

# Audio Classification of Capuchin Bird Clips

**TASK AT HAND :** *The Challenge is to build a Machine Learning model and code to count the number of Capuchinbird calls within a given clip*



# 1. Loading Packages

```
In [1]: import os
import csv
import tensorflow as tf
import matplotlib.pyplot as plt
import tensorflow_io as tfio
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Input, Conv2D, Dense, Flatten
from tensorflow.keras.callbacks import EarlyStopping, ReduceLROnPlateau
```

```
/opt/conda/lib/python3.10/site-packages/scipy/__init__.py:146: UserWarning: A NumPy version >=1.16.5 and <1.23.0 is required for this version of SciPy (detected version 1.24.3)
warnings.warn(f"A NumPy version >={np_minversion} and <{np_maxversion}")
```

## 2. Method for Loading Data

```
In [2]: CAPUCHIN_FILE = os.path.join('data', '/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/Parsed_Capuchinbird_Clips', 'XC114131-0.wav')
NOT_CAPUCHIN_FILE = os.path.join('data', '/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/Parsed_Not_Capuchinbird_Clips', 'after_0.wav')
```

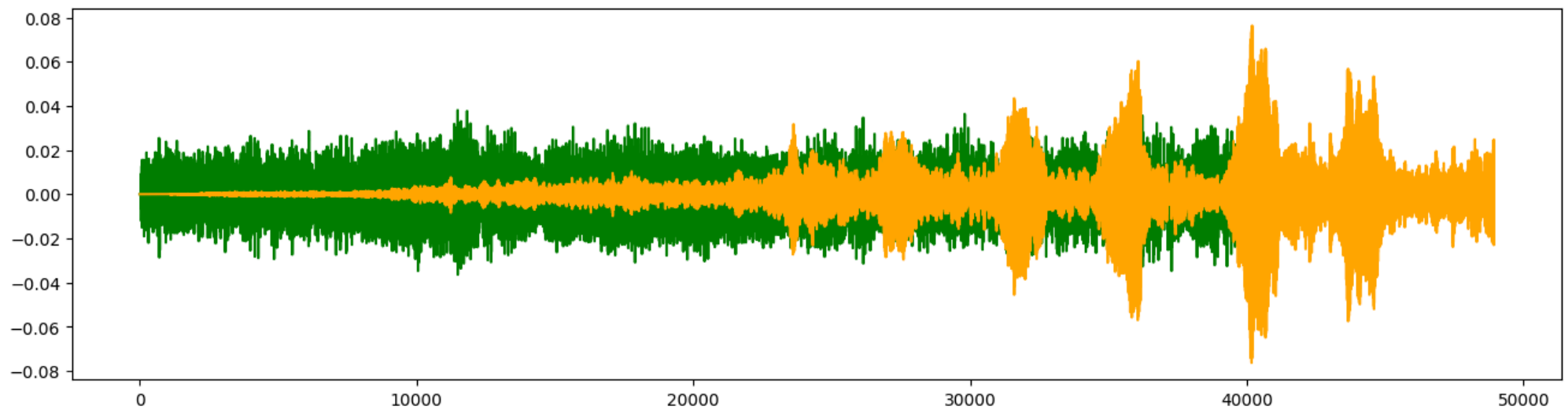
### 2.1 Paths to Data Folders

```
In [3]: def load_wav_16k_mono(filename):

    file_contents = tf.io.read_file(filename)
    wav, sample_rate = tf.audio.decode_wav(file_contents, desired_channels=1)
    wav = tf.squeeze(wav, axis=-1)
    sample_rate = tf.cast(sample_rate, dtype=tf.int64)
    wav = tfio.audio.resample(wav, rate_in=sample_rate, rate_out=16000)
    return wav
```

```
In [4]: capuchin = load_wav_16k_mono(CAPUCHIN_FILE)
not_capuchin = load_wav_16k_mono(NOT_CAPUCHIN_FILE)
```

```
In [5]: plt.figure(figsize=(16,4))
plt.plot(capuchin,color='green')
plt.plot(not_capuchin,color='orange')
plt.show()
```



### 3. Tensorflow Dataset

```
In [6]: Capuchin_folder = os.path.join('data', '/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/Parsed_Capuchinbird_Clips')
not_Capuchin_folder = os.path.join('data', '/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/Parsed_Not_Capuchinbird_Clips')
```

```
In [7]: capuchin = tf.data.Dataset.list_files(Capuchin_folder + '/*-*.wav')
not_capuchin = tf.data.Dataset.list_files(not_Capuchin_folder + '/*-*.wav')
```

