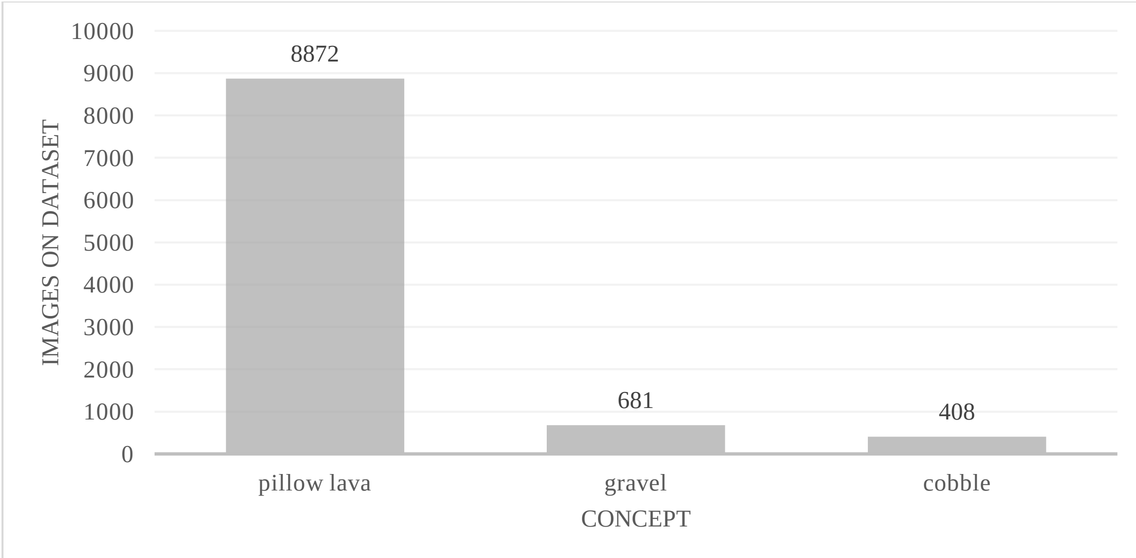

Figure 5. Geology concepts image distribution

The image quality across classes is non-uniform. As the VARS database is over thirty years old, the frame grabs added over time have been sourced from video tapes of varying quality, and thus the frame grabs stored on the MBARI file server are multi-resolution. On top of this, the *bobd-borders* application outputs a wide range of image resolutions when cropping the images. As a result, the images on the *FathomNet* dataset are considerably inconsistent in resolution per class, ranging from 640x364 pixels to 1920x1080 pixels.

4 DISCUSSION

4.1 QUANTITY AND QUALITY

There is significant non-uniformity in the image distribution across concepts in every category. This is a result of an imbalance in the concepts observed during underwater missions which have been annotated on the MBARI VARS database. Because of this skew, machine learning models trained on *FathomNet* could perform varying capacities in identifying different concepts. Though there is no global minimum number of images recommended for a training set, the average class within *FathomNet* contains 2029 images.

In its current state, *FathomNet* is best suited for transfer learning techniques. As transfer learning models train only a shallow portion of the network, they can apply a given pre-trained feature space to a different domain of interest. As a result, transfer learning models require far fewer images to learn a new domain of interest than needed to train a full network from scratch (Pan & Yang 2010).

Most of the popular pre-trained models for transfer learning require a smaller minimum resolution than the smallest-resolution image in the *FathomNet* dataset; for example, the *VGG16* and *ResNet* models require 224x224 pixel images as inputs (Simonyan & Zisserman 2014). Therefore, it can be concluded that all of the images on *FathomNet* are of sufficient spatial resolution for most of the current transfer learning models. It is important to note that the location and size of concepts within the images on *FathomNet* are currently unspecified. In addition, lighting conditions, pose, and orientation of concepts within the dataset are largely varied within each class. An example comparison for these differences can be seen in Fig. 6. Analysis beyond the scope of this paper would be necessary to determine the impact of these factors on different transfer learning models.

Resolution: 720 x 366
Box size: 32 x 38

Resolution: 1920 x 1080
Box size: 315 x 706

Figure 6. Size comparison between concepts in two images of *Bathochordaeus mcnutti*

4.2 DATASET COMPARISON

ImageNet is a large and well-known image dataset that contains over 14 million labeled images (*ImageNet* 2010). It is commonly used as training data in the machine learning community, particularly in conjunction with CNNs. Some of the most popular CNNs for transfer learning, such as *VGG16*, *InceptionV3*, and *ResNet*, have been pre-trained exclusively on the *ImageNet* dataset (Simonyan & Zisserman 2014). The average non-empty large concept, or *synset*, within *ImageNet* contains 650 images (*ImageNet* 2010), which is less than one third of the average image count per concept in *FathomNet*. On the other hand, *ImageNet* contains 21841 concepts (*ImageNet* 2010), compared to the 38 concepts currently distinguished by *FathomNet*.

