

Perspective

Confronting Deep-Learning and Biodiversity Challenges for Automatic Video-Monitoring of Marine Ecosystems

Sébastien Villon *, Corina Iovan, Morgan Mangeas  and Laurent Vigliola

Institut de Recherche pour le Développement (IRD), UMR ENTROPIE (IRD, University of New-Caledonia, University of La Reunion, CNRS, Ifremer), 101 Promenade Roger Laroque, 98848 Noumea, France; corina.iovan@ird.fr (C.I.); morgan.mangeas@ird.fr (M.M.); laurent.vigliola@ird.fr (L.V.)

* Correspondence: sebastien.villon@ird.fr

Abstract: With the availability of low-cost and efficient digital cameras, ecologists can now survey the world's biodiversity through image sensors, especially in the previously rather inaccessible marine realm. However, the data rapidly accumulates, and ecologists face a data processing bottleneck. While computer vision has long been used as a tool to speed up image processing, it is only since the breakthrough of deep learning (DL) algorithms that the revolution in the automatic assessment of biodiversity by video recording can be considered. However, current applications of DL models to biodiversity monitoring do not consider some universal rules of biodiversity, especially rules on the distribution of species abundance, species rarity and ecosystem openness. Yet, these rules imply three issues for deep learning applications: the imbalance of long-tail datasets biases the training of DL models; scarce data greatly lessens the performances of DL models for classes with few data. Finally, the open-world issue implies that objects that are absent from the training dataset are incorrectly classified in the application dataset. Promising solutions to these issues are discussed, including data augmentation, data generation, cross-entropy modification, few-shot learning and open set recognition. At a time when biodiversity faces the immense challenges of climate change and the Anthropocene defaunation, stronger collaboration between computer scientists and ecologists is urgently needed to unlock the automatic monitoring of biodiversity.

Keywords: ecology; deep learning; artificial Intelligence; ecosystem monitoring



Citation: Villon, S.; Iovan, C.; Mangeas, M.; Vigliola, L. Confronting Deep-Learning and Biodiversity Challenges for Automatic Video-Monitoring of Marine Ecosystems. *Sensors* **2022**, *22*, 497. <https://doi.org/10.3390/s22020497>

Academic Editor: Jiachen Yang

Received: 29 November 2021

Accepted: 29 December 2021

Published: 10 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the age of climate change and anthropogenic defaunation [1,2] innovative methodology is needed to monitor ecosystems at large-spatial scales and high-temporal frequencies. Since the beginning of time, humans have learned from nature through visual observation, gradually using drawings, paintings, then photographs and videos. With the advent of low-cost digital cameras, modern ecologists can now visually gather visual data all around the planet, but they face a data processing bottleneck. While computer vision has long been used to speed up image processing, it is only since the emergence of deep learning (DL) algorithms that the revolution in the automatic assessment of biodiversity by video recording can be considered [3]. In fact, this revolution is underway, as shown by the exponential number of publications combining the words “biodiversity” and “deep learning” in a web of science bibliographic searches from 1975 to 2021, returning, on 28/10/2021, 175 publications, including zero before 2015, 12 from 2015 to 2017 and 118 since 2020.

The automation of video processing for biodiversity monitoring purposes is even more pressing in the oceans. Indeed, with 361 million km², the oceans cover 71% of our planet and monitoring their biodiversity requires an immense effort only achievable through automation. Furthermore, due to notorious difficulties in observing underwater biodiversity (e.g., divers are limited by bottom time and by depth), video surveys have been increasingly used for decades in many habitats, with some examples in shallow

reefs [4], sandy lagoons [5], deep seas [6] and in the pelagic ecosystem [7]. Global video surveys have even already been conducted. For instance, the FinPrint initiative had more than 15,000 video stations deployed in 58 countries in just three years for the first global assessment of the conservation status of reef sharks [8]. If the manual processing of such a large amount of videos was achievable for sharks at the cost of a massive labour effort, identifying and counting the abundance of thousands of other species in 15,000 video stations appears virtually impossible without automation. Unfortunately, the field of deep learning applied to marine biodiversity remains at a preliminary stage. A web of science bibliographic search combining the words “biodiversity”, “deep learning” and “marine” from 1975 to 2021 only returned 33 publications, including zero before 2016 and an average of 5.5 publications per year since then.

While DL algorithms have the potential to unclog the data processing bottleneck of video surveys, intrinsic characteristics of biodiversity, especially in the oceans, are in fact challenging this field of artificial intelligence, requiring special attention from computer scientists and strong collaboration with ecologists. Indeed, the current work of deep learning applications to automatically detect and classify animals in imagery is based on two premises: (1) an important database of each class of interest (thereafter “species”), and (2) a balanced dataset. These hypotheses are not verified for unconstrained wildlife video census. In particular, we point out three issues inherent to biodiversity video census, as well as the state of the art of DL approaches to answer them.