#### 3.1 Adding labels to capuchin and not\_capuchin samples

```
In [8]: capuchins = tf.data.Dataset.zip((capuchin, tf.data.Dataset.from_tensor_slices(tf.ones(len(capuchin)))))
not_capuchins = tf.data.Dataset.zip((not_capuchin, tf.data.Dataset.from_tensor_slices(tf.zeros(len(not_capuchin)))))
```

```
In [9]: data = capuchins.concatenate(not_capuchins)
```

### 4. Calculating the Avg Length of a Clip

#### 4.1 calculate wave cycle length

```
In [10]: lengths = []
for file in os.listdir(os.path.join('data', '/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/Parsed_Capuchinbird_Clips')):
    file_path = os.path.join('data', '/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/Parsed_Capuchinbird_Clips', file)
```

```
tensor_wave = load_wav_16k_mono(file_path)
lengths.append(len(tensor_wave))
```

```
In [11]: max_length = tf.math.reduce_max(lengths)
mean_length = tf.math.reduce_mean(lengths)
min_length = tf.math.reduce_min(lengths)

print(f"Minimum Length : {min_length}")
print(f"Average Length : {mean_length}")
print(f"Maximum Length : {max_length}")
```

```
Minimum Length : 32000
Average Length : 54156
Maximum Length : 80000
```

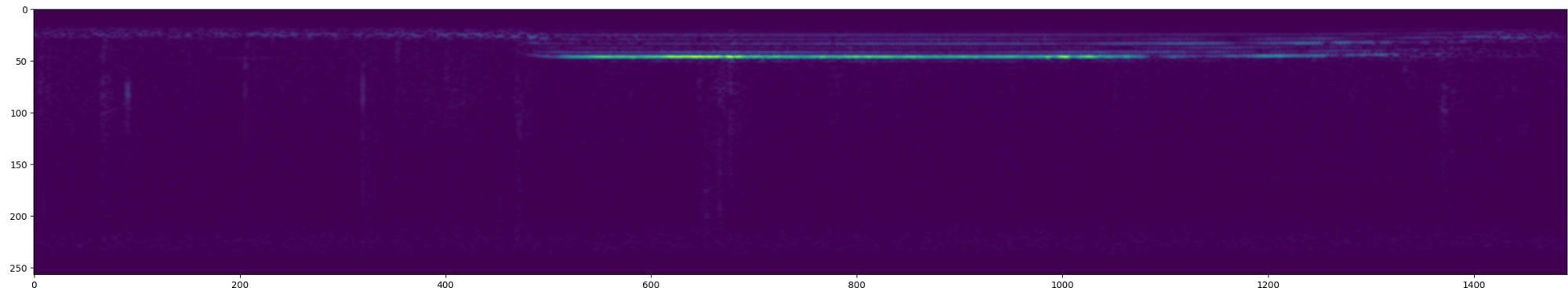
## 5. Building Function to Convert wave to a Spectrogram

```
In [12]: 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
```

### 5.1 Plot the Spectrogram

```
In [13]: filepath, label = capuchins.shuffle(buffer_size=10000).as_numpy_iterator().next()
spectrogram, label = preprocess(filepath, label)

plt.figure(figsize=(30,20))
plt.imshow(tf.transpose(spectrogram)[0])
plt.show()
```



## 6. Training and Testing Splits

```
In [14]: data = data.map(preprocess)
data = data.cache()
data = data.shuffle(buffer_size = 10000)
data = data.batch(16)
data = data.prefetch(8)
```

### 6.1. Split into training and testing data (70-30)

```
In [15]: train = data.take(36)
test = data.skip(36).take(15)
```

```
In [16]: samples, labels = train.as_numpy_iterator().next()
samples.shape
```

```
Out[16]: (16, 1491, 257, 1)
```

