Automated Detection of Paleoenvironmental Proxy, Eucampia Index, in a Microscopic Slide Using a Convolutional Neural Network System

Saki Ishino (ishino.saki@aist.go.jp)
Geological Survey of Japan / AIST, Institute of Geoinformation and Geology, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8567, Japan  https://orcid.org/0000-0003-1100-6515

Takuya Itaki
National Institute of Advanced Industrial Science and Technology, Geological Survey of Japan

Methodology

Keywords: Artificial intelligence, Automated classification, Diatoms, Microfossils, Paleoenvironment

DOI: https://doi.org/10.21203/rs.3.rs-88945/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The *Eucampia* Index, which is calculated from valve ratio of Antarctic diatom *Eucampia ainarctica* varieties, has been expected to be a useful indicator of sea ice coverage or/and sea surface temperature variation in the Southern Ocean. To verify the relationship between the index value and the environmental factors, considerable effort is needed to classify and count valves of *E. antarctica* in a very large number of samples. In this study, to realize automated detection of the *Eucampia* Index, we constructed a deep-learning (one of the learning methods of artificial intelligence) based models for identifying *Eucampia* valves from various particles in a diatom slide. The microfossil Classification and Rapid Accumulation Device (miCRAD) system, which can be used for scanning a slide and cropping images of particles automatically, was employed to collect images in training dataset for the model and test dataset for model verification. As a result of classifying particle images in the test dataset by the initial model "Eant_1000px_200616", accuracy was 78.8%. The *Eucampia* Index value prepared in the test dataset was 0.80, and the value predicted using the developed model from the same dataset was 0.76. The predicted value was in the range of the manual counting error. These results suggest that the classification performance of the model is similar to that of a human expert. This study revealed that a model capable of detecting the ratio of two diatom species can be constructed using the miCRAD system for the first time. The miCRAD system connected with the developed model in this study is capable of automatically classifying particle images at the same time of capturing images so that the system can be applied to a large-scale analysis of the *Eucampia* index in the Southern Ocean. Depending on the setting of the classification category, similar method is relevant to investigators who have to process a large number of diatom samples such as for detecting specific species for biostratigraphic and paleoenvironmental studies.

Introduction

Numerous studies of diatom valves have focused on assemblage structure, abundance, and morphometric changes. These analyses have provided useful information on, not only the ecological aspects of various diatom taxa, but also paleoenvironmental and chronological aspects of diatom proxies. Most of these studies require large amounts of data and considerable effort is needed to classify and count hundreds of diatom valves (e.g., Armand et al., 2005; Gersonde et al., 2005; Shukla et al., 2013, Esper and Gersonde, 2014). For example, although the *Eucampia* Index have been inferred to be a useful indicator of sea ice coverage or/and sea surface temperature (SST) variation in the Southern Ocean (Whitehead et al., 2005; Allen, 2014), verifying relationship between the index value and the environmental factors needs a very large number of the diatom valve observations. The *Eucampia* Index has been used to assess the ratio of intercalary valves to total valves (i.e., intercalary and terminal valves) of *Eucampia antarctica*, a species that is endemic to the Southern Ocean (Kaczmarska et al., 1993). The intercalary and terminal valves are identified by the shape of their 'horns' (Fig. 1); the horns of the intercalary valves are flattened while those of the terminal valves are pointed. Although these valves can be easily identified in a diatom assemblage, counting them in diatom slides requires using a field of view of more than ten
times when their relative abundances in a diatom assemblage are low. Therefore, it is expected that an automated identification system for diatom species would help researchers to collect and analyze large amounts of data in studies that involve counting the presence of specific diatoms such as *Eucampia antarctica*.

Recent studies have succeeded in automating the classification and morphometric analysis of some diatom species using techniques that combine artificial intelligence and microscopic imaging. For example, some studies have focused on automating morphometric measurements of diatom valves (e.g., Sepaulding et al., 2013; Kloster et al., 2014; 2017) and these methods have been demonstrated to be useful for analyses of certain species that are abundant in a sample. Other studies have presented classification techniques that automatically predict the correct taxon name from an image sample containing a single diatom (e.g., Schulze et al., 2013; Pappas et al., 2014; Bueno et al., 2017). These studies have typically employed machine learning methods that use handcrafted features, which describe the discriminant properties of diatoms, such as valve area, shape, and texture. Pedraza et al. (2017) approached the task of automatic identification using a deep-learning model and reported 99% accuracy for 80 species.

