ICES Journal of
Marine Science

ICES Journal of Marine Science (2020), 77(4), 1274–1285. doi:10.1093/icesjms/fsz057

Contribution to the Themed Section: '*Applications of machine learning and artificial intelligence in marine science*'

Quo Vadimus

Machine intelligence and the data-driven future of marine science

Ketil Malde [1,2]*, Nils Olav Handegard[1], Line Eikvil[3], and Arnt-Børre Salberg[3]

[1]*Institute of Marine Research, Bergen, Norway*
[2]*Department of Informatics, University of Bergen, Norway*
[3]*Norwegian Computing Center, Oslo, Norway*

**Corresponding author: tel: +47 98691834; e-mail: ketil.malde@imr.no.*

Malde, K., Handegard, N. O., Eikvil, L., and Salberg, A-B. Machine intelligence and the data-driven future of marine science. – ICES Journal of Marine Science, 77: 1274–1285.

Received 12 September 2018; revised 7 February 2019; accepted 20 February 2019; advance access publication 15 April 2019.

Oceans constitute over 70% of the earth's surface, and the marine environment and ecosystems are central to many global challenges. Not only are the oceans an important source of food and other resources, but they also play a important roles in the earth's climate and provide crucial ecosystem services. To monitor the environment and ensure sustainable exploitation of marine resources, extensive data collection and analysis efforts form the backbone of management programmes on global, regional, or national levels. Technological advances in sensor technology, autonomous platforms, and information and communications technology now allow marine scientists to collect data in larger volumes than ever before. But our capacity for data analysis has not progressed comparably, and the growing discrepancy is becoming a major bottleneck for effective use of the available data, as well as an obstacle to scaling up data collection further. Recent years have seen rapid advances in the fields of artificial intelligence and machine learning, and in particular, so-called deep learning systems are now able to solve complex tasks that previously required human expertise. This technology is directly applicable to many important data analysis problems and it will provide tools that are needed to solve many complex challenges in marine science and resource management. Here we give a brief review of recent developments in deep learning, and highlight the many opportunities and challenges for effective adoption of this technology across the marine sciences.

Keywords: analysis bottleneck, convolutional neural networks, data processing, deep learning, observations, resource management.

Introduction

In March 2016, Google DeepMind pitched their computer programme AlphaGo (Silver *et al.*, 2016) against expert go player (ranked 9th dan) Lee Sedol in a five-game match, and won. This happened 20 years after IBM's chess playing computer *Deep Blue* famously played to a draw against grand master Gary Kasparov (Campbell *et al.*, 2002). Go is considered a notoriously difficult game for computers, and the event was widely reported in the press as an important milestone in the development of artificial intelligence (Wood, n.d.), and it was listed in Science as runner-up for the title of *Breakthrough of the Year* (Science, 2016).

Yet, this is only one of a series of remarkable achievements brought forth by recent developments in the field of artificial intelligence, and the triumph was soon overshadowed by new successes, for instance when AlphaZero managed to surpass human level skill in go, chess, and shogi solely from the experience it gathers playing against itself (Silver *et al.*, 2017).

Systems that are becoming increasingly intelligent are now being deployed on every scale, from mobile phones to supercomputers, and they are involved in a diversity of tasks, including personalized ranking of search results, selecting relevant advertisements, assisting vehicle driving, recognizing handwriting, and understanding spoken commands. Common to these successes is the application of a new approach called *deep learning* (LeCun *et al.*, 2015).

Many of the high-profile uses of deep learning originate from corporations like Google, Facebook, Microsoft, and Amazon. These are consumer-oriented, technology-driven companies with access to large data repositories and computing resources (three of the four run commercial cloud services). Interestingly, these companies are also on the forefront of academic research, Google lists (Google, 2018) close to 1500 research papers on machine intelligence, perception, and translation, and another 380 on natural language processing. Microsoft reports publishing 239 papers on artificial intelligence in 2017 alone (Microsoft, 2018).

Technological progress has made data collection less costly, and this also affects the marine sciences. Large infrastructure projects are being developed to store and organize the data, and analysis is increasingly becoming a bottleneck. To meet many of the global challenges in marine science and management, it is necessary to realize the potential of collected data through automating more of the analysis. Here, we explore how new analysis technologies can be exploited to meet this goal.

Navigating an ocean of data

More than two-thirds of the planet is covered by oceans. The marine environment is a key component of the earth's climate, and its diverse ecosystems provide about half the global biological production and essential ecosystem services. The UN sustainability goals 2 (food security) and 3 (health) indirectly address the ocean, whereas goal 14 (use of oceans) explicitly acknowledges the need for sustainable development for the oceans and seas.

Marine science must rise to these challenges and provide the knowledge needed to ensure sustainable use of the marine environment. The necessity of an ecosystems approach to marine management is accepted worldwide (Pikitch *et al.*, 2004; Bianchi and Skjoldal, 2008; Koslow, 2009; Link and Browman, 2014) and is reflected in the [revised] European common fisheries policy and the marine strategy framework directive. Further development of models and observing systems is needed to meet these requirements, and a key challenge is how to extract relevant information when data volumes increases, data complexity increases, and data quality varies.

Increased data volumes

A direct consequence of improvements in sensor technology is an increase in data volume, usually accompanied by lower cost. This is brought about by several factors: higher data rates, decrease cost of sensor equipment, and for sensors operating *in situ*, advances in autonomous platforms technologies. New or upgraded sensors now allow us to observe essential ocean variables as well as other biological data, in both the field and the laboratory, at scales that were earlier beyond our ability. A few cases serve to illustrate this.

Acoustics is the primary sensor on acoustic–trawl surveys (MacLennan and Simmonds, 2005), and calibrated high-quality echo sounders are mounted on research vessels. These are now commonly installed on a wider range of platforms including vessels of opportunity, e.g. fishing vessels (Honkalehto *et al.*, 2011; Fassler *et al.*, 2016) and autonomous platforms, e.g. autonomous underwater vehicles (Fernandes *et al.*, 2003), gliders (Guihen*et al.*, 2014), observatories (Godø *et al.*, 2014), and autonomous surface vehicles (Mordy *et al.*, 2017). In concert these sensors could form an observation system that can inform ecosystem models (Handegard *et al.*, 2013), but the traditional manual data processing is a major bottleneck.

Research projects now routinely sequence the full genomes (Berthelot *et al.*, 2014; Lien *et al.*, 2016) or transcriptomes of tens or hundreds of individuals (Schunter *et al.*, 2014), resulting in several terabytes of data. Since the landmark Human Genome Project (Venter *et al.*, 2001), sequence costs have plummeted six orders of magnitude, and molecular methods are now used in new contexts like sequencing of marine communities to reveal its species composition or functional diversity (metagenomics) (Jackson *et al.*, 2015; Kodzius and Gojobori, 2015), or using genomic methods to investigate population structure, evolution, and migration patterns (Larson *et al.*, 2014; Malde *et al.*, 2017).

Camera equipment has become more advanced, robust, and inexpensive. Still and moving images are now used in a wide range of applications, including baited video surveys (Cappo *et al.*, 2007), benthic monitoring (Buhl-Mortensen *et al.*, 2015), in-trawl monitoring (Rosen *et al.*, 2013), plankton imaging (Stemmann and Boss, 2012). Processing the resulting wealth of image data still often requires manual or partially manual labelling to extract meaningful information. In some cases, training data can be simulated (Figure 1), but often the lack of good training data hampers exploitation of technological advances and limits mass deployment of cameras.

Increased data complexity

Besides increased data quantity, new methods, and technology often let us collect and derive increasingly more complex data and information. This is true for model outputs and observations alike, and combining and analysing complex data are challenging

Figure 1. Simulated image mimicking output from the Deep Vision trawl camera solution. The simulator produces infinite training data for a classifier by producing random collages of fish images pasted onto an empty background. Image courtesy of Thomas Mahiout and Tiffanie Schreyeck.

since the relationships are often non-linear. Like for data quantity, the increased complexity applies almost universally, and a few cases are presented for illustration.

