Happy Whale Classification Image Challenge Report

Alizibali Abdulmonem, Dániel Bence Papp

*Department of Computer Science, Old Dominion University  
Norfolk, VA, USA.*

***Abstract*— The classification and clustering of data is still an enormous challenge for the machine learning and data science community. The given project was carried out using the commonly known K-means algorithm, implemented in two different ways. In this project, we were tasked to cluster a data set based on their x and y coordinates and their respective distances to our adapting centroids.**

1. The libraries

Regardless of what approach we might use, some of these basic libraries are crucially important to have for ease of development. To visualize our data on plots and graphs, we are going to be using the Seaborn and the Matplotlib libraries. Both packages are very common in use cases across the internet, as they are easy to use and have a huge variety of idioms at our disposal. For parts of this project, I’ll be using Pandas and Numpy. Pandas will be used to

A picture containing aquatic mammal, mammal, water, dolphin

Description automatically generated

Fig. 1 Four chosen images from the given dataset

1. approach

To perform the classification task. the first step is to understand the dataset. It was observed that direct classification of classes would be less effective than staged classification. We have used a cascaded convolutional neural network to accomplish the given task.  Using a cascaded network will not only reduce the complexity of the model but will also work more efficiently as compared to direct classification models. As seen in Table 1 there are two major classes dolphin and whale, which are further classified into various subclasses so, our approach is using the same hierarchy. There are two stages in the approach, one is the initial stage where binary classification is performed and the second is the final stage where multiple classifications are performed to determine the exact class. Figure 2 shows the flow diagram of our model.

1. algorithm

As discussed above in the dataset section the whole dataset can be classified into two basic classes i.e., whale and dolphin. In this initial stage, we are performing binary classification to predict whether the given data belongs to the whale class or the dolphin class.

Diagram

Description automatically generated

Fig. 2 Flow diagram of our model

To perform binary classification, we have used a convolutional neural network. Figure 3 shows the convolutional neural network architecture used for binary classification.

Text

Description automatically generated

Fig. 3 The convolutional neural network architecture of the initial stage

As this stage is based on binary classification, the activation function for the above-mentioned function is sigmoid and the loss function used is the binary cross-entropy function. The output of the initial stage is in float numbers ranging from 0 to 1 where values lesser than 0.5 represent whales and more than 0.5 represent dolphins.

1. sklearn

Text

Description automatically generated

Fig. 4 The convolutional neural network architecture of the dolphin classification

In the previous stage, the model predicted the major class (species), and in this stage, the model predicts the actual subclass (the type of species). This stage has two convolutional neural networks, one is for whale classification and the other is for dolphin classification. Out of two, only one is used at a time depending upon the predictions made in the previous stage. The architecture of the convolutional neural network for whale classification and dolphin classification is nearly the same except for the output layer as a dolphin has 10 classes whereas a whale has 18 classes. Figure 4 shows the convolutional neural network architecture of dolphin classification and Figure 5 shows the convolutional neural network architecture of whale classification.

Text

Description automatically generated

Fig. 5 The convolutional neural network architecture of the whale classification

As this stage is based on multiple classifications, the activation function for the above-mentioned function is the soft-max function and the loss function used is the sparse categorical cross-entropy. The output of the final stage is in float numbers ranging from 0 to 1 and the class which obtains the maximum value is considered the predicted class for both whale and dolphin classification convolutional neural networks.

Figure 6 and figure 7 show the output results of the dolphin and whale classification convolutional neural networks.  In the images given below each row is one image and the columns represent the output classes. In the case of dolphins, we have 10 columns i.e., 10 classes whereas in the case of whales, we have 18 columns i.e., 18 classes.

A screenshot of a computer

Description automatically generated with medium confidence

Fig. 6 The output of the convolutional neural network representing dolphins

A screenshot of a computer

Description automatically generated with medium confidence

Fig. 7.1 The output of the convolutional neural network representing the first 10 columns of whales

A screenshot of a computer

Description automatically generated with low confidence

Fig. 7.2 The output of the convolutional neural network representing the last 8 columns of whales

This output is encoded so after inverse encoding and arranging output is presented in form of a dictionary where keys are the indexes of the images and values are the predicted class. Figure 8 shows the arranged output.

Text

Description automatically generated

Fig. 8 The final output

1. numpy

The above-explained model saves encoded weights and trained models. To test the model initially, all the encoded weights along with three trained convolutional neural networks should be loaded. After this image can be fed and the output is received in form of a dictionary.

The training dataset was split before training to calculate performance metrics and evaluate the performance. This was done because the ground truth for testing images was not present. The split dataset consists of 695 images and using these images’ confusion matrix the *f1 score*, and the *accuracy* for all of the convolutional neural networks was calculated. The size of the confusion matrix is dependent on the number of classes to be predicted by the convolutional neural network. Figure 9 shows the confusion matrix.