As automation systems for diatom identification advance, it has become apparent that the system design varies depending on the methods and data used. A system that automatically count particular species such as *Eucampia antarctica*, even if their abundance in the assemblage is low, would be a powerful tool in paleoenvironmental research using sediments. However, the devices and programs that have been developed to date are generally not suitable for these studies, because these workflows typically require hundreds of manually cropped images of each object of interest to construct the datasets for the model. Preparing hundreds of images of particular species and/or valve types also requires an amount of effort when the species is found rarely. The classification models that have been developed to date have focused on performing highly accurate and precise identification of major modern diatoms, therefore not focused on ease to construct automatic classification models.

In this study, to automate detection of *Eucampia* Index, we employed a newly developed system, the microfossil Classification and Rapid Accumulation Device (miCRAD), which was used for automatic detection of radiolarians (Itaki et al., 2020). We constructed a classification model for identifying *Eucampia* valves in a diatom slide and verified the usefulness of the model for paleoenvironmental studies. The system can be used to construct deep-learning based models for classifying some species easily and rapidly by automated collecting images of particles present in a normal microscopic slide (Itaki et al., 2020). Since the resolution of the original imaging system was insufficient for classifying diatom valves to species, we increased the magnification of the objective lens and the resolution of the charge-coupled device (CCD) camera used for microscopic observations.

**Methods**
Preparing samples for automatic detection involved the following three steps: (1) microscopic slides were prepared for diatom observation, (2) a dataset of digital images of particles in the slides was compiled, and (3) a deep learning model was created using software. These steps and verification of the created model are described below.

**Slide preparation**

Normal slides for light microscopy observation were prepared using surface sediments. The sediments were collected using gravity and piston cores by the Technology Research Center of Japan National Oil Corporation (JNOC) on TH83 cruises undertaken in 1983. The sample names used in this study and its core sites are listed in Supplementary Table 1.

Methods for sediment treatment and slide preparation were the same those used for manual counts of fossil diatoms as follows. Approximately 0.1 g of dried sediments were placed in 200 ml beakers containing approximately 1 ml hydrogen peroxide (H2O2, 10%) and hydrochloric acid (HCl, 10%), and boiled to remove organic and calcareous materials. Distilled water was then added to a volume of 200 ml and left for 5 h to separate the residues and acidic water. The residue was separated by decanting the supernatant, and the beaker was refilled again with distilled water. This process was repeated four times to neutralize suspension. Approximately 100 µl was then taken from a 100-fold dilution of the agitated suspension and dried on a 24 × 32 mm cover glass and mounting media was added (Norland optical adhesive No. 61, refractive index: 1.56) before curing under UV light. The reason for using cover glasses measuring 24 × 32 mm, which are bigger than ones typically used for manual observations, was to ensure that the particles were sparsely distributed.

**Image acquisition of E. antarctica varieties and construction of training dataset**

Images of all particles including *E. antarctica* valves in the prepared slides were captured using the Image Collection Unit of the miCRAD system described in Itaki et al. (2020). The Image Collection Unit, which is based on “Collection Pro” from Micro Support Co., Ltd., automatically acquires digital microscopic images of particles scattered in the observation field using an electric X-Y stage microscope controlled by a computer. The field of view was projected on a display using a · 50 objective lens, a 5 million-pixel CCD camera, and · 6 in transmitted light mode. After scanning a slide, individual particle images were clipped to a size 1,000 · 1,000 pixels. In cases when particles overlap, they are erroneously recognized as a single individual by the image processing software. Therefore, by adapting the settings for particle separation and contrast recognition in the software and making sparse slides, most particles are isolated successfully and captured singly in a clipped image (Figs. 1-(b) and (c)).

**Construction of the classification model**

The classification model for distinguishing intercalary and terminal valves of *E. antarctica* from other particles in a diatom slide was constructed using the Classification Unit of the miCRAD system described by Itaki et al. (2020). The Classification Unit consists of deep learning software “RAPID machine learning”
(NEC Corp.), which incorporates a convolutional neural network (CNN). To construct a model, images in a training dataset were manually labeled when they were imported to the unit. The learning repetitions were typically set to 30 epochs. Errors for each epoch were calculated as a loss function using cross-entropy.