Early echo sounders emitted a single frequency, and received an intensity representing the reflected signal, conveniently plotted in a 2D diagram with time and depth (Sund, 1935). Multi-frequency equipment emits several frequencies simultaneously, and the difference in signal response provides valuable information about parameters like fish species, sizes, and orientations (Kloser *et al.*, 2002; Korneliussen and Ona, 2003). But the multiple diagrams are more demanding to interpret. Broadband equipment (Stanton *et al.*, 2010) replaces the multiple frequencies with continuous spectra, adding further complexity. Methods that can deal with these data have the potential to increase the information we get from the observations.

Similarly, most cameras capture visible light in the three primary colours corresponding to the photoreceptors in the human eye. In many cases, information is conveyed outside this spectrum, as evidenced by species like the mantis shrimps (*Stomatopoda* spp.), whose eyes have 16 different photoreceptors and the ability to detect both ultraviolet and polarized light (Marshall and Oberwinkler, 1999). Hyperspectral or multispectral photography that can record images both within and beyond the visible spectrum are likely to be useful in many settings, since light absorption and reflection of many substances strongly depend on the wavelength. For instance, the "colour" of the ocean is determined by the interactions of incident light with substances or particles present in the water. By exploiting multispectral data with fine spectral resolution several services provide frequent updates of a wide range of products based on the ocean colour (NASA, 2018). Methods to further exploit the increased data complexity are needed.

End to end ecosystem models have been proposed to be a key tool in integrated fisheries assessments (Fulton *et al.*, 2014). These models include components from physical forcing, geochemistry, primary production, and higher trophic levels, and the resulting model framework and model states are complex. Methods to extract relevant information, and often combining information from several sources are required, e.g. through ensemble modelling (Olsen *et al.*, 2016) or combining information from different data types. The state space from these models can be considered a complex data set and analysed as such. Methods to be able to find patterns and signals in the model states are needed.

Data quality

Improved technology generally leads to higher quality data, but occasionally increased data volumes are obtained by trading off quality for quantity. An example of this is research vessel surveys, which are costly to scale up. An alternative could be to collect data from the commercial fishing fleet, but with loss of rigid sampling design employed on research vessel surveys (Fassler *et al.*, 2016). Alternatively, relatively simple autonomous platforms could collect acoustics data, but without trawl sampling that has key information on age structure and species composition. Similarly, ARGO floats (Roemmich *et al.*, 2009) collect oceanographic data at a fraction of the cost of surveys using research vessels, but they can only drift with ocean currents, and we lose the ability to actively set up sampling designs or collect water samples. The information from increased data quantities may compensate for a loss of quality, but the lack of rigid designs will often introduce biases which pose new challenges for analysis.

While the cases we highlight here exemplify the growing data volumes, increasing data complexity, and deteriorating data quality, they are not exhaustive. Rather they demonstrate how analysis increasingly is becoming a bottleneck for effective use of collected data across diverse fields and technologies. Relying on manual scrutiny by human experts does not scale well, and automatic analysis of data is necessary to alleviate a rapidly narrowing analysis bottleneck.

The deep learning revolution

Machine learning at a glance

A classical computer programme is an executable expression of an algorithm. That is, the programmer formulates a precise stepwise description of how to produce the desired result from the input. In contrast, a machine learning programme requires the programmer to specify only a more general model or architecture for the solution. The model is then *trained* using available data. Typically, training consists of gradually adjusting the parameters of the model, causing the programme to produce increasingly accurate results. By definition, a machine learning programme is a programme that is able to improve its performance from experience (Mitchell, 1997).

In principle, statistical methods like linear regression and estimation of probability distributions can be considered machine learning methods, but here we use the term to refer to more complex systems, like artificial neural networks, random forests, and support vector machines. And in contrast to statistical methods where the parameters are inherently meaningful, the parameters of more complex machine learning systems often capture some general pattern in the data in an opaque way, and the interpretation of the individual parameters can be difficult.

Neural networks

One of the archetypal machine learning systems, and a cornerstone of the recent revolution in machine learning, is the artificial neural network (Parker, 1985; Rumelhart *et al.*, 1986). It is conceptually simple, yet can solve complex problems, in fact, by the *universal approximation theorem* any function can be modelled by a neural network (Cybenko, 1989; Hornik *et al.*, 1989). A neural network consists of layers of simple computational units (or *neurons*), arranged so that the output of the units in one layer feed into the inputs of the next layer's units (Figure 2). Each unit calculates a weighted sum of it inputs, and applies a function (the *activation function*), $f(\cdot)$, that introduces non-linearity into the system. The weights, w_{ij}, of the inputs to each unit constitute the parameters to be learned. This is usually achieved using *back propagation* (Rumelhart *et al.*, 1986) to calculate the gradient for a cost function, which is then minimized iteratively using some variant of gradient descent.

Deep learning and the renaissance of neural networks

Work on neural networks in the 1980s and 1990s (Parker, 1985; Rumelhart *et al.*, 1986) was limited by computational power, lack of sufficiently large labelled data sets for training, and limitations in the learning algorithms. Hence, the dominant approach to machine learning was to use application dependent hand-designed features to describe the data in a compact form, reducing its dimensionality. For instance, computer vision would typically

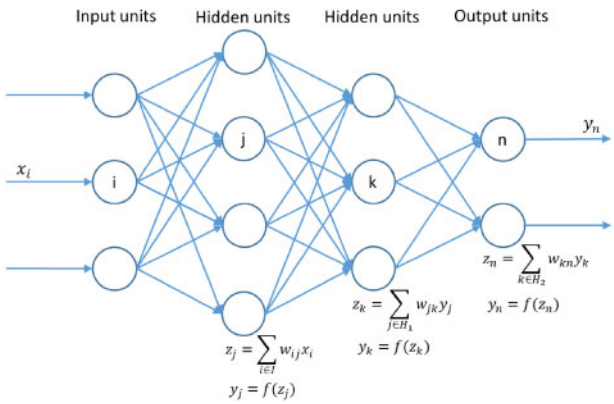

Figure 2. An artificial neural network typically consists of one input layer, several hidden layers, and one output layer. Each unit calculates a weighted sum of the inputs, and applies an activation function, (*f*). For simplicity, we have omitted bias terms.

preprocess input images with a manually designed programme to detect features like edges and corners (Lowe, 2004; Dalal and Triggs, 2005). Classification algorithms like decision trees, shallow neural networks, and support vector machines (Boser *et al.*, 1992) would then be applied to learn patterns from the features, rather than from the raw image data. The input features would typically be hand-crafted to each problem, and standard feature sets like scale-invariant feature transform (Lowe, 1999) or histogram of oriented gradients (Dalal and Triggs, 2005) were developed to be reusable and applicable to many image classification problems. Although these approaches were successful for many applications, there is a necessary trade-off between generality and the specific task at hand, and reusable but static features cannot capture the inherent complexity of many objects, nor translate easily to non-image or higher-dimensional data.

In recent years, the availability of computational power from the use of graphics processing units (GPUs) (Chellapilla *et al.*, 2006; Bergstra *et al.*, 2010) and distributed computing (Dean *et al.*, 2012), large annotated data sets like ImageNet (Russakovsky *et al.*, 2015) as well as algorithmic improvements (Nair and Hinton, 2010; Hinton, Srivastava, *et al.*, 2012; Ioffe and Szegedy, 2015; He *et al.*, 2016) has allowed the construction of much larger and deeper neural networks than before. The added complexity allows a network to learn relevant features in the data automatically, which is a defining element of *deep learning* (LeCun *et al.*, 2015; Schmidhuber, 2015). For instance in computer vision, the lower layers in the network learn to recognize primitive, general features like edges and corners in an image. Higher layers learn to identify more abstract features as combinations of features (e.g. object parts formed by primitive features). Finally, the highest layers learn to identify abstract classes as combinations of object parts. This hierarchical structure of the deep convolutional neural networks (CNNs) thus naturally models the hierarchical composition of the objects to be recognized.

