Convolutional neural network application on a new middle Eocene radiolarian dataset

Veronica Carlsson, Taniel Danelian, Martin Tetard, Mathias Meunier, Pierre Boulet, Philippe Devienne, Sandra Ventalon

Convolutional neural network application on a new middle Eocene radiolarian dataset

Veronica Carlsson[1,2]*, Taniel Danelian[1], Martin Tetard[3], Mathias Meunier[1], Pierre Boulet[2], Philippe Devienne[2], Sandra Ventalon[4]

[1]Univ. Lille, CNRS, UMR 8198, Evo-Eco-Paleo, F-59000 Lille, France.

[2]Univ. Lille, CNRS, CRIStAL – Centre de Recherche en Informatique Signal et Automatique de Lille, UMR 9189, F-59000 Lille, France.

[3]GNS Science, NZ-5040, Lower Hutt, New Zealand.

[4]Univ. Lille, CNRS, Univ. Littoral Cote d'Opale, UMR 8187, LOG, F-59000 Lille, France.

* Corresponding author. E-mail address: veronica.carlsson@univ-lille.fr

Abstract

A new radiolarian image database was used to train a Convolutional Neural Network (CNN) for automatic image classification. The focus was on 39 commonly occurring nassellarian species, which are important for biostratigraphy.

The database consisted of tropical radiolarian assemblages from 129 middle Eocene samples retrieved from ODP Holes 1258A, 1259A, and 1260A (Demerara Rise). A total of 116 taxonomic classes were established, with 96 classes used for training a ResNet50 CNN. To represent the diverse radiolarian assemblage, some classes were formed by grouping forms based on external morphological criteria. This approach resulted in an 86.6% training accuracy.

A test set of 800 images from new samples obtained from Hole 1260A was used to validate the CNN, achieving a 75.69% accuracy. The focus then shifted to 39 well-known nassellarian

species, using a total of 15 932 images from the new samples. The goal was to determine if the targeted species were correctly classified and explore potential real-world applications of the trained CNN.

Different prediction threshold values were experimented with. In most cases, a lower threshold value was preferred to ensure that all species were captured in the correct groups, even if it resulted in lower accuracies within the classes.

Keywords: middle Eocene, radiolaria, convolutional neural network, image database, automated identification, image recognition

1 Introduction

Polycystine radiolaria are microscopic unicellular protists living currently in all modern oceans; they are characterized by an aesthetically pleasing siliceous skeleton known in the fossil record since the Cambrian (Pouille et al., 2011; Aitchison et al., 2017). Their fossil record is thus of much interest for evolutionary studies (i.e., Danelian and Johnson, 2001; Renaudie and Lazarus, 2013; Tetard et al., 2017; Trubovitz et al., 2020). They are usually the only fossils capable of dating siliceous sedimentary sequences (i.e., Danelian et al., 2012; Vrielynck et al., 2003) and are commonly used in paleoceanography (Matsuzaki et al., 2018; Itaki et al., 2020). Due to the small number of experts, radiolarian taxonomy is less well-elaborated than the one of other microfossil groups, such as foraminifera and nannofossils.

Today, most of the studies involving identification and counting of microfossils, such as radiolarians, are conducted manually and require substantial taxonomic expertise. This process is known to be time-consuming, particularly when microfossils are used for paleoceanography. Moreover, consistency in species classification may be difficult to achieve

between different taxonomic experts. Therefore, artificial intelligence (AI) has been introduced to this field to simplify or automate the work done by micropaleontologists, as for example through automatic image recognition and counting. Several applications of CNNs for automatic image recognition were developed over the last 20 years, since the introduction of SYRACO by Dollfus and Beaufort (1999). Nowadays, CNNs are used on various microfossil groups, such as foraminifera (ex. Mitra et al., 2019, Hsiang et al., 2019; Marchant et al., 2020), coccoliths (ex. Dollfus and Beaufort, 1999; Beaufort and Dollfus, 2004), pollen (ex. Bourel et al., 2020), or even radiolarians (Itaki et al., 2020; Renaudie et al., 2018; Tetard et al., 2020).

Sediments recovered from the Demerara Rise (tropical Atlantic Ocean) during the Leg 207 are rich in middle Eocene radiolarians, preserved in a continuous and expanded carbonate sequence. The encountered radiolarian diversity is immense and based on our estimates it accounts for ca. 500 species, many of which are not described yet. Indeed, although Eocene radiolarians have been studied for about 150 years (since Ehrenberg, 1874) and more in depth for the last 50 years (Riedel and Sanfilippo, 1970, 1978), past research was mainly focused on their biostratigraphic applications (see Meunier and Danelian, 2022 and references therein).

Taking advantage of today's technological achievements, our objective was to design a reliable approach to automatically classify middle Eocene tropical radiolarians from Demerara Rise. The main question addressed in this study is whether a CNN can accurately classify 39 commonly observed nassellarian radiolarian species, most of which have an established biostratigraphic significance. We thus trained a CNN on a newly established image database of middle Eocene radiolarians with a focus on some common nassellarian species. To do this we classified every single object appearing on prepared radiolarian slides.

We were inspired by a recent similar study, conducted by Tetard et al. (2020), who studied middle Miocene to Quaternary radiolarians from the West Pacific Warm Pool. We also included additional *Podocyrtis* species prepared for an earlier dataset (Carlsson, 2022). Finally, a new set of a small number of samples was imaged to test the consistency of our trained CNN, which was double checked with manual identifications made by a taxonomist.

2 Materials and methods

2.1 Core setting and sample preparation

The middle Eocene samples used in this study were collected during ODP Leg 207 from Demerara Rise, situated off the coast of Suriname (Erbacher et al., 2004, Danelian et al. 2005). This dataset includes samples from several cores recovered from sites 1258, 1259 and 1260. It is noteworthy that the middle Eocene sequence at site 1260 is thick and contains radiolarians of an excellent state of preservation (Danelian et al., 2007). The full sample list used in this study may be found in Supplements 1, Table S1. The sediment samples consist primarily of nannofossil and foraminifera chalk, but also contain abundant and well-preserved siliceous microfossils, composed essentially of radiolarians (Danelian et al. 2007, Meunier and Danelian, 2022), as well as diatoms (Danelian et al. 2007, Renaudie et al. 2010).

Sediment samples from ODP holes 1258A, 1259A, and 1260A were first processed to disaggregate organic matter and dissolve carbonates and were then sieved through a 45 µm mesh to remove smaller particles. Thereafter, the samples were prepared using a recent random settling protocol described in Tetard et al. (2020). About 1/3 to ¼ of a microspoon spatula was used from the dried residues, corresponding to approximately 0.2-1.0 g for each sample. Samples were uniformly settled onto 12 mm x 12 mm cover slides using a 3D-printed decanter, as in Tetard et al. (2020), preventing contact between radiolarians and other

remaining objects on the slide. The cover slides were allowed to dry overnight before being mounted with NOA81 optical glue. A total of eight different cover slips, all from the same samples, were placed onto one 76 mm x 26 mm glass slide, forming one sample.

Finally, a new set of four samples (see Supplements 1, Table S1) was prepared for manually testing the actual accuracy of the trained CNN. An improved cleaning technique was used, which kept only siliceous particles that are larger than 63 µm, and completely removed all clay, calcite and smaller or broken radiolarians (Sanfilippo et al., 1985, Tetard et al., 2020). In fact, most radiolarians are larger, so using a 63 µm sieve will probably just remove smaller broken pieces or radiolarians rather than the radiolarians themselves. At first, about 2-3 cm of raw sediment sample was placed in a plastic beaker. Thereafter 30 ml of 30 % hydrochloric acid was added and left for two hours and until there was no more reaction. Furthermore, 200 ml of distilled water was added to the beaker, which was stirred gently and left to settle for two additional hours. The supernatant was removed and exchanged with 30 ml of 10 % hydrogen peroxide and was left to rest for another two hours. The residue was thereafter washed with a 63 µm sieve into a 100 ml beaker. To remove remaining clay particles, the samples were processed in ultrasonic bath for ten minutes; they were later sieved again with a 63 µm mesh and collected afterwards in a filter. Next, the residues were dried in an oven at 50 °C and transferred into a glass vial.

2.2 Image collection and processing

All samples were photographed using a Nikon Eclipse Ni automatic microscope equipped with a Nikon DS-Ri2 microscope camera and Nikon NIS Element software, using a 20 x objective, allowing a 200 x magnification and 0.36 µm/pixel resolution. The lowest and

highest focal points on the Z-axis were manually determined for each sample since the depth varied depending on the thickness of the glass, coverslip, optical glue, and individual radiolarian specimens. For each field of view (FOV), an image was taken at different focal depths, every 10 µm, depending on the minimum and maximum focal points determined for each sample. The microscope then automatically stacked all images, taken at different focal points for each FOV, in order to create a composite image. The microscope was programmed to capture small images in 20 x 20 (400) FOV, covering about 10 x 10 mm out of the 12 x 12 mm available surface for each individual sample. The individual FOV images were subsequently merged into one large mosaic image (see Fig. 1), which has as a result to lose less images of specimens located on the edges of FOVs.