To distinguish intercalary and terminal valves of *E. antarctica* valves from other sediment particles, three categories were used for the image classification ([Terminal], [Intercalary], and [Other particles]) in an initial model, which we referred to as “Eant_1000px_200616”. The images showing one terminal and intercalary valves of *E. antarctica* were assigned to the categories of [Terminal] and [Intercalary], respectively. *Eucampia antarctica* observed in the Southern Ocean contains two varieties that are *E. antarctica* var. *recta* (Mangin) Fryxell & Prasad and *E. antarctica* var. *antarctica* (Castracane) Mangin. In this study, the two varieties were classified as one group of *E. antarctica* and criteria of identification were adapted from descriptions in Fryxell & Prasad (1990). In the [Other particles] category, images displaying other diatom species, fragments of diatom valves, more than two valves, and other particles that were either not diatoms or that contained unclear shapes were selected randomly. [Other particles] also included valves that were smaller than one-half the size of a typical *E. antarctica* valve.

We manually selected 505, 969, and 5000 images for [Terminal], [Intercalary], and [Other particles], respectively, from more than 100,000 images collected from 15 training samples (Supplementary Table 1). A training dataset for the model was built with the original images and generated images that were rotated, flipped, brightened or darkened, and shifted horizontally. These data expansions were prepared using the Keras package, which is a free Python library (https://keras.io). The total number of images in the training dataset were 10650, 10659, and 5000 for [Terminal], [Intercalary], and [Other particles], respectively.

### Verification tests of the classification model

Differences between the *Eucampia* Index detected manually and automatically in a test dataset were verified to confirm the usefulness of the created model. Core sediments from site G501 off Prydz Bay in East Antarctica were used for the test dataset (Supplementary Table 1). The site of the test dataset was differentiated from those of the training dataset because the application of the constructed model for unknown samples was considered.

In the test dataset, 264, 1047, and 998 images of [Terminal], [Intercalary], and [Other particles], respectively, were prepared using the Image Collection Unit with the same parameters as those used for the training dataset. To confirm the model aptitude for classifying *E. antarctica* valves that were variable in shape, the test dataset included generated images of the original terminal and intercalary valves prepared using Keras package. A total of 120 original *E. antarctica* images were used, which is sufficient for representing the *Eucampia* Index of each sample (Whitehead et al., 2005).

The *Eucampia* Index was calculated using a method that was similar to that of Kaczmarska et al. (1993). The equation is as follows:
The Eucampia Index

\[
\text{No. of intercalary valves counted per a slide} = \frac{\text{No. of terminal and intercalary valves counted per a slide}}{\text{No. of terminal and intercalary valves counted per a slide}}
\]

Note that only the oblique and girdle views of *E. antarctica* valves were counted. Images of the valve view were not included in the training and test datasets, and they were not used to calculate the *Eucampia* Index, because valve horns in each image were often broken or out of focus.

**Results And Discussions**

The initial model, Eant_1000px_200616, was constructed without overfitting and the error was 0.136 (Supplementary Fig. 2). Verification tests for the model were performed with a test dataset prepared using sample JNOC-G501 and confidence values for each category in all images, which were calculated by the Classification Unit using the softmax function (Supplementary Table 2). All test images that were predicted either correctly or incorrectly were compiled in the Supplementary Image Dataset. A prediction of the category for each image was made by selecting the ones assigned the highest confidence value. In this study, of the three categories, the ones that are assigned the first or second highest confidence values are referred to as the 1st and 2nd categories, respectively. For convenience, a manually classified category that is assigned to an object for preparing test datasets is referred to as a “True-category” in this paper. The evaluation of the model using a test dataset and the accuracy of the *Eucampia* Index calculated by the model are described below.

