Residual Networks based Classification of Right Whales in the Ocean

Ghazanfar Latif [1*], Faisal Yousif Al Anezi [2], Mohammad O. Butt [3], Jaafar Alghazo [4]

[1] Department of Computer Science, Prince Mohammad bin Fahd University, Al-Khobar, Saudi Arabia
[2] Department of Management Information Systems, Prince Mohammad bin Fahd University, Al-Khobar, Saudi Arabia
[3] Department of Electrical Engineering, Prince Mohammad bin Fahd University, Al-Khobar, Saudi Arabia
[4] Department of Electrical Engineering, Prince Mohammad bin Fahd University, Al-Khobar, Saudi Arabia
* glatif@pmu.edu.sa

Keywords: Deep ResNet, CNN, Whale Classification, Whale Recognition, Deep Learning

Abstract

Automatic classification and recognition of various species of whales is useful in the close monitoring of various species of whales for scientific purposes and for population count as well. Marine animals including whales play an integral part in balancing the ecosystem and if left unsupervised whaling and other fishing activities (legal or illegal) can lead certain species to extinction, thus causing an imbalance in the ecosystem. In this work, a modified Residual Network (ResNet) has been proposed for the classification of right whales amongst 50 different classes. ResNets with their skip-connections feature bypasses the problem of vanishing gradient which results from the use of CNN with many convolutional layers. The deep ResNet is composed of 72 layers and involves from 5 up to a maximum of 30 iterations for the purpose of classifying the right whales. The use of the proposed ResNet achieved a recognition rate of 92.15%, showing a very high accuracy of correct whale classification. The accuracy achieved in this paper is better than those reported in previous literature.

1 Introduction

Whales are cold water animals and along with various species of fish, their existence has been greatly affected due to the anthropogenic activities leading to global warming. Furthermore, due to extensive whaling [1] in 20th century, the population of this mammal was reduced significantly and some species were endangered to extinction. Whales also play an important role in balancing the ecosystem [2, 3] by absorbing the carbon that emanates from oceans, consequently leading to increased greenhouse gases in atmosphere and regulating the climatic temperature. Over the past several decades scientists have taken serious interest in understanding the whale lifecycle and their migration trends.

One type of whale is the right whales or referred to as the black whales which are three species of the baleen whales. These three species are: the north pacific right whale, the north Atlantic right whale, and the southern whales. The white whale have distinctive features and the most distinguishing feature is a head full of rough patches of skin appearing white due to whale lice. These right whales are the most preferred by whales to hunt and thus put them on the top chart of the most endangered species. Another type of whale is the Humpback whale, named because they have a distinctive hump on their back [4]. The whale-watchers tend to show attraction to this type of whales because of their tact of jumping out of water and slapping the water surface with tails. A statistical analysis [5] shows that 13 million whale-watchers belonging to 119 countries were registered in 2008, consequently leading to a global economic activity summing up to USD 2.1 billion.

The humpback whales have unique flukes (tail patterns), fin sizes and scars and their flukes can be used for identification purposes. Until recently, humpback whales were considered to be an endangered species. A recent statistical analysis [6] shows that due to protective measures, the population of humpback whales is increasing and it is not an endangered species anymore. The shapes of flukes and unique markings found in footage is utilized to identify the species, and precisely logging the whale pod dynamics and movements. Scientists use image surveillance systems in order to aid the whale conservation efforts. Over the past 40 years, most of the image acquisition and classification work has been done manually by scientists, leaving a huge trove of data untapped and underutilized. The accuracy of such analyses is also low because there are numerous types of whales. The later studies have incorporated a partial labelling of images containing flukes [7].