The aim of this work is (1) to point out ecological questions that can or cannot yet be tackled through DL applications by understanding its possibilities and limitations, and (2) to highlight recent advances of DL to unclog unconstrained wildlife video census.

2. Deep Learning for Biodiversity Monitoring

In recent years, a number of studies have examined the use of deep learning applied to ecological questions [9,10]. As shown in the most recent papers, the application of DL for species identification or detection relied on a limited number of species to process, as well as an important and balanced dataset. For instance, [11] discriminated 20 mammals species with an accuracy of 87.5% thanks to a dataset composed of 111,467 images; [12] detected and counted one species in videos with an accuracy of 93.4% with a dataset composed of 4020 images; [13] identified eight moth species with an F-measure (a common metric combining recall and precision) of 93% with a dataset of 1800 images, artificially augmented to 57,600 during the training of the model; [14] discriminated 109 plant species with an accuracy of 93.9% thanks to a dataset of 28,046 images. In the rare applications of deep learning on underwater videos [15], discriminated 20 coral reef fish species with an accuracy of 78% using a dataset of 69,169 images. Most proposals were adding information to DL models in order to reinforce the identification/detection, such as image enhancement, object tracking and class hierarchy. Historically, DL models were trained on benchmarks, composed of hundreds to thousands of images per class, and applied to a closed testing dataset [16]. A closed dataset is defined as classes in the testing dataset that are the same as classes in the training dataset. Thus, deep learning models trained for biodiversity monitoring still rely on closed, relatively large and balanced collection of images, following the framework developed in the field of computer science without considering intrinsic properties of biodiversity.

3. Biodiversity Rules and Deep Learning Limits

Species are no simple objects to classify. Their distribution and abundance follow a few universal rules that need to be accounted for in order to unlock the automatic assessment of biodiversity in underwater videos.

Since the early work of [17–19], ample evidence across world ecosystems show that in nearly every community in which species have been counted, the distribution of species abundance is highly skewed, such that a few species are very abundant, and many species are present in relatively low numbers. For deep learning applications, this first universal

rule of ecology implies heavily unbalanced training datasets, while balanced datasets are a crucial part of robust and accurate models. This issue, hereafter referenced as the “long-tail dataset issue”, is especially acute for speciose communities, such as coral reef fishes, where several hundred species, of which only a few dozen are abundant, can co-occur at a single site and at a single time in a video or another sampling station [20,21]. Intimately linked to the first rule, a second universal rule of ecology was proposed by [22] based on the early work of [23,24]. It states that species abundance is highest near the centre of their geographic range or environmental niche and then declines towards the boundaries. Thus, species tend to be scarce near the limits of their distribution. More generally, rarity is an intrinsic characteristic of biodiversity, with most communities composed of a large number of rare species. For deep learning, species rarity implies a lack of training images for a large part of species, where only a few or one individual can be seen in hundreds of hours of videos. This issue, hereafter referenced as the “scarce data issue”, is significantly marked in species-rich assemblages, such as coral reef fishes, where most species are demographically rare [20,21].

The third rule of ecology that seems relevant for deep learning stems from the openness of ecosystems [25,26]. The flow of energy, material, individuals and species across ecosystem boundaries is ubiquitous and plays a key role in ecosystem functioning. The degree of openness of marine ecosystems is particularly high because of the aquatic environment that facilitates the movement of species and the existence of most marine species of a planktonic larval stage favouring dispersal [27]. Ecosystem openness implies the issue of applying a deep learning model to an “open world problem”. According to [28], this issue is intrinsic to DL and is defined by a greater number of classes (in our case, species or conditions) in the application dataset than in the training dataset. In the context of biodiversity monitoring, the application dataset is composed of unconstrained recordings of wildlife and ecosystems, and due to ecosystem openness and the limits of sampling and annotation efforts, it cannot be considered as a closed-world application.

By comparing the current state-of-the-art of deep learning applications for biodiversity monitoring with some universal rules of biodiversity, we highlighted three problems inherent in such methods that remain to be solved in order to unlock the automatic assessment of biodiversity in underwater videos. We now discuss some potential solutions to these issues (Figure 1).



Figure 1. Ecological rules, their impacts on machine learning and state-of-the-art proposal to answer it.

