Multimedia Systems
https://doi.org/10.1007/s00530-020-00733-x

SPECIAL ISSUE PAPER

Application of machine learning in ocean data

Ranran Lou[1] · Zhihan Lv[1] · Shuping Dang[5] · Tianyun Su[2,3,4] · Xinfang Li[2,3,4]

Received: 9 October 2020 / Accepted: 5 December 2020

Abstract

In recent years, machine learning has become a hot research method in various fields and has been applied to every aspect of our life, providing an intelligent solution to problems that could not be solved or difficult to be solved before. Machine learning is driven by data. It learns from a part of the input data and builds a model. The model is used to predict and analyze another part of the data to get the results people want. With the continuous advancement of ocean observation technology, the amount of ocean data and data dimensions are rising sharply. The use of traditional data analysis methods to analyze massive amounts of data has revealed many shortcomings. The development of machine learning has solved these shortcomings. Nowadays, the use of machine learning technology to analyze and apply ocean data becomes the focus of scientific research. This method has important practical and long-term significance for protecting the ocean environment, predicting ocean elements, exploring the unknown, and responding to extreme weather. This paper focuses on the analysis of the state of the art and specific practices of machine learning in ocean data, review the application examples of machine learning in various fields such as ocean sound source identification and positioning, ocean element prediction, ocean biodiversity monitoring, and deep-sea resource monitoring. We also point out some constraints that still exist in the research and put forward the future development direction and application prospects.

Keywords Ocean · Data · Ocean data · Machine learning

✉ Zhihan Lv
lvzhihan@gmail.com

Ranran Lou
louranran1113@gmail.com

Shuping Dang
shuping.dang@kaust.edu.sa

Tianyun Su
sutiany@fio.org.cn

Xinfang Li
lixinfang@fio.org.cn

1 School of Data Science and Software Engineering, Qingdao University (QDU), Qingdao, China

2 Laboratory for Regional Oceanography and Numerical Modeling, Pilot National Laboratory for Marine Science and Technology, Qingdao, China

3 Marine Data and Information Center, The First Institute of Oceanography, MNR, Qingdao, China

4 National Engineering Laboratory for Integrated, Aero-Space-Ground-Ocean Big Data Application Technology, Qingdao, China

5 Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

1 Introduction

Nowadays, big data has become the focal discussion and attention in various industries [1]. Through various data generated in production and life, we can summarize the laws of nature and society and predict future trends. Benefitting from advances in ocean observation technology and system processing capabilities, different types of ocean data exceeding TB level are collected from various kinds of sensor equipment every day [2–4]. The ocean itself is a huge and complex ecosystem, including many disciplines and fields, involving marine chemistry, marine geology, physical oceanography, marine biology, etc. For ocean data research, collecting data is only one aspect, while how to process these ocean big data with different types and sources for scientific research is an urgent problem that we need to solve. In the past, people were still in the stage of acquiring and analyzing ocean data, and the application of ocean data was not extensive.

Making full use of big ocean data can help human beings achieve better development in the research of responding to climate change, protecting the ecological environment, and preventing natural disasters [5–9]. Traditional ocean data processing and analysis mostly use manual classification and recognition, traditional statistical analysis, ocean model simulation and other methods. These methods are often affected by subjective factors and cannot truly describe the hidden information in the data. Moreover, most of the big data in the ocean are unstructured or semi-structured data, with complex or unrelated relationships among the data, which also poses challenges to the traditional statistical analysis and ocean model simulation [10]. The numerical model itself needs to be cautious and professional in its realization. When a large amount of external information cannot be obtained and computing resources and expertise are limited, the results are not satisfactory. With the development and popularization of machine learning technology, the use of machine learning algorithms to analyze the application of big data have become a hot research fields. Big data provides sufficient data support for machine learning to extract patterns and build models [11]. Machine learning algorithms learn from a large amount of data and build models. Compared with traditional ocean data analysis methods, it has the advantages of high accuracy, low complexity, and less calculation, and in some cases, it reduces data requirements. At present, machine learning is used in ocean big data to identify and locate under the sea, predict marine elements, marine life distribution, and climate analysis [12–15]. This paper will discuss the definition of machine learning and the latest development of machine learning application in ocean data, summarize the problems involved, and analyze the future research directions.

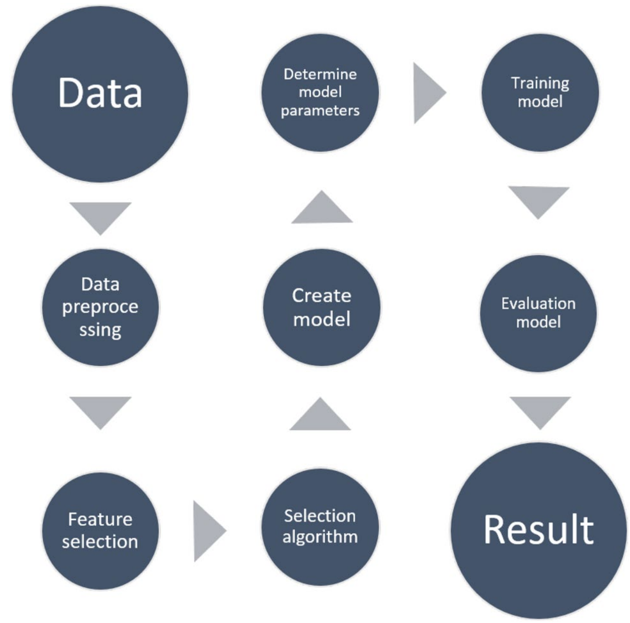

Fig. 1 Machine learning process

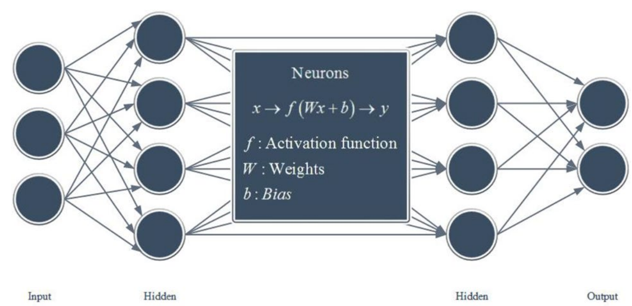

Fig. 2 Deep learning of multi-layer perceptron

2 Machine learning

A learning problem can be defined as a problem that improves the performance of a task through some type of training experience [16]. What is machine learning? Machine learning is an interdisciplinary subject involving mathematics, statistics, and computer science, it uses instance data or past experience to train computers to optimize certain performance standards [17]. With the rapid increase in the amount of data in various industries, the use of appropriate machine learning algorithms can improve the efficiency of data analysis and processing, and solve some practical problems. Figure 1 shows the process of machine learning.