The advancements in domain of image processing and machine learning have enabled automated image analysis schemes and eliminated the time-constraints imposed due to traditional manual procedures. The machine learning algorithms develop a mathematical model by utilizing the training data, and then perform predictive analysis on the data to be examined [8]. The image recognition and classification is usually performed by a class of multilayered feed forward neural networks, also referred as Convolutional Neural Networks (CNN). Although it has been observed that deep learning based techniques tend to provide a means of accurate and swift image analysis, especially in the assessment of the population of different types of whales but as the number of layers increases the accuracy of such CNN based systems degrades and CNN has been applied vast verity of imaging based applications including but not limited to handwriting recognition, traffic sign recognition and diabetic retinopathy detection [9-11]. Unlike the automatic learning schemes, deep learning realizes the tasks as a conceptual hierarchy i.e. it dilutes information into several layers using the modules that transform their representation into a more abstract level [12]. Thus, deep residual networks provide a better alternative for the automatic recognition of whale type. The skip connection feature in deep residual networks overcomes the problem of degraded accuracy with increase of layers.

The motivation of this work is to design an automatic system that is able to recognize the different species of whales. Such systems can be deployed globally to monitor the population of whales especially whale species that are considered to be endangered and at the verge of extinction. As mentioned previously, the ecosystem has a certain design or place for each living thing including whales. Therefore, it essential for humankind to ensure that species of any kind do not become extinct. In this paper, we will target only the recognition of whale species based on images which will allow the automatic deployment of such systems globally for monitoring whale populations.

The rest of this paper is organized as follows: section 2 is about introduction, section 3 describes the Residual networks and section 4 describes the proposed Residual Model. The experimental results are explained in the section 5 and at the end section 6 contains the conclusion.

2 Literature Review

An automated and accurate image processing scheme for classifying different species of whales is the key to identify the population of each species. In [13], the humpback whale identification process is explained in lieu of observing their migration patterns. A Hot Spotter algorithm based on matching of animals against a labelled database has been presented in [14], achieving a 77.81% top-1 ranking on Splash dataset. An approach based on Curve Matching that evaluates the integral curvature along the trailing edge of flukes is explained in [15], which has shown an accuracy of 87.7% over the Splash dataset. A cohesive approach utilizing both the Hot Spotter and Curve Matching algorithm has been presented in [16], which shows superior performance of achieving a top-1 accuracy of 94.9%. In [17], an approach for dog breed classification has been presented by employing the part localization and recognition rates have shown 90% accuracy. The Kaggle competition data has been used in [18], including 80% data for training and 20% for testing purposes. The complexities related to application of shallow CNNs (two layers) have been highlighted and an accuracy of is 32% achieved. The application of an Ensemble Network model showed an improved accuracy of 46.74% after running ten epochs on the same dataset.

An accurate automated image classification technique can aid in devising a targeted plan to protect the endangered species of whales. In [19], a collaborative work between data scientists and biologists has been presented which incorporates series CNNs for automated identification of region of interest along with several image adjustment procedures. The technique has achieved an accuracy of 87% and such inter-domain collaborative works can aid in bridging the gaps among conservation science and data science communities. The significance of homogenizing and data augmentation in deep learning models has been explained in [20]. In data augmentation, various derived images are generated by employing certain modifications to the original image such as rotation, translation and noise reduction. This technique is particularly useful if various classes contain only one sample. In [21], a deep learning based technique for classification texture has been presented and evaluated on various benchmark datasets pertaining to the texture including Brodatz, Kylberg v1.0 and Outex_TC_00012. Each dataset is subdivided into three sets i.e. training, testing and validation with 80%, 10% and 10% images allocated for each set respectively. In [19], an optimization technique based on Whale Optimization Algorithm (WOA) [22] has been presented. The optimization of filter values and weights is performed at two levels including convolutional and fully-connected layer. The predictive outcomes of this scheme have shown better accuracy in comparison to benchmark schemes.

In a similar manner, several studies have been conducted for accurate identification of whales based on the sounds they produce. A reliable detector implementation has been presented in [23] that has the ability to accurately perform detection of right whale calls and minimize the false alarm rate. The detection process starts by smoothing of spectrogram to extract the parameters such as duration, bandwidth and frequency contours and afterwards extracted parameters are used for classification of sound. Another cohesive approach [24] utilizes the outputs of various finely tuned CNNs in order to detect whale sounds in the spectrograms. Explainable Artificial Intelligence is utilized for assessment of relevance among features and also leads to creation of masking schemes. In [25], the construction of echolocation detectors for performing classification of spectrograms has been discussed based on the occurrence or lack of click. This scheme attained an accuracy of 99.5% in classification of 650 spectrograms of whales. Another approach to detect creaks and clicks for various whales has been developed in [26] which utilizes Teager-Kaiser energy operator. Based on the test data utilized in this study, the correct detection rate was 94.05% on average. In [27], a cost-effective and globally applicable approach for automated detection and counting of whales has been presented. Imagery of Google Earth from various whale-watching hotspots are tested and an accuracy of 84% and 97% in detection and counting of 80 whales respectively.