## 7. Building Deep Learning Model

### 7.1. Convolutional Neural Network

```
In [17]: model = Sequential()
model.add(Input(shape=(1491, 257, 1)))
model.add(Conv2D(16, (3,3), activation='relu'))
model.add(Conv2D(16, (3,3), activation='relu'))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

In [18]: model.compile('Adam', loss='BinaryCrossentropy', metrics=[tf.keras.metrics.Recall(),tf.keras.metrics.Precision()])
model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 1489, 255, 16)	160
conv2d_1 (Conv2D)	(None, 1487, 253, 16)	2320
flatten (Flatten)	(None, 6019376)	0
dense (Dense)	(None, 128)	770480256
dense_1 (Dense)	(None, 1)	129

=====  
Total params: 770482865 (2.87 GB)  
Trainable params: 770482865 (2.87 GB)  
Non-trainable params: 0 (0.00 Byte)

```
In [19]: earlystop = EarlyStopping(monitor='val_loss',
                                min_delta=0,
                                patience=5,
                                verbose=1,
                                restore_best_weights=True)

# Callback to reduce learning rate
reduce_lr = ReduceLRonPlateau(monitor='val_loss',
                              factor=0.2,
                              patience=6,
                              verbose=1,
                              min_delta=0.0001)

callbacks = [earlystop, reduce_lr]
```

## 7.2. Model Fit

```
In [20]: hist = model.fit(  
    train,  
    epochs=15,  
    validation_data=test,  
    callbacks=callbacks)
```

```

Epoch 1/15
36/36 [=====] - 21s 424ms/step - loss: 2.0907 - recall: 0.8750 - precision: 0.7824 - val_loss: 0.0312 - val_recall: 0.9692 - val_precision: 1.0000 - lr: 0.0010
Epoch 2/15
36/36 [=====] - 10s 292ms/step - loss: 0.1695 - recall: 0.9675 - precision: 0.9675 - val_loss: 0.0028 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 3/15
36/36 [=====] - 6s 179ms/step - loss: 0.0281 - recall: 0.9801 - precision: 0.9933 - val_loss: 0.1185 - val_recall: 1.0000 - val_precision: 0.9831 - lr: 0.0010
Epoch 4/15
36/36 [=====] - 10s 292ms/step - loss: 0.1161 - recall: 0.9809 - precision: 0.9872 - val_loss: 0.0014 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 5/15
36/36 [=====] - 10s 290ms/step - loss: 0.0024 - recall: 1.0000 - precision: 0.9937 - val_loss: 5.6674e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 6/15
36/36 [=====] - 10s 294ms/step - loss: 3.6977e-04 - recall: 1.0000 - precision: 1.0000 - val_loss: 1.9160e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 7/15
36/36 [=====] - 6s 180ms/step - loss: 1.0143e-04 - recall: 1.0000 - precision: 1.0000 - val_loss: 2.6583e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 8/15
36/36 [=====] - 10s 290ms/step - loss: 1.3442e-04 - recall: 1.0000 - precision: 1.0000 - val_loss: 1.1126e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 9/15
36/36 [=====] - 6s 179ms/step - loss: 5.1385e-05 - recall: 1.0000 - precision: 1.0000 - val_loss: 1.6098e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 10/15
36/36 [=====] - 10s 295ms/step - loss: 3.0267e-05 - recall: 1.0000 - precision: 1.0000 - val_loss: 3.2333e-05 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 11/15
36/36 [=====] - 10s 292ms/step - loss: 2.2182e-05 - recall: 1.0000 - precision: 1.0000 - val_loss: 2.7736e-05 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 12/15
36/36 [=====] - 7s 181ms/step - loss: 1.8129e-05 - recall: 1.0000 - precision: 1.0000 - val_loss: 2.7587e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 13/15
36/36 [=====] - 10s 292ms/step - loss: 1.2589e-05 - recall: 1.0000 - precision: 1.0000 - val_loss: 9.5401e-06 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 14/15
36/36 [=====] - 6s 179ms/step - loss: 1.5323e-04 - recall: 1.0000 - precision: 1.0000 - val_loss: 2.2469e-04 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010
Epoch 15/15
36/36 [=====] - 6s 180ms/step - loss: 1.8962e-04 - recall: 1.0000 - precision: 1.0000 - val_loss: 9.9462e-05 - val_recall: 1.0000 - val_precision: 1.0000 - lr: 0.0010

```

```
In [22]: def plot_training_history(history):
```

```
        """
```



Plots the training and validation accuracy and loss.

"""

```
prec = history.history['precision']
val_prec = history.history['val_precision']
recall = history.history['recall']
val_recall = history.history['val_recall']
loss = history.history['loss']
val_loss = history.history['val_loss']
```