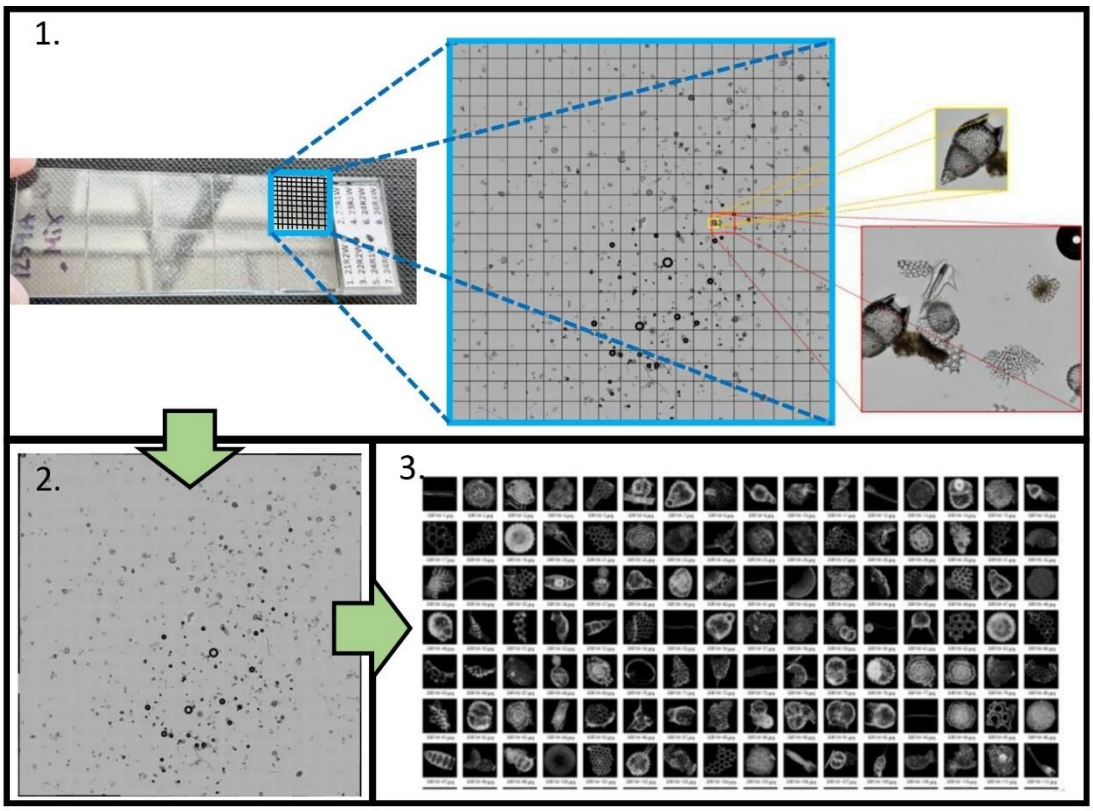


Figure 1. **1**) For each sample, 20 × 20 images are automatically photographed in a convolving way and merged into one large "mosaic composed" image, which enables preserving more complete specimens which are not cut in half. **2**) The mosaic

composed image thereafter went through grayscale conversion. **3)** Segmentation of each unique particle into vignettes and image conversion into 8-bit black and white with black background.

The composite mosaic pattern images received from the automatic microscope were first converted from RGB to 8-bit grayscale using Adobe Photoshop to decrease its size. The images were then segmented into ImageJ using the ImageJ BioVoxxel plugin (Brocher, 2022), and a modified script of the Autoradio_Segmenter plugin (Tetard et al., 2020), which enabled each individual particle to become its own individual image, or vignette. For more details, the reader is referred to Tetard et al. (2020).

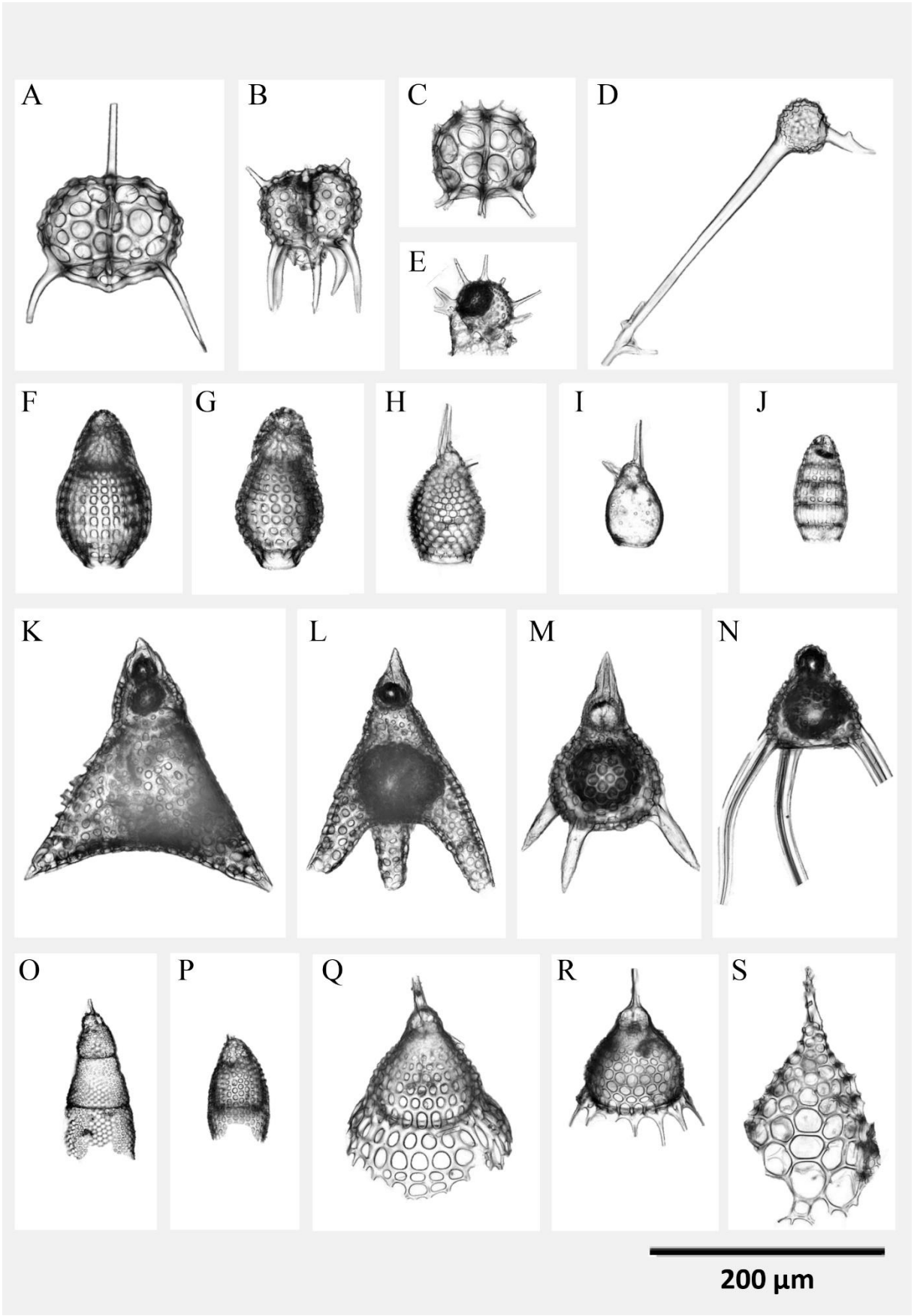
Finally, we also included additional *Podocyrtis* species images prepared for an earlier dataset (Carlsson, 2022).

2.3 Taxa selection and dataset

For this study we decided to focus on 39 species (Plates 1 and 2), which are the most common in the Middle Eocene interval of Demerara Rise and most of which are used in biostratigraphy (Riedel and Sanfilippo, 1970, 1978; Sanfilippo and Nigrini, 1998; Meunier and Danelian 2022).

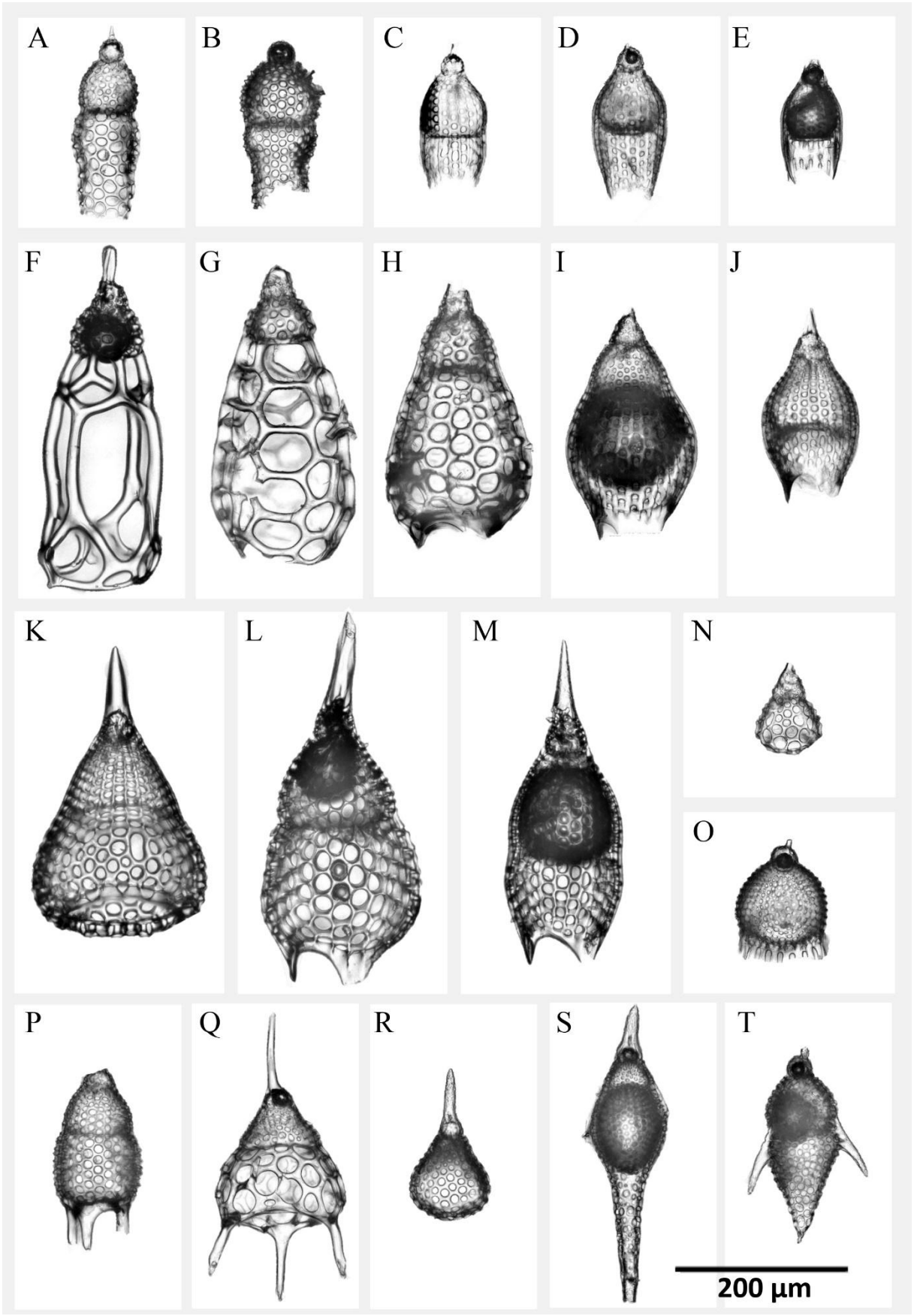
Plate 1. Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of Demerara Rise; species names are followed by the ODP site and hole, core, section and sampled level from which it comes from. A) *Dendrospyris stylophora* (Ehrenberg, 1874) from 1259A-20R-4 W, 53–55 cm, B) *Elaphospyris didiceros* (Ehrenberg, 1874) group from 1258A-2R-4 W, 55–57 cm, C) *Liriospyris clathrata* (Ehrenberg, 1874) group from 1259A-20R-4 W, 53–55 cm, D) *Rhabdolithis pipa* Ehrenberg 1854 from 1260A-15R-1 W, 55–57 cm, E) *Lophophaena radians* (Ehrenberg, 1874) group from 1259A-16R-2 W, 55–57 cm, F) *Dictyoprora mongolfieri* (Ehrenberg, 1854) from 1260A-6R-2 W, 55–57 cm, G) *Dictyoprora amphora* (Haeckel, 1887) group from 1258A-2R-4 W, 55–57 cm,

