

```
In [11]: import librosa
import librosa.display
import numpy as np
import scipy.signal
import matplotlib.pyplot as plt
%matplotlib inline
from IPython.display import Audio, IFrame, display
import glob
from scipy.ndimage import zoom
zoom_factors = (128, 128)

def resize_spectrogram(spectrogram, desired_shape):
    factors = (desired_shape[0] // spectrogram.shape[0], desired_shape[1] // spectrogram.shape[1])
    return zoom(spectrogram, zoom_factors)

birds=[
    labels=[]
    sr = 16000
    file_path=glob.glob('./birdSounds/**')
    #creating spectrograms of bird sounds and append them to bird list and their labels to label list.
    number=0
    for file in file_path:
        g, s = librosa.load(file, mono=True, sr=sr, offset=0)
        g = librosa.feature.melspectrogram(y=s, sr=sr, n_mels=128)
        g_db = librosa.power_to_db(g, ref=np.max)
        g_db = resize_spectrogram(g_db, desired_shape)
        birds.append(g_db)
        if 'Bent-Beak-Riffraff' in file:
            labels.append('Bent Beak Riffraff')
        elif 'Blue-collared-Zipper' in file:
            labels.append('Blue collared Zipper')
        elif 'Bombadil' in file:
            labels.append('Bombadil')
        elif 'Broad-winged-Jojo' in file:
            labels.append('Broad winged Jojo')
        elif 'Canadian-Cootamum' in file:
            labels.append('Canadian Cootamum')
        elif 'Carries-Champagne-Pipit' in file:
            labels.append('Carries Champagne Pipit')
        elif 'Darkwing-Sparrow' in file:
            labels.append('Darkwing Sparrow')
        elif 'Eastern-Corn-Skeet' in file:
            labels.append('Eastern Corn Skeet')
        elif 'Green-Tipped-Scarlet-Pipit' in file:
            labels.append('Green Tipped Scarlet Pipit')
        elif 'Lesser-Birchbeere' in file:
            labels.append('Lesser Birchbeere')
        elif 'Orange-Pine-Flower' in file:
            labels.append('Orange Pine Flower')
        elif 'Ordinary-Snape' in file:
            labels.append('Ordinary Snape')
        elif 'Pinkfinch' in file:
            labels.append('Pinkfinch')
        elif 'Purple-Tooing-Tout' in file:
            labels.append('Purple Tooing Tout')
        elif 'Qax' in file:
            labels.append('Qax')
        elif 'Queenscoat' in file:
            labels.append('Queenscoat')
        elif 'Rose-Crested-Blue-Pipit' in file:
            labels.append('Rose Crested Blue Pipit')
        elif 'Scrawny-Jay' in file:
            labels.append('Scrawny Jay')
        elif 'Vermillion-Trillian' in file:
            labels.append('Vermillion Trillian')
        else:
            labels.append('Unknown')
    number+=1

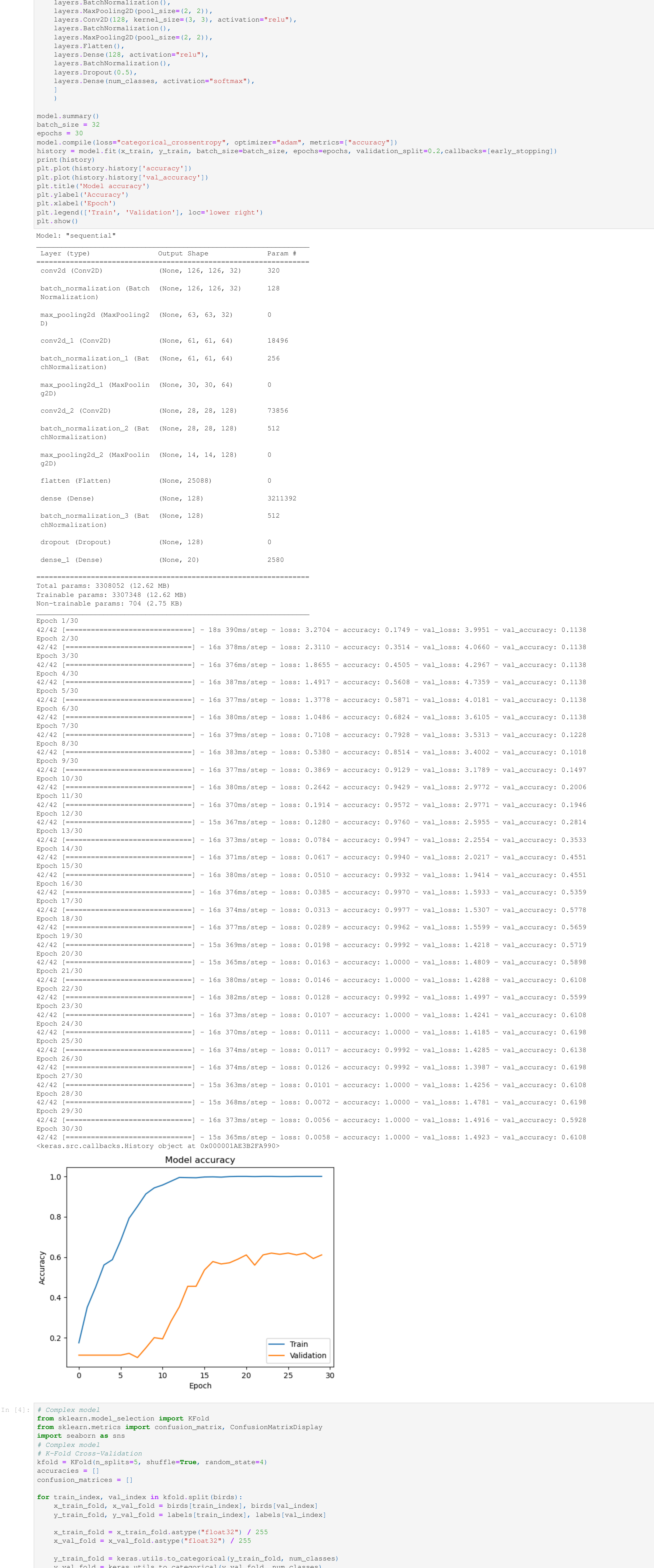
In [12]: from tensorflow.keras.callbacks import EarlyStopping
from tensorflow import keras
from tensorflow.keras import layers