In [28], the importance of underwater imagery is highlighted and is regarded as a non-destructive means of analyzing fish, flora and fauna. The underwater scenes are considered as variable and complex due to various parameters including the light intensity, variations in fish orientation, similarities in shapes and patterns among the fish from different species. Hence, a CNN model is trained using LifeCLEF14 and LifeCLEF15 datasets to learn the visual features depending on various species and 90% correct classification rate is achieved. The issues relating to data augmentation approaches are addressed in [29], where a small set of ImageNet dataset has been used to perform a comparative analysis of different data augmentation techniques. A methodology is devised in order to learn various augmentations that enhance the classification process. In case of dog versus fish, the neural augmentation has achieved 77.0% to 70.5%. An automated fish labelling technique has been proposed in [30], the classification is performed in a split technique where one block performs detection, alignment and the other block performs classification. A Fast Regions CNN has been used in [31], in order to perform the detection and recognition of various fish species by utilizing the underwater imagery. ImageCLEF dataset has been used which contains fish imagery from 12 classes, outcomes show faster detection rates and better mean average precision as compared to Deformable Parts Model.

3 Deep Residual Networks

The common approaches pertaining to classification of hyperspectral images are based on CNN models. However, it is observed that the accuracy of classification deteriorates as soon as the number of convolutional layers is increased over five layers [32]. The degradations in classification can be avoided by introducing certain modifications to the model i.e. shortcut connections amongst different layers, this technique is also referred to as deep residual learning [33]. The ResNets are typically implemented with multiple layer skips comprising of nonlinearities (ReLU) and batch normalization. Layers skip introduces a degree of simplification in the network i.e. use of lesser number of layers during initial training phase. Consequently, reduction in vanishing gradients leads to improvements in the learning rate. The skipped layers are restored gradually as the network learns its features space. However, in a normal neural network more feature space is explored and it shows vulnerability towards perturbations which indicates requiring more data for training purposes and recovery. In [34], a ResNet has been formulated according to Equation (1).

$$X_{k+1} = \max\{0, X_k + F(X_k, W_k)\} \quad (1)$$

where X_k denotes the outcomes of k^{th} layer, and W_k represents the residual structure parameters. The objective of nonlinearly stacked arrays is to generate the function $F(X_k, W_k)$ rather than a direct mapping of X_{k+1}.

In comparison to the CNNs, ResNets depict an ease in optimization, have higher representative capacities and better recognition accuracy while incorporating more layers. The general blocks in architecture of ResNet include batch normalization layers, residual blocks (each block containing several convolutional layers) and ReLU layers. Residual blocks allow bypassing of several convolutional layers, by addition of shortcut connections being aggregated at convolutional layer outputs [30].

4 Proposed ResNet Model

The proposed deep ResNet model is comprised of 72 layers in total. The layers include Input Layer (IL), Convolutional Layer (CL), Batch Normalization Layer (BNL), Activation Layer (AL), Addition Layer (ADL), Pooling Layer (PL), Flattening Layer (FL) and Dense Layer (DL). Each layer performs a specified task. The IL is composed of artificial neurons, providing the primary data to system for further processing. The CLs provide a features mapping by performing convolution of the data at their input. The BNLs perform normalization of input data, and are used between CLs and nonlinearities (skip connections), in order to reduce the sensitivity to initialization of network. The ALs tend to perform conversion of input signals into output signals, and are responsible for inducing the non-linearity in network. The ADLs are used to perform a summation of inputs originating from multiple layers. The PLs are used to perform down sampling of features set, thus reducing the computational complexity of system. The FL performs transformation of two dimensional data, and feeds forward for classification purpose. The DL can be regarded as a basic layer which is fed by all the outputs from prior layer and generates a separate output from each neuron.