H) *Rhopalosyringium*? *auriculaleporis* (Clark and Campbell, 1942) from 1260A-14R-6 W, 55–57 cm, I) *Rhopalosyringium*? *biauritum* (Ehrenberg, 1874) from 1260A-12R-3 W, 55–57 cm, J) *Dictyomitra parva* (Kim, 1992) from 1258A-2R-4 W, 55–57 cm, K) *Lithochytris vespertilio* Ehrenberg, 1874 from 1260A-10R-5 W, 55–57 cm, L) *Sethochytris triconiscus* (Haeckel 1887) from 1259A-20R-4 W, 53–55 cm, M) *Lychnocanium babylonis* (Clark and Campbell 1942) group from 1258A-3R-3 W, 56–58 cm, N) *Lychnocanoma bajunensis* (Renz, 1984) from 1258A-2R-4 W, 55–57 cm, O) *Stichopterygium microporum* (Ehrenberg, 1874*)* from 1260A-8R-4 W, 54–56 cm, P) *Carpocanopsis ornata* (Ehrenberg, 1874) group from 1260A-6R-5 W, 55–57 cm, Q) *Cycladophora spatiosa* (Ehrenberg, 1874) group from 1259A-20R-1 W, 55–57 cm, R) *Anthocyrtis mespilus* (Ehrenberg, 1847) group from 1259A-16R-2 W, 55–57 cm, S) *Zealithapium mitra* (Ehrenberg, 1874) from 1258A-2R-3 W, 55–57 cm.



173

Plate 2. Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of Demerara Rise; species names are followed by the ODP site and hole, core, section and sampled level from which it comes from. A) *Lophocyrtis alauda* (Ehrenberg, 1874) from 1260A-15R-1 W, 55–57 cm, B) *Aphetocyrtis zamenhofi* Meunier and Danelian, 2023 from 1259A-26R-5 W, 54–56 cm, C) *Theocyrtis scolopax* (Ehrenberg, 1874) from 1260A-15R-3 W, 55–57 cm, D) *Phormocyrtis embolum* (Ehrenberg, 1874) from 1258A-2R-4 W, 55–57 cm, E) *Phormocyrtis lazari* Meunier and Danelian, 2023 from 1260A-8R-6 W, 54–56 cm, F) *Podocyrtis (Lampterium) goetheana* (Haeckel, 1887) *from* 1259A-18R-1 W, 53–55 cm, G) *Podocyrtis (Lampterium) chalara* Riedel and Sanfilippo, 1970 from 1260A-6R-CC, 63–177 cm, H) *Podocyrtis (Lampterium) mitra* Ehrenberg group, 1854 from 1260A-9R-1 W, 55–57 cm, I) *Podocyrtis (Lampterium) sinuosa* Ehrenberg, 1874 from 1259A-15R-1 W, 55–57 cm, J) *Podocyrtis papalis* Ehrenberg, 1847 from 1258A-2R-4 W, 55–57 cm, K) *Podocyrtis (Podocyrtoges) ampla* Ehrenberg, 1874 from 1260A-10R-5 W, 55–57 cm, L) *Podocyrtis (Podocyrtoges) phyxis* Sanfilippo and Riedel, 1973 from 1259A-16R-1 W, 55–57 cm, M) *Podocyrtis (Podocyrtoges) diamesa* Riedel and Sanfilippo, 1970 from 1259A-26R-3 W, N) *Podocyrtis (Lampterium) puellasinensis* Ehrenberg, 1874 *from* 1259A-20R-4 W, 53–55 cm, O) *Calocyclas hispida* (Ehrenberg, 1874) from 1260A-6R-4 W, 55–57 cm, P) *Thyrsocyrtis (Thyrsocyrtis) rhizodon* Ehrenberg, 1874 from 1260A-6R-CC, 63–177 cm, Q) *Thyrsocyrtis (Pentalocorys) triacantha* (Ehrenberg, 1874) from 1260A-8R-6 W, 54–56 cm, R) *Eusyringium lagena* (Ehrenberg, 1874) from 1259A-25R-2 W, 55–57 cm, S) *Eusyringium fistuligerum* (Ehrenberg, 1874) group from 1259 to 18R-1 W, 53–55 cm, T) *Rhopalocanium ornatum* (Ehrenberg, 1874) from 1259A-22R-1 W, 55–57 cm.



191

192

Synonymy lists are given in the supplementary catalogue (Supplements 3), thus allowing the reader to understand the species concept followed in this study. Taxonomic information for all the other radiolarian classes used in the analysis is also presented in there; most of the other radiolarians were grouped in supraspecific taxa, with taxonomic information and some typical forms given in the catalogue (Supplements 3).

The ParticleTrieur software version 2.4.10, developed by Marchant et al. (2020), was used to label our dataset. It includes a built-in k-nearest neighbor (KNN) algorithm, which is a machine learning algorithm that can be used for supervising the classification tasks. It identifies the k-nearest training data points or neighbors for a new data point and predicts a label for the new data based on already labeled data. In the context of image classification, the data points correspond to the pixels in the image. Therefore, ParticleTrieur can recognize patterns in the images for individual classes and suggest classification for new or unclassified images, after some classes have already been added in a semi-supervised way but have of course been validated by a human expert.

We managed to build a dataset consisting of 12,217 images out of a total of ca. 50 000 images, distributed in 116 classes, including the 39 important key-classes of well-known nassellarian species (Plates 1 and 2). Some of the classes consisted of as few as one specimen per class, while others contained up to nearly a thousand images (i.e. the largest class). Classes with fewer than ten specimens were excluded from the CNN training, resulting in only 96 classes to be trained by the model. The taxonomic framework is in many cases classified based on Meunier and Danelian (2022 and 2023) at the species level, and higher taxonomic ranks are classified, mainly based on Suzuki et al. (2021).

2.4 CNN training

Earlier studies that attempted to compare the accuracy of several CNNs on microfossil assemblages found that Resnet50 usually performed very well for this purpose (Marchant et al., 2020; Tetard et al., 2020; Mitra et al., 2018); we therefore chose to apply this model for this particular study, instead of MobileNet v1 (Howard et al., 2017), which was we used previously (Carlsson et al. 2022) in study focusing on eight closely related species of the Eocene genus *Podocyrtis*. ResNet50 is a deep Convolutional Neural Network architecture (He et al., 2016), and is one of the variations of the ResNet (short for "Residual Network") family of models. The idea behind the formation of ResNet50 is to use residual learning to avoid disappearing gradients in very deep neural networks. Because when the networks get deeper, it becomes more difficult to update the weights of the earlier layers through backpropagation and by using residual learning; the network can thus propagate the gradient signal more easily, which improves the training of deeper networks. The weights of ResNet50 have also been pre-trained on a large dataset, for instance ImageNet, which includes millions of labelled images of about 1,000 classes (He et al., 2016). Our training set consisted of 80 % randomized images, chosen for each individual class present in the database, while the remaining 20 % was used for validation.

2.5 Performance metrics

The CNN training calculates automatically the classification accuracy and recall values based on the labeled dataset; both of them represent different ways of displaying the CNN performance (Fig. 2). For instance, if the aim is to detect all specimens belonging to a specific species (high recall value), sacrificing accuracy by including other objects not belonging to that species might be acceptable. This allows for easier tracking of the true abundance of that

specific species, and misclassified objects may be identified and ignored. However, if misclassified objects, which may belong to another important species, end up in another species class, this would lower their recall value. Therefore, it is important to have a high accuracy overall, but when examining individual classes separately, recall value holds a significant importance. Both indices are of great interest for different applications, more focused on biostratigraphy or paleoceanography for example.

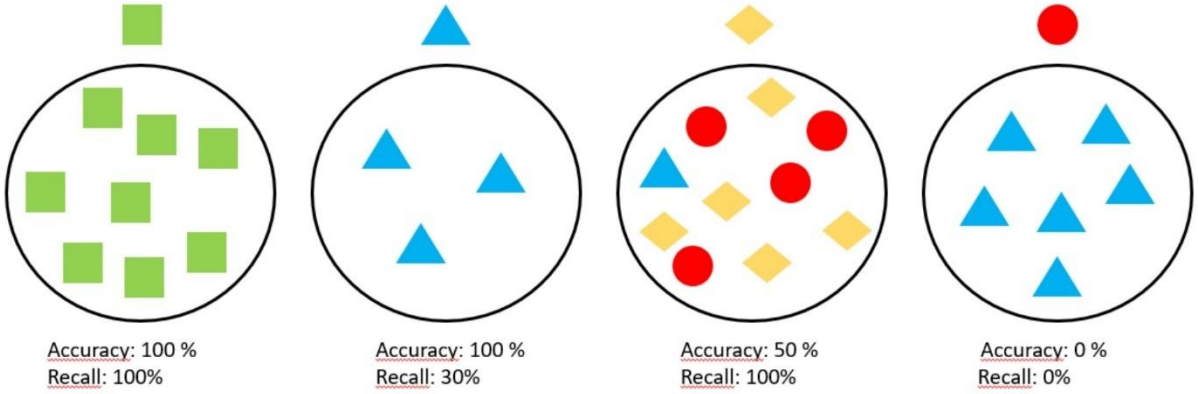


Figure 2. A theoretical example showing the importance of a high accuracy and a high recall value for individual classes.