4. Long-Tail Datasets

Long-tail datasets are problematic for deep learning model training. Classes with more samples in the training dataset have more impact on the final model. As a result, a model trained with an imbalanced dataset will have more success in predicting classes with more data, while predicting classes with fewer data will be hampered [29–31]. Furthermore, it was shown [32] that the degradation of predictions due to imbalanced datasets increases as the complexity of the task increases, which makes data imbalance of great impact in the case of complex ecosystems studies. Although a few studies suggested that training quality (e.g., sufficient data of all classes in datasets) can decrease the impact of imbalances [31,33], it is impossible to gather enough data for the large number of “rare” species composing marine ecosystems.

Two major ways of tackling this issue have emerged in the literature.

The first approach to address the long-tail dataset issue consists in balancing the dataset itself, for which the most popular methods are data augmentation or data generation. The technique of subsampling, which removes data from classes with large samples is not considered here because it wastes a large amount of useful information.

Data augmentation consists of artificially augmenting the number of images of classes with fewer data in order to increase their impact on the model training [34]. There are numerous methods to increase images datasets, such as resampling, geometric transformations, kernel filters (sharpness, colours and blurring) or feature space augmentation [35,36]. Apart from resampling, which simply uses the same image multiple times during the training phase, data augmentation consists in transforming existing images in the dataset and induces changes in order to mimic changing conditions and limit the overfit of DL models.

Data generation, through a generative adversarial network (GAN), variational autoencoder (VAE) or neural style transfer, is another way to increase a dataset’s size [35,37–41]. GANs are Deep Neural Networks (DNN), which, through learning thanks to existing data, are able to generate new images in the same representation space (i.e., images that “look like” those of the GAN training dataset). However, similarly to data augmentation, it is important to induce variations in the artificial dataset to prevent overfitting. Data can also be generated through visual engines, such as Unreal or Unity.

In the field of object detection, a few studies have directly cropped out objects of their original scenes and pasted them in new scenes [42,43]. Furthermore, a number of studies have simultaneously used both data augmentation and data generation [44].

The second approach to tackle the long-tail dataset issue consists in accounting for the imbalance of samples for each class in the training algorithms itself [45]. The focal loss, introduced by [46], adds to the Cross-Entropy (CE, a value used to improve deep model predictions during the training stage) two variables taking into account the ability of the networks to discriminate all classes and the proportion of each class in the dataset. In the same spirit, [47–49] propose to modify the CE with respect to dataset imbalances. Furthermore, [48] propose to control the classification space and margin between classes to boost the classification of classes with few data.

It has been shown that such methods can significantly improve the accuracy of DL model predictions [41].

5. Scarce Data

Deep learning models are efficient when a reasonable number of images per class is available during the training phase. However, ecosystems are composed of a large part of rare species, with only a few images in training datasets (e.g., the tail of species distribution).

The current popular proposal to unlock DL training with limited datasets is Few-Shot Learning (FSL), which builds algorithms able to discriminate classes with very few samples, classically from 1 to 20 images per class. Finn et al. [50] was a precursor of FSL and was based on a phase of meta-training, during which the model was trained on a different task at each iteration. To build a model able to discriminate five classes, at each training iteration, five classes were randomly selected from among the 64 possible training classes.

This phase enabled the model to “learn to learn” and to adapt itself to new tasks. Once the meta-training task was completed, the model could be adapted to a new task with very few images. More recently, [51,52] proposed to improve meta-learning by tweaking the meta-training or training batch compositions.

Apart from meta-training, matching network and metrics learning are other popular options [53–56]. Such methods aim to train a model to match a “Support Set” (i.e., a small training dataset composed of 1–5 images per class) with a “Query image” (i.e., the image to predict the label for). These approaches are two-fold: (1) they require a DL model able to manage the classification space with few images, and (2) they require the training of a robust metric to measure distances between the Query Image and the different Support Set clusters. More details on FSL can be found in the more exhaustive [57] review.

FSL is usually limited to 5–20 classes. However, recent papers are trying to overcome this limitation with approaches known as Many Classes Few-Shots [58,59] by leveraging a possible hierarchy among classes. This could be an opportunity for ecological applications, as there is a known hierarchy between species (i.e., taxonomy). Unfortunately, to date, most FSL algorithms are not able to discriminate more than 20 classes or with low accuracy [60]. Moreover, the study of methods for both imbalanced datasets and scarce data is for now limited to traditional benchmarks (such as imageNet [61] or mnist (<http://yann.lecun.com/exdb/mnist/>))(accessed on 1 June 2021).) without ecological questions. However, FSL’s application to underwater videos was first trialled in 2021 [62].