```
epochs_range = range(len(prec))
```

```
plt.figure(figsize=(20, 5))
```

*# Plot training and validation Precision*

```
plt.subplot(1, 3, 1)
plt.plot(epochs_range, prec, label='Training Precision')
plt.plot(epochs_range, val_prec, label='Validation Precision')
plt.legend(loc='lower right')
plt.title('Training and Validation Precision')
```

*# Plot training and validation Loss*

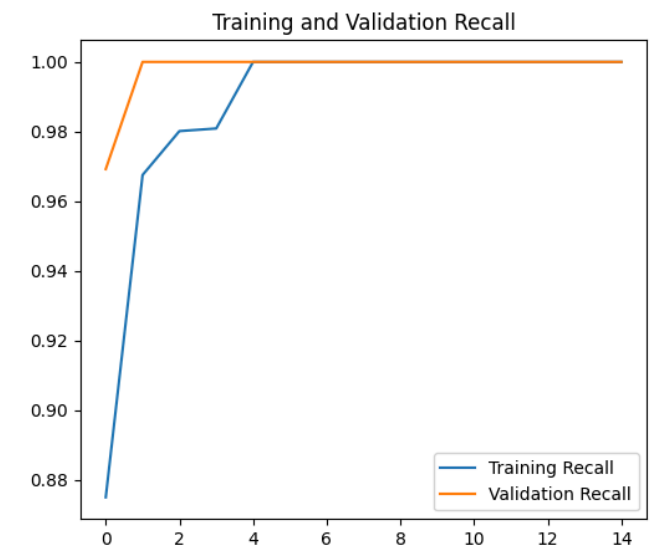
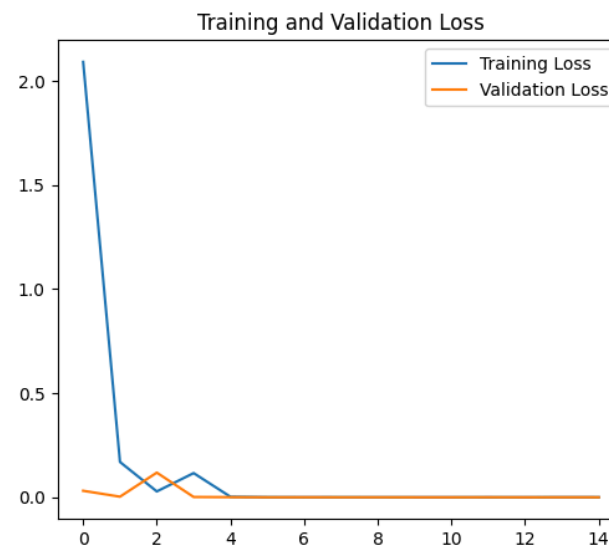
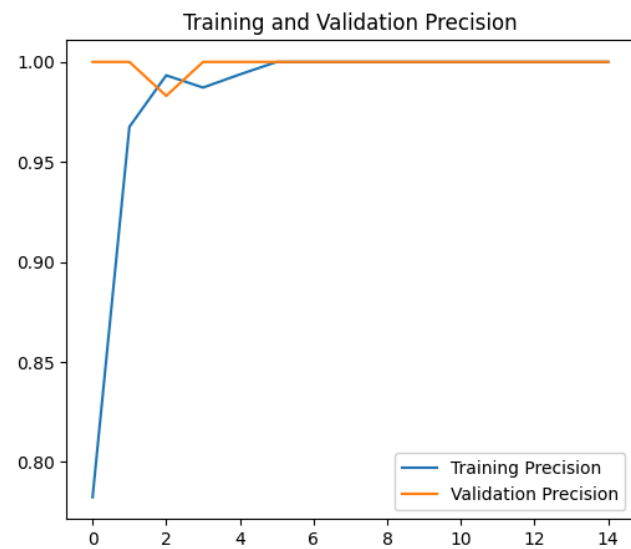
```
plt.subplot(1, 3, 2)
plt.plot(epochs_range, loss, label='Training Loss')
plt.plot(epochs_range, val_loss, label='Validation Loss')
plt.legend(loc='upper right')
plt.title('Training and Validation Loss')
```

*# Plot training and validation Recall*

```
plt.subplot(1, 3, 3)
plt.plot(epochs_range, recall, label='Training Recall')
plt.plot(epochs_range, val_recall, label='Validation Recall')
plt.legend(loc='lower right')
plt.title('Training and Validation Recall')
```