In 1950, Turing, the "Father of Artificial Intelligence", invented the famous "Turing Test", which initiated scientific research on artificial intelligence. In 1957, Frank Rosenblatt proposed the Perceptron concept and designed the first neural network. In 1969, Marvin Minsky and Seymour Papert raised the XOR problem in their book "Perceptron", which brought machine learning into a low ebb. In 1985, Rumelhart, Hinton and Williams proposed the well-known back propagation (BP) algorithm [18], which becomes the essential algorithm of neural networks. In 1990, various machine learning algorithms such as support vector machine (SVM), random forest (RF), and logistic regression (LR) came out, which are capable of completing basic recognition and classification tasks. In 2006, Hinton and Salakhutdinov invented the deep learning algorithm [19]. Figure 2 shows the algorithm for feature learning through a multi-layer perceptron (MLP) with multiple hidden layers. MLP consists of an input layer, an output layer, and a hidden layer. There can be multiple hidden layers. The input data will undergo a series of weighted sums in the hidden layer. After calculating the weighted summation of each hidden neuron, the result is applied to a non-linear function, the so-called activation function, and the result of this function is weighted and summed to obtain the output. The emergence of deep learning has also promoted the current upsurge in machine learning research. For example, in 2012, the Hinton research team used deep learning to win the most influential ImageNet competition in the field of computer vision. In the Go game held in March 2016, the AlphaGo robot, which applied the principles of deep learning, defeated Lee Sedol, the world champion of Go and a professional player with a nine-dan rank and a total score of 4-1.

Machine learning is divided into supervised learning and unsupervised learning according to whether the sample data contains label data. Supervised learning algorithms learn from the labeled train data and make label predictions on the test data. Unsupervised learning algorithms do not need to learn from labeled train data, they learn the structural features between the data from the unlabeled sample data, and classifies them according to the learned features.

3 Application status

3.1 Sound source identification and location

One of the main applications of ocean data is to use acoustic data to identify and locate at sea. Before applying machine learning, matching field processing (MFP) was the most common ocean positioning method. MFP is a general beamforming method that uses the spatial complexity of the sound field in ocean waveguides to locate sources in range, depth, and azimuth or to infer the parameters of the waveguide itself [20, 21]. However, the use of MFP for marine sound source localization relies on marine environmental information so this method can only be used in simple and stable environments [22, 23]. Machine learning provides a more reliable method for marine source location. It can learn directly from data without simulating the marine environment [22].

In the 1990s, researchers began to use machine learning to study marine acoustics [24, 25]. Commonly used machine learning algorithms for maritime positioning are SVM, RF, feedforward neural network (FNN), and convolutional neural network (CNN). Using these algorithms of machine learning, it is possible to conduct applied research on ship positioning, ship classification, and estimating seabed distance, and they have achieved better research results compared with MFP [22–27]. In 2017, Niu et al. preprocessed the normalized sample covariance matrix constructed by the pressure received by the vertical linear array and input it into three machine learning models: FNN, SVM and RF to learn the source range, proved the potential of machine learning in underwater source location [22]. At the same time, they used ship acoustic data at different speeds in the Santa Barbara Strait to prove the effectiveness of SVM and FNN machine learning classifiers for acoustic localization [26]. Van Komen et al. proved the feasibility of using deep learning CNN to simultaneously determine the seabed type and source range from the 1s pressure time series of impulsive sounds [27]. Although machine learning uses data-driven advantages to overcome the disadvantages of traditional marine sound source positioning that requires environmental factors, new problems have also followed. The existing marine acoustic data cannot meet the amount of data required for training models. For the research of ocean sound source positioning, this is the next problem that needs to be solved.

3.2 Ocean forecast

Sea surface temperature, sea waves, sea ice, etc. are all important ocean elements. The analysis and prediction of these elements are of great significance to disaster prevention, environmental protection, and weather forecasting. Numerous algorithms of machine learning provide accurate and efficient methods for analyzing and predicting ocean elements. However, the use of machine learning to predict ocean elements still has the problem of insufficiently clear characteristics.

Sea surface temperature Sea surface temperature (SST) depends on the heat budget of sea water. It has obvious diurnal change and seasonal change, especially the change of geographical distribution, and has important influence on climate and marine ecosystem [28–30]. The long short-term memory (LSTM) neural network has a strong learning and predictive ability for time series data such as SST, and its structure is shown in Fig. 3. The network selectively memorizes the input from the previous node through input gates, forget gates, and output gates, and determines the output of the current state. Based on LSTM, Xiao et al. built a 5-layer deep neural network model for SST anomaly (SSTA) prediction, as shown in Fig. 4, at the same time, the AdaBoost integrated learning model was used to solve the problem of overfitting. Machine performance was improved by combining these two powerful and heterogeneous models [31]. They also established a spatio-temporal deep learning model, using convolutional long short-term memory (ConvLSTM) as a building block for training, and had relatively accurate prediction results for the short-term and mid-term daily forecasts of SST [32]. Lins et al. proposed to predict seasonal SST through the SVM [33].

Waves Nowadays, wave forecasting provides great convenience for people's sea life, and is helpful for shipping, fishing, national defense, and offshore energy exploration; not only that, the prediction of ocean waves also helps to study the energy transmission and material exchange of marine ecology [35]. The formation of ocean waves is a complicated seawater movement process, which is the propagation of the undulating shape of the sea surface and a wave

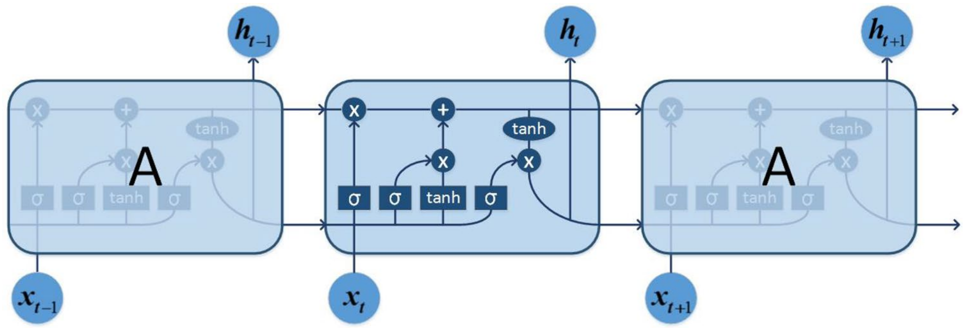

Fig. 3 Long short-term memory network structure [34]