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
early_stopping = EarlyStopping(monitor='val_loss', patience=5, restore_best_weights=True)
#Converting lists to Numpy array
birdemp.array(birds)
labels=np.array(labels)
birds=birds[... , np.newaxis]
birds = np.squeeze(birds)

num_classes = 20
input_shape = (128, 128, 1)
# Simple model
labelencoder = LabelEncoder()
labels = labelencoder.fit_transform(labels)
x_train, x_test, y_train, y_test = train_test_split(birds, labels, test_size=0.2, random_state=4)
x_train = x_train.astype("float32") / 255
x_test = x_test.astype("float32") / 255

y_train = keras.utils.to_categorical(y_train, num_classes)

model = keras.Sequential([
    layers.Input(shape=input_shape),
    layers.Conv2D(32, kernel_size=(3, 3), activation="relu"),
    layers.BatchNormalization(),
    layers.MaxPooling2D(pool_size=(2, 2)),
    layers.Conv2D(64, kernel_size=(3, 3), activation="relu"),
    layers.BatchNormalization(),
    layers.MaxPooling2D(pool_size=(2, 2)),
    layers.Conv2D(128, kernel_size=(3, 3), activation="relu"),
    layers.BatchNormalization(),
    layers.MaxPooling2D(pool_size=(2, 2)),
    layers.Flatten(),
    layers.Dense(128, activation="relu"),
    layers.BatchNormalization(),
    layers.Dropout(0.5),
    layers.Dense(num_classes, activation="softmax"),
])

model.summary()
batch_size = 32
epochs = 30
model.compile(loss="categorical_crossentropy", optimizer="adam", metrics=["accuracy"])
history = model.fit(x_train, y_train, batch_size=batch_size, epochs=epochs, validation_split=0.2, callbacks=[early_stopping])
print(history)
plt.plot(history.history['accuracy'])
plt.title('Model accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Train', 'Validation'], loc='lower right')
plt.show()
```



```
In [41]: # Complex Model
from sklearn.model_selection import KFold
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay
import seaborn as sns
# Complex model
# K-Fold Cross-Validation
x_train, x_test, y_train, y_test = train_test_split(birds, labels, test_size=0.2, random_state=4)
confusion_matrices = []

for train_index, val_index in KFold.split(birds):
    x_train_fold, x_val_fold = birds[train_index], birds[val_index]
    y_train_fold, y_val_fold = labels[train_index], labels[val_index]

    x_train_fold = x_train_fold.astype("float32") / 255
    x_val_fold = x_val_fold.astype("float32") / 255

    y_train_fold = keras.utils.to_categorical(y_train_fold, num_classes)
    y_val_fold = keras.utils.to_categorical(y_val_fold, num_classes)

    model = keras.Sequential([
        layers.Input(shape=input_shape),
        layers.Conv2D(32, kernel_size=(3, 3), activation="relu"),
        layers.BatchNormalization(),
        layers.MaxPooling2D(pool_size=(2, 2)),
        layers.Conv2D(64, kernel_size=(3, 3), activation="relu"),
        layers.BatchNormalization(),
        layers.MaxPooling2D(pool_size=(2, 2)),