2.6. Test set to validate the CNN

To validate the consistency of the CNN training and testing from our dataset, we once again estimated how accurately the trained CNN performed and we compared it with a human operator. The neural network training produced a prediction model that was inserted into ParticleTrieur version 3.0.0. A threshold value can be set directly in ParticleTrieur before classifying new images. The threshold value constrains the degree of accuracy desired for an image needed to be classified into a given class. If the probability for an image/specimen to be classified to a specific class is too low, this image will be left unclassified.

We randomly let the ParticleTrieur software pick 200 images selected from four new samples (800 in total), which were unlabeled and contained all particles segmented from composed images, from ODP site 1260, coming from different intervals within those used initially to train the CNN, which were 1260A-6R-3W, 18-20 cm; 1260A-8R-5W, 70-72 cm; 1260A-13R-5W, 66-68 cm; and 1260A-15R-4W, 69-71 cm. In ParticleTrieur we let the trained CNN identify all of them, using a threshold value of 0.5. We selected this low threshold value since it is better having more images classified, even if that will give a somewhat lower accuracy, recall, precision to classify more images rather than that they unclassified. We then examined afterwards what was correctly or incorrectly classified.

2.7 Application on 39 species

With the same four samples, we then this time entered all segmented particles from the entire mosaic composed images covering most parts of the coverslips, resulting in a total of 15 932 images, which were automatically classified with the CNN. Here we focused on the targeted 39 classes representing the selected well-known species. We applied different threshold (1-0.5) values for the pre-trained network and checked how the CNN could recognize these 39 most common nassellarian species. The aim of this test was to try future potential applications.

3 Results

3.1 Training of the initial dataset

Our trained CNN obtained 86.6 % in overall accuracy, with 75.6 % in precision/accuracy (Fig. 2); the latter measures the ability to avoid false positives and corresponds to the number

of specimens classified as a class and also belonging to that class, divided by all specimens classified to that class. Our CNN obtained 78 % of recall, which calculates the ability to detect all correct classifications, as it corresponds to the number of specimens in a class that were correctly classified divided by the total number of specimens in that class. The training and validation iterations are given in Supplements 2 figure S1, which confirms that the data is neither overfit, nor underfit.

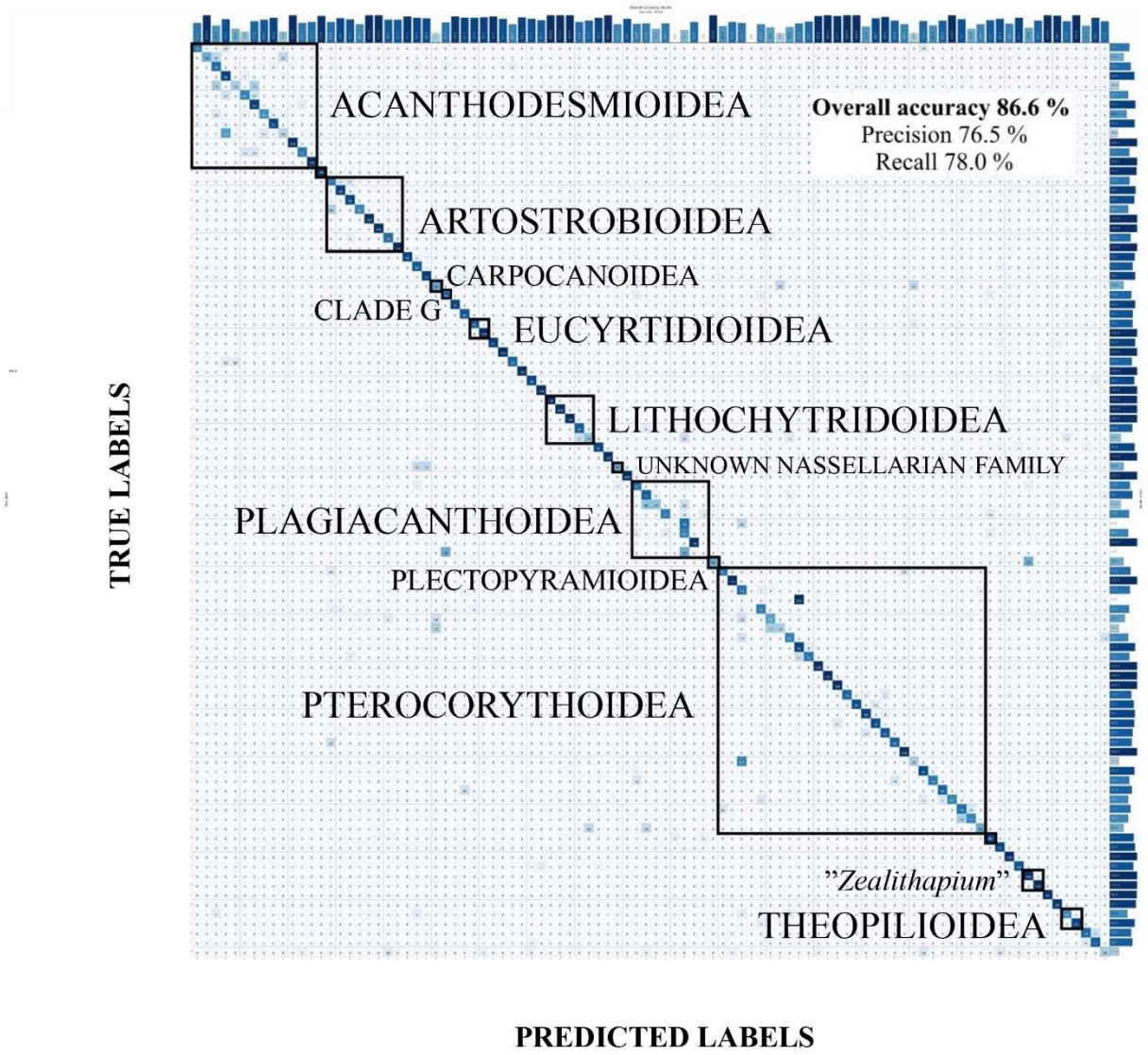


Figure 3. Simplified confusion matrix, showing the classification between different classes, with a focus on nassellarian super families. The x-axis on the left shows the true classes while the right axis shows recall values; the y-axis at the bottom shows the predicted classes while the top shows the precision value.

Some of the classes that are visible in the confusion matrix (Fig. 2 or Fig. S2 in Supplements 2 for a more detailed confusion matrix) show a low individual score (diagonal numbers), mainly due to the low number of available specimens (see number next to the label name), or to a high degree of resemblance between closely related or similar looking species. Most of the 39 important species of nassellarians we focused on were classified with a high accuracy (Fig. S2 in Supplements 2).

Table 1. Precision, recall and F1 score for the 39 most important nassellarian species focused in this study.

Species	**Training (#)**	**Test (#)**	**Precision**	**Recall**	**F1 score**
Dendrospyris stylophora	35	7	0.71	0.71	0.71
Elaphospyris didiceros group	295	59	0.85	0.90	0.88
Liriospyris clathrata group	77	15	0.78	0.93	0.85
Dictyomitra parva	229	46	0.98	0.98	0.98
Dictyoprora amphora group	124	25	0.75	0.72	0.73
Dictyoprora mongolfieri	271	54	0.89	0.94	0.92
Rhopalosyringium auriculaleporis	104	21	0.84	0.76	0.80
Rhopalosyringium biaurata	24	5	0.83	1.00	0.91
Carpocanopsis ornata group	21	4	0.33	0.50	0.40
Stichopterygium microporum	64	13	0.93	1.00	0.96
Sethochytris triconiscus	22	4	1.00	1.00	1.00
Lithochytris vespertilio	20	4	1.00	1.00	1.00

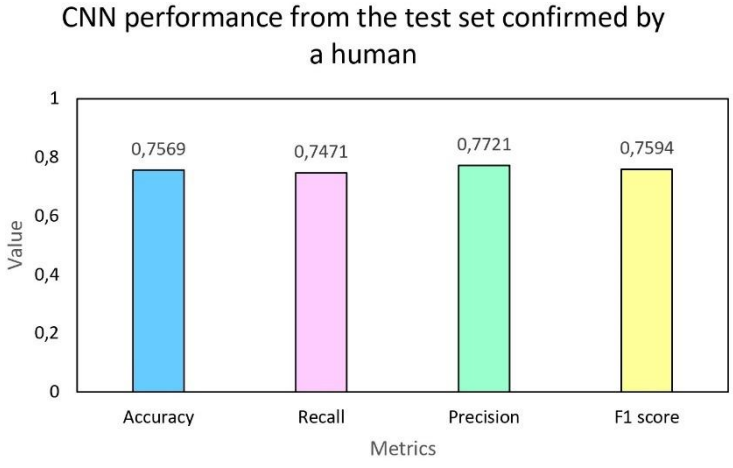
Species	Training (#)	Test (#)	Precision	Recall	F1 score
Lychnocanoma bajunensis	103	21	1.00	1.00	1.00
Lychnocanium babylonis group	56	11	0.75	0.82	0.78
Lophophaena radians group	48	10	0.70	0.70	0.70
Rhabdolithis pipa	28	6	0.75	1.00	0.86
Aphetocyrtis zamenhofi	40	8	0.60	0.75	0.67
Lophocyrtis alauda	17	3	0.75	1.00	0.86
Theocyrtis scolopax	22	4	0.00	0.00	0.00
Phormocyrtis embolum	130	26	0.71	0.92	0.80
Phormocyrtis lazari	68	14	0.77	0.71	0.74
Calocyclas hispida	29	6	0.57	0.67	
Podocyrtis (Lampterium) chalara	207	41	0.98	1.00	0.99
Podocyrtis (Lampterium) goetheana	115	23	1.00	1.00	1.00
Podocyrtis (Lampterium) mitra	184	37	0.97	1.00	0.99
Podocyrtis (Lampterium) sinuosa	93	19	1.00	0.79	0.88
Podocyrtis (Podocyrtoges) ampla	42	8	1.00	0.88	0.93
Podocyrtis (Podocyrtoges) diamesa	62	12	0.58	0.92	0.71
Podocyrtis (Podocyrtoges) phyxis	44	9	0.89	0.89	0.89
Podocyrtis (Podocyrtis) papalis	302	60	0.95	0.83	0.83
Podocyrtis (Lampterium) puellasinensis	24	5	0.80	0.80	0.80
Pentalocorys triacantha	100	20	0.78	0.90	0.84
Thyrsocyrtis rhizodon	43	9	1.00	0.67	0.80