6. Open World Application

Applying a DL model of animal detection and/or species identification in the wild necessarily implies an “Open World” application. Indeed, DL algorithms, by nature, optimise a global function capable of discriminating several known classes of interest [16]. For species detection and identification, the algorithms also have to discriminate such classes from the background [63]. Unfortunately, it is not possible to predict the behaviour of a DL model when facing objects unseen during the training phase (e.g., new species, new morphologies, weather conditions, seascapes).

Approaches to tackling open world issues are regrouped under the Open Set Recognition (OSR) proposal, whether applied to Machine Learning [64–73] or Deep [67,69,74–81]. OSR is a growing field of research that has only been studied since the last decade.

Proposals on OSR rely on three principles that can be associated. First, managing the classification space in order to maximise inter-class margin and minimise intra-class spaces occupation. Such managing will create dense clusters of classes representation. Second, distances are chosen or learned through machine learning to evaluate new images with respect to learned clusters. Third, thresholds are selected or learned through machine learning to discriminate “known classes” from “new classes” with respect to the chosen distance. The overarching assumption is that images of new classes will be classified in the unused classification space, away from learned classes’ clusters. Applied to DL, most efficient methods rely on “OpenMax”, proposed by [74] and extended by [75,76], which replaces the usual classification layer of deep architecture known as “SoftMax”. The SoftMax function transforms the activation vector (i.e., the last feature vector of a Deep Network input) into a vector of n values, with n being the number of classes to discriminate. OpenMax adds a rejection function to SoftMax. This rejection function relies on the distance computed between previously learned data and the new input. The two main risks of OSR approaches are (1) depending on a training dataset to learn something that is not present in the training dataset and (2) potential overfit through the minimisation of potential classification space for learned classes. We also noted that most approaches are applied to image classification issues, and very few works cover object detection.

To date, there is only one research article [82] trying to resolve the three issues of imbalanced datasets, scarce data and the open world at the same time for a classification problem. Yet, in order to be robust to real life applications, DL needs to move from its initial

challenge of discriminating a limited number of classes with balanced and numerous data to a more realistic imbalanced, scarce and open data distribution.

7. Conclusions

Deep learning applications for biodiversity monitoring have been increasingly explored since 2017 [9] and are still in their early stages in marine realm applications. However, most studies rely on methods designed for and tested on generic benchmarks, which restrains the field of applications. To become an efficient tool for unconstrained wildlife census and conservation monitoring, collaborative research in computing science and ecology shall account for some of the universal rules of biodiversity. In this perspective, we highlighted methods from the state-of-the-art of artificial intelligence. Such methods have the potential to overcome the current limits of automatic video processing by focusing more thoroughly on the topics of imbalanced datasets, scarce data and open-world application. As such, efficient deep networks working with few data, such as few-shot learners and one-shot learners, improvement of the robustness to data imbalances through a specifically built learning process, and the ability to treat information absent from the training datasets with Open Set Recognition paves the way for an interdisciplinary branch of science between computer sciences and ecology. Rather than merely transferring the DL methods originally developed to perform on benchmarks to ecological questions, ecologists and computer scientists should foster their collaborations at the interface of both disciplines. As such, DL algorithms would become question-driven instead of adapted, which could leverage the immense challenges that biodiversity is facing with climate change and the Anthropocene defaunation. Conversely, ecologists may have a great interest in understanding the full potential offered by artificial intelligence techniques in order to develop new indicators that require too many human resources to operate until now or for lack of available data.

Author Contributions: Conceptualization, S.V. and L.V.; methodology, S.V. and L.V.; investigation, S.V. and L.V.; resources, S.V. and L.V.; original draft preparation, S.V. and L.V.; supervision: S.V. and L.V.; Writing, review and edition; S.V., C.I., M.M., L.V.; Funding acquisition, L.V.; Project administration, S.V. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the French National Research Agency project ANR 18-CE02-0016 SEAMOUNTS.