In contrast to feature-specific machine learning, where generality of features is at odds with specificity to the problem, deep learning is a generalized *approach* for developing the solution simultaneously with the problem-specific features. Neural network architectures still benefit when tailored to specific data types and problems, but the ability of deep networks to learn the primitive features directly from the raw data makes the technology directly applicable to a wide range of problems.

CNNs and computer vision

CNNs (Fukushima, 1988; LeCun *et al.*, 1999) are structured as stacks of filters, each recognizing increasingly abstract features in the data. This approach is very effective for many image analysis problems, where objects are often recognized independent of their location. The convolutional network applies the same set of filters to all parts of the image, recognizing the same kinds of features regardless of their position. This leads to a dramatic reduction in the number of weights and consequently a reduction in training effort and data requirement.

In 2012, Krizhevsky *et al.* (2012) demonstrated that deep convolutional networks could obtain substantially higher image classification accuracy on the ImageNet Large Visual Recognition Challenge (ILSVRC) (Russakovsky *et al.*, 2015) than competing systems. Their success was a result of designing a deep CNN and training it using new and more efficient strategies, including rectifying non-linearities (ReLUs) (Nair and Hinton, 2010; He *et al.*, 2015; Xu *et al.*, 2015), dropout regularization (Srivastava *et al.*, 2014), and batch normalization (Ioffe and Szegedy, 2015). To train a CNN with performance metrics comparable to the ones reported by Krizhevsky *et al.* (2012), a substantial amount of labelled training images is needed, in addition to sufficient computational power (e.g. parallel computers or GPU accelerators).

The great improvements demonstrated by Krizhevsky *et al.* (2012) were followed by a sequence of increasingly successful ILSVRC contestants using deep neural networks (Zeiler and Fergus, 2014; Badrinarayanan *et al.*, 2015; Long *et al.*, 2014; Yu and Koltun, 2015; He *et al.*, 2016), and have placed image recognition tasks at the centre of an ongoing deep learning revolution. Similar techniques have been extended to object localization by identifying their coordinates and bounding boxes (Ren *et al.*, 2015; Redmon *et al.*, 2016). Related tasks are semantic segmentation, where individual pixels are mapped to classes representing different objects (Badrinarayanan *et al.*, 2015; Long *et al.*, 2014; Yu and Koltun, 2015; Chen *et al.*, 2018), and instance segmentation, where each instance of an object is identified in addition to being segmented (He *et al.*, 2017).

These challenges are important in their own right, and also pave the way towards complete scene understanding, a core computer vision problem that is important for a number of applications, including autonomous driving, human–machine interaction (Baccouche *et al.*, 2011), earth observation (Kampffmeyer *et al.*, 2016; Maggiori *et al.*, 2017), image search engines (Wan *et al.*, 2014), to name a few.

Beyond images

In many cases, machines exceed human level accuracy, e.g. for optical character recognition (Goodfellow *et al.*, 2013), face verification (Taigman *et al.*, 2014), and recognition of specialized object categories, like different breeds of dogs or species of birds (Xiao *et al.*, 2014). Even text obfuscated for the specific purpose of distinguishing humans from computers (so-called *captchas*) are ironically deciphered more accurately by computers than by humans (Goodfellow *et al.*, 2013). Deep learning has led to rapid advances in many other areas beside computer vision, and it has successfully been applied to problems like speech recognition (Hinton, Deng, *et al.*, 2012), machine translation (Sutskever *et al.*, 2014; Zhang *et al.*, 2015), and financial applications (Heaton *et al.*, 2017). The technology is starting to be applied to data analysis in many sciences, including high energy physics

(Baldi *et al.*, 2014), drug activity prediction (Dahl *et al.*, 2014), and visual processing of microscope data to reconstruct 3D models of brain tissue.

Similar methods can also be applied to unsupervised learning, where data are unlabelled. One of the most popular unsupervised deep learning approaches is the autoencoder (Hinton and Salakhutdinov, 2006), which typically learns a representation (encoding) of the data, by training the network to ignore signal "noise" (Vincent *et al.*, 2010). Another promising direction for clustering is to learn representations and simultaneously discover cluster structure in unlabelled data by optimizing a discriminative loss function. The deep embedded clustering (DEC) (Xie *et al.*, 2016) approach represents, to the best of our knowledge, the state-of-the-art. The DEC is based on an optimization strategy in which a neural network is pre-trained by means of an autoencoder and then fine-tuned by jointly optimizing cluster centroids in output space and the underlying feature representation. The restricted Boltzmann machine (RBM) is a type of neural network that may be used to model probability distributions. RBM algorithms are used in several applications, including dimensionality reduction (Hinton and Salakhutdinov, 2006), clustering (Beyan *et al.*, 2018), and collaborative filtering (Salakhutdinov *et al.*, 2007). Within marine science, Beyan *et al.* (2017) proposed an effective outlier detection algorithm that is based on cluster cardinality. Here, clusters were obtained applying a mean-covariance RBM to group the data such that data points in the same group are more similar to each other than to those in other groups.

Neural networks are often referred to as *feed* forward network, as each layer forwards its output to the inputs of the next. In contrast, recurrent neural networks (RNNs) (Pineda, 1987) incorporate one or more backward links, forming a cyclic architecture. This allows the network to retain information about previous states, and RNNs are therefore often applied to time series data. RNNs have been used successfully to model language (Hochreiter and Schmidhuber, 1997). Deeper models and special memory units called long short-term memory have allowed RNNs to achieve state of the art performance in e.g. speech recognition (Graves *et al.*, 2013).

Machine learning in marine science

The growing data volumes, increased data complexity, and reduced data quality pose challenges for the marine science discipline, but at the same time recent advances in machine learning offer new possibilities of addressing them. Systems for automatic data analysis can be considered on several levels, from making manual work more efficient to novel analyses of complex and heterogeneous data.

Emulating basic human expertise

Machine learning systems are typically trained to emulate human curation, and thus a natural application is to use such systems to replace labor intensive steps in existing analysis pipelines. Reliance of manual curation is currently limiting effective data use, and automatic systems can reduce cost or increase throughput, for instance identifying fish species from images (Allken *et al.*, 2019; Siddiqui *et al.*, 2018; Villon *et al.*, 2018), fish trajectory estimation (Beyan *et al.*, 2018), or automatic age reading of otoliths (Moen *et al.*, 2018). The latter is perhaps of particular interest, as it demonstrates that a deep learning can obtain an accuracy comparable to human curators. This is in contrast to Fisher and Hunter (2018), who reviewed traditional, shallow machine learning approaches and concluded that they provided no substantial advantage over human curation.

A fully automated system with accuracy comparable to a human curator is ideal, but more limited systems have also merit. The ability to sort out irrelevant data (e.g. frames with no objects of interest in them) can reduce manual work by orders of magnitude, and rudimentary classifiers with limited accuracy can reduce it further. As a bonus, with an automatic system taking care of tedious routine and trivial cases, the curation work remaining for the human expert is likely to be more interesting and rewarding.

In many cases, less than perfect accuracy may be sufficient. For instance, in cases where the sampling variance is large, a small bias may be acceptable if a larger number of observations can be exploited. Analysis of plankton images often have many and variable categories and be confounded by detritus and variation in visibility and lightning conditions, and machine learning methods are often used to guide or assist the human curator (Uusitalo *et al.*, 2016). Furthermore, where judgement of human experts varies, automated systems are consistent and can be duplicated as needed. They are likely to be cheaper and easier to deploy in hostile conditions. And although initial systems may have an unsatisfactory accuracy, technology improves over time. With improved systems, data can be reanalysed with little effort.