        layers.Conv2D(128, kernel_size=(3, 3), activation="relu"),
        layers.BatchNormalization(),
        layers.MaxPooling2D(pool_size=(2, 2)),
        layers.Flatten(),
        layers.Dense(128, activation="relu"),
        layers.BatchNormalization(),
        layers.Dropout(0.5),
        layers.Dense(num_classes, activation="softmax"),
    ])

    model.compile(loss="categorical_crossentropy", optimizer="adam", metrics=["accuracy"])

    history = model.fit(x_train_fold, y_train_fold, batch_size=32, epochs=30, validation_data=(x_val_fold, y_val_fold), callbacks=[early_stopping])

    scores = model.evaluate(x_val_fold, y_val_fold, verbose=0)
    accuracies.append(scores[1])

    # Predict the labels for the validation set
    y_pred_fold = model.predict(x_val_fold)
    y_pred_fold = np.argmax(y_pred_fold, axis=1)
    y_val_fold = np.argmax(y_val_fold, axis=1)

    # Compute the confusion matrix
    cm = confusion_matrix(y_val_fold, y_pred_fold)
    confusion_matrices.append(cm)
    print(f"Fold accuracy: {scores[1]}")

print(f"Mean cross-validated accuracy: {np.mean(accuracies)}")

# Plot the training history of the last fold
plt.plot(history.history['accuracy'])
plt.plot(history.history['val_accuracy'])
plt.title('Model accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Train', 'Validation'], loc='lower right')
plt.show()

# Ensure labelencoder.classes_ contains the correct bird names
bird_names = [
    'Bent Beak Riffraff', 'Blue collared Zipper', 'Bombadil', 'Broad winged Jojo',
    'Canadian Cootamum', 'Carries Champagne Pipit', 'Darkwing Sparrow', 'Eastern Corn Skeet',
    'Green Tipped Scarlet Pipit', 'Lesser Birchbeere', 'Orange Pine Flower', 'Ordinary Snape',
    'Pinkfinch', 'Purple Tooing Tout', 'Qax', 'Queenscoat', 'Rose Crested Blue Pipit',
    'Scrawny Jay', 'Vermillion Trillian', 'Unknown'
]

# Assign bird names to labelencoder.classes_
labelencoder.classes_ = np.array(bird_names)

# Plot the confusion matrix of the last fold using seaborn
sns.heatmap(confusion_matrices[-1], annot=True, fmt='d', cmap='Blues', xticklabels=labelencoder.classes_, yticklabels=labelencoder.classes_)
plt.title('Confusion Matrix')
plt.xlabel('Predicted Label')
plt.ylabel('True Label')
plt.show()

Epoch 1/30
53/53 [=====] - 21s 385ms/step - loss: 3.1343 - accuracy: 0.2011 - val_loss: 3.8668 - val_accuracy: 0.1415
Epoch 2/30
53/53 [=====] - 19s 368ms/step - loss: 2.2408 - accuracy: 0.3565 - val_loss: 4.2851 - val_accuracy: 0.1415
Epoch 3/30
53/53 [=====] - 20s 374ms/step - loss: 1.7415 - accuracy: 0.4940 - val_loss: 4.1240 - val_accuracy: 0.1415
Epoch 4/30
53/53 [=====] - 20s 373ms/step - loss: 1.4310 - accuracy: 0.5900 - val_loss: 3.7771 - val_accuracy: 0.1439
Epoch 5/30
53/53 [=====] - 21s 390ms/step - loss: 1.4873 - accuracy: 0.5498 - val_loss: 3.2694 - val_accuracy: 0.1775