Acknowledgments: We want to thank our reviewers for their insights and their additions to our perspective. The authors want to thank Florence Bayard for the design and the realization of the main figure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dirzo, R.; Young, H.S.; Galetti, M.; Ceballos, G.; Isaac, N.J.B.; Collen, B. Defaunation in the Anthropocene. *Science* **2014**, *345*, 401–406. [[CrossRef](#)] [[PubMed](#)]
2. Young, H.S.; McCauley, D.J.; Galetti, M.; Dirzo, R. Patterns, Causes, and Consequences of Anthropocene Defaunation. *Annu. Rev. Ecol. Evol. Syst.* **2016**, *47*, 333–358. [[CrossRef](#)]
3. Lürig, M.D.; Donoughe, S.; Svensson, E.I.; Porto, A.; Tsuboi, M. Computer Vision, Machine Learning, and the Promise of Phenomics in Ecology and Evolutionary Biology. *Front. Ecol. Evol.* **2021**, *9*. [[CrossRef](#)]
4. Juhel, J.B.; Vigliola, L.; Wantiez, L.; Letessier, T.B.; Meeuwig, J.J.; Mouillot, D. Isolation and no-entry marine reserves mitigate anthropogenic impacts on grey reef shark behavior. *Sci. Rep.* **2019**, *91*, 1–11. [[CrossRef](#)] [[PubMed](#)]
5. Cappel, M.; De’ath, G.; Speare, P. Inter-reef vertebrate communities of the Great Barrier Reef Marine Park determined by baited remote underwater video stations. *Mar. Ecol. Prog. Ser.* **2007**, *350*, 209–221. [[CrossRef](#)]
6. Zintzen, M.J.; Anderson, C.D.; Roberts, E.S.; Harvey, E.S.; Stewart, A.L. Effects of latitude and depth on the beta diversity of New Zealand fish communities. *Sci. Rep.* **2017**, *7*, 1–10. [[CrossRef](#)] [[PubMed](#)]
7. Letessier, T.B.; Mouillot, D.; Bouchet, P.J.; Vigliola, L.; Fernandes, M.C.; Thompson, C.; Boussarie, G.; Turner, J.; Juhel, J.B.; Maire, E. Remote reefs and seamounts are the last refuges for marine predators across the Indo-Pacific. *PLoS Biol.* **2019**, *17*, 1–20. [[CrossRef](#)]
8. MacNeil, M.A.; Chapman, D.D.; Heupel, M.; Simpfendorfer, C.A.; Heithaus, H.; Meekan, M.; Harvey, E.; Goetze, J.; Kiszka, J.; Bond, M.E. Global status and conservation potential of reef sharks. *Nature* **2020**, *583*, 801–806. [[CrossRef](#)]
9. Christin, S.; Hervet, E.; Lecomte, N. Applications for deep learning in ecology. *Methods Ecol. Evol.* **2019**, 1632–1644. [[CrossRef](#)]