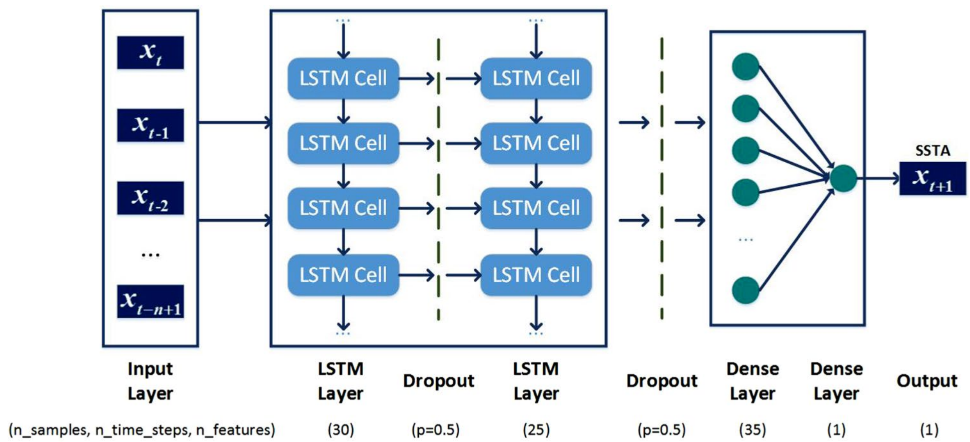

Fig. 4 Neural network structure for predicting SSTA [31]

formed by the periodic vibration of water quality points when they leave the equilibrium position and propagate in a certain direction. Traditional ocean wave forecasting is to establish a numerical model by simulating the wave evolution process generated by the wind field acting on the ocean surface. Currently, the third-generation wave forecasting model is usually used, including the WAM (Wave Model) established by the WAMDI team in 1988 [36], the SWAN (Simulating Wave Nearshore) model developed by Booij et al. [37] and the WAVEWATCH III model developed by Tolman et al. [38]. However, these forecasting models require more and accurate input data, the forecasting time is long, the complexity is high, and the forecasting effect is not satisfactory. In recent years, the use of machine learning to predict ocean waves has become the focus of attention of researchers and has been widely used. Artificial neural network(ANN) is widely used to predict wave parameters [39–42]. Rao and Mandal in 2005 used the neural network approach to estimate wave parameters from the wind field generated by cyclones [43]. Compared with traditional numerical prediction models, neural networks improve accuracy, reduce complexity, reduce the amount of calculation, and in some cases, reduce the need for data. As a commonly used classification algorithm, SVM is also one of the machine learning algorithms for predicting ocean waves [44, 45]. Mahjoubi and Mosabbeb used the current wind speed and the hourly wind speed in the previous six hours collected from the deep water of Lake Michigan as input data, used the SVM to predict the wave height, and used the same data with the ANN and Radial Basis Function (RBF). Multiple evaluation indexes [deviation, correlation coefficient (R), root mean square error (RMSE) and scatter index (SI)] show that SVM can be successfully used in wave height prediction and the error of SVM is slightly better than that of ANN [44]. Compared with ANN, SVM does not overfit, requires fewer parameters, shorter calculation time, and higher accuracy. For the time series interpolation problem of buoy missing data, these software algorithms are also applicable [46–48]. To meet the constantly changing data flow, Durán-Rosal et al. proposed to use the evolutionary unit neural network (EPUNN) and use the linear model as the input part to reconstruct the data. This method has good performance in the real case of reconstruction of a large number of lost data on 6 wave buoys in the Alaska Bay [48].

Ocean eddy Eddy is a vortex-type water vortex, also known as a black hole in the ocean. Ocean eddies are usually caused by tides. The global ocean circulation is also largely affected by mesoscale ocean eddies [49]. These eddies exist in all sea areas around the world and play a role in the transmission of kinetic energy in the ocean circulation. To monitor and track eddy currents, Franz et al. proposed a framework that combines CNN with the image processing tool Kanade–Lucas–Tomasi (KLT), and compared with the LSTM. This method achieves a high recognition rate and accuracy [49]. Bai et al. developed a deep learning method called streampath-based region-based convolutional neural networks (SP-RCNN) for automatically identifying ocean vortices from flow field data. First, a large-scale eddy dataset is constructed from ocean current data through a streampath-based method. Then the multi-layer features in the neural network are combined with the features of the eddy, and more particles are placed in the eddy domain image to enhance the display of the eddy; the mean average precision (mAP) of the monitoring results is 90.64%; the success of detection rate (SDR) is 98.91%, which solves the problem that it is difficult to detect the eddy in the sparse flow path area. This is also the first method to apply deep learning technology to identifying eddy currents in flow field data [50]. In addition, the use of machine learning can also predict turbulence processes and ocean flow fields, and classify eddies [51, 52].

3.3 Ocean biodiversity monitoring

Biodiversity is a broad concept that describes the degree of diversity in the natural world. May believes that from the genetic diversity within the native population of a species to the genetic diversity between geographically different populations of the same species, to communities or ecosystems, biodiversity exists at many different levels [53]. Compared with land, there are more types of marine life, but people pay less attention [54]. Machine learning is expected to replace methods such as equation fitting and manual monitoring, making marine biological monitoring more accurate, convenient and efficient.