Epoch 6/30
53/53 [=====] - 20s 387ms/step - loss: 1.1894 - accuracy: 0.6303 - val_loss: 3.3488 - val_accuracy: 0.1439
Epoch 7/30
53/53 [=====] - 20s 382ms/step - loss: 0.9433 - accuracy: 0.7179 - val_loss: 3.1965 - val_accuracy: 0.1799
Epoch 8/30
53/53 [=====] - 20s 383ms/step - loss: 0.6622 - accuracy: 0.8145 - val_loss: 3.0112 - val_accuracy: 0.1871
Epoch 9/30
53/53 [=====] - 20s 383ms/step - loss: 0.5910 - accuracy: 0.8379 - val_loss: 2.6114 - val_accuracy: 0.2326
Epoch 10/30
53/53 [=====] - 20s 384ms/step - loss: 0.4215 - accuracy: 0.8830 - val_loss: 2.2815 - val_accuracy: 0.2950
Epoch 11/30
53/53 [=====] - 20s 374ms/step - loss: 0.3008 - accuracy: 0.9352 - val_loss: 2.0379 - val_accuracy: 0.3765
Epoch 12/30
53/53 [=====] - 20s 374ms/step - loss: 0.3449 - accuracy: 0.9082 - val_loss: 1.5125 - val_accuracy: 0.5360
Epoch 13/30
53/53 [=====] - 20s 373ms/step - loss: 0.2270 - accuracy: 0.9490 - val_loss: 1.4253 - val_accuracy: 0.5683
Epoch 14/30
53/53 [=====] - 22s 423ms/step - loss: 0.2311 - accuracy: 0.9474 - val_loss: 1.1941 - val_accuracy: 0.5088
Epoch 15/30
53/53 [=====] - 25s 470ms/step - loss: 0.1448 - accuracy: 0.9724 - val_loss: 1.5116 - val_accuracy: 0.4384
Epoch 16/30
53/53 [=====] - 21s 403ms/step - loss: 0.1072 - accuracy: 0.9802 - val_loss: 1.6813 - val_accuracy: 0.5276
Epoch 17/30
53/53 [=====] - 23s 434ms/step - loss: 0.1276 - accuracy: 0.9784 - val_loss: 1.3221 - val_accuracy: 0.5971
Epoch 18/30
53/53 [=====] - 23s 429ms/step - loss: 0.0952 - accuracy: 0.9838 - val_loss: 1.3323 - val_accuracy: 0.6115
Epoch 19/30
53/53 [=====] - 21s 403ms/step - loss: 0.0631 - accuracy: 0.9898 - val_loss: 1.2876 - val_accuracy: 0.6259
Epoch 20/30
53/53 [=====] - 20s 385ms/step - loss: 0.0422 - accuracy: 0.9964 - val_loss: 1.2738 - val_accuracy: 0.6403
Epoch 21/30
53/53 [=====] - 22s 408ms/step - loss: 0.0330 - accuracy: 0.9982 - val_loss: 1.4268 - val_accuracy: 0.6259
Epoch 22/30
53/53 [=====] - 21s 399ms/step - loss: 0.0308 - accuracy: 0.9982 - val_loss: 1.4492 - val_accuracy: 0.6307
Epoch 23/30
53/53 [=====] - 21s 391ms/step - loss: 0.0432 - accuracy: 0.9940 - val_loss: 1.5293 - val_accuracy: 0.5947
Epoch 24/30
53/53 [=====] - 22s 408ms/step - loss: 0.3757 - accuracy: 0.8908 - val_loss: 1.4266 - val_accuracy: 0.2782
Epoch 25/30
53/53 [=====] - 20s 384ms/step - loss: 0.2151 - accuracy: 0.9418 - val_loss: 1.7390 - val_accuracy: 0.4724
Epoch 26/30
14/14 [=====] - 1s 66ms/step
Fold accuracy: 0.6402877569198608
Epoch 1/30
53/53 [=====] - 22s 403ms/step - loss: 3.0469 - accuracy: 0.2209 - val_loss: 4.8622 - val_accuracy: 0.0911
Epoch 2/30
53/53 [=====] - 22s 408ms/step - loss: 2.0441 - accuracy: 0.4088 - val_loss: 7.6056 - val_accuracy: 0.0911
Epoch 3/30
53/53 [=====] - 22s 423ms/step - loss: 1.6866 - accuracy: 0.5006 - val_loss: 1.1091 - val_accuracy: 0.0911
Epoch 4/30