Advancing beyond the human expert

In many cases, overwhelming data volumes means that automatic systems are necessary for analysis. But for an increasing number of tasks, machine learning systems can surpass human experts in quality as well as quantity.

Some tasks that can be solved in principle are still too complex in practice, even for human experts. Analysis can be elusive when systems consist of many different factors which interact in many different ways, ecosystems being a typical example. We may have knowledge of each species involved, their migratory behaviour, predators, and prey relationships, reproductive biology, and so on, and a species can be isolated in the lab and its behaviour and responses studied. However, aggregating this information and deriving the behaviour of complex systems in the wild is challenging. Instead, we often rely on complex ecosystem models based on assumed interactions between the various components, and make inferences about the system from the model results (Fulton *et al.*, 2003). This assumes that we have successfully included the key processes in our model and that we have correctly parameterized them. A common critique is that we rely too much on the assumptions (Planque, 2016). Another, more parsimonious, approach is to use conventional statistical models to fit the data, but these models may be too simplistic since non-linear effects are difficult to handle. The deep learning approach may offer a third approach, where the analysis is still based on observed data, but the system is more capable detect and model non-linearities. However, it is prudent to note that the information that we can extract from the data is limited by the information content in the first place. Even so, deep learning methods may be able to tease out patterns the other methods fail to do.

Gaining new scientific insights

A common criticism of many machine learning methods is that the resulting model is opaque: although it can be shown empirically to work, it is often not clear *how* the model works, or what knowledge the model captures. For instance, the learned parameters of a linear regression have clear interpretations as slope and intercept. In contrast, the individual weights in a trained neural network do not carry any obvious meaning and can have very different significance for different inputs. This is analogous to human knowledge. As observed by Polanyi (2009), many tasks require knowledge that we are unable to express explicitly. For instance, we can recognize a face instantly, yet we are at a loss for describing the exact process of doing so. In science the goal is often to *understand* a phenomenon. This is often achieved by exploring model dynamics, but is less transparent in typical deep learning models.

Despite this opacity, it is nevertheless possible to get a glimpse of the knowledge embedded in a machine learning system. For instance, convolutional layers in deep neural networks often recognize specific features of the input. By identifying regions of the data (parts of an image, say) where specific neurons are triggered, we can observe the feature recognized by that neuron (Montavon *et al.*, 2018). Such an approach could for instance reveal whether a system of automatic otolith reading (Moen *et al.*, 2018) is counting rings, or whether it is using other geometric features, like shape or size, and to what extent each feature is informative.

A slightly different method consists of feeding the network noise, and then using a variant of back propagation to amplify elements of the input data that cause a particular classification result (Erhan *et al.*, 2009). Several variations of this method have been developed (Bach *et al.*, 2015; Yosinski *et al.*, 2015), producing synthetic images that illustrate the type of features used by the network to identify a certain class. While recognizable, the resulting image is not necessarily representative for actual data.

Reproducibility of science and improved processes

Marine science and management advice for marine resources go hand in hand. A data processing pipeline for management, starting with data collection, going through various analyses and simulations, and ending with stock forecasts and management advice, are central to many marine science institutions. Currently, this process contains several interpretation steps, where a human expert must examine data to extract information for use as input to subsequent steps.

Automating these interpretation steps gives us several advantages. First, the whole process becomes deterministic and reproducible. Verifying the model output from the input data can be done by simply rerunning the pipeline, and this helps build confidence in the results. More importantly, it lets researchers experiment with the model, adjusting its parameters and inputs to discover how they affect the output, and let us quantify the consequences of changes. For instance, one can estimate the effect of reducing cruise activities in favour of less expensive floats or autonomous stations, or whether deployment of more advanced equipment is justifiable. This knowledge will be important for optimizing resource usage and reduce uncertainty in the results.

Heterogeneous data and integrative analysis

Ecosystems are complex networks of biological, chemical, and physical factors which also includes human activities. It is unclear to what extent such systems can be understood from a reductionist approach of isolating and studying each component. That a more holistic approach is necessary is a key tenet of transdisciplinary science (Nicolescu, 2008). But multi- and interdisciplinary approaches could also benefit marine science to a larger extent. For instance, molecular methods could complement traditional surveys for detecting the presence of species (Foote *et al.*, 2012; Thomsen *et al.*, 2012), cameras can detect fragile species that are destroyed by more intrusive methods (Remsen *et al.*, 2004), and autonomous platforms (Mordy *et al.*, 2017) could augment data from more traditional surveys. Integrative approaches could collect data from multiple databases representing a variety of collection regimes and scientific disciplines and reanalyse these data in new ways to derive new information. Making data interoperable is a key step for effective integrative analysis, and several large efforts aim at providing centralized infrastructures and standardized organization for data collected by third parties.

An advantage of machine learning methods is their ability to work well with ambiguous data. Deep learning methods incorporate multiple levels of representation (LeCun *et al.*, 2015). Lower layers learn less abstract representations of the input, and these methods can therefore be applied directly to data without preprocessing (e.g. to images represented as pixel arrays, or free-form text), identifying and extract salient features automatically. In contrast to shallower systems which depend on hand-crafted features, the relevant structure and information content in the data is captured implicitly by the model. This has allowed e.g. natural language processing systems using deep learning methods to deal with ambiguities and imprecision in human languages. This robustness is not limited to language, and allows us to construct compound systems with the ability to deal usefully with existing data that may be incomplete, inconsistent, ambiguous, and weakly structured (Raghupathi and Raghupathi, 2014).

Of course, deep learning systems can also be applied to more abstract features, or to combinations of features and raw data. For instance, a popular task in computer vision is automatic image caption generation, where image features (extracted directly from raw data) is fed into a RNN that generates the appropriate caption describing the image content (Vinyals *et al.*, 2014)

Challenges

To realize the potential of automatic analysis, we need effective methods capable of handling the large amounts of data generated. Although successful projects that apply deep learning in the marine sciences exist (ICES, 2018), the technology has not yet seen widespread deployment, and several obstacles must be overcome for successful development and implementation.

Data availability in a form suitable for analysis

One obstacle is the lack of large and well-structured data sets suitable for training machine learning models. There is considerable third party interest in machine learning, and online competitions like (Kaggle, 2018) show that the availability of clearly defined problems and curated data sets attracts expertise and effort. Current efforts to aggregate data in central data servers and to standardize formats and metadata are steps in the right direction, but it is important that such efforts are developed in concert with intended analysis. In many cases, new methods for unsupervised or semi-supervised analysis of data need to be developed.

Perhaps the most common problem is the lack of adequate metadata (in this context referring to response variables, classes, annotations, or labels). Large volumes of raw data are collected and stored, but the specific and detailed results from analysis are not systematically recorded (Harris *et al.*, 2010), leaving the data essentially unannotated. In other cases, annotation is available, but made in an *ad hoc* manner. So where one annotator might label a plankton image "copepod, large," another might label it "large copepod." Often classes are poorly defined and inconsistent, and do not make use of available standards. And even when both data and metadata are available, in some cases the link between them is unreliable.

For applications with sparsely labelled training data, the discovery of the deep CNNs' ability to generalize and the usefulness of transfer learning have been a break-through (Razavian *et al.*, 2014; Yosinski *et al.*, 2014; Azizpour *et al.*, 2014). Transfer learning concerns the concept of transferring knowledge from one area to another (usually related) domain, and fully pre-trained nets trained from large databases with large label spaces (e.g. ImageNet) have shown good performance on several tasks (Razavian *et al.*, 2014). The transfer learning for CNNs is typically performed by either using the pre-trained network as a feature extractor (Razavian *et al.*, 2014). This approach was used by Siddiqui *et al.* (2018), who used trained a support vector machine to classify fish species based on the output from a pre-trained neural network. An alternative is to fine-tune the pre-trained network on the new target data. Fine-tuning can be restricted to higher layers, as the nets here tend to become more specific to details of the original labels. Less abstract features from lower layers are often useful for new tasks without modification (Azizpour *et al.*, 2014).