**Verification tests of the classification model**

The components of the 1st category are shown in Table 1. For each group, i.e., [Terminal], [Intercalary], and [Other particles], 57%, 77%, and 87% of the images were predicted correctly, respectively. The overall accuracy evaluated from these results was 78.8%. The accuracy was not as high as that for the CNN model reported by Bueno et al. (2017), which described the first classification model of diatom valves. However, overall, the correct predictions made by the model showed a tendency towards having higher confidence values than incorrect predictions. The predicted number of each confidence value range estimated for 1st category is shown in Fig. 2. The histogram of the number of incorrectly identified images is uniform, indicating that tens of the images occur almost constantly throughout the confidence value range (0.300–1.00) calculated for the 1st category. The number of correctly classified images increases markedly in the confidence range of 0.800–1.00, and is much higher than the number of incorrectly classified ones. In the confidence value range of 0.300–0.599, 124 images were incorrectly classified and 86 were correctly identified (counted from Supplementary Table 2). These findings imply that there are many incorrect predictions when the confidence value for the 1st category is 0.599 or less, and that the images with relatively high values in the range 0.00–0.499 for 2nd category contain the other two categories including true-categories.
Table 1 also shows the type and number of images assigned to the 2nd category when the 1st category was incorrectly assigned. Approximately 70% of the incorrectly predicted images were correctly identified at the stage of selecting the 2nd category. Of 114 true-[Terminal] images predicted incorrectly as [Intercalary], 113 images were assigned to [Terminal] of the 2nd categories. For all of 54 of the true-[Intercalary] images incorrectly predicted as [Terminal] of the 1st category, [Intercalary] was assigned to the 2nd category. Of 109 true-[Intercalary] images predicted as [Other particles], 103 images were identified as belonging to [Intercalary] of the 2nd category. The reason for more images in the true-[Intercalary] category being recognized as [Other particles] compared to the true-[Terminal] category (Table 1) is because of the greater variety in the shape of intercalary valves. *Eucampia antarctica* valves are asymmetrical and therefore have a wide range of aspect ratios (Allen et al. 2014).

Figure 3 shows examples that represent trends recognized visually from the classification results, including original images taken by “Collection Pro” and images generated using the Keras package. The confidence values for all three categories are also described in each image in Fig. 3. The confidence value of the image generated by Keras and the original image are slightly different, so the images are found to be identified as different objects by the model. The correctly predicted objects in the [Terminal] and [Intercalary] categories with markedly higher confidence values in the 1st category had better-preserved valves and the images were in focus (Figs. 3(a)-1–3, 3(b)-1–3). Conversely, some objects in the true-[Terminal] category, which were incorrectly predicted as [Intercalary], had horns that were unclear (Figs. 3(c) and (d)). The true-[Intercalary] objects with longer and unclear horns were incorrectly predicted to be [Terminal] (Figs. 3(e) and 3(f)). Furthermore, the differences in confidence values between the [Terminal] and [Intercalary] categories in these images are smaller than those obtained for the images in Figs. 3(a) and 3(b), indicating that the classification was uncertain. For the out of focus images and images containing more than two particles, [Other particles] was selected as the 1st category (Figs. 3(g) and 3(h)). This classification tendency is probably caused by a criterion in the [Other particles] training dataset, which included out-of-focus and/or two or more particles in an image.

**Eucampia Index comparison between automatic and manual counting**

The results of the model evaluation revealed that confidence values calculated for the test images reflect the degree of similarity between the three categories, i.e., whether the shape of particles in an image resembles a terminal or intercalary valve, or neither. Moreover, the confidence values can indicate the information that includes even the difficulty in classification because of poor preservation of diatom valves, and out-of-focus of images. Shoji et al. (2018) reported that the relative abundances of each category could be shown as the average confidence values using CNN model that learned outline similarity of particles categorized into four. Thus, the average confidence values obtained for the [Terminal] and [Intercalary] categories in this study must reflect similarly a ratio between the two valve types in an image dataset.

To compare the *Eucampia* Indexes that counted manually and predicted by the model, each index was calculated based on the abundances of true-[Intercalary] and true-[Terminal] in the test dataset, and the
average confidence values obtained for them by the model, respectively (Table 2). The Eucampia Index value derived from the average confidence values is 0.76, and the index value estimated from number true-category images was 0.80. Considering the counting probability error was < ± 0.053 when total of 100 E. antarctica valves were used (Whitehead et al., 2005), this result shows that the Eucampia Index detected automatically using the developed model is comparable to those obtained manually.

Future perspectives for automatic diatom detection using the miCRAD system

This study revealed that a model capable of detecting the ratio of two diatom species can be constructed using the miCRAD system for the first time. The Image Collection Unit in the miCRAD system enables researchers not only to obtain cropped object images for training datasets to construct CNN models, but also to conduct automated classification of particle images at the same time of capturing after constructing CNN models (Itaki et al., 2020). Using the model constructed in this study, automatic detection of Eucampia Index from a diatom slide can be applied to a large-scale investigation of the index variation and geographical distribution in the Southern Ocean. Depending on the setting of the classification category, similar method is relevant to investigators who have to process a large number of diatom samples such as for detecting specific species for biostratigraphic and paleoenvironmental studies.