10. Weinstein, B.G. A computer vision for animal ecology. *J. Anim. Ecol.* **2017**, *87*, 533–545. [\[CrossRef\]](#)
11. Miao, Z.; Gaynor, K.M.; Wang, J.; Liu, Z.; Muellerklein, O.; Norouzzadeh, M.S.; Getz, W.M. Insights and approaches using deep learning to classify wildlife. *Sci. Rep.* **2019**, *9*, 1–9. [\[CrossRef\]](#)
12. Ditria, E.M.; Lopez-marcano, S.; Sievers, M.; Jinks, E.L.; Brown, C.J.; Connolly, R.M. Automating the Analysis of Fish Abundance Using Object Detection: Optimizing Animal Ecology With Deep Learning. *Front. Mar. Sci.* **2020**, *7*, 1–9. [\[CrossRef\]](#)
13. Bjerge, K.; Nielsen, J.B.; Sepstrup, M.V.; Helsing-Nielsen, F.; Høye, T.T. An automated light trap to monitor moths (Lepidoptera) using computer vision-based tracking and deep learning. *Sensors (Switzerland)* **2021**, *21*, 343. [\[CrossRef\]](#)
14. Hieu, N.V.; Hien, N.L.H. Automatic plant image identification of Vietnamese species using deep learning models. *arXiv* **2020**, arXiv:2005.02832. Available online: <https://arxiv.org/abs/2005.02832> (accessed on 10 December 2021).
15. Villon, S.; Mouillot, D.; Chaumont, M.; Darling, E.C.; Subsol, G.; Claverie, T.; Villéger, S. A Deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecol. Inform.* **2018**, *48*, 238–244. [\[CrossRef\]](#)
16. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
17. Gleason, H.A. The Significance of Raunkiaer's Law of Frequency. *Ecology* **1929**, *10*, 406–408. [\[CrossRef\]](#)
18. Fisher, R.A.; Corbet, A.S.; Williams, C.B. The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. *J. Anim. Ecol.* **1943**, *12*, 42. [\[CrossRef\]](#)
19. Preston, F.W. The Commonness, And Rarity, of Species. *Ecol. Soc. Am.* **1948**, *29*, 254–283. [\[CrossRef\]](#)
20. Hercos, A.P.; Sobansky, M.; Queiroz, H.L. Magurran AE. Local and regional rarity in a diverse tropical fish assemblage. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*. [\[CrossRef\]](#)
21. Jones, G.P.; Munday, P.L.; Caley, M.J. Rarity in Coral Reef Fish Communities. In *Coral Reef Fishes*; Sale, P.F., Ed.; Academic Press: Cambridge, MA, USA, 2002; pp. 81–101.
22. Brown, J.H. On the Relationship between Abundance and Distribution of Species. *Am. Nat.* **1984**, *124*, 255–279. [\[CrossRef\]](#)
23. Whittaker, R.H. Dominance and Diversity in Land Plant Communities. *Am. Assoc. Adv. Sci. Stable* **1965**, *147*, 250–260. [\[CrossRef\]](#)
24. Whittaker, R.H. Vegetation of the Great Smoky Mountains. *Ecol. Monogr.* **1956**, *26*, 1–80. [\[CrossRef\]](#)
25. Loreau, M.; Holt, R.D. Spatial flows and the regulation of ecosystems. *Am. Nat.* **2004**, *163*, 606–615. [\[CrossRef\]](#)
26. Holt, R.D.; Loreau, M. Biodiversity and Ecosystem Functioning: The Role of Trophic Interactions and the Importance of System Openness. In *The Functional Consequences of Biodiversity*; Princeton University Press: Princeton, NJ, USA, 2013; pp. 246–262.
27. Carr, M.H.; Neigel, J.E.; Estes, J.A.; Andelman, S.; Warner, R.R.; Largier, J.L. Comparing Marine and Terrestrial Ecosystems: Implications for the Design of Coastal Marine Reserves. *Ecol. Appl.* **2003**, *13*, 90–107. [\[CrossRef\]](#)
28. Scheirer, W.J.; de Rezende Rocha, A.; Sapkota, A.; Boulton, T.E. Toward Open Set Recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*; IEEE: New York, NY, USA, 2013. [\[CrossRef\]](#)
29. Vluymans, S. *Learning from Imbalanced Data In IEEE Transactions on Knowledge and Data Engineering*; IEEE: New York, NY, USA, 2009; pp. 1263–1284. [\[CrossRef\]](#)
30. Aggarwal, U.; Popescu, A.; Hudelot, C. Active learning for imbalanced datasets. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 1428–1437. [\[CrossRef\]](#)
31. Krawczyk, B. Learning from imbalanced data: Open challenges and future directions. *Prog. Artif. Intell.* **2016**, *5*, 221–232. [\[CrossRef\]](#)
32. Japkowicz, N. The Class Imbalance Problem: Significance and Strategies. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.35.1693&rep=rep1&type=pdf> (accessed on 10 December 2021).
33. Yates, K.L.; Bouchet, P.J.; Caley, M.J.; Mengersen, K.; Randin, C.F.; Parnell, S.; Fielding, A.H.; Bamford, A.J.; Ban, S.; Márcia, A.M. Outstanding Challenges in the Transferability of Ecological Models. *Trends Ecol. Evol.* **2018**, *33*, 790–802. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Van Dyk, D.A.; Meng, X.L. The Art of Data Augmentation. *J. Comput. Graph. Stat.* **2012**, *8600*, 1–50. [\[CrossRef\]](#)
35. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*. [\[CrossRef\]](#)
36. Wong, S.C.; McDonnell, M.D.; Adam, G.; Victor, S. Understanding data augmentation for classification: When to warp ? In Proceedings of the 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, Australia, 30 November–2 December 2016; pp. 1–6.
37. Mariani, G.; Scheidegger, F.; Istrate, R.; Bekas, C.; Malossi, C. BAGAN: Data Augmentation with Balancing GAN. *arXiv* **2018**, arXiv:1803.09655. Available online: <https://arxiv.org/abs/1803.09655> (accessed on 10 December 2021).
38. Bowles, C.; Chen, L.; Guerrero, R.; Bentley, P.; Gunn, R.; Hammers, A.; Dickie, D.A.; Hernández, M.V.; Wardlaw, J.; Rueckert, D. GAN Augmentation: Augmenting Training Data using Generative Adversarial Networks. *arXiv* **2018**, arXiv:1810.10863. Available online: <https://arxiv.org/abs/1810.10863> (accessed on 10 December 2021).
39. Frid-Adar, M.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. Synthetic data augmentation using GAN for improved liver lesion classification. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging, Washington, DC, USA, 4–7 April 2018; pp. 289–293.
40. Doersch, C. Tutorial on Variational Autoencoders. *arXiv* **2016**, arXiv:1606.05908. Available online: <http://arxiv.org/abs/1606.05908> (accessed on 10 December 2021).
41. Beery, S.; Liu, Y.; Morris, D.; Piavis, J.; Kapoor, A.; Joshi, N.; Meister, M.; Perona, P. Synthetic examples improve generalization for rare classes. In Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020, Snowmass, CO, USA, 2–5 March 2020; pp. 852–862. [\[CrossRef\]](#)