To better study marine life, researchers use various data collected in the ocean and combine machine learning algorithms to carry out various experiments. Researchers such as Wei and others used the RF to predict global seabed biomass in the Census of Marine Life (CoML) project; this method models the complex and potentially nonlinear relationships between ocean attributes and seabed conventional populations, analyzing the cycle of organic matter effectively predicts the biomass and abundance of the global seabed [55]. Fish species classification is one of the important studies of marine life. In the 1990s, scientists tried to use Principal Component Analysis (PCA) and linear discriminant analysis to extract the main characteristics of fish are classified, but the accuracy is not high [56, 57]. Huang et al. proposed a new balanced optimization tree (BEOTR) classifier with rejection options for live fish recognition. After the recognition phase, a rejection system based on Gaussian Mixture Model (GMM) is added to the classifier. The rejection function evaluates the posterior probability of the test sample, which can overcome some misclassifications in the BEOTR classifier. The classifier tested 24,150 artificially labeled images containing 15 common fish species in the waters of Taiwan, with an accuracy rate of 74.8% [58]. In 2018, Siddiqui et al. used a CNN model pre-trained in a public image set to extract features from 16 different fish images in the temperate and subtropical coastal waters of Western Australia, and finally applied the linear one-to-many strategy SVM performs classification with an accuracy of 94.3% [59]. In addition, for more marine biological research, Reus et al. established the first publicly available seagrass image dataset and proposed a machine learning method to automatically estimate seagrass coverage on the seafloor, and studied the use of CNN to describe seagrass patches and superpixels [60]; Glotin et al. used machine learning to study sperm whale bioacoustics [61, 62]; Al-Barazanchi et al. used CNN not only to classify plankton images, but also to extend to new classifications [63].

3.4 Deep-sea resource monitoring

Human development is inseparable from the development and utilization of various resources. Today, when land resources are gradually depleted, people have turned their attention to the deep ocean. There are huge reserves of various energy and minerals in the deep sea. Deep-sea resource development will not only significantly increase the world's resource base, but also bring considerable economic benefits to the world in the future. Measuring distribution is an indispensable task in the early stage of seabed resource development. Traditional marine resource exploration requires various types of observation equipment to sample and use mathematical methods for modeling, which is very time-consuming, labor-intensive and has low accuracy. Machine learning can quickly and accurately measure and model deep resources.

The deep-sea iron-manganese nodules found in the Clarion–Clipperton Zone (CCZ) of the Pacific Ocean are a huge potential source of metals such as nickel, cobalt, and manganese. To obtain data on the quantity and quality distribution of nodules in the CCZ, Hari et al. proposed a method based on artificial neural network is used to model the nodule parameters in CCZ using limited data available in the open domain [64]. Similarly, to measure the coverage of nodules, Jie used side scan sonar and Automatic Underwater Vehicle (AUV) collected data on the Clarion and Clipperton Fracture Zone (CCFZ), and proposed an ANN based evaluation The PMN abundance of metal nodules has a test accuracy of 84% [65].

4 The problem

4.1 Data

First of all, in terms of ocean data standards, with the development of ocean observation technology, the data standards collected by various observation methods are also different, which is a challenge for data-driven machine learning. Ocean metadata is the main means to solve ocean data management. The current ocean metadata standards include marine environmental data inventory (MEDI), European directory of the initial ocean-observing system (EDIOS), array for real-time geostrophic oceanography (ARGO) and many more. These standards apply to different scopes. Therefore, establishing a unified ocean data standard is one of the important prerequisites for improving the use rate of machine learning in marine data and the accuracy of the model. Second, in terms of data preprocessing, due to different data sources and diverse data types, sometimes it is necessary to apply different types and different sources of data to the machine learning model. It takes a long time to analyze the data before building the model. For preprocessing, the fusion technology of various marine data will be improved in the future to improve the efficiency of model building. Finally, in terms of data volume, although a large amount of ocean observation data is generated every day, the types of these data are not evenly distributed, and some types of data cannot meet the needs of machine learning, such as source identification and positioning in the ocean. Among them, the amount of marine acoustic data is not enough to meet the training needs of the model. In the future, it is necessary to expand the scope of marine monitoring and increase the amount of various data collection.

4.2 Scope of application

At present, most of the application of machine learning in ocean data is to select data of a specific geographic range for experimental research. There are no comparative experiments for different sea areas. The environment of each sea area is different, and the input data of the model is quite different. Therefore, whether the research method can be applied in other regions is a problem that researchers need to solve in the research process.

4.3 Algorithms comparison

Machine learning has many algorithms, and the application scenarios in the ocean are increasing. In research, most researchers use existing algorithm transformation or multi-algorithm comparison for training, and it is not clear which algorithm should be used in the research field. To be suitable, an optimal solution should be found by constantly adjusting the algorithm, which requires a huge time cost.

5 Application prospects

In the future, machine learning will be widely used in ocean data to prevent natural disasters, marine environment monitoring, marine resource development, marine transportation research, and other fields. Through the continuous expansion of marine data, it will promote the development of the marine industry.

In the prevention of natural disasters, the use of data collected by marine sensors, meteorological satellites and other observation methods, through the analysis of machine learning algorithms, improves the level of forecasting and early warning of severe maritime weather in coastal areas and reduces the loss of life and property. A typhoon is a tropical cyclone that carries huge amounts of energy. Wherever it goes, it may bring natural disasters such as squalls and rains to people. The emergence of machine learning will improve the traditional typhoon prediction model and make the prediction more accurate [66–68].

In the field of marine environmental monitoring, the three-dimensional monitoring network composed of sea, land and air is used to monitor the entire sea area, and machine learning is used to better identify marine environmental problems such as marine red tides, storm surges, sea waves, sea ice, and marine oil spills. In recent years, with the increasing frequency of offshore oil exploration and development and marine transportation activities, and frequent oil spills, marine oil spill pollution has become one of the most important threats to the marine environment. The analysis and processing of ocean image data using CNN can effectively classify and identify oil spills. [69–72]. As shown in Fig. 5, the output of each layer of CNN is the input of the next layer. Through the convolutional layer, the features are extracted. Through the pooling layer, similar features are merged to reduce the amount of data and generalize general features. The fully connected layer merges the results after convolution and pooling. With the rapid development of the aquaculture industry, the degree of eutrophication has intensified, and the concentration and structure of nutrients in the water body have changed, resulting in the frequent occurrence of harmful algal blooms. In the future, the water environment can be predicted and identified through machine learning. Prevent red tide from polluting sea water [73–75].

In the field of Marine resource development, machine learning in marine fish identification and monitoring technology is applied to fishery development, so as to make fishery fishing more efficient and create higher economic benefits [76, 77]. In addition, it can regulate fishing behaviors, prevent ecological damage, and facilitate the supervision of law enforcement personnel [78, 79].

