Massive Data Storage and Retrieval : Course Project Intermediate report

Ranjitha Korrapati – rk850

Whale Call Acoustics : A Transfer learning approach to classify whale calls.

Section I : Project goals:

Planet earth is bestowed with myriad of species of both flora and fauna which is not totally known to humans and is almost unreachable. It is very important for a lot of biologists and marine scientists to understand various species. One such area that has always been a mystery is marine life. The goal is to use deep learning and machine learning techniques to solve or automate the process of understanding the diversity of species in earth.

As an application to extend this idea, whale call classification was chosen , to decipher how whales live socially. Whale make songs / calls to communicate among themselves. They live together in social groups called “pods” and carry out their day-to-day activities like hunting. So whales communicate through calls/songs within the pods. The idea is to classify a whale call into these pods as well as their respective species based on the audio clips of whale calls available to us.

Problem statement: To find the type of species and pod of a particular whale based on the audio clip of whale call.

The problem is interesting because

1. It is not possible manually to classify whale recordings on a large scale.
2. For marine biologists to estimate the population numbers of whales , to understand the social groups of whales and different types of species , such models can help automate the process .
3. This will also help in comparing whale populations over years and take measures if there are signs of extinction.

Transfer learning can help us achieve training a neural network with small amounts of data using pre-trained networks and adding personalized layers. Such techniques are especially useful in real world scenarios where data is not always available as we require. So with whale call data set not being so large , pre-trained neural networks can be used to efficiently classify the whale calls. Though the pre-trained neural networks are trained on images, the subtleties can be decided only in the higher levels. So using pre trained networks based on images can still be extended to audio data by customizing the higher layers according to our requirements. Some of the examples of the pre trained networks are three versions of Inception by google, ResNets , etc.

Section II : Data:

Data Collection and Exploratory Data Analysis:

Source: Data has been collected from the source called Whale FM.

Whale FM is a project to organize Pilot whale calls and Ocra whale calls. A lot of volunteers helped identify the recordings collected as a particular species based on fixed and approved recordings.

Data Format:

Whale call Data is available in the form of :

1. Audio clips (.mp3)
2. Spectrograms (.jpg)

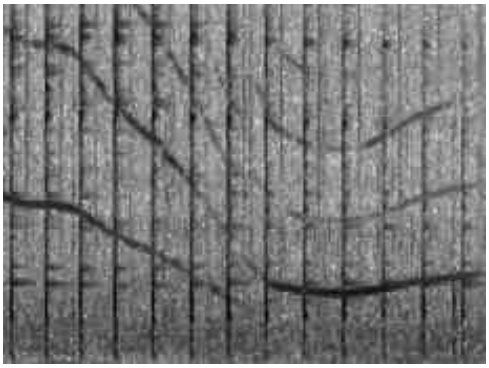
We can see that the same kind of whale call is available in terms of audio as well as spectrograms, both of these formats can be made use of in the analysis and for training the model.

Sample data in the formats mentioned above:

Audio clip 1 : Whale call

▶ Click on the icon to listen to a whale call.

Figure 1: Spectrogram of the same whale call



The idea is to use both the images and sound to train pre trained neural networks to use transfer learning in our application.

Data Preprocessing:

The data archive has a lot of information about how the whale calls were classified, how many volunteers could confidently classify it as one type and also the different types of whales, their location.

These images, audio clips and a lot of relevant information is available in both the following formats with links to those image and audio files:

- .csv
- .sql
- .json

For the purpose of this project, csv was used to get the data.

The following are some of the important columns whale_fm_anon_04-03-2015_assets.csv:

Table 1: Description of some of the columns in whale_fm_anon_04-03-2015_assets.csv

Name of the column	Description
'Name'	Name of the audio file
'Location'	Location of the audio file
'Classification_Count'	No of times the clip was classified as a particular type of whale
'Spectrogram'	Location of spectrogram
'Lat'	Latitude of the place where this recording was found.
'Lng'	Longitude of the place where this recording was found.
'Whale_Id'	Id of the whale: essentially indicating the pod a whale belongs to.
'Whale_Type	Species of the whale.

There are 26 pods in the given data set:

Here is the description of the 26 pods that give us characteristics of a particular pod:

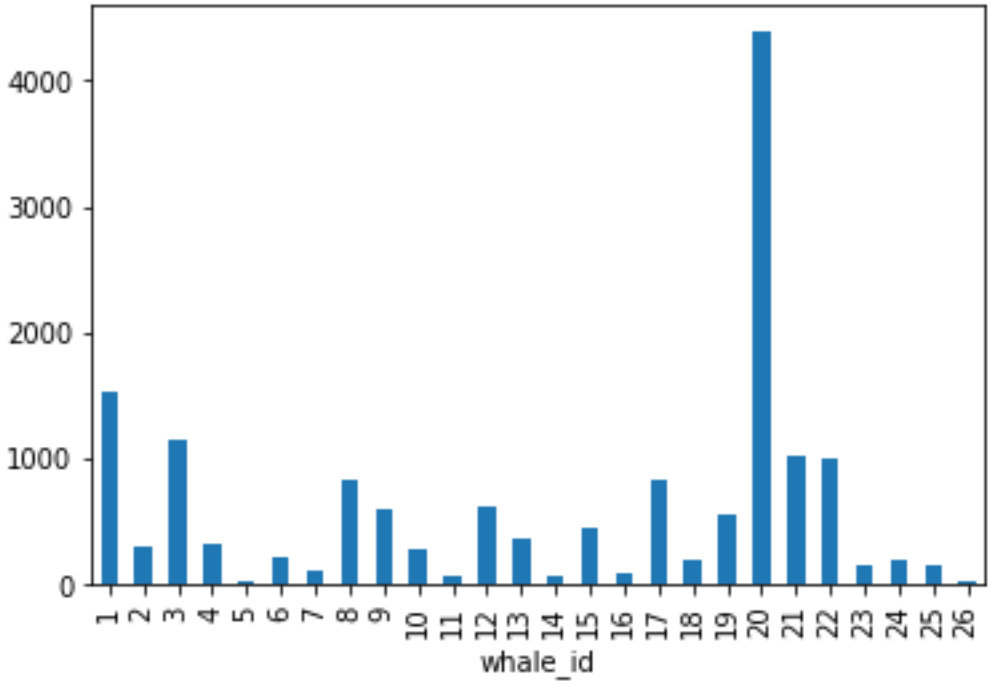
Table 2: Pod Description

Pod	Description
1	A short-finned pilot whale. This is adult male was recorded near the Bahamas.
2	A short-finned pilot whale. This is either an adult female or a younger whale. It was recorded near the Bahamas.
3	A short-finned pilot whale. This is adult male was recorded near the Bahamas.
4	A short-finned pilot whale. This is adult male was recorded near the Bahamas.
5	An adult female whale tagged off the coast of Iceland.
6	A juvenile male whale off the coast of Iceland.
7	A juvenile whale of unknown gender, tagged off the coast of Iceland.
8	We don't have much information on this particular animal. It was tagged near Norway.
9	We don't have much information on this particular animal. It was tagged near Norway.
10	We don't have much information on this particular animal. It was tagged near Norway.
11	We don't have much information on this particular animal. It was tagged near Norway.
12	We don't have much information on this particular animal. It was tagged near Norway.
13	We don't have much information on this particular animal. It was tagged near Norway.
14	We don't have much information on this particular animal. It was tagged near Norway.
15	A medium-size, female with a calf, travelling as part of a group of 30 to 40 whales.
16	A medium-sized whale of unknown age and gender in a group of 10 to 20 animals.
17	A female pilot whale in a group of 40, including calves. She was often seen traveling with a calve.
18	A large male pilot whale. Part of a group of 80-100 whales
19	An adult male travelling in a group of 15 animals, with calves both on adult males. Calls from all the animals may be recorded by this tag.

Pod	Description
20	Long-finned pilot whale calls record from the first expedition of the Delphinus hydrophone array.
21	Long-finned pilot whale calls record from the second expedition of the Delphinus hydrophone array.
22	NULL
23	We don't have much information on this particular animal. It was tagged near Norway.
24	This whale was tagged near Norway.
25	We don't have much information on this particular animal. It was tagged near Norway.
26	We don't have much information on this particular animal. It was tagged near Norway.

So our aim is to classify a given whale call into one of the 26 pods and also the species but before that we will have to see if we have enough data from each of the pods.

Figure 2: Number of samples in each pod



The x axis represents whale id which indicates the pod number and the y axis represents We can see that some pods have samples as high as 4000 whereas others in few tens. In order to train network accurately , only pods with number of samples greater than a threshold will be selected.

Here are some of the stats:

1. The total number of recordings = 15,531 rows
2. As seen in the figure, the number of samples of each type of pod is different in each case. A threshold of 200 is set and only pods with number of samples greater than 200 are considered. The following are the whale pods (17 pods) that have greater than or equal to 200 samples of recordings:

Table 3: Pods with number of recording samples greater than threshold 200 in descending order

Pod Number	Number of occurrences:
20	4390
1	1536
3	1148
21	1020
22	989
17	824
8	823
12	611
9	600
19	561
15	449
13	358
4	329
2	303
10	288
6	216
24	200

So our task boiled down to:

1. Classifying the whales as killer and pilot whales and into one of the seventeen pods mentioned above.
2. The data from these pods will be considered for the analysis.
3. The number of recordings from these 17 pods: 14, 645