```
plt.show()
```

```
plot_training_history(hist)
```



## 8. Predictions Time

```
In [23]: X_test, y_test = test.as_numpy_iterator().next()
         yhat = model.predict(X_test)

1/1 [=====] - 0s 140ms/step
```

```
In [24]: # Converting Probabilities to Classes
         yhat = [1 if prediction > 0.5 else 0 for prediction in yhat]
```

```
In [30]: y_test
```

```
Out[30]: array([1., 0., 0., 1., 1., 0., 0., 0., 0., 1., 0., 1., 0., 1., 0., 0.],
              dtype=float32)
```

```
In [25]: yhat
```

```
Out[25]: [1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0]
```

## 9. Forest Clips Parsing

```
In [31]: def load_mp3_16k_mono(filename):
         """ Load a WAV file, convert it to a float tensor, resample to 16 kHz single-channel audio. """
         res = tfio.audio.AudioIOTensor(filename)
         # Convert to tensor and combine channels
         tensor = res.to_tensor()
```

```

tensor = tf.math.reduce_sum(tensor, axis=1) / 2
# Extract sample rate and cast
sample_rate = res.rate
sample_rate = tf.cast(sample_rate, dtype=tf.int64)
# Resample to 16 kHz
wav = tfio.audio.resample(tensor, rate_in=sample_rate, rate_out=16000)
return wav

```

## 9.2 Build Function to convert clips into windowed spectrogram

```

In [37]: def preprocess_mp3(sample, index):
        sample = sample[0]
        zero_padding = tf.zeros([48000] - tf.shape(sample), dtype=tf.float32)
        wav = tf.concat([zero_padding, sample], 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

```

## 10. Forest Clips Predictions

```

In [38]: results = {}
        for file in os.listdir(os.path.join('/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/', 'Forest Recordings')):
            FILEPATH = os.path.join('/kaggle/input/z-by-hp-unlocked-challenge-3-signal-processing/', 'Forest Recordings', file)

            wav = load_mp3_16k_mono(FILEPATH)
            audio_slices = tf.keras.utils.timeseries_dataset_from_array(wav, wav, sequence_length=48000, sequence_stride=48000, batch_size=1)
            audio_slices = audio_slices.map(preprocess_mp3)
            audio_slices = audio_slices.batch(64)

            yhat = model.predict(audio_slices)