53/53 [=====] - 21s 404ms/step - loss: 1.2783 - accuracy: 0.6261 - val_loss: 11.4796 - val_accuracy: 0.0911
Epoch 5/30
53/53 [=====] - 22s 416ms/step - loss: 0.9617 - accuracy: 0.7033 - val_loss: 11.0911 - val_accuracy: 0.0911
Epoch 6/30
53/53 [=====] - 22s 417ms/step - loss: 0.1016 - accuracy: 0.6785 - val_loss: 11.4910 - val_accuracy: 0.0911
Epoch 7/30
14/14 [=====] - 1s 74ms/step
Fold accuracy: 0.0911270976065894
Epoch 1/30
53/53 [=====] - 21s 438ms/step - loss: 3.0332 - accuracy: 0.2233 - val_loss: 3.6903 - val_accuracy: 0.1055
Epoch 2/30
53/53 [=====] - 21s 387ms/step - loss: 2.0818 - accuracy: 0.4142 - val_loss: 3.5971 - val_accuracy: 0.1055
Epoch 3/30
53/53 [=====] - 21s 392ms/step - loss: 1.7268 - accuracy: 0.4982 - val_loss: 3.8307 - val_accuracy: 0.1055
Epoch 4/30
53/53 [=====] - 21s 393ms/step - loss: 1.3892 - accuracy: 0.5810 - val_loss: 3.5618 - val_accuracy: 0.1055
Epoch 5/30
53/53 [=====] - 22s 419ms/step - loss: 1.0686 - accuracy: 0.6687 - val_loss: 3.2439 - val_accuracy: 0.1055
Epoch 6/30
53/53 [=====] - 21s 392ms/step - loss: 0.9131 - accuracy: 0.7215 - val_loss: 2.8272 - val_accuracy: 0.2278
Epoch 7/30
53/53 [=====] - 22s 421ms/step - loss: 1.2428 - accuracy: 0.6134 - val_loss: 2.8040 - val_accuracy: 0.1535
Epoch 8/30
53/53 [=====] - 21s 391ms/step - loss: 1.1949 - accuracy: 0.6519 - val_loss: 2.6380 - val_accuracy: 0.1391
Epoch 9/30
53/53 [=====] - 21s 397ms/step - loss: 0.9023 - accuracy: 0.7329 - val_loss: 2.1847 - val_accuracy: 0.3237
Epoch 10/30
53/53 [=====] - 22s 410ms/step - loss: 0.6209 - accuracy: 0.8205 - val_loss: 1.9559 - val_accuracy: 0.3813
Epoch 11/30
53/53 [=====] - 21s 397ms/step - loss: 0.4713 - accuracy: 0.8744 - val_loss: 1.6133 - val_accuracy: 0.4503
Epoch 12/30
53/53 [=====] - 20s 383ms/step - loss: 0.3362 - accuracy: 0.9245 - val_loss: 1.7396 - val_accuracy: 0.4058
Epoch 13/30
53/53 [=====] - 21s 390ms/step - loss: 0.2698 - accuracy: 0.9412 - val_loss: 1.5629 - val_accuracy: 0.5132
Epoch 14/30
53/53 [=====] - 21s 391ms/step - loss: 0.2063 - accuracy: 0.9580 - val_loss: 1.4898 - val_accuracy: 0.5108
Epoch 15/30
53/53 [=====] - 22s 417ms/step - loss: 0.1863 - accuracy: 0.9628 - val_loss: 1.4781 - val_accuracy: 0.5376
Epoch 16/30
53/53 [=====] - 21s 405ms/step - loss: 0.1226 - accuracy: 0.9804 - val_loss: 1.4490 - val_accuracy: 0.5875
Epoch 17/30
53/53 [=====] - 22s 424ms/step - loss: 0.0902 - accuracy: 0.9844 - val_loss: 1.4477 - val_accuracy: 0.5755
Epoch 18/30
53/53 [=====] - 21s 397ms/step - loss: 0.0722 - accuracy: 0.9928 - val_loss: 1.4890 - val_accuracy: 0.5755
Epoch 19/30
53/53 [=====] - 22s 409ms/step - loss: 0.5020 - accuracy: 0.8487 - val_loss: 4.5295 - val_accuracy: 0.2542
Epoch 20/30
53/53 [=====] - 21s 403ms/step - loss: 0.2280 - accuracy: 0.9484 - val_loss: 1.8259 - val_accuracy: 0.3022
Epoch 21/30
53/53 [=====] - 20s 371ms/step - loss: 0.1656 - accuracy: 0.9604 - val_loss: 1.4523 - val_accuracy: 0.5204