With the development of economic globalization, more and more ships are participating in maritime trade. Unlike

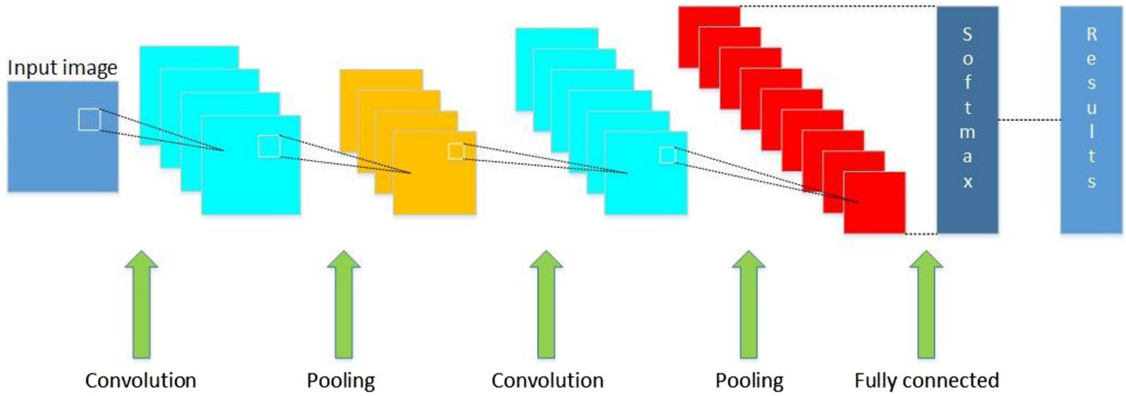

Fig. 5 Convolutional neural network

land transportation, maritime transportation has the characteristics of uncertain ship routes, which makes maritime traffic monitoring increasingly difficult. In the near future, it is expected to use machine learning to improve maritime traffic conditions and calculate ship density through ocean big data combined with ship traffic information and port and waterway information; reduce traffic accidents caused by harsh ocean environments; assess the risk of sailing routes through harsh environments and remote areas [80–84].

Over the past period of time, wireless communication technology has made considerable progress [85–89]. A large amount of ocean observation data is wireless transmission data. It is especially important to properly manage these wireless sensors to ensure that they can reliably and continuously transmit data [90]. In the future, while using machine learning to efficiently analyze wireless transmission data, wireless sensors can also be reasonably controlled [91].

Today, neuromorphic computing is very popular. Silicon neurons provide a medium that can simulate neural networks directly in hardware, and they are more suitable for real-time large-scale neural simulations than those performed on general purpose computers [92]. Multicompartment emulation is an important step to enhance the biological realism of neuromorphic system and to further understand the computational power of neurons. It can accurately reproduce the biodynamics of a single neuron. So far, scientists have proposed a neuromorphic structure that can be used to realize a large-scale biologically meaningful neural network with one million multicompartment neurons [93]. By combining work on event-based neuromorph systems, activity-driven event-based vision sensors can quickly output compressed digital data in the form of events [94]. Based on this, underwater identification will become more efficient in the future.

In addition, the analysis of ocean data through machine learning may also solve scientific issues such as global warming, sea level rise, and “La Madre” [95, 96]. “La Madre” is also known as the “Pacific Decade Oscillation”, which alternately appears over the Pacific in two forms of “warm phase” and “cold phase”. Each phenomenon lasts for 20 to 30 years. When the “La Madre” phenomenon appears in the form of “warm phase”, the water temperature of the sea near the North American continent will rise abnormally, while the temperature of the North Pacific ocean surface will drop abnormally; when the “cold phase” appears, the situation is just the opposite. The cold phase period is a period of concentrated outbreaks of global strong earthquakes. The development of machine learning and ocean data mining technology may provide a new idea for the study and prediction of “La Madre” phenomenon.

6 Conclusions

From the various applications of machine learning in ocean data at this stage, it can be seen that machine learning has changed the traditional way of manually performing ocean data analysis, improved the efficiency of data analysis, and provided solutions for specific scientific research problems in this field. The new method is of great significance for revealing the laws of the ocean, protecting the ocean ecological environment, and developing the marine economy. At the same time, when machine learning and ocean data are combined, there are still problems such as inconsistent data standards, low data utilization, small application scope, and unclear algorithm usage. With the continuous development of machine learning and ocean observation technology, it is believed that in the future, the application range of machine learning and ocean data will be wider, the application will become cheaper and the application efficiency will be higher.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant no. 61902203, Key Research and Development Plan—Major

Scientific and Technological Innovation Projects of ShanDong Province (2019JZZY020101).