Species	Training (#)	Test (#)	Precision	Recall	F1 score
Eusyringium fistuligerum group	38	8	0.86	0.75	0.80
Eusyringium lagena	16	3	0.50	0.67	0.57
Rhopalocanium ornatum group	21	4	0.67	0.67	0.57
Zealithapium mitra	33	7	0.54	1.00	0.70
Anthocyrtis mespilus group	53	11	0.78	0.64	0.70
Cycladophora spatiosa group	53	11	1.00	0.91	0.95

306

307 **3.2 Performance validation from the test set**

308 By using a threshold value of 0.5 the CNN classification resulted in that 769 images, out of a
309 total of 800, were correctly classified. All classes were individually examined and the
310 precision and recall values were calculated for each detected class (see Supplements 4). The
311 CNN could classify these images into 76 different classes, while the human classifier
312 considered that these groups belonged to 63 classes, excluding rare species which could not
313 be classified into a proper class and were therefore referred to the class "others". Finally, an
314 overall accuracy, recall, precision and F1 score were calculated (see Fig. 4).



315

Figure 4. CNN performance metrics based on a test set consisting of a total of 800 images, which were validated by a human expert.

3.3 Application on new samples

The manual validation of the classification made by a CNN on the four new samples using a threshold value of 1, always provided a correct interpretation without having any misclassified species, although it was far from detecting all classes and all true specimens for each class. Interestingly, with a threshold value of 0.9, we could in some sense detect almost all classes present in the dataset with a 69-73 % accuracy (see Supplements 5) and get some estimates for the number of taxa present in the samples. Some possible misclassifications can be easily reviewed at a later stage. Lowering the threshold values increased indeed the number of truly correct specimens in the right species, but it also increased the number of false classifications (see Supplements 5). Regarding the average score of predicted key-species using several threshold values (Table 2), the CNN was usually able to correctly identify between 24 to 34 out of the 39 targeted species and also not falsely detect species which do not exist (see Supplements 5).

Table 2. Average prediction results of the 39 key-species and its accuracy for different threshold values along with the total number of images.

Threshold value	Predicted key-species	Correctly predicted key species	Accuracy key-species	Predicted images	Total amount of images
1.00	77	77	1.00	1664	15,932
0.90	1144	784	0.69	11,074	15,932
0.80	1336	865	0.65	12,438	15,932
0.70	1501	943	0.63	13,462	15,932
0.60	1653	998	0.60	14,370	15,932

Threshold value	Predicted key-species	Correctly predicted key species	Accuracy key-species	Predicted images	Total amount of images
0.50	1799	1043	0.58	15,232	15,932

4 Discussion

4.1 Classification

Due to the high radiolarian diversity preserved in the studied samples obtained from the equatorial Atlantic, the correct assignment of every single image to a class may be challenging. One of the particularities of the studied radiolarian fauna is that it contains a lot of rare and undescribed species. In addition, the current state of the art does not allow confident taxonomic divisions in higher classes, as there are often unclear taxonomic boundaries. A lot of taxa are also morphologically similar to each other, and a lot of similarities sometimes exist between different families, giving us often hard time to find for them a proper affinity and even acceptable taxonomic names. As an example, we may state the confusion of *Carpocanopsis ornata* group with juvenile/broken Nassellaria group B and Pterocorythoid group (see the catalogue in Supplements 3 and the confusion matrix in Fig. S2, Supplements 2).

There have been challenges in determining the most commonly occurring nassellarian species. We aimed to ensure taxonomic clarity within these groups and only included taxa for which we are very confident for their assignment to that specific class. Consequently, numerous similar-looking species have been excluded from the classified dataset. This is because they did not fit into other groups or we judged that they would be confused with the taxon they most resemble with. However, this may be challenging as we strive to represent as much as possible from these samples, while also collecting new samples for automatic

classification. As is the case of an example mentioned above, we encountered many specimens that resemble *Carpocanopsis ornata* group, but which do not always display the discriminatory characters of the species. Those specimens would also always be misidentified as *C. ornata* group and were therefore removed from the dataset. It is possible that these forms may be misidentified as *C. ornata* group in future samples, but they have to be looked upon individually afterwards. The important point is to obtain a dataset that consists of clean *C. ornata* group specimens identified with high confidence, so that there can be a clear reference of what a *C. ornata* group looks like (see the catalogue in Supplements 3 and the confusion matrix in Fig. S2, Supplements 2).

In addition, difficulties have also been met when we attempted to consistently classify a high number of images. Also, differences in individual specimen orientation and bubble inclusions played a big role in getting the CNN to work and to find the proper classes.

Since all objects appearing on a slide are trained and given a class, some of the included classes may be artificially defined and therefore correspond to taxonomically "bin" classes. We focused mainly on nassellarian radiolarians, trying to include as many classes as possible neatly defined at the species level. Although, for some nassellarian classes presented at higher taxonomic levels, we were obliged to accept a very large taxonomic concept. Regarding spumellarians, as the recognition of their inner structures is important, but difficult to detect with computer vision, identifications are even more challenging.

4.2 CNN training and new test set score

It is not a shock that the test accuracy, which randomly selected 800 images from four new samples, performed less than the test of the 20 % of the labeled dataset, from which 80 %

were used to train the CNN. The accuracy of the test is 75.69 %, whereas the training accuracy is 85.6 %. From our labeled dataset, we have purposely removed a lot of "trash" particles, that will say particles which are broken fragments of radiolarians, blurry background particles etc. because if we kept these images, the CNN would rather be overtrained by the thousands of trash images and perform less. Besides that, rare species which consisted of too few species were left untrained by the neural network and therefore the 20 % of test did not include that many "trash" images or any rare species, in contrast to the new test set, which were completely randomly selected among any kind of particle that had been segmented. This last test was just to confirm how well the CNN was generally trained. For our last application we tried to just focus on the 39 well known radiolarian species, since it is the radiolarians that are of interest.

4.3 Feedback on individual species

By examining the 39 targeted species individually in every single sample, we observed that some species were easily identified correctly, while others performed poorly during the CNN training iteration and ended up in different classes. Table 3 presents the examined species and samples, along with their training performance. This arrangement facilitates a better understanding of the high or low number of correctly predicted species based on the training performance of the CNN.

Table 3. Training accuracy, recall and number of correctly identified specimens for each one of the 39 targeted species in this study and sample using a threshold value of 0.5, which is the lowest value we used in the test to identify all species which have an identification correctness equal or higher than 0.5. Hyphens "-" correspond to species not found in the samples (Meunier and Danelian, 2022; Meunier and Danelian, 2023).

Species	**Accuracy**	**Recall**	**1260A 6R-3 W, 18-20 cm**	**1260A 8R-5 W, 70-72 cm**	**1260A 13R-5 W, 66-68 cm**	**1260A 15R-4 W, 69-71 cm**
Dendrospyris stylophora	0.71	0.71	1	4	1	0
Elaphospyris didiceros group	0.85	0.90	16	27	10	48
Liriospyris clathrata group	0.78	0.93	12	2	7	10
Dictyomitra parva	0.98	0.98	–	–	–	108
Dictyoprora mongolfieri	0.89	0.94	36	9	38	14
Dictyoprora amphora group	0.75	0.72	1	2	8	68
Rhopalosyringium? *auriculaleporis*	0.84	0.76	–	2	3	3
Rhopalosyringium? *biaurata*	0.83	1.00	–	0	1	0
Carpocanopsis ornata group	0.33	0.50	12	0	1	–
Stichopterygium microporum	0.93	1.00	2	3	6	3
Sethochytris triconiscus	1.00	1.00	0	10	–	–
Lithochytris vespertilio	1.00	1.00	–	3	3	6
Lychnocanoma bajunensis	1.00	1.00	46	25	23	10
Lychnocanium babylonis group	0.75	0.82	–	5	14	8
Lophophaena radians group	0.70	0.70	2	0	2	–
Rhabdolithis pipa	0.75	1.00	2	6	1	3
Lophocyrtis alauda	0.75	1.00	–	–	8	9
Aphetocyrtis zamenhofi	0.60	0.75	–	–	0	20
Theocyrtis scolopax	0.00	0.00	–	–	2	1
Calocyclas hispida	0.57	0.67	4	12	0	0

Species	**Accuracy**	**Recall**	**1260A 6R-3 W, 18-20 cm**	**1260A 8R-5 W, 70-72 cm**	**1260A 13R-5 W, 66-68 cm**	**1260A 15R-4 W, 69-71 cm**
Phormocyrtis embolum	0.71	0.92	–	–	16	3
Phormocyrtis lazari	0.77	0.71	9	4	–	–
Podocyrtis (Lampterium) chalara	0.98	1.00	7	20	–	–
Podocyrtis (Lampterium) goetheana	1.00	1.00	2	–	–	–
Podocyrtis (Lampterium) mitra	0.97	1.00	–	2	1	–
Podocyrtis (Lampterium) sinuosa	1.00	0.79	–	–	1	2
Podocyrtis (Podocyrtoges) ampla	1.00	0.88	–	–	4	0
Podocyrtis (Podocyrtoges) phyxis	0.89	0.89	–	–	–	0
Podocyrtis (Podocyrtoges) diamesa	0.58	0.92	–	–	–	0
Podocyrtis (Podocyrtis) papalis	0.94	0.83	8	6	1	6
Podocyrtis (Lampterium) puellasinensis	0.80	0.80	0	0	–	–
Thyrsocyrtis rhizodon	1.00	0.67	15	2	4	11
Pentalocorys triacantha	0.78	0.90	12	43	7	4
Eusyringium lagena	0.50	0.67	–	–	3	7
Eusyringium fistuligerum group	0.86	0.75	9	5	2	0
Rhopalocanium ornatum group	0.67	0.67	3	0	6	1
Cycladophora spatiosa group	1.00	0.91	31	20	2	1
Anthocyrtis mespilus group	0.78	0.64	14	3	10	3
Zealithapium mitra	0.54	1.00	7	7	7	1

405

In general, the classes with good performance are *Elaphospyris didiceros* group (Plate 1.B), *Dictyomitra parva* (Plate 1.J), *Sethochytris triconiscus* (Plate 1.L), *Lithochytris vespertilio* (Plate 1.K), *Lychnocanoma bajunensis* (Plate 1.N), *Lychnocanium babylonis* group (Plate 1.M) and *Thyrsocyrts (Pentalocorys) triacantha* (Plate 1.Q). They are well-classified with few misclassified objects in their respective classes and they rarely appear in other classes.