Most lines of work study and solve the problem of transfer learning within the same modality. The cross-modality transfer problem has received less attention, but approaches considering this typically rely on the existence of paired modalities (Gupta *et al.*, 2016) or shared label spaces, for example by hallucinating modalities during training time (Hoffman *et al.*, 2016; Kampffmeyer *et al.*, 2017), or jointly embedding or learning representations from multiple modalities into a shared feature space. The existence of shared label spaces or images paired with labelled natural images is, however, not the rule for applications involving non-standard data, which is often the case within marine science.

Anchoring projects in existing infrastructure and pipelines

The value of data is in its use, and for marine data to be useful, it must be analysed and the output used in science, for resource management, or by industry. With data sets available, methods can readily be developed, but without integration into existing processes, the impact is small or non-existent. To reap the benefits of new methods, it is crucial to involve the whole value chain, from data collection, to data storage and management, to analysis, and final use of the information. Projects must seek to involve existing stakeholders and have long-term implementation as a central goal, i.e. technology on its own has no merit in this context.

Remote electronic monitoring systems like AIS (which broadcasts position and other information) have been in use for some time, and can be analysed to identify vessel activities. Such systems have been used successfully for effective enforcement of fisheries policies and marine protected areas (McCauley *et al.*, 2016). Vessel monitoring systems can provide more detailed information, e.g. monitoring catch and bycatch from video surveillance (Joo *et al.*, 2011; French *et al.*, 2015). Electronic monitoring will enable effective and more specific policies for sustainable operations, but depends on automatic analysis to be cost effective (van Helmond *et al.*, 2017), and stakeholder commitment is crucial for implementation (ICES, 2018).

Developing new expertise and methods

Since Krizhevsky *et al.* (2012), machine learning has seen a tremendous increase in interest. In particular, many large, data-oriented corporations, including Google, Facebook, Amazon, Microsoft, IBM, and Baidu, are aggressively recruiting people with machine learning expertise. The academic sector is struggling to compete with enterprises for competence, and recruitment of experienced academic personnel to the commercial sector is likely to impede development of solutions needed for scientific progress; as well as having negative consequences for the education and training that the commercial sector itself depends on.

Structures are needed that encourage development and retention of machine learning expertise in the marine sciences. There is a need to provide motivation and opportunities for people with this background to work closely with stakeholders in the marine domains. For standard problems like image classification, it may be sufficient to adopt methods from other fields, but when dealing with data types and problems that are more particular to marine sciences, interdisciplinary approaches are needed, and scientists need to understand both machine learning and the relevant disciplines like biology or oceanography.

Software tools and frameworks

Deep learning has proven to be an effective tool in many similar situations and fields, and several popular software packages now exist that can be downloaded, adapted, and deployed quickly and easily. TensorFlow (Abadi *et al.*, 2015) is a flexible framework that abstracts computing hardware, but which has a steep learning curve. Keras (Chollet *et al.*, 2015) builds on top of TensorFlow or Theano (Bergstra *et al.*, 2010), providing an easier to use, but less flexible interface. PyTorch (Paszke *et al.*, 2017) is another popular framework combining ease of use with expressive power. These frameworks are general and can be adapted to challenges in the marine domain with relative ease (Allken *et al.*, 2019; Moen *et al.*, 2018; Siddiqui *et al.*, 2018; Villon *et al.*, 2018). The vast number of online tutorials and documentation is a major advantage, and pre-trained models are available from public repositories (often referred to as *model zoos*). Although these are usually aimed at generic tasks like classification of standard image data sets, they accelerate development of specific solutions by providing well-tested architectures and initial parameters that are useful as a starting point (Orenstein and Beijbom, 2017) for further training.

Until recently, developing and applying advanced analysis methods required programming skills as well as a good understanding of methods and software frameworks. A variety of programming languages—Fortran, MatLab, C++, Java, and R, to name a few—are used in marine science, but the bulk of commercial and academic development of new machine learning methods targets Python. A lack of familiarity with Python could limit uptake of new technologies, or restrict developers to an inferior selection of tools and frameworks available in their preferred language.

We are also seeing the introduction of tools and libraries that target the marine sciences specifically. Such domain-specific solutions provide solutions that are tailored to common use cases and with intuitive interfaces. This can help to make the technology much more accessible for non-experts. One recent example is the VIAME toolkit (Dawkins *et al.*, 2017), which is an ambitious project that integrates data processing and analyses in a comprehensive framework, and supports multiple programming languages.

In conclusion there are several levels for which the user can use and deploy these techniques. In general, there is a trade-off between ease of use and flexibility, and choice of framework and methods must be tailored to the competence and ambitions of each individual project. The authors of this paper use Keras and Theano daily and have found they serve as a reasonable balance between flexibility and ease of use.

Conclusions

In the near future, the volume and complexity of marine data are expected to increase by orders of magnitude. Autonomous platforms already drift, float, sail, and glide across the ocean surface and below it, collecting large amounts of data at relatively low cost. Additional data are collected from commercial and other non-scientific vessels, and from stationary observatories. Simultaneously, sensor technology is advancing rapidly, increasing resolution and detail level of the collected information.

Deep learning and CNNs have made impressive advances, and is likely to change the way we interpret, analyse, and collect data. For classification or regression over large, regularly structured data, existing methods can be (and is) applied more or less directly. Similarly, methods exist that can deal with time series and textual data. More speculatively, techniques from deep learning aimed at dealing with large numbers of parameters may bring insights in how to better model complex adaptive systems.

Nevertheless, some moderation is warranted, and it is not sufficient merely to accumulate vast amounts of data and expect a clever enough algorithm to readily extract valuable insights. All data are not created equal, and no analysis will be able to extract information that is not present in the data. Careful design of surveys and experiments is and will remain important. Also, deep learning methods often perform well within its domain, but can give unpredictable results on unfamiliar data. When such methods are deployed, a regime of careful monitoring of performance and subsequent adjustments will be necessary.

The transition into a data rich science is a paradigm shift with important implications. Current sparse sampling regimes and population-based models can be replaced with comprehensive monitoring at high resolution, sometimes down to the individual level. For locations of particular interest, like rivers or spawning grounds, it is already within our reach to register the presence of each individual fish, and classifying its species as well as behaviour and interactions. But data collection on this scale requires data analysis capabilities well beyond current manual methods, and will only be realized when the analysis bottleneck is solved.

Acknowledgements

The authors would like to thank Robert Jenssen for valuable comments and discussion.

Funding

This work was supported by the COGMAR project, Research Council of Norway grant no 270966/O70, and by the Norwegian Ministry of Trade, Industry and Fisheries.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S. *et al.* 2015. TensorFlow: Large-scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/ (last accessed 29 March 2019).

Allken, V., Handegard, N. O., Rosen, S., Schreyeck, T., Mahiout, T., and Malde, K. 2019. Fish species identification using a convolutional neural network trained on synthetic data. ICES Journal of Marine Science, 76: 342–349.

Azizpour, H., Razavian, A. S., Sullivan, J., Maki, A., and Carlsson, S. 2014. Factors of Transferability for a Generic ConvNet Representation. arXiv:1406.5774 [cs].

Baccouche, M., Mamalet, F., Wolf, C., Garcia, C., and Baskurt, A. 2011. Sequential deep learning for human action recognition. *In* International Workshop on Human Behavior Understanding, pp. 29–39. Ed. by A. A. Salah and B. Lepri. Springer, Heidelberg.