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

References

1. Jin, X., Wah, B.W., Cheng, X., Wang, Y.: Significance and challenges of big data research. Big Data Res. **2**(2), 59–64 (2015)
2. Shuai, L., Ge, C., Ying-Jie, L., Feng-Lin, T.: Research and analysis on marine big data applied technology. Periodical of Ocean University of China (2020)
3. Riser, S.C., Freeland, H.J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., Gilbert, D., Xu, J., Pouliquen, S., Thresher, A., et al.: Fifteen years of ocean observations with the global Argo array. Nat. Clim. Change **6**(2), 145–153 (2016)
4. Shi, R., Gan, Y., Wang, Y.: Evaluating scalability bottlenecks by workload extrapolation. In: 2018 IEEE 26th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), pp. 333–347 (2018). https://doi.org/10.1109/MASCOTS.2018.00039
5. Deo, R.C., Şahin, M.: Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern australia. Atmos. Res. **153**, 512–525 (2015)
6. Rasouli, K., Hsieh, W.W., Cannon, A.J.: Daily streamflow forecasting by machine learning methods with weather and climate inputs. J. Hydrol. **414**, 284–293 (2012)
7. Kim, Y.H., Im, J., Ha, H.K., Choi, J.K., Ha, S.: Machine learning approaches to coastal water quality monitoring using GOCI satellite data. GISci. Remote Sens. **51**(2), 158–174 (2014)
8. Rosso, I., Mazloff, M.R., Talley, L.D., Purkey, S.G., Freeman, N.M., Maze, G.: Water mass and biogeochemical variability in the kerguelen sector of the southern ocean: A machine learning approach for a mixing hot spot. J. Geophys. Res. Oceans **125**(3), e2019JC015877 (2020)
9. Mosavi, A., Ozturk, P., Chau, K.W.: Flood prediction using machine learning models. Lit. Rev. Water **10**(11), 1536 (2018)
10. Sun, M., Yu, F.U., Chongjing, L., Jiang, X.: Deep learning application in marine big data mining. Sci. Technol. Rev. **36**(17), 83–90 (2018). http://www.kjdb.org/CN/10.3981/j.issn.1000-7857.2018.17.010
11. Zhou, L., Pan, S., Wang, J., Vasilakos, A.V.: Machine learning on big data: opportunities and challenges. Neurocomputing **237**, 350–361 (2017)
12. Asefa, T., Kemblowski, M., McKee, M., Khalil, A.: Multi-time scale stream flow predictions: the support vector machines approach. J. Hydrol. **318**(1–4), 7–16 (2006)
13. Guilford, T., Meade, J., Willis, J., Phillips, R.A., Boyle, D., Roberts, S., Collett, M., Freeman, R., Perrins, C.: Migration and stopover in a small pelagic seabird, the manx shearwater *Puffinus puffinus*: insights from machine learning. Proc. R. Soc. B Biol. Sci. **276**(1660), 1215–1223 (2009)
14. Krinitskiy, M.: Application of machine learning methods to the solar disk state detection by all-sky images over the ocean. Oceanology **57**(2), 265–269 (2017)
15. Deo, M.: Artificial neural networks in coastal and ocean engineering. Indian J. Geo-Mar. Sci. **39**(4), 589–596 (2010)
16. Jordan, M.I., Mitchell, T.M.: Machine learning: trends, perspectives, and prospects. Science **349**(6245), 255–260 (2015)
17. Alpaydin, E.: Introduction to Machine Learning. MIT Press, Cambridge (2020)
18. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature **323**(6088), 533–536 (1986)
19. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006)
20. Baggeroer, A.B., Kuperman, W.A., Mikhalevsky, P.N.: An overview of matched field methods in ocean acoustics. IEEE J. Ocean. Eng. **18**(4), 401–424 (1993)
21. Baggeroer A.B., Kuperman W.A.: Matched field processing in ocean acoustics. In: Moura J.M.F., Lourtie I.M.G. (eds.) Acoustic Signal Processing for Ocean Exploration. NATO ASI Series (Series C: Mathematical and Physical Sciences), vol 388. Springer, Dordrecht (1993). https://doi.org/10.1007/978-94-011-1604-6_8
22. Niu, H., Reeves, E., Gerstoft, P.: Source localization in an ocean waveguide using supervised machine learning. J. Acoust. Soc. Am. **142**(3), 1176–1188 (2017)
23. Choi, J., Choo, Y., Lee, K.: Acoustic classification of surface and underwater vessels in the ocean using supervised machine learning. Sensors **19**(16), 3492 (2019)
24. Steinberg, B.Z., Beran, M.J., Chin, S.H., Howard Jr., J.H.: A neural network approach to source localization. J. Acoust. Soc. Am. **90**(4), 2081–2090 (1991)
25. Caiti, A., Parisini, T.: Mapping ocean sediments by RBF networks. IEEE J. Ocean. Eng. **19**(4), 577–582 (1994)
26. Niu, H., Ozanich, E., Gerstoft, P.: Ship localization in Santa Barbara channel using machine learning classifiers. J. Acoust. Soc. Am. **142**(5), EL455–EL460 (2017)
27. Van Komen, D.F., Neilsen, T.B., Howarth, K., Knobles, D.P., Dahl, P.H.: Seabed and range estimation of impulsive time series using a convolutional neural network. J. Acoust. Soc. Am. **147**(5), EL403–EL408 (2020)
28. Cane, M.A., Clement, A.C., Kaplan, A., Kushnir, Y., Pozdnyakov, D., Seager, R., Zebiak, S.E., Murtugudde, R.: Twentieth-century sea surface temperature trends. Science **275**(5302), 957–960 (1997)
29. Castro, S.L., Wick, G.A., Steele, M.: Validation of satellite sea surface temperature analyses in the Beaufort Sea using UpTempO buoys. Remote Sens. Environ. **187**, 458–475 (2016)
30. Chaidez, V., Dreano, D., Agusti, S., Duarte, C.M., Hoteit, I.: Decadal trends in red sea maximum surface temperature. Sci. Rep. **7**(1), 1–8 (2017)
31. Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., Chen, Z.: Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. Remote Sens. Environ. **233**, 111358 (2019)
32. Xiao, C., Chen, N., Hu, C., Wang, K., Xu, Z., Cai, Y., Xu, L., Chen, Z., Gong, J.: A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data. Environ. Model. Softw. **120**, 104502 (2019)
33. Lins, I.D., Araujo, M., das Chagas Moura, M., Silva, M.A., Droguett, E.L.: Prediction of sea surface temperature in the tropical Atlantic by support vector machines. Comput. Stat. Data Anal. **61**, 187–198 (2013)
34. Olah, C.: Understanding LSTM Networks, August 2015. http://colah.github.io/posts/2015-08-Understanding-LSTMs/
35. Savitha, R., Al Mamun, A., et al.: Regional ocean wave height prediction using sequential learning neural networks. Ocean Eng. **129**, 605–612 (2017)
36. Group, T.W.: The wam model—a third generation ocean wave prediction model. J. Phys. Oceanogr. **18**(12), 1775–1810 (1988)
37. Booij, N., Ris, R.C., Holthuijsen, L.H.: A third-generation wave model for coastal regions: 1. Model description and validation. J. Geophys. Res. Oceans **104**(C4), 7649–7666 (1999)
38. Tolman, H.L., Chalikov, D.: Source terms in a third-generation wind wave model. J. Phys. Oceanogr. **26**(11), 2497–2518 (1996)