The CNN was able to detect some of the true specimens of *Dendrospyris stylophora* (Plate 1.A). However, in many cases some trissocyclids/cephalospyrids with long feet were also misclassified as *D. stylophora*. Since this class is quite rare, it is difficult to make any accurate estimate about the application accuracy. For the training iteration the CNN obtained an accuracy of about 70 %.

Liriospyris clathrata group is a simple single-segmented nassellarian with large pores on its cephalis (Plate 1.C). Occasionally, some specimens may be misclassified into higher-ranked taxonomic classes, but overall, it performs well. It has about 80 % accuracy in the CNN training.

Overall, the CNN was able to correctly identify all forms of *Dictyoprora amphora* (Plate 1.G) group, but many other broken and unusual radiolarian forms, including some *Dictyoprora* spp. and *Dictyoprora mongolfieri* (Plate 1.F), were confused with this species group. Although the majority *of D. mongolfieri*) were identified in their true class, a smaller number of radiolarians or objects in the class were misclassified, indicating that these particular classes have been well-trained with a 90 % accuracy.

Although some *Rhopalosyringium* ? *auriculaleporis* specimens (Plate 1.H) were identified correctly, there are still several other objects that are misclassified as this species. The same goes for *R.* ? *biauratum* (Plate 1.I), but since not many have been observed in our test samples, it is difficult to make any fair estimates for this particular species.

Carpocanopsis ornata group (Plate 1.P) is poorly trained and is largely misclassified in the training process, with a training accuracy of only ca. 30 %. This is likely due to its very simple, smooth outline that is similar to many other taxa.

Stichopterigyum microporum (Plate 1.O) is classified well, despite having many similarities with *Euctyrtidium levisaltarix*, a species that was not individualized in this study, but was included in the *Eucyrtidium* genus class. Occasionally, some of these species may be mixed up if there are no morphologically distinct morphotypes. However, in those cases where there are distinct morphotypes, they are classified correctly.

The training accuracy of *Lophophaena radians* group (Plate 1.E) was ca. 70%, although there are not many estimates on how well this species is classified in the new samples obtained.

Rhabdolithis pipa (Plate 1.D) was detected frequently in our samples. However, a lot of other particles also ended up being included in this class, alike some types of spumellarians, since *R. pipa* has only one simple segment and it does not display any radial symmetry and bears two very long spines.

Lophocyrtis alauda (Plate 2.A) is well detected in samples coming from 1260A-13R-5W, 66-68 cm and 1260A-15R-4W, 69-71 cm. The somewhat similar looking species, *Aphetocyrtis*

zamenhofi (Meunier and Danelian 2023) was only found present in 1260A-15R-4W, 69-71 cm but other objects or specimens were also mistakenly classified as *A. zamenhofi*, even at samples in which they do not exist. The training accuracies are 60 % for *A. zamenhofi* and 75 % for *L. alauda*.

Not many specimens of *Calocyclas hispida* (Plate 2.O) have been trained by the CNN, and it is only in sample 1260A-8R-5W, 70-72 cm that they appeared more often; they were classified well, without having a lot of misidentified radiolaria or other objects appearing in that class.

We did not encounter many typical specimens of *Podocyrtis goetheana* (Plate 2.F), but mainly early/transitional forms that do not display the elongated abdomen with the typical long straight bars on the median row of pores. In any case, most transitional forms were classified as *Podocyrtis chalara* (Plate 2.G) and exceptionally as *P. goetheana*. Otherwise, *P. goetheana* has a unique morphology compared to the rest of radiolarians present in our samples and it was therefore trained very well with a perfect F1 score; both accuracy and precision were 100 %.

Podocyrtis chalara (Plate 2.G) is well classified and recognized by the CNN and is easy to detect in our test samples. However, when it comes to *Podocyrtis mitra* (Plate 2.H), our samples contained transitional forms between *Podocyrtis sinuosa* (Plate 2.I) and *P. mitra* or *P. mitra* and *P. chalara*. In the latter case, most specimens we captured were actually closer to *P. chalara* than *P. mitra* and were therefore classified as *P. chalara* rather than *P. mitra*. In the studied material there were too few typical specimens of *P. sinuosa* and *P. mitra* to make up a clear mind, but most of them were transitional forms and the CNN had two specimens

classified as *P. sinuosa* and one as *P. mitra*. Individually, *P. sinuosa* from other samples are rather well-detected, even though it happened to have samples with a lower abundance of *P. sinuosa*.

The CNN can detect well *Podocyrtis papalis* (Plate 2.J). However, some forms that do not belong to *P. papalis* were incorrectly classified, quite often as *Podocyrtis ampla* (Plate 2.K) or *Podocyrtis diamesa* (Plate 2.M). *Theocyrtis scolopax* was also found classified as *P. papalis*. There were not many specimens of *P. diamesa* in these samples; therefore, none was classified as *P. diamesa*, and the few specimens of *P. diamesa* were actually classified as *P. papalis*, which is logical since they are very similar (see also Carlsson et al., 2022). Finally, *P. ampla* was always confused with *P. papalis*, as regrettably the CNN could not correctly detect any single *P. ampla*.

Podocyrtis phyxis (Plate 2.L), an important biostratigraphic index species occurring only within a short interval, was trained in the CNN with an 89 % accuracy and a recall value of 88 %; however, it was never classified correctly into its own class in the new test set of four samples. Instead, it was frequently misidentified as *Thyrsocyrtis rhizodon* (Plate 2.P), which is understandable, given that both species have an equal number of segments and consist of a horn and feet (although they differ in size and shape) and are more or less barrel shaped. Unfortunately, the CNN cannot detect size differences because all images are resized to the same dimensions.

In conclusion, the *Podocyrtoges* lineage, which includes *P. ampla*, *P. phyxis*, and *P. diamesa*, cannot be reliably detected in the new set of test samples using our currently trained CNN. The different morphospecies of this lineage are difficult to be identified correctly due to the

frequent occurrence of transitional forms that look very similar to other taxa in our dataset. Although we have a sufficient dataset of these species, more data and adjustments to similar-looking classes may be necessary to allow the CNN to more clearly distinguish them with a high degree of accuracy, as humans are able to do.

We did not encounter any *Podocyrtis puellasinensis* (Plate 2.N) species but it was originally trained with an accuracy and precision of 80 %.

Eusyringium lagena (Plate 2.R) can be easily detected by the CNN, and *Eusyringium fistuligerum* group (Plate 2.S) is occasionally misclassified as *L. vespertilio* (Plate. 1.K) or *S. triconiscus* (Plate 1.L). This is understandable since their proximal parts (thorax, cephalis, and thick conical horn) look almost identical.

Not many specimens of *Rhopalocanium ornatum* group (Plate 2.T) were detected. This species was trained on a small number of specimens and therefore only obtained a training accuracy of about 70 %. Some specimens of the *R. ornatum* group were found in its true class but were also appearing in other species and higher taxonomic leveled classes, which implies a lower recall number.

Both *Cycladophora spatiosa* group (Plate 1.Q) and *Anthocyrtis mespilus* group (Plate 1.R) could be detected well with a high accuracy in the samples in which they existed. In other samples, they had a lower prediction accuracy with more specimens incorrectly classified as either *A. spatiosa* group or *A. mespilus* group.

The classification of *Zealithapium mitra* (Plate 1.S) is not reliable due to its insufficient training dataset, which comprises only a small number of images. As a result, many broken radiolarians with large pores are frequently misidentified as *Z. mitra*, despite the fact that this species is characterized by large pores with a more conical shape.

The results we obtained suggest that many of the classes we distinguished may be confidently used in future applications (biostratigraphic or paleoceanographic). Well distinct species that the CNN can easily detect in whole assemblage analyses have a low recall value and appear rarely in wrong classes. As in many cases, the presence/absence of an index species is sufficient for biostratigraphy, the automated classification of targeted species in whole assemblage studies described in this paper, enables us to quickly confirm the correct classification of species and thus opens new perspectives for the application of Artificial Intelligence to radiolarian biostratigraphic studies. Apart from the 39 targeted species, many of the other classes had a worse performance; indeed, many half-complete or blurry specimens were classified as other objects. This makes it difficult for the moment to fully trust the CNN classification for counting all radiolarian species in order to get information about their relative abundances.

5 Conclusions

The newly established dataset of middle Eocene tropical radiolarians is well adjusted to fit a CNN. We obtained a high training accuracy of 86.6 % for its training in a CNN.

We evaluated the performance of our trained Convolutional Neural Network (CNN) on new tests and compared it to human performance, and obtained a testing accuracy of about 75.69 %. We thereafter specifically focused on 39 different species which the CNN demonstrated notable success in accurately identifying those species that had been well-trained.

In order to obtain an acceptable accuracy of the CNN for further studies, the labeling of classes was also revised to groups or separate taxa and reached the best compromise between CNN accuracy and consistent taxonomy. For example, when two morphologically close species or subspecies where often confused by the CNN, we found it better to fuse them together in an acceptable taxonomic framework, unless they were individually of biostratigraphic importance, instead of artificially biasing the CNN accuracy by often mistaking these two taxa with each other, in the same way as they may be confused by an operator.