The proposed deep ResNet model comprises of 72 layers in total, out of which there is 01 OL, 21 CLs, 19 BNLs, 19 ALs, 09 ADLs, 01 PL, 01 FL and 01 DL. The classification performance has been evaluated on iterations ranging from 5 to 30 iterations. The proposed ResNet model has been compared with CNN model consisting 16 layers.

4 Experimental Results

The dataset of right whales comprising of 1137 images provided by National Oceanic and Atmospheric Administration (NOAA) has been used in this work [35]. The dataset is comprised of 50 classes of whales. In order to emphasize on the right whale classification, the metadata has been removed from images of dataset. A total of 937 images have been used for the learning purposes and remaining 200 images have been used for testing/classification purpose. The images are of 64×100 size, with RGB coloring scheme. Fig. 1 shows some Whale dataset samples in the ocean used for Whale identity classification. As the dataset size is very small so the data is expended using the augmentation based angular rotations and shift whale images based on random flips. The training data has been augmented eighteen times using nine angular rotations and nine random flips which increases the dataset size to 21603 whale images.

The performance of proposed deep ResNet based scheme have been evaluated in Table 2 while the comparison is made with the CNN as shown in Table 1. Table 1 shows that CNN achieves maximum training accuracy 82.45% with loss value of 0.801 while the RestNet model achieves higher accuracies. As shown in Table 2, the degree of correct classification improved as the number of iterations have increased i.e. for 5 iterations a classification accuracy of 51.75% was achieved, whereas for 30 iterations an accuracy of 92.15% has been achieved. Furthermore, it has been observed that the performance of classification tends to saturate at the evaluated maximum value of 92.15%, even if the number of iterations is increased.

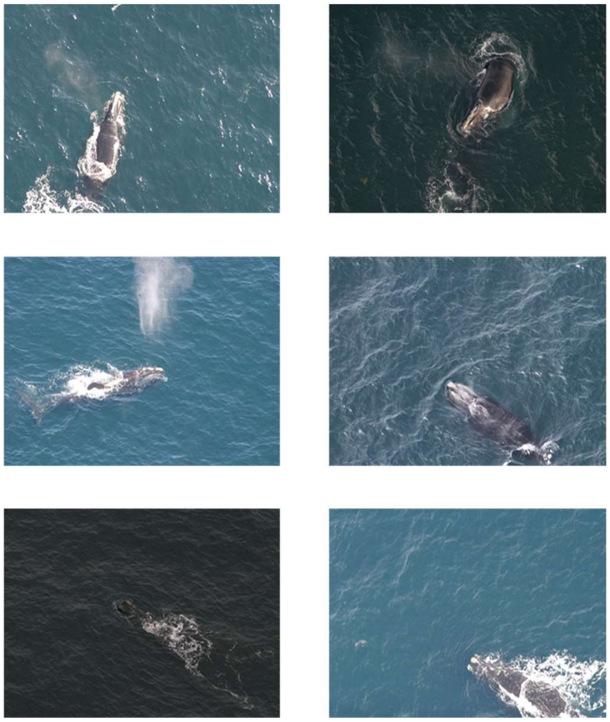

Figure 1. Sample of Whale Images for Whale Identity classification

Table 1. Right whale classification results using CNN

Iterations	Val. Acc. %	Val. Loss	Test Acc. %	Test Loss
5	33.16	2.564	38.15	2.403
10	52.77	1.787	46.60	2.015
15	58.52	1.583	52.05	1.801
20	64.01	1.289	62.60	1.610
25	67.59	1.292	69.70	1.190
30	79.67	1.002	**82.45**	0.801

Table 2. Right whale classification results using ResNet

Iterations	Val. Acc. %	Val. Loss	Test_Acc. %	Test-Loss
5	46.92	2.0088	51.75	1.9444
10	71.18	1.1733	67.20	1.2858
15	82.20	0.8292	84.05	0.7557
20	86.88	0.6840	89.15	0.6000
25	89.23	0.6201	91.55	0.5595
30	90.62	0.5714	**92.15**	0.5434

The results indicate that the deep residual network has increased the recognition accuracy to the maximum of 92.15 reaching a saturation at the point irrelevant of layer increase. However, in future work, we will enhance the deep residual network design in order to enhance the accuracy to be closer to 100%.