            results[file] = yhat

```

1/1 [=====] - 3s 3s/step  
1/1 [=====] - 1s 706ms/step  
1/1 [=====] - 1s 716ms/step  
1/1 [=====] - 1s 702ms/step  
1/1 [=====] - 1s 708ms/step  
1/1 [=====] - 1s 688ms/step  
1/1 [=====] - 1s 910ms/step  
1/1 [=====] - 1s 701ms/step  
1/1 [=====] - 1s 706ms/step  
1/1 [=====] - 1s 708ms/step  
1/1 [=====] - 1s 686ms/step  
1/1 [=====] - 1s 704ms/step  
1/1 [=====] - 1s 702ms/step  
1/1 [=====] - 1s 707ms/step  
1/1 [=====] - 1s 710ms/step  
1/1 [=====] - 1s 712ms/step  
1/1 [=====] - 1s 705ms/step  
1/1 [=====] - 1s 722ms/step  
1/1 [=====] - 1s 691ms/step  
1/1 [=====] - 1s 728ms/step  
1/1 [=====] - 1s 705ms/step  
1/1 [=====] - 1s 717ms/step  
1/1 [=====] - 1s 700ms/step  
1/1 [=====] - 1s 724ms/step  
1/1 [=====] - 1s 687ms/step  
1/1 [=====] - 1s 696ms/step  
1/1 [=====] - 1s 708ms/step  
1/1 [=====] - 1s 816ms/step  
1/1 [=====] - 1s 706ms/step  
1/1 [=====] - 1s 726ms/step  
1/1 [=====] - 1s 699ms/step  
1/1 [=====] - 1s 702ms/step  
1/1 [=====] - 1s 722ms/step  
1/1 [=====] - 1s 728ms/step  
1/1 [=====] - 1s 717ms/step  
1/1 [=====] - 1s 711ms/step  
1/1 [=====] - 1s 704ms/step  
1/1 [=====] - 1s 709ms/step  
1/1 [=====] - 1s 702ms/step  
1/1 [=====] - 1s 695ms/step  
1/1 [=====] - 1s 764ms/step  
1/1 [=====] - 1s 709ms/step  
1/1 [=====] - 1s 683ms/step  
1/1 [=====] - 1s 706ms/step  
1/1 [=====] - 1s 700ms/step  
1/1 [=====] - 1s 702ms/step  
1/1 [=====] - 1s 703ms/step  
1/1 [=====] - 1s 802ms/step

1/1 [=====] - 1s 732ms/step  
1/1 [=====] - 1s 702ms/step  
1/1 [=====] - 1s 704ms/step  
1/1 [=====] - 1s 675ms/step  
1/1 [=====] - 1s 688ms/step  
1/1 [=====] - 1s 682ms/step  
1/1 [=====] - 1s 692ms/step  
1/1 [=====] - 1s 708ms/step  
1/1 [=====] - 1s 695ms/step  
1/1 [=====] - 1s 676ms/step  
1/1 [=====] - 1s 672ms/step  
1/1 [=====] - 1s 674ms/step  
1/1 [=====] - 1s 721ms/step  
1/1 [=====] - 1s 704ms/step  
1/1 [=====] - 1s 684ms/step  
1/1 [=====] - 1s 686ms/step  
1/1 [=====] - 1s 700ms/step  
1/1 [=====] - 1s 688ms/step  
1/1 [=====] - 1s 697ms/step  
1/1 [=====] - 1s 738ms/step  
1/1 [=====] - 1s 683ms/step  
1/1 [=====] - 1s 815ms/step  
1/1 [=====] - 1s 695ms/step  
1/1 [=====] - 1s 686ms/step  
1/1 [=====] - 1s 699ms/step  
1/1 [=====] - 1s 679ms/step  
1/1 [=====] - 1s 683ms/step  
1/1 [=====] - 1s 671ms/step  
1/1 [=====] - 1s 679ms/step  
1/1 [=====] - 1s 687ms/step  
1/1 [=====] - 1s 670ms/step  
1/1 [=====] - 1s 672ms/step  
1/1 [=====] - 1s 673ms/step  
1/1 [=====] - 1s 695ms/step  
1/1 [=====] - 1s 706ms/step  
1/1 [=====] - 1s 680ms/step  
1/1 [=====] - 1s 695ms/step  
1/1 [=====] - 1s 698ms/step  
1/1 [=====] - 1s 693ms/step  
1/1 [=====] - 1s 695ms/step  
1/1 [=====] - 1s 683ms/step  
1/1 [=====] - 1s 680ms/step  
1/1 [=====] - 1s 847ms/step  
1/1 [=====] - 1s 680ms/step  
1/1 [=====] - 1s 664ms/step  
1/1 [=====] - 1s 669ms/step  
1/1 [=====] - 1s 752ms/step  
1/1 [=====] - 1s 683ms/step

```
1/1 [=====] - 1s 688ms/step
1/1 [=====] - 1s 705ms/step
1/1 [=====] - 1s 688ms/step
1/1 [=====] - 1s 684ms/step
```

In [43]: *# Convert Predictions into Classes*

```
class_preds = {}
for file, logits in results.items():
    class_preds[file] = [1 if prediction > 0.99 else 0 for prediction in logits]
class_preds
```

```
Out[43]: {'recording_76.mp3': [0,
```

[illegible]





[illegible]

[illegible]

[illegible]

[illegible]

```
0,
0,
1,
0,
0,
0,
1,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_43.mp3': [0,
0,
0,
0,
0,
0,
1,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
1,
0,
0,
0,
0,
```

```
0,
0,
0,
0,
1,
1,
0,
0,
0,
0,
0,
1,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
1,
1,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_68.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```

[illegible]



```
'recording_95.mp3': [0,
```

 $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$

[illegible]

[illegible]

```
0,
0,
0,
0,
0,
0,
0,
0,
1,
0,
0,
1,
0,
1,
1,
0,
1,
0,
0,
0,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_74.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```

[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

```
'recording_64.mp3': [0,
```

 $\theta,$  $\theta,$  $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$  $\theta,$ 0,  
0.0,  
0.0,  
0. $\theta,$   
 $\theta$ 0,  
00,  
00,  
0

0, 0

0,  
0.

0.  
0.0.  
0.

0.

0.

 $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$  $\theta,$ 

1,

 $\theta,$  $\theta,$  $\theta,$

[illegible]

[illegible]