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., and Samek, W. 2015. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS One, 10: e0130140.

Badrinarayanan, V., Kendall, A., and Cipolla, R. 2015. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. arXiv, 1511.00561 [cs]. http://arxiv.org/abs/1511.00561 (last accessed 31 August 2016). preprint: not peer reviewed.

Baldi, P., Sadowski, P., and Whiteson, D. 2014. Searching for exotic particles in high-energy physics with deep learning. Nature Communications, 5: http://www.nature.com/ncomms/2014/140702/ncomms5308/full/ncomms5308.html (last accessed 4 September 2015).

Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J. *et al.* 2010. Theano: a CPU and GPU math compiler in Python. *In* Proc. 9th Python in Science Conf.

Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., Bento, P. *et al.* 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nature Communications, 5: 3657.

Beyan, C., Katsageorgiou, V.-M., and Fisher, R. B. 2018. Extracting statistically significant behaviour from fish tracking data with and without large dataset cleaning. IET Computer Vision, 12: 162–170.

Bianchi, G., and Skjoldal, H. R. 2008. The Ecosystem Approach to Fisheries. CABI, Oxfordshire. 379 pp.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. 1992. A training algorithm for optimal margin classifiers. *In* Proceedings of the Fifth Annual Workshop on Computational Learning Theory, pp. 144–152. Ed. by D. Haussler. ACM, New York, NY.

Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M. F. J., and Holte, B. 2015. The MAREANO programme—a full coverage mapping of the Norwegian off-shore benthic environment and fauna. Marine Biology Research, 11: 4–17.

Campbell, M., Hoane, A. J., and Hsu, F. 2002. Deep Blue. Artificial Intelligence, 134: 57–83.

Cappo, M., Harvey, E. S., and Shortis, M. R. 2007. Counting and measuring fish with baited video techniques—an overview. *In* Australian Society for Fish Biology 2006 Workshop Proceedings, pp. 101–114. http://epubs.aims.gov.au//handle/11068/7468 (last accessed 14 February 2018).

Chellapilla, K., Puri, S., and Simard, P. 2006. High Performance Convolutional Neural Networks for Document Processing. *In* Tenth International Workshop on Frontiers in Handwriting Recognition. Suvisoft.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. 2018. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40: 834–848. [10.1109/TPAMI.2017.2699184]

Chollet, F. *et al.* 2015. Keras. https://keras.io (last accessed 29 March 2019).

Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems, 2: 303–314.

Dahl, G. E., Jaitly, N., and Salakhutdinov, R. 2014. Multi-task neural networks for QSAR predictions. arXiv, 1406.1231 [cs, stat]. preprint: not peer reviewed.

Dalal, N., and Triggs, B. 2005. Histograms of oriented gradients for human detection. *In* 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 886–893. IEEE.

Dawkins, M., Sherrill, L., Fieldhouse, K., Hoogs, A., Richards, B., Zhang, D., Prasad, L. *et al.* 2017. An open-source platform for underwater image and video analytics. *In* 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 898–906. IEEE.

Dean, J., Corrado, G., Monga, R., Chen, K., Devin, M., Mao, M., Ranzato, M. *et al.* 2012. Large scale distributed deep networks. *In* Advances in Neural Information Processing Systems 25, pp. 1223–1231. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger. Curran Associates, Inc., Red Hook. http://papers.nips.cc/paper/4687-large-scale-distributed-deep-networks.pdf (last accessed 4 September 2015).

Erhan, D., Bengio, Y., Courville, A., and Vincent, P. 2009. Visualizing higher-layer features of a deep network. University of Montreal, 1341: 1.

Fassler, S. M. M., Brunel, T., Gastauer, S., and Burggraaf, D. 2016. Acoustic data collected on pelagic fishing vessels throughout an annual cycle: operational framework, interpretation of observations, and future perspectives. Fisheries Research, 178: 39–46.

Fernandes, P. G., Stevenson, P., Brierley, A. S., Armstrong, F., and Simmonds, E. J. 2003. Autonomous underwater vehicles: future platforms for fisheries acoustics. ICES Journal of Marine Science: Journal du Conseil, 60: 684–691.

Fisher, M., and Hunter, E. 2018. Digital imaging techniques in otolith data capture, analysis and interpretation. Marine Ecology Progress Series, 598: 213–231.

Foote, A. D., Thomsen, P. F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L. A., Salling, A. B. *et al.* 2012. Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. PLoS One, 7: e41781.

French, G., Fisher, M., Mackiewicz, M., and Needle, C. 2015. Convolutional Neural Networks for Counting Fish in Fisheries Surveillance Video. *In* Proceedings of the Machine Vision of Animals and their Behaviour (MVAB), pp. 7.1–7.10. Ed. by S. Amaral, T. Matthews, T. Plötz, S. McKenna, and R. Fisher. BMVA Press, University of Swansea. https://ueaeprints.uea.ac.uk/55574/ (last accessed 31 January 2019).

Fukushima, K. 1988. Neocognitron: a hierarchical neural network capable of visual pattern recognition. Neural Networks, 1: 119–130.

Fulton, E. A., Smith, A. D. M., and Johnson, C. R. 2003. Effect of complexity on marine ecosystem models. Marine Ecology Progress Series, 253: 1–16.

Fulton, E. A., Smith, A. D. M., Smith, D. C., and Johnson, P. 2014. An integrated approach is needed for ecosystem based fisheries management: insights from ecosystem-level management strategy evaluation. PLoS One, 9: e84242.

Godø, O. R., Johnsen, S., and Torkelsen, T. 2014. The LoVe ocean observatory is in operation. Marine Technology Society Journal, 48: 24–30.

Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., and Shet, V. 2013. Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv, 1312.6082 [cs]. http://arxiv.org/abs/1312.6082 (last accessed 4 September 2015). preprint: not peer reviewed.

Google. 2018. Publications. https://ai.google/research/pubs/ (last accessed 9 May 2018).

Graves, A., Mohamed, A., and Hinton, G. 2013. Speech recognition with deep recurrent neural networks. *In* 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6645–6649. IEEE.

Guihen, D., Fielding, S., Murphy, E. J., Heywood, K. J., and Griffiths, G. 2014. An assessment of the use of ocean gliders to undertake acoustic measurements of zooplankton: the distribution and density of Antarctic krill (*Euphausia superba*) in the Weddell Sea. Limnology and Oceanography: Methods, 12: 373–389.

Gupta, S., Hoffman, J., and Malik, J. 2016. Cross modal distillation for supervision transfer. *In* 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2827–2836. IEEE.

Handegard, N. O., Buisson, L., du Brehmer, P., Chalmers, S. J., De Robertis, A., Huse, G., Kloser, R. *et al.* 2013. Towards an acoustic-based coupled observation and modelling system for monitoring and predicting ecosystem dynamics of the open ocean. Fish and Fisheries, 14: 605–615.

Harris, G., Thompson, R., Childs, J. L., and Sanderson, J. G. 2010. Automatic storage and analysis of camera trap data. The Bulletin of the Ecological Society of America, 91: 352–360.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. 2017. Mask R-CNN. arXiv, 1703.06870 [cs]. preprint: not peer reviewed.

He, K., Zhang, X., Ren, S., and Sun, J. 2016. Deep residual learning for image recognition. *In* Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778. IEEE.

He, K., Zhang, X., Ren, S., and Sun, J. 2016. Deep residual learning for image recognition. *In* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. IEEE.

Heaton, J. B., Polson, N. G., and Witte, J. H. 2017. Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33: 3–12.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A., Jaitly, N., Senior, A. *et al.* 2012. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Processing Magazine, 29: 82–97.