2 Conclusion

In this research paper, a method of residual deep learning has been proposed using the ResNet. It has the ability to outperform the traditional CNN based approach involving many convolutional layers. The proposed methodology is implemented for learning discriminatory features of right whales and classifying them. The skip-connections used in ResNets address the issue of gradient vanishing that occurs while using the CNN with many layers. The proposed methodology has resulted in an accuracy of 92.15% for the testing set. The use of ResNets reveals true potential of utilizing deep learning in the classification of right whales in the real environment. As a future work, different deep learning architectures would be explored in order to enhance the performance. The deep learning residual network models will be utilized in other research areas, such as prediction, segmentation of diseases, palm trees detection and other very exciting projects.

References

[1]. Smith TD, Reeves RR, Josephson EA, Lund JN. 2012 Spatial and seasonal distribution of American whaling and whales in the age of sail. PLoS ONE 7, e34905. (doi:10.1371/journal. pone.0034905)

[2]. O'Connor S, Campbell R, Cortez H, Knowles T (2009) Whale Watching Worldwide: tourism numbers, expenditures and expanding economic benefits, a special report from the International Fund for Animal Welfare. Yarmouth MA, USA, prepared by Economists at Large 228.

[3]. Bowen WD (1997) Role of marine mammals in aquatic ecosystems. Mar. Ecol. Prog. Ser. 158, 267-274.

[4]. Li, Changyu, Jiayu He, Huang Lin, and Wenda Chen. "Humpback Whale Identification Challenge."

[5]. Smith LV, et al. (2013) Preliminary investigation into the stimulation of phytoplankton photophysiology and growth by whale faeces. J. Exp. Mar. Biol. Ecol. 446, 1-9.

[6]. Zerbini, Alexandre N., Grant Adams, John Best, Phillip J. Clapham, Jennifer A. Jackson, and Andre E. Punt. "Assessing the recovery of an Antarctic predator from historical exploitation." Royal Society Open Science 6, no. 10 (2019): 190368.

[7]. E. Kniest, D. Burns, and P. Harrison, Fluke matcher: A computer-aided matching system for humpback whale (megaptera novaeangliae) Flukes." Marine Mammal Sci., no. 3, p. 744, Dec. 2010.

[8]. Bishop, Christopher M. Pattern recognition and machine learning. Springer Science+ Business Media, 2006.

[9]. E. Shaikh, I. Mohiuddin, A. Manzoor, G. Latif, and N. Mohammad. "Automated Grading for Handwritten Answer Sheets using Convolutional Neural Networks." In 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS), pp. 1-6. IEEE, 2019.

[10]. D. A. Alghmgham, G. Latif, J. Alghazo, and L. Alzubaidi. "Autonomous Traffic Sign (ATSR) Detection and Recognition using Deep CNN." Procedia Computer Science 163 (2019): 266-274.

[11]. M. M. Butt, G. Latif, DNF A. Iskandar, J. Alghazo, and A. H. Khan. "Multi-channel Convolutions Neural Network Based Diabetic Retinopathy Detection from Fundus Images." Procedia Computer Science 163 (2019): 283-291.

[12]. Latif, G., Alghazo, J., Alzubaidi, L., Naseer, M. M., & Alghazo, Y. (2018, March). Deep Convolutional Neural Network for Recognition of Unified Multi-Language Handwritten Numerals. In 2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR) (pp. 90-95). IEEE.

[13]. Batbouta, Andrew. "Computer Assisted Labeling of Humpback Whales and Whale Sharks." PhD diss., Rensselaer Polytechnic Institute, 2017.

[14]. J. P. Crall, C. V. Stewart, T. Y. Berger-Wolf, D. I. Rubenstein and S. R. Sundaresan, "HotSpotter — Patterned species instance recognition," 2013 IEEE Workshop on Applications of Computer Vision (WACV), Tampa, FL, 2013, pp. 230-237.