When samples that differ in age and/or sedimentary environment from the specimens used to construct the training dataset are used practically to detect Eucampia Index, loss of model accuracy is presumed to occur if there are a largely different number of images in each category. The test dataset used in this evaluation differed from a normal diatom slide in the number of [Other particles] images. The test dataset used in this evaluation differed from a normal diatom slide in the number of [Other particles] images. The test dataset contained 1311 E. antarctica and 998 [Other particles] images. From a normal diatom slide, for example, from a slide prepared using the site G501 sediment sample, 154 and 1991 images of E. antarctica and [Other particles] were obtained using the Image Collection Unit, respectively. When the images of [Other particles] is detected at overwhelmingly abundant than E. antarctica valves in a slide, then it is predicted that the average confidence values of [Intercalary] and [Terminal] decrease significantly. As a result, the difference between the manually counted Eucampia Index value and the detected value inferred using the average confidence value may be larger because of the relatively larger errors in the average confidence values.

To increase the accuracy of the diatom species detection, it is necessary that many other particles are not captured. Some studies have employed deep-learning-based automatic segmentation techniques to detect each diatom in a field of view (Pedraza et al., 2018; Tang et al., 2018; Ruiz-Santaquiteria et al., 2020). In addition, many studies have been developed the CNN models with sufficient accuracy for diatom classification (e.g., Bueno et al., 2017; Pedraza et al., 2017). These knowledges have contributed the software utility, on the other side, the results of diatom classification using the miCRAD system will contribute the development of devices for practical use. It is expected that new practical and accurate
automatic identification and detection techniques will be realized by further development of the miCRAD system once automatic segmentation is implemented or CNN models constructed by other programs of the Classification Unit can be used.

Conclusion

A classification model for distinguishing intercalary and terminal valves of *E. antarctica* from other particles in a slide was constructed using the miCRAD system, which incorporates a CNN. The training dataset was prepared using the Collection Unit of the miCRAD system, which automatically captures images of micro particles from a normal slide. The *Eucampia* Index value (i.e., the ratio of the number of intercalary valves to the total number of terminal and intercalary valves) estimated using the developed model (Eant_1000px_200616) was comparable to the value calculated manually. The findings suggest that the classification performance of the model is similar to that of a human expert. The model constructed in this study combined with the miCRAD system will be powerful tools to be used in a large-scale analysis of the *Eucampia* Index in the Southern Ocean. This experimental result can be applied to practical use of detecting some diatom species such as environmental and age specific species in huge number of sediment samples.

Declarations

Availability of data and material

The training datasets and the constructed CNN model used in this study are available upon reasonable request from the corresponding author.

Competing interests

The authors declare that they have no competing interest.

Funding

This work was supported by JSPS KAKENHI Grant Numbers 17H06318 and 18H01329.

Authors' contributions

SI proposed the topic, conceived and carried out the experimental study. TI led the development of the microscope system, helped in the interpretation of the experimental study, and acquired fundings for the experiment. Both authors read and approved the final manuscript.

Acknowledgements
The authors are grateful to Minoru Ikehara of Kochi University and Saiko Sugisaki of Geological Survey of Japan for compiling information of JNOC sediment cores, and Masato Ito of Japan Agency for Marine-earth Science and Technology for obtaining samples and data at 59th Japanese Antarctic Research Expedition. We also would like to thank Hitomi Yamazaki for their assistance in the laboratory experiments.