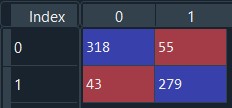


Fig 9.1 Confusion Matrix of initial convolution neural network

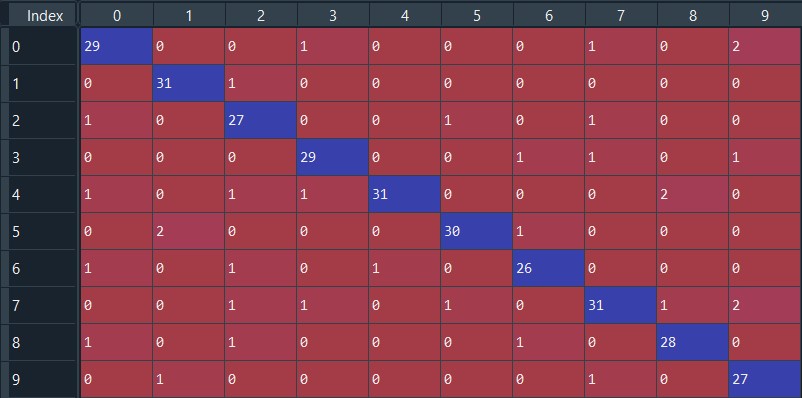


Fig 9.2 Confusion Matrix of dolphin classification convolutional neural network

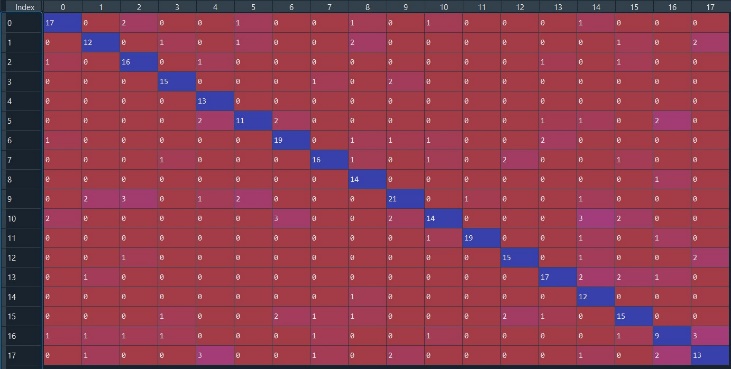


Fig 9.3 Confusion Matrix of whale classification convolutional neural network

Figure 10 shows the *f1 score*. The *f1 score* is calculated for each class that is to be predicted. Values of F1 score range between 0 and 1 where 1 is considered to be the best output.

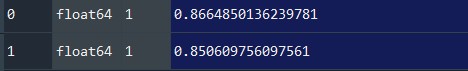


Fig 10.1 F1 score of initial convolution neural network

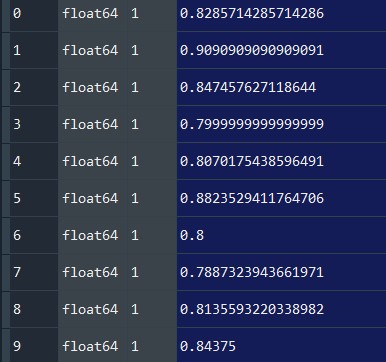


Fig 10.2 F1 score of dolphin classification convolutional neural network

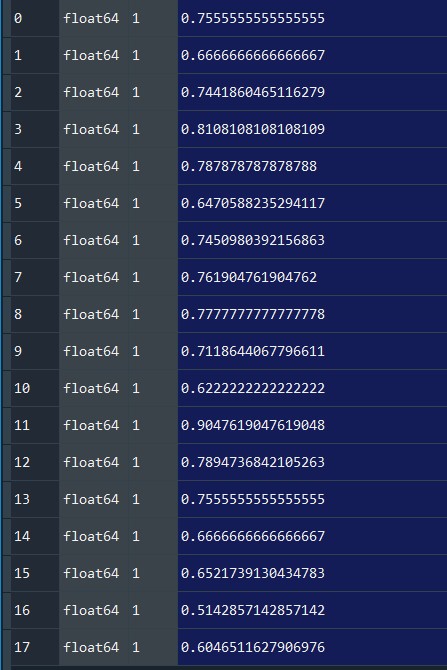


Fig 10.3 F1 score of whale classification convolutional neural network

Apart from the confusion matrix and the *f1 score,* the *accuracy* for all the convolutional neural networks was calculated. The accuracy of the initial convolutional neural network, the dolphin classification convolutional neural network, and the whale classification convolutional neural network are 85.89%, 83.22%, and 71.84% respectively. It was observed that with the increase in the number of classes the accuracy of the convolutional neural network is decreased. The overall network accuracy was calculated to be 80.31%.

The model mean accuracy formula that was used to calculate the overall accuracy of the convolutional neural network

1. conclusion

The classification of marine beings is still a challenge for researchers across the globe. In the given model, a cascaded convolution neural network was used to solve the problem at hand. Our model is solved in two steps: firstly, a binary classification is used to classify between parent classes which are whale and dolphin, and in the second step, a multi-class classification is done to determine the exact class of whale and dolphin. This approach is less complex and better than direct classification. The training accuracy of this model resulted in over 80%. In the future, there is scope to improve this model by adding more convolutional neural networks in training or by introducing segmentation in preprocessing stage.

References

1. “Happywhale - Whale and dolphin identification,” Kaggle. [Online]. Available: https://www.kaggle.com/c/happy-whale-and-dolphin. [Accessed: 03-Apr-2022]. ­
2. “ML Practicum: Image Classification  |  google developers,” Google. [Online]. Available: https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks. [Accessed: 03-Apr-2022].

Run your algorithm with K=3, calculating the Sum of Squared Error (SSE) at the end of each iteration of cluster membership assigning followed by centroid updating in your algorithm. Plot the SSE against the number of iterations, i.e., the convergence curve (as in the 6th slide from lecture 15). Does the SSE monotonically decrease when the number of iterations grows? If not, why?

(2) Run your algorithm with different choices of K’s (ranging from 2 to 10 with step size 1). Calculate the SSE for clustering resulted from each run, and plot SSE against K in a figure as shown on 8th slide from lecture 15. Note for each K, multiple runs of the algorithm with different initial centroids may be needed, in order to have a curve that is monotonically decreasing.

(3) Run your algorithm with K=3. Show the clustering results in a figure with varying colors or shapes of data points to represent different cluster assignment (as in those figures in 7th slide from lecture 15). Note multiple runs may be needed in order to obtain a satisfactory clustering result.