[15]. H. J. Weideman et al., Integral curvature representation and matching algorithms for identification of dolphins and whales," in 2017 IEEE Int. Conf. on Computer Vision (ICCV), Dec 2017, pp. 2831{2839}.

[16]. Batbouta, Andrew. "Computer Assisted Labeling of Humpback Whales and Whale Sharks." PhD diss., Rensselaer Polytechnic Institute, 2017.

[17]. Liu, Jiongxin, Angjoo Kanazawa, David Jacobs, and Peter Belhumeur. "Dog breed classification using part localization." In European conference on computer vision, pp. 172-185. Springer, Berlin, Heidelberg, 2012.

[18]. Zhang, Ruoqi, Kai Wang, Yuekuan Luo, and Tianran Zhang. "The Identification of the Endangered Whale."

[19]. Bogucki, Robert, Marek Cygan, Christin Brangwynne Khan, Maciej Klimek, Jan Kanty Milczek, and Marcin Mucha. "Applying deep learning to right whale photo identification." Conservation Biology 33, no. 3 (2019): 676-684.

[20]. Lopez, Luis Fernando De Mingo, Clemencio Morales Lucas, Nuria Gómez Blas, and Krassimira Ivanova. "Pattern Recognition with Convolutional Neural Networks: Humpback Whale Tails." (2019).

[21]. [15] Dixit, Ujjawal, Apoorva Mishra, Anupam Shukla, and Ritu Tiwari. "Texture classification using convolutional neural network optimized with whale optimization algorithm." SN Applied Sciences 1, no. 6 (2019): 655.

[22]. Mirjalili, Seyedali, and Andrew Lewis. "The whale optimization algorithm." Advances in engineering software 95 (2016): 51-67.

[23]. Gillespie, Douglas. "Detection and classification of right whale calls using an'edge'detector operating on a smoothed spectrogram." Canadian Acoustics 32, no. 2 (2004): 39-47.

[24]. Schiller, Dominik, Tobias Huber, Florian Lingenfelser, Michael Dietz, Andreas Seiderer, and Elisabeth André. "Relevance-based feature masking: Improving neural network based whale classification through explainable artificial intelligence." In 20th Annual Conference of the International Speech Communication Association INTERSPEECH (2019, in press). 2019.

[25]. Bermant, Peter C., Michael M. Bronstein, Robert J. Wood, Shane Gero, and David F. Gruber. "Deep Machine Learning Techniques for the Detection and Classification of Sperm Whale Bioacoustics." Scientific reports 9, no. 1 (2019): 1-10.

[26]. Kandia, V., and Y. Stylianou. "Detection of sperm whale clicks based on the Teager–Kaiser energy operator." Applied Acoustics 67, no. 11-12 (2006): 1144-1163.

[27]. Guirado, Emilio, Siham Tabik, Marga L. Rivas, Domingo Alcaraz-Segura, and Francisco Herrera. "Automatic whale counting in satellite images with deep learning." bioRxiv (2018): 443671.

[28]. Salman, Ahmad, Ahsan Jalal, Faisal Shafait, Ajmal Mian, Mark Shortis, James Seager, and Euan Harvey. "Fish species classification in unconstrained underwater environments based on deep learning." Limnology and Oceanography: Methods 14, no. 9 (2016): 570-585.

[29]. Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." arXiv preprint arXiv:1712.04621 (2017).

[30]. Chen, Guang, Peng Sun, and Yi Shang. "Automatic Fish Classification System Using Deep Learning." In 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 24-29. IEEE, 2017.

[31]. Li, Xiu, Min Shang, Hongwei Qin, and Liansheng Chen. "Fast accurate fish detection and recognition of underwater images with fast r-cnn." In OCEANS 2015-MTS/IEEE Washington, pp. 1-5. IEEE, 2015.

[32]. Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," IEEE Trans. Geosci. Remote Sens., vol. 54, no. 10, pp. 6232-6251, 2016.

[33]. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.

[34]. Zhong, Zilong, Jonathan Li, Lingfei Ma, Han Jiang, and He Zhao. "Deep residual networks for hyperspectral image classification." In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1824-1827. IEEE, 2017.

[35]. NOAA Right Whales Dataset: https://www.kaggle.com/c/noaa-right-whale-recognition