Hinton, G. E., and Salakhutdinov, R. R. 2006. Reducing the dimensionality of data with neural networks. Science, 313: 504–507.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. 2012. Improving neural networks by preventing co-adaptation of feature detectors. arXiv:1207.0580 [cs]. preprint: not peer reviewed.

Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. Neural Computation, 9: 1735–1780.

Hoffman, J., Gupta, S., and Darrell, T. 2016. Learning with side information through modality hallucination. *In* 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 826–834. IEEE.

Honkalehto, T., Ressler, P. H., Towler, R. H., and Wilson, C. D. 2011. Using acoustic data from fishing vessels to estimate walleye pollock (*Theragra chalcogramma*) abundance in the eastern Bering Sea. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1231–1242.

Hornik, K., Stinchcombe, M., and White, H. 1989. Multilayer feedforward networks are universal approximators. Neural Networks, 2: 359–366.

ICES. 2018. Report of the Workshop on Machine Learning in Marine Science (WKMLEARN), 16–20 April 2018, ICES Headquarters, Copenhagen, Denmark. ICES CM 2018/EOSG: 28 pp.

Ioffe, S., and Szegedy, C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv: 1502.03167 [cs]. preprint: not peer reviewed.

Jackson, S. A., Borchert, E., O'Gara, F., and Dobson, A. D. 2015. Metagenomics for the discovery of novel biosurfactants of environmental interest from marine ecosystems. Current Opinion in Biotechnology, 33: 176–182.

Joo, R., Bertrand, S., Chaigneau, A., and Ñiquen, M. 2011. Optimization of an artificial neural network for identifying fishing set positions from VMS data: an example from the Peruvian anchovy purse seine fishery. Ecological Modelling, 222: 1048–1059.

Kaggle. 2018. https://www.kaggle.com/ (last accessed 11 July 2018).

Kampffmeyer, M., Salberg, A.-B., and Jenssen, R. 2016. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. *In* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 1–9. IEEE. http://www.cv-foundation.org/openaccess/content_cvpr_2016_workshops/w19/html/Kampffmeyer_Semantic_Segmentation_of_CVPR_2016_paper.html (last accessed 31 August 2016).

Kampffmeyer, M., Salberg, A.-B., and Jenssen, R. 2017. Urban land cover classification with missing data using deep convolutional neural networks. *In* 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5161–5164. IEEE.

Kloser, R. J., Ryan, T., Sakov, P., Williams, A., and Koslow, J. A. 2002. Species identification in deep water using multiple acoustic frequencies. Canadian Journal of Fisheries and Aquatic Sciences, 59: 1065–1077.

Kodzius, R., and Gojobori, T. 2015. Marine metagenomics as a source for bioprospecting. Marine Genomics, 24: 21–30.

Korneliussen, R. J., and Ona, E. 2003. Synthetic echograms generated from the relative frequency response. ICES Journal of Marine Science: Journal du Conseil, 60: 636–640.

Koslow, J. A. 2009. The role of acoustics in ecosystem-based fishery management. ICES Journal of Marine Science, 66: 966–973.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. 2012. ImageNet classification with deep convolutional neural networks. *In* Advances in Neural Information Processing Systems 25, pp. 1097–1105. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger. Curran Associates, Inc., Red Hook. http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf (last accessed 4 September 2015).

Larson, W. A., Seeb, L. W., Everett, M. V., Waples, R. K., Templin, W. D., and Seeb, J. E. 2014. Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). Evolutionary Applications, 7: 355–369.

LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. Nature, 521: 436–444.

LeCun, Y., Haffner, P., Bottou, L., and Bengio, Y. 1999. Object recognition with gradient-based learning. *In* Shape, Contour and Grouping in Computer Vision, pp. 319–345. Ed. by D. A. Forsyth, J. L. Mundy, V. di Gesú, and R. Cipolla. Springer, Berlin, Heidelberg. https://link.springer.com/chapter/10.1007/3-540-46805-6_19 (last accessed 10 July 2018).

Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., Hvidsten, T. R. *et al.* 2016. The Atlantic salmon genome provides insights into rediploidization. Nature, 533: 200–205.

Link, J., and Browman, H. 2014. Integrating what? Levels of marine ecosystem-based assessment and management. ICES Journal of Marine Science, 71: 1170–1173.

Long, J., Shelhamer, E., and Darrell, T. 2014. Fully convolutional networks for semantic segmentation. arXiv:1411.4038 [cs].

Lowe, D. G. 1999. Object recognition from local scale-invariant features. *In* Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, pp. 1150–1157. IEEE.

Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60: 91–110.

MacLennan, D., and Simmonds, E. J. 2005. Fisheries Acoustics. Fish and Aquatic Resources Series 10. Chapman & Hall, London.

Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P. 2017. Convolutional neural networks for large-scale remote-sensing image classification. IEEE Transactions on Geoscience and Remote Sensing, 55: 645–657. [10.1109/TGRS.2016.2612821]

Malde, K., Seliussen, B. B., Quintela, M., Dahle, G., Besnier, F., Skaug, H. J., Øien, N. *et al.* 2017. Whole genome resequencing reveals diagnostic markers for investigating global migration and hybridization between minke whale species. BMC Genomics, 18: 76.

Marshall, J., and Oberwinkler, J. 1999. Ultraviolet vision: the colourful world of the mantis shrimp. Nature, 401: 873–874.

McCauley, D. J., Woods, P., Sullivan, B., Bergman, B., Jablonicky, C., Roan, A., Hirshfield, M. *et al.* 2016. Ending hide and seek at sea. Science, 351: 1148–1150.

Microsoft. 2018. Search. https://www.microsoft.com/en-us/research/search/ (last accessed 10 May 2018).

Mitchell, T. M. 1997. Machine Learning. WCB. McGraw-Hill, Boston, MA.

Moen, E., Handegard, N. O., Allken, V., Albert, O. T., Harbitz, A., and Malde, K. 2018. Automatic interpretation of otoliths using deep learning. PLoS One, 13: e0204713.

Montavon, G., Samek, W., and Müller, K.-R. 2018. Methods for interpreting and understanding deep neural networks. Digital Signal Processing, 73: 1–15.

Mordy, C. W., Cokelet, E. D., De Robertis, A., Jenkins, R., Kuhn, C. E., Lawrence-Slavas, N., Berchok, C. L. *et al.* 2017. Advances in ecosystem research: Saildrone surveys of oceanography, fish, and marine mammals in the Bering Sea. Oceanography, 30: 113–115.

Nair, V., and Hinton, G. E. 2010. Rectified linear units improve restricted Boltzmann machines. *In* Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 807–814. Omnipress.

NASA. 2018. NASA Ocean Color. https://oceancolor.gsfc.nasa.gov/ (last accessed 10 July 2018).

Nicolescu, B. 2008. Transdisciplinarity: Theory and Practice. Hampton Press, New York. 332 pp.

Olsen, E., Fay, G., Gaichas, S., Gamble, R., Lucey, S., and Link, J. S. 2016. Ecosystem model skill assessment. Yes We Can! PLoS One, 11: e0146467.

Orenstein, E. C., and Beijbom, O. 2017. Transfer learning and deep feature extraction for planktonic image data sets. *In* 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1082–1088. IEEE.

Parker, D. B. 1985. Learning-logic: learning-logic: casting the cortex of the human brain in silicon. Technical Report Tr-47, Center for Computational Research in Economics and Management Science. MIT Cambridge, MA.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z. *et al.* 2017. Automatic differentiation in PyTorch. https://openreview.net/forum? id=BJJsrmfCZ (last accessed 20 November 2018).

Pikitch, E. K., Santora, C., Babcock, E. A., Bakun, A., Bonfil, R., Conover, D. O., Dayton, P. *et al.* 2004. Ecosystem-based fishery management. Science, 305: 346–347.