39. Makarynskyy, O.: Improving wave predictions with artificial neural networks. Ocean Eng. **31**(5–6), 709–724 (2004)
40. Agrawal, J., Deo, M.: On-line wave prediction. Mar. Struct. **15**(1), 57–74 (2002)
41. Jain, P., Deo, M.: Artificial intelligence tools to forecast ocean waves in real time. Open Ocean Eng. J. **1**, 13–20 (2008)
42. James, S.C., Zhang, Y., O'Donncha, F.: A machine learning framework to forecast wave conditions. Coast. Eng. **137**, 1–10 (2018)
43. Rao, S., Mandal, S.: Hindcasting of storm waves using neural networks. Ocean Eng. **32**(5–6), 667–684 (2005)
44. Mahjoobi, J., Mosabbeb, E.A.: Prediction of significant wave height using regressive support vector machines. Ocean Eng. **36**(5), 339–347 (2009)
45. Quan, J., Feng, H., Yong-Zeng, Y.: Prediction of the significant wave height based on the support vector machine. Adv. Mar. Sci. **37**(2), 199–209 (2019)
46. Alexandre, E., Cuadra, L., Nieto-Borge, J., Candil-Garcia, G., Del Pino, M., Salcedo-Sanz, S.: A hybrid genetic algorithm-extreme learning machine approach for accurate significant wave height reconstruction. Ocean Model. **92**, 115–123 (2015)
47. Salcedo-Sanz, S., Borge, J.N., Carro-Calvo, L., Cuadra, L., Hessner, K., Alexandre, E.: Significant wave height estimation using SVR algorithms and shadowing information from simulated and real measured X-band radar images of the sea surface. Ocean Eng. **101**, 244–253 (2015)
48. Durán-Rosal, A., Hervás-Martínez, C., Tallón-Ballesteros, A., Martínez-Estudillo, A., Salcedo-Sanz, S.: Massive missing data reconstruction in ocean buoys with evolutionary product unit neural networks. Ocean Eng. **117**, 292–301 (2016)
49. Franz, K., Roscher, R., Milioto, A., Wenzel, S., Kusche, J.: Ocean eddy identification and tracking using neural networks. In: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 6887–6890. IEEE (2018)
50. Bai, X., Wang, C., Li, C.: A streampath-based RCNN approach to ocean eddy detection. IEEE Access **7**, 106336–106345 (2019)
51. Lguensat, R., Sun, M., Fablet, R., Tandeo, P., Mason, E., Chen, G.: Eddynet: A deep neural network for pixel-wise classification of oceanic eddies. In: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 1764–1767. IEEE (2018)
52. Bolton, T., Zanna, L.: Applications of deep learning to ocean data inference and subgrid parameterization. J. Adv. Model. Earth Syst. **11**(1), 376–399 (2019)
53. May, R.M.: Conceptual aspects of the quantification of the extent of biological diversity. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. **345**(1311), 13–20 (1994)
54. Ormond, R.: Marine biodiversity: causes and consequences. J. Mar. Biol. Assoc. U. K. **76**(1), 151–152 (1996)
55. Wei, C.L., Rowe, G.T., Escobar-Briones, E., Boetius, A., Soltwedel, T., Caley, M.J., Soliman, Y., Huettmann, F., Qu, F., Yu, Z., et al.: Global patterns and predictions of seafloor biomass using random forests. PloS One **5**(12), e15323 (2010)
56. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cognit. Neurosci. **3**(1), 71–86 (1991)
57. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Mullers, K.R.: Fisher discriminant analysis with kernels. In: Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (cat. no. 98th8468), pp. 41–48. IEEE (1999)
58. Huang, P.X.: Hierarchical classification system with reject option for live fish recognition. In: Fish4Knowledge: Collecting and Analyzing Massive Coral Reef Fish Video Data, pp. 141–159. Springer (2016)
59. Siddiqui, S.A., Salman, A., Malik, M.I., Shafait, F., Mian, A., Shortis, M.R., Harvey, E.S.: Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. ICES J. Mar. Sci. **75**(1), 374–389 (2018)
60. Reus, G., Möller, T., Jäger, J., Schultz, S.T., Kruschel, C., Hasenauer, J., Wolff, V., Fricke-Neuderth, K.: Looking for seagrass: deep learning for visual coverage estimation. In: 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO), pp. 1–6. IEEE (2018)
61. Glotin, H., Spong, P., Symonds, H., Roger, V., Balestriero, R., Ferrari, M., Poupard, M., Towers, J., Veirs, S., Marxer, R., et al.: Deep learning for ethoacoustical mapping: application to a single cachalot long term recording on joint observatories in vancouver island. J. Acoust. Soc. Am. **144**(3), 1776–1777 (2018)
62. Bermant, P.C., Bronstein, M.M., Wood, R.J., Gero, S., Gruber, D.F.: Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. Sci. Rep. **9**(1), 1–10 (2019)
63. Al-Barazanchi, H., Verma, A., Wang, S.X.: Intelligent plankton image classification with deep learning. Int. J. Comput. Vis. Robot. **8**(6), 561–571 (2018)
64. Hari, V.N., Kalyan, B., Chitre, M., Ganesan, V.: Spatial modeling of deep-sea ferromanganese nodules with limited data using neural networks. IEEE J. Ocean. Eng. **43**(4), 997–1014 (2017)
65. Jie, W.L., Kalyan, B., Chitre, M., Vishnu, H.: Polymetallic nodules abundance estimation using sidescan sonar: a quantitative approach using artificial neural network. In: OCEANS 2017-Aberdeen, pp. 1–6. IEEE (2017)
66. Jiang, G.Q., Xu, J., Wei, J.: A deep learning algorithm of neural network for the parameterization of typhoon-ocean feedback in typhoon forecast models. Geophys. Res. Lett. **45**(8), 3706–3716 (2018)
67. Hashemi, M.R., Spaulding, M.L., Shaw, A., Farhadi, H., Lewis, M.: An efficient artificial intelligence model for prediction of tropical storm surge. Nat. Hazards **82**(1), 471–491 (2016)
68. Zhang, C., Durgan, S.D., Lagomasino, D.: Modeling risk of mangroves to tropical cyclones: a case study of hurricane IRMA. Estuar. Coast. Shelf Sci. **224**, 108–116 (2019)
69. Khlongkhoi, P., Chayantrakom, K., Kanbua, W.: Application of a deep learning technique to the problem of oil spreading in the Gulf of Thailand. Adv. Differ. Equ. **2019**(1), 306 (2019)
70. Topouzelis, K., Psyllos, A.: Oil spill feature selection and classification using decision tree forest on SAR image data. ISPRS J. Photogramm. Remote Sens. **68**, 135–143 (2012)
71. Xu, L., Li, J., Brenning, A.: A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. Remote Sens. Environ. **141**, 14–23 (2014)
72. Brekke, C., Solberg, A.H.: Classifiers and confidence estimation for oil spill detection in ENVISAT ASAR images. IEEE Geosci. Remote Sens. Lett. **5**(1), 65–69 (2008)
73. Grasso, I., Archer, S.D., Burnell, C., Tupper, B., Rauschenberg, C., Kanwit, K., Record, N.R.: The hunt for red tides: deep learning algorithm forecasts shellfish toxicity at site scales in coastal maine. Ecosphere **10**(12), e02960 (2019)
74. Bak, S.H., Hwang, D.H., Kim, H.M., Kim, B.K., Enkgjargal, U., Oh, S.Y., Yoon, H.J.: A study on red tide detection technique by using multi-layer perceptron. Int. J. Grid Distrib. Comput. **11**(9), 93–102 (2018)
75. Fdez-Riverola, F., Corchado, J.M.: Fsfrt: forecasting system for red tides. a hybrid autonomous ai model. Appl. Artif. Intell. **17**(10), 955–982 (2003)
76. Sala, E., Mayorga, J., Costello, C., Kroodsma, D., Palomares, M.L., Pauly, D., Sumaila, U.R., Zeller, D.: The economics of fishing the high seas. Sci. Adv. **4**(6), 2504 (2018)
77. Fernandes, J.A., Irigoien, X., Goikoetxea, N., Lozano, J.A., Inza, I., Pérez, A., Bode, A.: Fish recruitment prediction, using robust

