted class will be a backhand.

```
1 if y_hat > 0.5:
2     print(f'Predddicted class is Forehand')
3 else:
4     print(f'Predicted class is Backhand')
   Predicted class is Backhand
```

✓ Model Evaluation

My best performing model was the CNN with an 'Adam' optimizer and loss function set to 'Binary Cross-Entropy' trained over 3 epochs. I adjusted the model from 20 epochs to 3 because the model was quickly achieving a 100% accuracy score so this approximately saved about 1 hour of training time. The model was showing 100% accuracy scores on training and testing data. The model is most likely overfitting due to the low variance in my data making it over confident.

✓ Conclusions

The model showed scores of 100% on accuracy, precision, and recall on testing and validation data. The model is performing perfectly most likely due to the low variance in my dataset leading to its overconfidence. To build a more robust model 10x more data with high variance needs to be inputted into the model.

data with high variance needs to be inputted into the model.

✓ Limitations

Limited amount of data was used. The data consisted only of myself from one camera angle in a few different locations with minimal background "noise". Additionally the model only learned about two-handed backhand shots and not one-handed backhand shots.

✓ Recommendations

We need to increase the amount of variance in the dataset. Since the current model is performing so well we can start to add images from the start of swing preparation until the completion of the swing and not just up until the point of contact with the ball.

✓ Next Steps

Obtain more 10x more data with more variance of different people of all ages hitting shots from different camera angles in different backgrounds. The real world is "noisier" so we need to start training our model this way.

