Whale sound classification using deep neural networks

1 Description

Convolutional neural networks [1, 2, 3, 4, 5], recurrent neural networks [6, 7], and deep belief networks [8] can be applied to acoustic modeling and speech recognition. One interesting acoustic modeling application is to detect whale calls versus noise in recorded data.

2 Objectives

The goal of the project is to design a deep neural network (convolutional, recurrent, or deep belief net) that outputs the probability that a given sound file is a whale call. The dataset was used for one of the Kaggle challenges[1], as commented in class.

The student should: 1) Preprocess the data. 2) Design the network architecture and train it. 3) Validate the network.

As in other projects, a report should describe the characteristics of the design, implementation, and results. A Jupyter notebook should include calls to the implemented function that illustrate the way it works.

3 Suggestions

- See literature in previous applications of DNNs to acoustic data.
- Implementations can use any Python library that implements DNNs.

References

[1] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and Jürgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, page 1237. Barcelona, Spain, 2011.

[2] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 675–678. ACM, 2014.

[3] Yoon Kim. Convolutional neural networks for sentence classification. *CoRR*, abs/1408.5882, 2014.

[4] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[1] The data is available from `https://www.kaggle.com/c/the-icml-2013-whale-challenge-right-whale-redux/data`

[5] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[6] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[7] Tomas Mikolov, Martin Karafiát, Lukas Burget, Jan Cernockỳ, and Sanjeev Khudanpur. Recurrent neural network based language model. In *Proceedings of the 2010 Interspeech Conference*, volume 2, page 3, 2010.

[8] Abdel-rahman Mohamed, George E Dahl, and Geoffrey Hinton. Acoustic modeling using deep belief networks. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(1):14–22, 2012.