Pineda, F. J. 1987. Generalization of back-propagation to recurrent neural networks. Physical Review Letters, 59: 2229–2232.

Planque, B. 2016. Projecting the future state of marine ecosystems, "la grande illusion"? ICES Journal of Marine Science, 73: 204–208.

Polanyi, M. 2009. The Tacit Dimension. University of Chicago Press, Chicago. 129 pp.

Raghupathi, W., and Raghupathi, V. 2014. Big data analytics in healthcare: promise and potential. Health Information Science and Systems, 2: 3.

Razavian, A. S., Azizpour, H., Sullivan, J., and Carlsson, S. 2014. CNN features off-the-shelf: an astounding baseline for recognition. *In* Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 512–519. IEEE Computer Society, Washington, DC. http://dx.doi.org/10.1109/CVPRW.2014.131 (last accessed 14 February 2018).

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. 2016. You only look once: unified, real-time object detection. *In* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788. IEEE.

Remsen, A., Hopkins, T. L., and Samson, S. 2004. What you see is not what you catch: a comparison of concurrently collected net, Optical Plankton Counter, and Shadowed Image Particle Profiling Evaluation Recorder data from the northeast Gulf of Mexico. Deep Sea Research Part I: Oceanographic Research Papers, 51: 129–151.

Ren, S., He, K., Girshick, R., and Sun, J. 2015. Faster R-CNN: towards real-time object detection with region proposal networks. arXiv: 1506.01497 [cs].

Roemmich, D., Johnson, G. C., Riser, S., Davis, R., Gilson, J., Owens, W. B., Garzoli, S. L. *et al.* 2009. The Argo Program: observing the global ocean with profiling floats. Oceanography, 22: 34–43.

Rosen, S., Jörgensen, T., Hammersland-White, D., and Holst, J. C. 2013. DeepVision: a stereo camera system provides highly accurate counts and lengths of fish passing inside a trawl. Canadian Journal of Fisheries and Aquatic Sciences, 70: 1456–1467.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning representations by back-propagating errors. Nature, 323: 533–536.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z. *et al.* 2015. Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 115: 211–252.

Salakhutdinov, R., Mnih, A., and Hinton, G. 2007. Restricted Boltzmann machines for collaborative filtering. *In* Proceedings of the 24th international conference on Machine learning, ACM, pp. 791–798.

Schmidhuber, J. 2015. Deep learning in neural networks: an overview. Neural Networks, 61: 85–117.

Schunter, C., Vollmer, S. V., Macpherson, E., and Pascual, M. 2014. Transcriptome analyses and differential gene expression in a non-model fish species with alternative mating tactics. BMC Genomics, 15: 167.

Science. 2016. From AI to Protein Folding: Our Breakthrough Runners-up. http://www.sciencemag.org/news/2016/12/ai-protein-folding-our-breakthrough-runners (last accessed 9 May 2018).

Siddiqui, S. A., Salman, A., Malik, M. I., Shafait, F., Mian, A., Shortis, M. R., and Harvey, E. S. 2018. Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. ICES Journal of Marine Science, 75: 374–389.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G., van den Schrittwieser, J. *et al.* 2016. Mastering the game of Go with deep neural networks and tree search. Nature, 529: 484–489.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T. *et al.* 2017. Mastering the game of Go without human knowledge. Nature, 550: 354–359.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. 2014. Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15: 1929–1958.

Stanton, T. K., Chu, D., Jech, J. M., and Irish, J. D. 2010. New broadband methods for resonance classification and high-resolution imagery of fish with swimbladders using a modified commercial broadband echosounder. ICES Journal of Marine Science: Journal du Conseil, 67: 365–378.

Stemmann, L., and Boss, E. 2012. Plankton and particle size and packaging: from determining optical properties to driving the biological pump. Annual Review of Marine Science, 4: 263–290.

Sund, O. 1935. Echo sounding in fishery research. Nature, 135: 953.

Sutskever, I., Vinyals, O., and Le, Q. V. V. 2014. Sequence to sequence learning with neural networks. *In* Advances in Neural Information Processing Systems 27, pp. 3104–3112. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger. Curran Associates, Inc., Red Hook. http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf (last accessed 4 September 2015).

Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. 2014. DeepFace: closing the gap to human-level performance in face verification. *In* 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1701–1708. IEEE.

Thomsen, P. F., Kielgast, J., Iversen, L. L., Møller, P. R., Rasmussen, M., and Willerslev, E. 2012. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS One, 7: e41732.

Uusitalo, L., Fernandes, J. A., Bachiller, E., Tasala, S., and Lehtiniemi, M. 2016. Semi-automated classification method addressing marine strategy framework directive (MSFD) zooplankton indicators. Ecological Indicators, 71: 398–405.

van Helmond, A. T. M., Chen, C., and Poos, J. J. 2017. Using electronic monitoring to record catches of sole (*Solea solea*) in a bottom trawl fishery. ICES Journal of Marine Science, 74: 1421–1427.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O. *et al.* 2001. The sequence of the human genome. Science (New York, NY), 291: 1304–1351.

Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G., Claverie, T., and Villéger, S. 2018. A deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecological Informatics, 48: 238–244.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P.-A. 2010. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. The Journal of Machine Learning Research, 11: 3371–3408.

Vinyals, O., Toshev, A., Bengio, S., and Erhan, D. 2014. Show and tell: a neural image caption generator. arXiv:1411.4555 [cs]. preprint: not peer reviewed.

Wan, L., Zeiler, M., Zhang, S., Cun, Y. L., and Fergus, R. 2013. Regularization of Neural Networks using DropConnect. In PMLR, pp. 1058–1066. http://proceedings.mlr.press/v28/wan13.html (last accessed 14 February 2018).

Wood, G. 2018. Google's AI Wins Fifth and Final Game Against Go Genius Lee Sedol. https://www.wired.com/2016/03/googles-ai-wins-fifth-final-game-go-genius-lee-sedol/ (last accessed 8 May 2018).

Xiao, T., Xu, Y., Yang, K., Zhang, J., Peng, Y., and Zhang, Z. 2014. The application of two-level attention models in deep convolutional neural network for fine-grained image classification. arXiv:1411.6447 [cs].

Xie, J., Girshick, R., and Farhadi, A. 2016. Unsupervised deep embedding for clustering analysis. *In* Proceedings of the 33rd International Conference on International Conference on Machine Learning, 48, pp. 478–487. Ed. by M. F. Balcan and K. Q. Weinberger. JMLR.org, New York, NY. http://dl.acm.org/citation.cfm? id=3045390.3045442 (last accessed 31 January 2019).

Xu, B., Wang, N., Chen, T., and Li, M. 2015. Empirical evaluation of rectified activations in convolutional network. arXiv:1505.00853 [cs, stat]. preprint: not peer reviewed.

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. 2014. How transferable are features in deep neural networks? *In* Advances in Neural Information Processing Systems 27, pp. 3320–3328. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q.

Weinberger. Curran Associates, Inc., Red Hook. http://papers.nips.cc/paper/5347-how-transferable-are-features-in-deep-neural-networks.pdf (last accessed 14 February 2018).

Yosinski, J., Clune, J., Nguyen, A., Fuchs, T., and Lipson, H. 2015. Understanding neural networks through deep visualization. arXiv: 1506.06579 [cs]. http://arxiv.org/abs/1506.06579 (last accessed 27 November 2018). preprint: not peer reviewed.

Yu, F., and Koltun, V. 2015. Multi-scale context aggregation by dilated convolutions. arXiv:1511.07122 [cs]. preprint: not peer reviewed.

Zeiler, M. D., and Fergus, R. 2014. Visualizing and understanding convolutional networks. *In* European Conference on Computer Vision, pp. 818–833. Springer.

Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., and Shen, D. 2015. Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation. NeuroImage, 108: 214–224.

Handling editor: David Demer