42. Allken, V.; Handegard, N.O.; Rosen, S.; Schreyeck, T.; Mahiout, T.; Malde, K. Fish species identification using a convolutional neural network trained on synthetic data. *ICES J. Mar. Sci.* **2019**, *76*, 342–349. [\[CrossRef\]](#)
43. Ekbatani, H.K.; Pujol, O.; Segui, S. Synthetic data generation for deep learning in counting pedestrians. In Proceedings of the ICPRAM 2017–6th International Conference on Pattern Recognition Applications and Methods, Porto, Portugal, 24–26 January 2017; pp. 318–323. [\[CrossRef\]](#)
44. Perez, L.; Wang, J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *arXiv* **2017**, arXiv:1712.04621. Available online: <https://arxiv.org/abs/1712.04621> (accessed on 10 December 2021).
45. Buda, M.; Maki, A.; Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* **2018**, *106*, 249–259. [\[CrossRef\]](#)
46. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
47. Cui, Y.; Jia, M.; Lin, T.; Tech, C. Class-Balanced Loss Based on Effective Number of Samples. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
48. Cao, K.; Wei, C.; Gaidon, A.; Arechiga, N.; Ma, T. Learning imbalanced datasets with label-distribution-aware margin loss. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 1–18.
49. Tan, J.R.; Wang, C.B.; Li, B.Y.; Li, Q.Q.; Ouyang, W.L.; Yin, C.Q.; Yan, J.J. Equalization loss for long-tailed object recognition. In Proceedings of the IEEE Comput Soc Conf Comput Vis Pattern Recognit, Seattle, WA, USA, 13–19 June 2020; pp. 11659–11668. [\[CrossRef\]](#)
50. Finn, C.; Abbeel, P.; Levine, S. Model-agnostic meta-learning for fast adaptation of deep networks. In Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, Australia, 6–11 August 2017; pp. 1856–1868.
51. Jamal, M.A.; Cloud, H. Task Agnostic Meta-Learning for Few-Shot Learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.
52. Sun, Q.; Chua, Y.L.T. Meta-Transfer Learning for Few-Shot Learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 403–412.
53. Li, H.; Eigen, D.; Dodge, S.; Zeiler, M.; Wang, X. Finding task-relevant features for few-shot learning by category traversal. *arXiv* **2019**, arXiv:1905.11116. Available online: <https://arxiv.org/abs/1905.11116> (accessed on 10 December 2021).
54. Sung, F.; Yang, Y.; Zhang, L. Learning to Compare: Relation Network for Few-Shot Learning Queen Mary University of London. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake, UT, USA, 18–23 June 2018; pp. 1199–1208.
55. Oreshkin, B.N.; Rodriguez, P.; Lacoste, A. Tadam: Task dependent adaptive metric for improved few-shot learning. *Adv. Neural Inf. Process. Syst.* **2018**, 721–731.
56. Zhang, X.; Hospedales, T. RelationNet2: Deep Comparison Columns for Few-Shot Learning. *arXiv Prepr.* **2017**, arXiv:181107100. Available online: <https://arxiv.org/abs/1811.07100> (accessed on 10 December 2021).
57. Wang, Y.; Yao, Q.; Ni, L.M. Generalizing from a Few Examples: A Survey on Few-shot. *ACM Comput. Surv.* **2020**, *53*. [\[CrossRef\]](#)
58. Li, A.; Luo, T.; Lu, Z.; Xiang, T.; Wang, L. Large-Scale Few-Shot Learning: Knowledge Transfer With Class Hierarchy. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 7212–7220.
59. Liu, L.; Zhou, T.; LONG, G.; Jiang, J.; Zhang, C. Many-Class Few-Shot Learning on Multi-Granularity Class Hierarchy. *arXiv* **2020**, arXiv:2006.15479. Available online: <https://arxiv.org/abs/2006.15479> (accessed on 10 December 2021). [\[CrossRef\]](#)
60. Wang, Y.; Yao, Q.; Kwok, J.T.; Ni, L.M. Generalizing from a Few Examples: A Survey on Few-Shot Learning. *arXiv* **2019**, arXiv:1904.05046. Available online: <https://arxiv.org/abs/1904.05046> (accessed on 10 December 2021). [\[CrossRef\]](#)
61. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, H.Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int J. Comput. Vis.* **2015**, *115*, 211–252. [\[CrossRef\]](#)
62. Villon, S.; Iovan, C.; Mangeas, M.; Claverie, T. Automatic underwater fish species classification with limited data using few-shot learning. *Ecol. Inform.* **2021**, *63*, 101320. [\[CrossRef\]](#)
63. Zhao, Z.; Zheng, P. *Object Detection with Deep Learning: A Review In IEEE Transactions on Neural Networks and Learning Systems*; IEEE: New York, NY, USA, 2019; pp. 3212–3232.
64. Scheirer, W.J. Probability Models for Open Set Recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*; IEEE: New York, NY, USA, 2014. [\[CrossRef\]](#)
65. Jain, L.P.; Scheirer, W.J.; Boulton, T.E. Multi-class Open Set Recognition Using Probability of Inclusion. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 393–409.
66. Zhang, H.; Member, S.; Patel, V.M.; Member, S. Sparse Representation-based Open Set Recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*; IEEE: New York, NY, USA, 2016; pp. 1–8.
67. Lonij, V.P.A.; Rawat, A.; Nicolae, M. Open-World Visual Recognition Using Knowledge Graphs. *arXiv* **2017**, arXiv:1708.08310. Available online: <https://arxiv.org/abs/1708.08310> (accessed on 10 December 2021).
68. Geng, C.; Huang, S.; Chen, S. Recent Advances in Open Set Recognition: A Survey. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*; IEEE: New York, NY, USA, 2020. [\[CrossRef\]](#)
69. Hassen, M.; Chan, P.K. Learning a Neural-network-based Representation for Open Set Recognition. In Proceedings of the 2020 SIAM International Conference on Data Mining, Cincinnati, OH, USA, 5–8 May 2020.