Section III: Processing the audio data

Figure 3: Raw wave form of a pilot whale

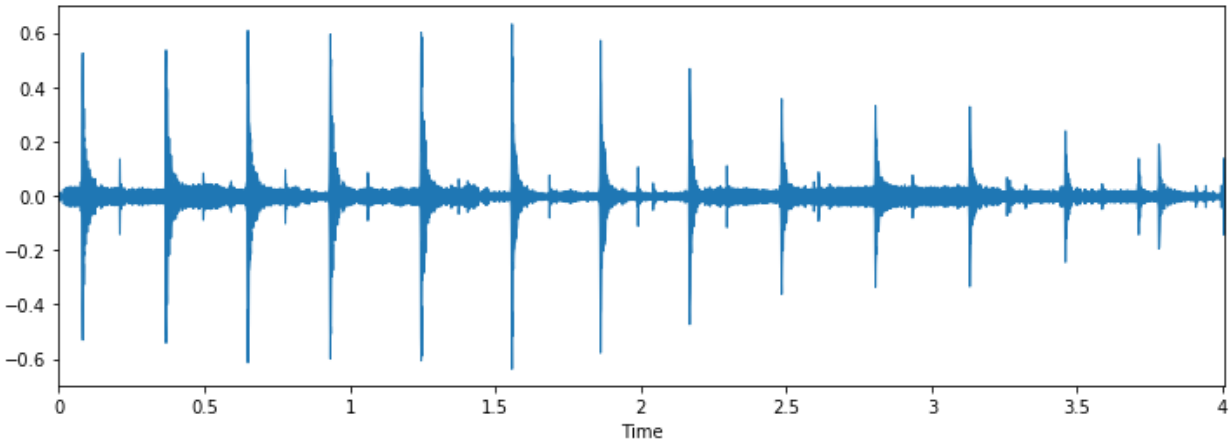
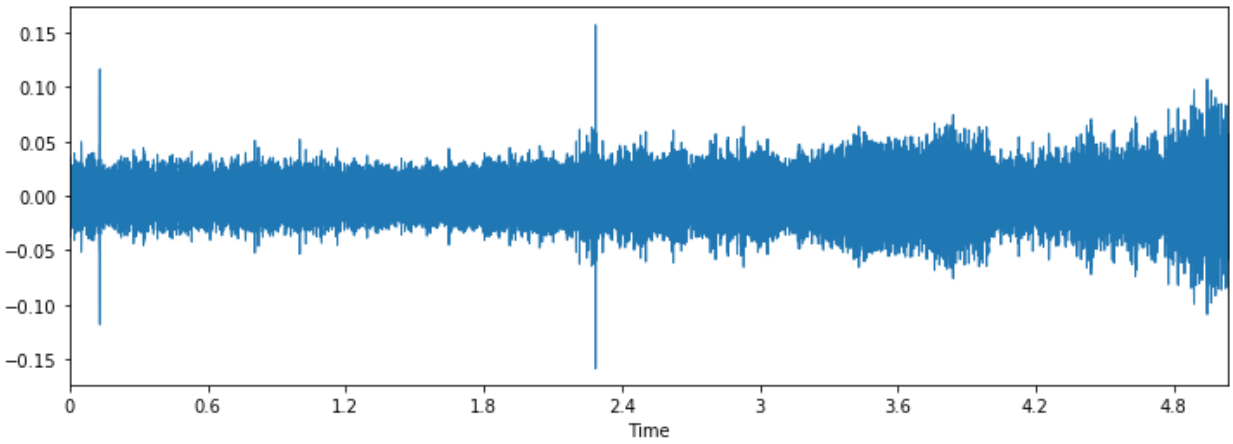


Figure 4: Raw wave form of a killer whale



We can see that the raw wave forms can show how different a pilot whale and killer whale audio clips are different, but this step is not alone sufficient as they are not ready to be compared coming from different scales, durations, sampling rates and also noise in the background. So more data processing is required to understand the problem better. The audio files range from duration of one second to eight seconds

Important pre-processing steps:

- Resampling of all the waveforms to 22050 Hz to ensure uniformity and normalization.
- Choosing time duration windows(three) to capture various sketches of a whale call.

All the audio files are not directly available in the csv files, but the links are available. So, instead of loading the data repeatedly, we can pickle the mp3 files to be used for later .

Section IV: Applications of the data

- Identification of underwater species.
- Estimation of population of whales in a given region
- Understanding whale population interactions and their social life.
- Measures can be taken in times of extinction prediction.

Section V: Next Phase

- Using the preprocessed data to train pre-trained Convolutional Neural networks to classify sounds.
- Extend the analysis to log-mel spectrograms as well.
- Use the existing given spectrograms for analysis

Section VI : References

1. https://whale.fm/
2. https://github.com/zooniverse/WhaleFM
3. Large-Scale Whale-Call Classification by Transfer Learning on Multi-Scale Waveforms and Time-Frequency Features Lilun Zhang , Dezhi Wang , Changchun Bao , Yongxian Wang and Kele Xu : Link
4. Classification of large acoustic datasets using machine learning and crowdsourcing: Application to whale calls : Link
5. https://www.analyticsvidhya.com/blog/2017/08/audio-voice-processing-deep-learning/
6. https://towardsdatascience.com/getting-to-know-the-mel-spectrogram-31bca3e2d9d0