Epoch 22/30
53/53 [=====] - 21s 394ms/step - loss: 0.1225 - accuracy: 0.9754 - val_loss: 1.7309 - val_accuracy: 0.5516
Epoch 23/30
14/14 [=====] - 1s 72ms/step
Fold accuracy: 0.575539589822227
Epoch 1/30
53/53 [=====] - 23s 406ms/step - loss: 3.0513 - accuracy: 0.2166 - val_loss: 3.9397 - val_accuracy: 0.0385
Epoch 2/30
53/53 [=====] - 21s 399ms/step - loss: 2.1692 - accuracy: 0.3905 - val_loss: 3.8521 - val_accuracy: 0.1322
Epoch 3/30
53/53 [=====] - 21s 397ms/step - loss: 1.7901 - accuracy: 0.4841 - val_loss: 3.4302 - val_accuracy: 0.1322
Epoch 4/30
53/53 [=====] - 20s 382ms/step - loss: 1.6180 - accuracy: 0.5369 - val_loss: 3.2080 - val_accuracy: 0.1322
Epoch 5/30
53/53 [=====] - 21s 404ms/step - loss: 1.2637 - accuracy: 0.6275 - val_loss: 3.0486 - val_accuracy: 0.1322
Epoch 6/30
53/53 [=====] - 21s 395ms/step - loss: 1.1393 - accuracy: 0.6665 - val_loss: 2.9735 - val_accuracy: 0.1587
Epoch 7/30
53/53 [=====] - 21s 393ms/step - loss: 1.1291 - accuracy: 0.6239 - val_loss: 2.4840 - val_accuracy: 0.2716
Epoch 8/30
53/53 [=====] - 20s 379ms/step - loss: 0.8941 - accuracy: 0.7029 - val_loss: 2.1686 - val_accuracy: 0.2067
Epoch 9/30
53/53 [=====] - 21s 390ms/step - loss: 0.7041 - accuracy: 0.8038 - val_loss: 2.2256 - val_accuracy: 0.3510
Epoch 10/30
53/53 [=====] - 21s 397ms/step - loss: 0.5380 - accuracy: 0.8767 - val_loss: 2.1813 - val_accuracy: 0.3598
Epoch 11/30
53/53 [=====] - 20s 383ms/step - loss: 0.9482 - accuracy: 0.7052 - val_loss: 2.4015 - val_accuracy: 0.2981
Epoch 12/30
53/53 [=====] - 21s 392ms/step - loss: 0.6021 - accuracy: 0.8266 - val_loss: 2.1386 - val_accuracy: 0.3582
Epoch 13/30
53/53 [=====] - 20s 379ms/step - loss: 0.4674 - accuracy: 0.8830 - val_loss: 1.5763 - val_accuracy: 0.4591
Epoch 14/30
53/53 [=====] - 19s 362ms/step - loss: 0.3454 - accuracy: 0.9262 - val_loss: 1.5328 - val_accuracy: 0.5245
Epoch 15/30
53/53 [=====] - 20s 379ms/step - loss: 0.2463 - accuracy: 0.9406 - val_loss: 1.6530 - val_accuracy: 0.5240
Epoch 16/30
53/53 [=====] - 20s 369ms/step - loss: 0.1893 - accuracy: 0.9664 - val_loss: 1.5038 - val_accuracy: 0.5649
Epoch 17/30
53/53 [=====] - 20s 383ms/step - loss: 0.2034 - accuracy: 0.9538 - val_loss: 1.7858 - val_accuracy: 0.5120
Epoch 18/30
53/53 [=====] - 20s 378ms/step - loss: 0.2033 - accuracy: 0.9562 - val_loss: 1.8658 - val_accuracy: 0.4928
Epoch 19/30
53/53 [=====] - 21s 395ms/step - loss: 0.1494 - accuracy: 0.9706 - val_loss: 1.7280 - val_accuracy: 0.5072
Epoch 20/30
53/53 [=====] - 20s 385ms/step - loss: 0.1196 - accuracy: 0.9790 - val_loss: 1.6173 - val_accuracy: 0.5361
Epoch 21/30
53/53 [=====] - 21s 393ms/step - loss: 0.0862 - accuracy: 0.9874 - val_loss: 1.7479 - val_accuracy: 0.5385
Epoch 22/30
13/13 [=====] - 1s 69ms/step
Fold accuracy: 0.5913461446762085
Mean cross-validated accuracy: 0.49264088669094846
```