References

1. Allen, C. S. (2014). Proxy development: a new facet of morphological diversity in the marine diatom *Eucampia antarctica* (Castracane) Mangin. J. Micropalaeontol., 33(2), 131–142. doi: 10.1144/jmpaleo2013-025
2. Armand, L. K., Crosta, X., Romero, O., Pichon, J. J. (2005). The biogeography of major diatom taxa in Southern Ocean sediments: 1. Sea ice related species. Palaeogeogr., Palaeoclimatol., Palaeoecol., 223(1–2), 93–126. doi:10.1016/j.palaeo.2005.02.015
3. Bueno, G., Deniz, O., Pedraza, A., Ruiz-Santaquiteria, J., Salido, J., Cristóbal, G., Borrego-Ramos M., Blanco, S. (2017). Automated diatom classification (Part A): handcrafted feature approaches. Appl. Sci., 7(8), 753. doi:10.3390/app7080753
4. Chollot, F. Keras: Deep learning library for theano and tensorflow. https://keras.io (2015).
5. Esper, O., Gersonde, R. (2014). Quaternary surface water temperature estimations: New diatom transfer functions for the Southern Ocean. Palaeogeogr., Palaeoclimatol., Palaeoecol., 414, 1–19. doi:10.1016/j.palaeo.2014.08.008
6. Fryxell, G. A., Prasad, A. K. S. K. (1990). *Eucampia antarctica* var. *recta* (Mangin) stat. nov. (Biddulphiaceae, Bacillariophyceae): life stages at the Weddell Sea ice edge. Phycologia, 29(1), 27–38. doi:10.2216/i0031-8884-29-1-27.1
7. Gersonde, R., Crosta, X., Abelmann, A., Armand, L. (2005). Sea-surface temperature and sea ice distribution of the Southern Ocean at the EPILOG Last Glacial Maximum—a circum-Antarctic view based on siliceous microfossil records. Quaternary Science Reviews, 24(7–9), 869–896. doi: 10.1016/j.quascirev.2004.07.015
8. Itaki, T., Taira, Y., Kuwamori, N., Maebayashi, T., Takeshima, S., Toya, K. (2020). Automated collection of single species of microfossils using a deep learning–micromanipulator system. PEPS, 7, 1–7. doi:10.1186/s40645-020-00332-4
9. Kaczmarska, I., Barbrick, N. E., Ehrman, J. M., Cant, G. P. (1993). *Eucampia* Index as an indicator of the Late Pleistocene oscillations of the winter sea-ice extent at the ODP Leg 119 Site 745B at the Kerguelen Plateau. In: van Dam H. (eds) Twelfth International Diatom Symposium. Developments in Hydrobiology, Springer, Dordrecht. doi:10.1007/978-94-017-3622-0_13
10. Kloster, M., Kauer, G., Beszteri, B. (2014). SHERPA: an image segmentation and outline feature extraction tool for diatoms and other objects. BMC bioinform., 15(1), 218. doi:10.1186/1471-2105-15-218
11. Kloster, M., Esper, O., Kauer, G., Beszteri, B. (2017). Large-scale permanent slide imaging and image analysis for diatom morphometrics. Appl. Sci., 7(4), 330. doi:10.3390/app7040330

12. Pappas, J., Kociolek, P., Stoermer, E. F. (2014). Quantitative morphometric methods in diatom research. Nova Hedwig., 143, 281–306. doi:10.1127/1436-7270/2014/015

13. Pedraza, A., Bueno, G., Deniz, O., Cristóbal, G., Blanco, S., Borrego-Ramos, M. (2017). Automated diatom classification (Part B): a deep learning approach. Appl. Sci., 7(5), 460. doi:10.3390/app7050460

14. Pedraza, A., Bueno, G., Deniz, O., Ruiz-Santaquiteria, J., Sanchez, C., Blanco, S., Borrego-Ramos M., Olenici A., Cristobal, G. (2018). Lights and pitfalls of convolutional neural networks for diatom identification. Opt.Photonics Digit. Technol. or Imaging Appl. V 10679, 106790G. doi:10.1117/12.2309488

15. Ruiz-Santaquiteria, J., Bueno, G., Deniz, O., Vallez, N., Cristobal, G. (2020). Semantic versus instance segmentation in microscopic algae detection. Eng. Appl. Artif. Intell., 87, 103271. doi:10.1016/j.engappai.2019.103271

16. Schulze, K., Tillich, U. M., Dandekar, T., Frohme, M. (2013). PlanktoVision—an automated analysis system for the identification of phytoplankton. BMC bioinform., 14(1), 1–10. doi:10.1186/1471-2105-14-11

17. Spaulding, S. A., Jewson, D. H., Bixby, R. J., Nelson, H., McKnight, D. M. (2012). Automated measurement of diatom size. Limnol. Oceanogr.: Methods, 10(11), 882–890. doi:10.4319/lom.2012.10.882

18. Shoji, D., Noguchi, R., Otsuki, S., Hino, H. (2018). Classification of volcanic ash particles using a convolutional neural network and probability. Sci. Rep., 8(1), 8111. doi:10.1038/s41598-018-26200-2

19. Tang, N., Zhou, F., Gu, Z., Zheng, H., Yu, Z., Zheng, B. (2018). Unsupervised pixel-wise classification for Chaetoceros image segmentation. Neurocomputing, 318, 261–270. doi:10.1016/j.neucom.2018.08.064

20. Whitehead, J. M., Wotherspoon, S., Bohaty, S. M. (2005). Minimal Antarctic sea ice during the Pliocene. Geology, 33(2), 137–140. doi:10.1130/G21013.1

Tables

Due to technical limitations, table 1-2 is only available as a download in the Supplemental Files section.