This has proved to be an efficient way, both in speed and easiness, to quickly see what kind of radiolarian species exist and how many of them. However, since we focused only on a few classes, we cannot compare the relative abundance with certain taxa in relation to all radiolarians yet, but with improved methods and building a stronger dataset, it will be possible to get a closer estimate of the relative abundance of many taxa. This also highlights the importance of building good taxonomic datasets.

Overall, applying automatic image classification to the studied samples is time-saving, particularly for detecting the presence of the selected nassellarian species. This approach eliminates the need to manually count and track by an operator the targeted taxa present in a sample and avoids the risk of identification bias between different operators.

Data availability

Microscope slides from Leg 207, Hole 1258A, 1259A and 1260A, which were used for our dataset and application to a trained CNN, are stored at the University of Lille, France. The

dataset (https://doi.org/10.57745/E9YXW6, Carlsson, 2023) is published in the University of Lille repository at Recherche Data Gouv.

Acknowledgments

This study was partly funded by the French government through the program "Investissements d'avenir" (I-ULNE SITE/ANR-16-IDEX-0004 ULNE) managed by the National Research Agency. It also received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement no. 847568. It was also supported by UMR 8198 Evo-Eco-Paléo and IRCICA (CNRS and Univ. Lille USR-3380).

A special thanks to Sylvie Regnier for helping to prepare the new samples by removal of non-siliceous materials, and also many thanks to Ross Marchant for the quick reparation of bugs in ParticleTrieur.

Declaration of interest

None.

References

Adobe Systems Incorporated. (2022). Photoshop (Version 23.0) [Computer software]. Retrieved from https://www.adobe.com/products/photoshop.html

Aitchison, J.C., Suzuki, N., Caridroit, M., Danelian, T., and Noble, P. 2017. Paleozoic radiolarian biostratigraphy. In T. Danelian, M. Caridroit, P. Noble, and J. C. Aitchison (Eds.), Catalogue of Paleozoic Radiolarian Genera, pp. 503-531. Geodiversitas, Vol. 39. Scientific Publications of the Muséum National d'Histoire Naturelle, Paris. doi:10.5252/g2017n3a5.

Beaufort, L., and Dollfus, D. 2004. Automatic recognition of coccoliths by dynamical neural networks. Marine Micropaleontology, 51, 57-73. doi:10.1016/j.marmicro.2003.09.003.

Bourel, B., Marchant, R., de Garidel-Thoron, T., Tetard, M., Barboni, D., Gally, Y., and Beaufort, L. 2020. Automated recognition by multiple convolutional neural networks of modern, fossil, intact and damaged pollen grains. Computers and Geosciences, 140, 104498. doi:10.1016/j.cageo.2020.104498.

Brocher, J. 2022. biovoxxel/BioVoxxel-Toolbox: BioVoxxel Toolbox v2.5.3. Zenodo. doi:10.5281/zenodo.5986129.

Carlsson, V. 2022. Podocyrtis Image Dataset [data set]. Recherche Data Gouv. doi:10.57745/G7CHQL.

Carlsson, V. 2023. Middle Eocene Radiolarian image dataset from ODP Leg 207 (Demerara Rise) [data set]. Recherche Data Gouv. doi: 10.57745/E9YXW6.

Carlsson, V., Danelian, T., Boulet, P., Devienne, P., Laforge, A., and Renaudie, J. 2022. Artificial intelligence applied to the classification of eight middle Eocene species of the genus *Podocyrtis* (polycystine radiolaria). Journal of Micropalaeontology, 41, 165-182. doi:10.5194/jm-41-165-2022.

Danelian, T., and Johnson, K.G. 2001. Patterns of biotic change in Middle Jurassic to Early Cretaceous Tethyan radiolaria. Marine Micropaleontology, 43, 239-260. doi:10.1016/S0377-8398(01)00029-9.

626

627 Danelian, T., Le Callonec, L., Erbacher, J., Mosher, D., Malone, M., Berti, D., Bice, K.,
628 Bostock, H., Brumsack, H.-J., Forster, A., Heidersdorf, F., Henderiks, J., Janecek, T., Junium,
629 C., MacLeod, K., Meyers, P., Mutterlose, J., Nishi, H., Norris, R., Ogg, J., O'Regan, M., Rea,
630 B., Sexton, P., Sturt, H., Suganuma, Y., Thurow, J., Wilson, P., Wise, S., and Glatz, C. 2005.
631 Preliminary results on Cretaceous-Tertiary tropical Atlantic pelagic sedimentation (Demerara
632 Rise, ODP Leg 207). Comptes Rendus Geoscience, 337(6), 609-616.
633 doi:10.1016/j.crte.2005.01.011.

634

635 Danelian, T., Saint Martin, S., and Blanc-Valleron, M.-M. 2007. Middle Eocene radiolarian
636 and diatom accumulation in the equatorial Atlantic (Demerara Rise, ODP Leg 207): Possible
637 links with climatic and palaeoceanographic changes. Comptes Rendus Palevol, 6, 103-114.
638 doi:10.1016/j.crpv.2006.08.002.

639

640 Danelian, T., Asatryan, G., Galoyan, G., Sosson, M., Sahakyan, L., Caridroit, M., Avagyan,
641 A. 2012. Geological history of ophiolites in the Lesser Caucasus and correlation with the
642 Izmir-Ankara-Erzincan suture zone: insights from radiolarian biochronology. Bull. Soc. Géol.
643 France 183, 331–342.

644

645 Danelian, T., and Monnet, C. 2021. Early Paleozoic radiolarian plankton diversity and the
646 Great Ordovician Biodiversification Event. Earth-Science Reviews, 218, no. 103672.
647 doi:/10.1016/j.earscirev.2021.103672.

648 Dollfus, D., and Beaufort, L. 1999. Fat neural network for recognition of position-normalised
649 objects. Neural Networks, 12, 553-560. doi:10.1016/S0893-6080(99)00011-8.

650

Ehrenberg, C.G., 1874. Grössere Felsproben des Polycystinen-Mergels von Barbados mit weiteren Erläuterungen. Monat. Königl. Akad. Wiss. Berlin, 1873: 213–262.

Erbacher, J., Mosher, D.C., Malone, M.J., et al. 2004. Proceedings of the Ocean Drilling Program, Initial Reports, 207. College Station, TX: Ocean Drilling Program. doi:10.2973/odp.proc.ir.207.2004.

Hay, W.W., DeConto, R., Wold, C.N., Wilson, K M., Voigt, S., Schulz, M., Wold-Rossby, A., Dullo, W.-C., Ronov, A.B., Balukhovsky, A.N., and Soeding, E. 1999. Alternative global Cretaceous paleogeography. In Barrera, E., and Johnson, C. (Eds.), The Evolution of Cretaceous Ocean/Climate Systems, Geological Society of America Special Paper 332, 1-47. doi:10.1130/0-8137-2332-9.1.

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 (CVPR), 770-778. doi:10.48550/arXiv.1512.03385

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861, doi:10.48550/arXiv.1704.04861.

Hsiang, A.Y., Brombacher, A., Rillo, M.C., Mleneck-Vautravers, M.J., Conn, S., Lordsmith, S., Jentzen, A., Henehan, M.J., Metcalfe, B., Fenton, I.S., Wade, B.S., Fox, L., Meilland, J., Davis, C. V., Baranowski, U., Groeneveld, J., Edgar, K.M., Movellan, A., Aze, T., Dowsett,

H. J., Giles Miller, C., Rios, N., and Hull, P. M. 2019. Endless Forams: >34,000 modern planktonic foraminiferal images for taxonomic training and automated species recognition using convolutional neural networks. Paleoceanography and Paleoclimatology, 34, 1157-1177. doi:10.1029/2019PA003612.

Itaki, T., Taira, Y., Kuwamori, N., Saito, H., Ikehara, M., and Hoshino, T. 2020. Innovative microfossil (radiolarian) analysis using a system for automated image collection and AI-based classification of species. Scientific Reports. doi:10.1038/s41598-020-77812-6.

Itaki, T., Utsuki, S., Haneda, Y., Izumi, K., Kubota, Y., Suganuma, Y., and Okada, M. 2022. Millennial-scale oscillations in the Kuroshio-Oyashio boundary during MIS 19 based on the radiolarian record from the Chiba composite section, central Japan. Progress in Earth and Planetary Science, 9(1), Article No. 5. doi:10.1186/s40645-021-00465-0

Marchant, R., Tetard, M., Pratiwi, A., Adebayo, M., and de Garidel-Thoron, T. 2020. Automated analysis of foraminifera fossil records by image classification using a convolutional neural network. Journal of Micropalaeontology, 39, 183-202. doi:10.1101/840926.

Matsuzaki, K.M., Itaki, T., Tada, R., and Kamikuri, S. 2018. Paleoceanographic history of the Japan Sea over the last 9.5 million years inferred from radiolarian assemblages (IODP Expedition 346 Sites U1425 and U1430). Progress in Earth and Planetary Science, 5(1), 54.