5 CONCLUSIONS / RECOMMENDATIONS

The *FathomNet* dataset has been given a strong foundation through the process described by this paper. Though it is incomplete, every concept currently included in the dataset has a significant quantity of relatively high-resolution images. For these reasons, the *FathomNet* dataset is sufficient for training transfer learning models for image classification.

To improve the dataset, more images must be added. In its current state, *FathomNet* is built solely from individual frame grabs annotated in the MBARI VARS database. As MBARI's video assets contain over 23,000 hours of archived tapes, MBARI possesses much more data that could potentially be digitized and added to the *FathomNet* dataset (Katija et al. 2018). As mentioned above, only 26.4% of the annotations currently listed on VARS have been digitized and added to the dataset; the remaining 73.6% are lacking frame grabs. Currently, however, gathering these frames individually is a manual, time-consuming process; for each annotation, the tape must be located, fast-forwarded to the annotated timestamp, then frame grabbed and uploaded. A solution to digitize and add more data to *FathomNet* in an automated fashion should be explored further. Finally, additional imagery from other oceanographic institutions (e.g., NOAA, Ocean Exploration Trust, WHOI) could also further improve the size and variety of the dataset.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentors, Kakani Katija and Joost Daniels, for guiding me through this process and providing me with invaluable assistance. I would not have been able to make the progress that I have made on this project without their help and support.

I would especially like to thank George Matsumoto and Linda Kuhnz, the MBARI Summer Internship coordinators, for organizing the 2018 program and presenting the interns with countless opportunities throughout.

I would also like to thank Lonny Lundsten and Brian Schlining for mentoring me in the operation of the VARS database. Thanks to Brian Schlining for his development of the *bobd-watermark* application.

Finally, I would like to thank NSF-EAGER #1812535 (GEO-OCE-OTIC to K. Katija), the David and Lucile Packard Foundation, and the Monterey Bay Aquarium Research Institute for affording me the opportunity and resources to work on this project.

References

Gomes-Pereira J.N., V. Auger, K. Beisiegel, R. Benjamin, M. Bergmann, D. Bowden, P. Buhl-Mortensen, F.C. De Leo, G. Dionísio, J.M. Durden, L. Edwards, A. Friedman, J. Greinert, N. Jacobsen-Stout, S. Lerner, M. Leslie, T.W. Nattkemper, J.A. Sameoto, T. Schoening, R. Schouten, J. Seager, H. Singh, O. Soubigou, I. Tojeira, I. van den Beld, F. Dias, F. Tempera, R.S. Santos (2016). Current and future trends in marine image annotation software. *Progress in Oceanography*, 149: 106-120. Available at: https://doi.org/10.1016/j.pocean.2016.07.005

ImageNet. (2010). *Summary and Statistics*. [online] Available at: http://image-net.org/about-stats

Katija, K., G. Montague, G. Young, G. Flaspohler, B. Woodward, J. Gyllinsky, and K.C. Bell (2018). Big Ocean, Big Data. *Journal of Open Exploration*. Retrieved from https://explore.pubpub.org/pub/big-ocean-big-data.

Pan S.J. and Q. Yang (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22: 1345-1359. Available at: https://doi.org/10.1109/TKDE.2009.191

Simonyan K. and A. Zisserman (2014). Very Deep Convolution Networks for Large-Scale Image Recognition. Available at: https://arxiv.org/abs/1409.1556

Appendix A. VARS Concept Lists

Concept	Images	Type
Peniagone	11343	Benthic
Coryphaenoides	3212	Benthic
Keratoisis	2291	Benthic
Paragorgia	1802	Benthic
Sebastes	1758	Benthic
Bathyraja	1282	Benthic
Careproctus	520	Benthic
Anoplopoma	466	Benthic
Chionoecetes	1013	Benthic
Funiculina	465	Benthic
Merluccius	803	Benthic
Microstomus	477	Benthic
Pannychia	787	Benthic
Rathbunaster	775	Benthic
Sebastolobus	1601	Benthic
Strongylocentrotus	3358	Benthic
Umbellula	462	Benthic

Concept	Images	Type
pillow lava	8872	Geology
cobble	408	Geology
gravel	681	Geology

Concept	Images	Type
Bathochordaeus	6564	Midwater
Apolemia	4572	Midwater
Solmissus	2464	Midwater
Nanomia	2005	Midwater
Aegina	1954	Midwater
Lampocteis	2012	Midwater
Bathocyroe	1973	Midwater
Atolla	1808	Midwater
Tomopteris	1801	Midwater
Poralia	1676	Midwater
Eusergestes	981	Midwater
Erenna	1501	Midwater
Poeobius	1421	Midwater
Beroe	1257	Midwater
Llyria	1075	Midwater
Mitrocoma	596	Midwater
Benthocodon	545	Midwater
Paraphyllina	519	Midwater