70. Parmar, J.; Chouhan, S.S.; Rathore, S.S. Open-world Machine Learning: Applications, Challenges, and Opportunities. *arXiv* **2021**, arXiv:2105.13448. Available online: <https://arxiv.org/abs/2105.13448> (accessed on 10 December 2021).
71. Song, L.; Sehwal, V.; Bhagoji, A.N.; Mittal, P. A Critical Evaluation of Open-World Machine Learning. *arXiv* **2020**, arXiv:2007.04391. Available online: <https://arxiv.org/abs/2007.04391> (accessed on 10 December 2021).
72. Leng, Q.; Ye, M.; Tian, Q. A Survey of Open-World Person Re-identification. In *IEEE Transactions on Circuits and Systems for Video Technology*; IEEE: New York, NY, USA, 2019; pp. 1092–1108. [[CrossRef](#)]
73. Mendes, P.R.; Roberto, J.; Rafael, M.D.S.; Werneck, R.D.O.; Stein, B.V.; Pazinato, D.V.; de Almeida, W.R.; Penatti, O.A.B.; Torres, R.D.S.; Rocha, A. Nearest neighbors distance ratio open-set classifier. *Mach. Learn.* **2017**, *106*, 359–386. [[CrossRef](#)]
74. Bendale, A.; Boulton, T.E. Towards open set deep networks. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1563–1572. [[CrossRef](#)]
75. Dhamija, A.R.; Günther, M.; Boulton, T.E. Reducing network agnostophobia. *arXiv* **2018**, arXiv:1811.04110. Available online: <https://arxiv.org/abs/1811.04110> (accessed on 10 December 2021).
76. Ge, Z.; Chen, Z. Generative OpenMax for Multi-Class Open Set Classification. *arXiv* **2017**, arXiv:1707.07418. Available online: <https://arxiv.org/abs/1707.07418> (accessed on 10 December 2021).
77. Rosa, R.D.; Mensink, T.; Caputo, B. Online Open World Recognition. *arXiv* **2016**, arXiv:1604.02275. Available online: <https://arxiv.org/abs/1604.02275> (accessed on 10 December 2021).
78. Shu, L.; Xu, H.; Liu, B. Unseen Class Discovery in Open-World Classification. *arXiv* **2018**, arXiv:1801.05609. Available online: <https://arxiv.org/abs/1801.05609> (accessed on 10 December 2021).
79. Oza, P.; Patel, V.M. Deep CNN-based Multi-task Learning for Open-Set Recognition. *arXiv* **2019**, arXiv:1903.03161. Available online: <https://arxiv.org/abs/1903.03161> (accessed on 10 December 2021).
80. Guo, X.; Chen, X.; Zeng, K. Multi-stage Deep Classifier Cascades for Open World Recognition. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 179–188.
81. Miller, D.; Niko, S.; Milford, M.; Dayoub, F. Class Anchor Clustering: A Loss for Distance-based Open Set Recognition. *arXiv* **2021**, arXiv:2004.02434. Available online: <https://arxiv.org/abs/2004.02434> (accessed on 10 December 2021).
82. Liu, Z.; Miao, Z.; Zhan, X.; Wang, J.; Gong, B.; Yu, S.X. Large-scale long-tailed recognition in an open world. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2532–2541. [[CrossRef](#)]