Mitra, R., Marchitto, T.M., Ge, Q., Zhong, B., Kanakiya, B., Cook, M.S., Fehrenbacher, J.S., Ortiz, J.D., Tripati, A., and Lobaton, E. 2019. Automated species-level identification of

698 planktic foraminifera using convolutional neural networks, with comparison to human
699 performance. Marine Micropaleontology, 148, 1-14. doi:10.1016/j.marmicro.2019.01.005.
700 Meunier, M., and Danelian, T. 2022. Astronomical calibration of late middle Eocene
701 radiolarian bioevents from ODP Site 1260 (equatorial Atlantic, Leg 207) and refinement of
702 the global tropical radiolarian biozonation. Journal of Micropalaeontology, 41(1), 1-27.
703
704 Meunier, M., and Danelian, T. 2023. Progress in understanding middle Eocene nassellarian
705 (Radiolaria, Polycystinea) diversity; new insights from the western equatorial Atlantic Ocean.
706 Journal of Paleontology, 97(1), 1-25. doi:10.1017/jpa.2022.82
707
708 Moore Jr., T.C. 1971. Radiolaria. In: Initial Reports of the Deep Sea Drilling Project, Volume
709 8, edited by J.I. Tracey Jr., G.H. Sutton, W.D. Nesteroff, J. Galehouse, C.C. Von der Borch,
710 T. Moore, J. Lipps, U.Z.B.U. Haq, and J.P. Beckmann, U.S. Govt. Print. Office, Washington,
711 DC, USA, pp. 727-775. doi:10.2973/dsdp.proc.8.112.1971.
712
713 Pouille, L., Obut, O., Danelian, T., Sennikov, N. 2011. Lower Cambrian (Botomian)
714 policystine Radiolaria from the Altai Mountains (southern Siberia, Russia). C. R. Palevol, 10,
715 627-633. doi:10.1016/j.crpv.2011.05.004.
716
717 Renaudie, J., Danelian, T., Saint Martin, S., Le Callonnec, L., and Tribovillard, N. 2010.
718 Siliceous phytoplankton response to a Middle Eocene warming event recorded in the tropical
719 Atlantic (Demerara Rise, ODP Site 1260A). Palaeogeography, Palaeoclimatology,
720 Palaeoecology, 286, 121-134. doi:10.1016/j.palaeo.2009.12.004.
721

722 Renaudie, J., Lazarus, D. 2013. On the accuracy of paleodiversity reconstructions: A case
723 study in Antarctic Neogene radiolarians. Paleobiology, 39(3), 491-509.
724
725 Renaudie, J., Gray, R., and Lazarus, D.B. 2018. Accuracy of a neural net classification of
726 closely-related species of microfossils from a sparse dataset of unedited images. PeerJ
727 Preprints, 6:e27328v1.

728 Riedel, W.R., Sanfilippo, A. 1970. Radiolaria, Leg 4, Deep Sea Drilling Project. In Bader,
729 R.G., Gerard, R.D., Benson, W.E., Bolli, H.M., Hay, W.W., Rothwell, Jr., T., Ruef, M.H.,
730 Riedel, W.R., Sayles, F.L. (Eds.), Initial Reports of the Deep Sea Drilling Project, 4. U.S.
731 Govt. Print. Office, Washington, DC, USA, 503-575. doi:10.2973/dsdp.proc.4.124.1970.
732
733 Riedel, W.R., Sanfilippo, A. 1978. Stratigraphy and evolution of tropical Cenozoic
734 radiolarians. Micropaleontology, 24, 61-96. doi:10.2307/1485420.
735
736 Sanfilippo, A., Westberg-Smith, M.J., Riedel, W.R. 1985. Cenozoic Radiolaria. In Bolli,
737 H.M., Saunders, J.B., Perch-Nielsen, K. (Eds.), Plankton Stratigraphy. Cambridge University
738 Press, Cambridge, UK, 631-712.
739
740 Sanfilippo, A., Nigrini, C. 1998. Code numbers for Cenozoic low latitude radiolarian
741 biostratigraphic zones and GPTS conversion tables. Marine Micropaleontology, 33, 109-156.
742 doi:10.1016/S0377-8398(97)00030-3.
743

744 Suzuki, N., O'Dogherty, L., Caulet, J.-P., Dumitrica, P. 2021. A new integrated morpho- and
745 molecular systematic classification of Cenozoic radiolarians (Class Polycystinea) -
746 suprageneric taxonomy and logical nomenclatorial acts. In O'Dogherty, L. (Ed.), Catalog of
747 Cenozoic radiolarians. Geodiversitas, 43(15), 405-573. doi:10.5252/geodiversitas2021v43a15.

748

749 Tetard, M., Marchant, R., Cortese, G., Gally, Y., Thibault de Garidel-Thoron, et al. 2020.
750 Technical note: A new automated radiolarian image acquisition, stacking, processing,
751 segmentation and identification workflow. Climate of the Past, 16, 2415-2429.
752 doi:10.5194/cp-16-2415-2020.

753 Trubovitz, S., Lazarus, D., Renaudie, J., Noble, P. 2020. Marine plankton show threshold
754 extinction response to Neogene climate change. Nature Communications, 11, Article no.
755 5069. doi:10.1038/s41467-020-18879-7.

756

757 Vrielynck, B., Bonneau, M., Danelian, T., Cadet, J.P., Poisson, A. 2003. New insights on the
758 Antalya Nappes in the apex of the Isparta angle: The Isparta Cay unit revisited. Geol. J. 38,
759 283-293.

760

Figure 1. **1**) For each sample, 20 x 20 images are automatically photographed in a convolving way and merged into one large "mosaic composed" image, which enables preserving more complete specimens which are not cut in half. **2**) The mosaic composed image thereafter went through grayscale conversion. **3**) Segmentation of each unique particle into vignettes and image conversion into 8-bit black and white with black background.

Figure 2. A theoretical example showing the importance of a high accuracy and a high recall value for individual classes.

Figure 3. Simplified confusion matrix, showing the classification between different classes, with a focus on nassellarian super families. The x-axis on the left shows the true classes while the right axis shows recall values; the y-axis at the bottom shows the predicted classes while the top shows the precision value.

Figure 4. CNN performance metrics based on a test set consisting of a total of 800 images, which were validated by a human expert.

Plate 1. Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of Demerara Rise; species names are followed by the ODP site and hole, core, section and sampled level from which it comes from. **A)** *Dendrospyris stylophora* (Ehrenberg 1874) from 1259A-20R-4W, 53-55 cm, **B)** *Elaphospyris didiceros* (Ehrenberg 1874) group from 1258A-2R-4W, 55-57 cm, **C)** *Liriospyris clathrata* (Ehrenberg, 1854) group from 1259A-20R-4W, 53-55 cm, **D)** *Rhabdolithis pipa* Ehrenberg 1854 from 1260A-15R-1W, 55-57 cm, **E)**

Lophophaena radians (Ehrenberg, 1874) group from 1259A-16R-2W, 55-57 cm, **F)** *Dictyoprora mongolfieri* (Ehrenberg 1854) from 1260A-6R-2W, 55-57 cm, **G)** *Dictyoprora amphora* (Haeckel, 1887) group from 1258A-2R-4W, 55-57 cm, **H)** *Rhopalosyringium* ? *auriculaleporis* (Clark and Campbell, 1942) from 1260A-14R-6W, 55-57 cm, **I)** *Rhopalosyringium* ? *biauritum* (Ehrenberg, 1874) from 1260A-12R-3W, 55-57 cm, **J)** *Dictyomitra parva* (Kim 1992) from 1258A-2R-4W, 55-57 cm, **K)** *Lithochytris vespertilio* Ehrenberg 1874 from 1260A-10R-5W, 55-57 cm, **L)** *Sethochytris triconiscus* (Haeckel 1887) from 1259A-20R-4W, 53-55 cm, **M)** *Lychnocanium babylonis* (Clark and Campbell 1942) group from 1258A-3R-3W, 56-58 cm, **N)** *Lychnocanoma bajunensis* (Renz 1984) from 1258A-2R-4W, 55-57 cm, **O)** *Stichopterygium microporum* (Ehrenberg 1874*)* from 1260A-8R-4W, 54-56 cm, **P)** *Carpocanopsis ornata* (Ehrenberg, 1874) group from 1260A-6R-5W, 55-57 cm, **Q)** *Cycladophora spatiosa* (Ehrenberg 1874) group from 1259A-20R-1W, 55-57 cm, **R)** *Anthocyrtis mespilus* (Ehrenberg 1847) group from 1259A-16R-2W, 55-57 cm, **S)** *Zealithapium mitra* (Ehrenberg 1874) from 1258A-2R-3W, 55-57 cm.

Plate 2. Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of Demerara Rise; species names are followed by the ODP site and hole, core, section and sampled level from which it comes from. **A)** *Lophocyrtis alauda* (Ehrenberg, 1874) from 1260A-15R-1W, 55-57 cm, **B)** *Aphetocyrtis zamenhofi* Meunier and Danelian, 2023 from 1259A-26R-5W, 54-56 cm, **C)** *Theocyrtis scolopax* (Ehrenberg, 1874) from 1260A-15R-3W, 55-57 cm, **D)** *Phormocyrtis embolum* (Ehrenberg, 1874) from 1258A-2R-4W, 55-57 cm, **E)** *Phormocyrtis lazari* Meunier and Danelian, 2023 from 1260A-8R-6W, 54-56 cm, **F)** *Podocyrtis (Lampterium) goetheana* (Haeckel, 1887) *from* 1259A-18R-1W, 53-55 cm, **G)** *Podocyrtis (Lampterium) chalara* Riedel and Sanfilippo 1970 from 1260A-6R-CC, 63-177 cm, **H)** *Podocyrtis (Lampterium) mitra* Ehrenberg group, 1854 from 1260A-9R-1W, 55-57

cm, **I)** *Podocyrtis (Lampterium) sinuosa* Ehrenberg 1874 from 1259A-15R-1W, 55-57 cm, **J)** *Podocyrtis papalis* Ehrenberg, 1847 from 1258A-2R-4W, 55-57 cm, **K)** *Podocyrtis (Podocyrtoges) ampla* Ehrenberg 1874 from 1260A-10R-5W, 55-57 cm, **L)** *Podocyrtis (Podocyrtoges) phyxis* Sanfilippo and Riedel, 1973 from 1259A-16R-1W, 55-57 cm, **M)** *Podocyrtis (Podocyrtoges) diamesa* Riedel and Sanfilippo, 1970 from 1259A-26R-3W, **N)** *Podocyrtis (Lampterium) puellasinensis* Ehrenberg 1874 *from* 1259A-20R-4W, 53-55 cm, **O)** *Calocyclas hispida* (Ehrenberg, 1874) from 1260A-6R-4W, 55-57 cm, **P)** *Thyrsocyrtis (Thyrsocyrtis) rhizodon* Ehrenberg 1874 from 1260A-6R-CC, 63-177 cm, **Q)** *Thyrsocyrtis (Pentalocorys) triacantha* (Ehrenberg 1874) from 1260A-8R-6W, 54-56 cm, **R)** *Eusyringium lagena* (Ehrenberg 1874) from 1259A-25R-2W, 55-57 cm, **S)** *Eusyringium fistuligerum* (Ehrenberg 1874) group from 1259-18R-1W, 53-55 cm, **T)** *Rhopalocanium ornatum* (Ehrenberg 1874) from 1259A-22R-1W, 55-57 cm.