supervised classification methods. Ecol. Model. **221**(2), 338–352 (2010)
78. Stamoulis, K.A., Delevaux, J.M., Williams, I.D., Poti, M., Lecky, J., Costa, B., Kendall, M.S., Pittman, S.J., Donovan, M.K., Wedding, L.M., et al.: Seascape models reveal places to focus coastal fisheries management. Ecol. Appl. **28**(4), 910–925 (2018)
79. de Souza, E.N., Boerder, K., Matwin, S., Worm, B.: Improving fishing pattern detection from satellite AIS using data mining and machine learning. PloS One **11**(7), e0158248 (2016)
80. Ning, J., Huang, T., Diao, B., et al.: A fine grained grid-based maritime traffic density algorithm for mass ship trajectory data. Comput. Eng. Sci. **37**(12), 2242–2249 (2015)
81. Kim, D., Park, M.S., Park, Y.J., Kim, W.: Geostationary ocean color imager (GOCI) marine fog detection in combination with Himawari-8 based on the decision tree. Remote Sens. **12**(1), 149 (2020)
82. Tang, J., Deng, C., Huang, G.B., Zhao, B.: Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine. IEEE Trans. Geosci. Remote Sens. **53**(3), 1174–1185 (2014)
83. Khan, B., Khan, F., Veitch, B., Yang, M.: An operational risk analysis tool to analyze marine transportation in arctic waters. Reliabil. Eng. Syst. Saf. **169**, 485–502 (2018)
84. Trucco, P., Cagno, E., Ruggeri, F., Grande, O.: A bayesian belief network modelling of organisational factors in risk analysis: a case study in maritime transportation. Reliabil. Eng. Syst. Saf. **93**(6), 845–856 (2008)
85. Wen, M., Chen, X., Li, Q., Basar, E., Wu, Y.C., Zhang, W.: Index modulation aided subcarrier mapping for dual-hop OFDM relaying. IEEE Trans. Commun. **67**(9), 6012–6024 (2019)
86. Wen, M., Zheng, B., Kim, K.J., Di Renzo, M., Tsiftsis, T.A., Chen, K.C., Al-Dhahir, N.: A survey on spatial modulation in emerging wireless systems: Research progresses and applications. IEEE J. Sel. Areas Commun. **37**(9), 1949–1972 (2019)
87. Wen, M., Li, Q., Basar, E., Zhang, W.: Generalized multiple-mode OFDM with index modulation. IEEE Trans. Wirel. Commun. **17**(10), 6531–6543 (2018)
88. Wen, M., Basar, E., Li, Q., Zheng, B., Zhang, M.: Multiple-mode orthogonal frequency division multiplexing with index modulation. IEEE Trans. Commun. **65**(9), 3892–3906 (2017)
89. Wen, M., Ye, B., Basar, E., Li, Q., Ji, F.: Enhanced orthogonal frequency division multiplexing with index modulation. IEEE Trans. Wirel. Commun. **16**(7), 4786–4801 (2017)
90. Li, Y., Zhang, Y., Li, W., Jiang, T.: Marine wireless big data: efficient transmission, related applications, and challenges. IEEE Wirel. Commun. **25**(1), 19–25 (2018)
91. Park, S., Byun, J., Shin, K.S., Jo, O.: Ocean current prediction based on machine learning for deciding handover priority in underwater wireless sensor networks. In: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 505–509. IEEE (2020)
92. Indiveri, G., Linares-Barranco, B., Hamilton, T., van Schaik, A., Etienne-Cummings, R., Delbruck, T., Liu, S.C., Dudek, P., Häfliger, P., Renaud, S., Schemmel, J., Cauwenberghs, G., Arthur, J., Hynna, K., Folowosele, F., SAÏGHI, S., Serrano-Gotarredona, T., Wijekoon, J., Wang, Y., Boahen, K.: Neuromorphic silicon neuron circuits. Front. Neurosci. **5**, 73 (2011). https://doi.org/10.3389/fnins.2011.00073
93. Yang, S., Deng, B., Wang, J., Li, H., Lu, M., Che, Y., Wei, X., Loparo, K.A.: Scalable digital neuromorphic architecture for large-scale biophysically meaningful neural network with multi-compartment neurons. IEEE Trans. Neural Netw. Learn. Syst. **31**(1), 148–162 (2020). https://doi.org/10.1109/TNNLS.2019.2899936
94. Delbrück, T., Linares-Barranco, B., Culurciello, E., Posch, C.: Activity-driven, event-based vision sensors. In: Proceedings of 2010 IEEE International Symposium on Circuits and Systems, pp. 2426–2429 (2010). https://doi.org/10.1109/ISCAS.2010.5537149
95. D'Alelio, D., Rampone, S., Cusano, L.M., Morfino, V., Russo, L., Sanseverino, N., Cloern, J.E., Lomas, M.W.: Machine learning identifies a strong association between warming and reduced primary productivity in an oligotrophic ocean gyre. Sci. Rep. **10**(1), 1–12 (2020)
96. Su, H., Li, W., Yan, X.H.: Retrieving temperature anomaly in the global subsurface and deeper ocean from satellite observations. J. Geophys. Res. Oceans **123**(1), 399–410 (2018)