```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
1,
0,
0,
0,
1,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
1,
1,
0,
0,
0,
0,
0],
'recording_28.mp3': [1,
1,
0,
1,
1,
1,
0,
1,
1,
0,
1,
1,
0,
1,
1,
0,
1,
```

[illegible]

'recording\_05.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,



[illegible]

[illegible]

```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
1,
1,
0,
0,
0,
0,
1,
0,
0,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_12.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```

[illegible]

'recording\_59.mp3': [0,

1,

0,

1,

1,

0,

1,

1,

1,

1,

1,

1,

0,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

0,

1,

1,

1,

1,

1,

0,

1,

1,

1,

1,

1,

0,

1,

0,

1,

1,

1,

1,

1,

1,

0,

1,

1,

1,

[illegible]

[illegible]

```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_79.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```



[illegible]

'recording\_14.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

[illegible]

[illegible]

[illegible]

[illegible]

'recording\_03.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

[illegible]







[illegible]

'recording\_35.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,



[illegible]



[illegible]



'recording\_75.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

1,

1,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

[illegible]

[illegible]



[illegible]

'recording\_69.mp3': [0,

0,

0,

0,

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,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

[illegible]

[illegible]



```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_31.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```

[illegible]

'recording\_73.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

[illegible]





[illegible]

'recording\_66.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,









[illegible]

'recording\_25.mp3': [1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

0,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,





[illegible]



[illegible]

'recording\_87.mp3': [1,

1,

0,

1,

0,

1,

0,

1,

1,

0,

1,

0,

1,

0,

1,

1,

0,

1,

0,

1,

1,

1,

1,

1,

1,

0,

1,

1,

0,

1,

0,

1,

1,

1,

1,

1,

0,

1,

0,

1,

1,

0,

1,

0,

1,

1,

0,

[illegible]

[illegible]



[illegible]

[illegible]







```
0,
0,
0,
1,
1,
0,
0,
0,
0,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_91.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```

[illegible]

'recording\_13.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,







[illegible]



'recording\_55.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,





[illegible]

[illegible]

'recording\_72.mp3': [0,

1,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

1,

0,

0,

0,

1,

1,

0,

[illegible]







[illegible]

1,  
1,  
1,  
1,  
1,  
0,  
1,  
0,  
0,  
1,  
0,  
1,  
0,  
1,  
1,  
0,  
1,  
0,  
1,  
1,  
0,  
1,  
0,  
0,  
0,  
1,  
0,  
1,  
0,  
1,  
0,  
0,  
1,  
0,  
0,  
0,  
1,  
0,  
0,  
1,  
0,  
1,  
1,





```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_89.mp3': [0,
0,
0,
0,
1,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```

[illegible]

'recording\_01.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,



[illegible]

[illegible]



[illegible]

'recording\_77.mp3': [1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

1,

0,

1,

0,

1,

1,

1,

1,

1,

1,

[illegible]



```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_82.mp3': [0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
```



[illegible]

[illegible]

[illegible]

```
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0,
0],
'recording_99.mp3': [0,
0,
0,
0,
0,
0,
1,
1,
0,
0,
0,
1,
0,
0,
1,
1,
0,
0,
0,
0,
0,
1,
0,
0,
```



[illegible]

'recording\_50.mp3': [0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

0,

```
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0],  
'recording_71.mp3': [0,  
0,  
0,  
1,  
1,  
0,  
0,  
0,  
0,  
0,  
0,  
1,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
0,  
1,  
0,  
0,  
1,  
0,  
0,  
0,  
0,  
0,  
0,  
1,  
1,  
0,  
0,  
0,  
0,  
0,
```







[illegible]

In [48]: *# Grouping consecutive Detections*

```
from itertools import groupby

postprocessed = {}
for file, scores in class_preds.items():
    postprocessed[file] = tf.math.reduce_sum([key for key, group in groupby(scores)]).numpy()

sorted_postprocessed = dict(sorted(postprocessed.items(), key=lambda item: item[1], reverse=True))

top_5 = dict(list(sorted_postprocessed.items())[:5])

print("Top 5 Recordings with most Capuchin Bird Sounds are : ",top_5)
```

Top 5 Recordings with most Capuchin Bird Sounds are : {'recording\_08.mp3': 24, 'recording\_98.mp3': 21, 'recording\_87.mp3': 20, 'recording\_28.mp3': 14, 'recording\_59.mp3': 9}