



HAL
open science

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

► **To cite this version:**

Veronica Carlsson, Taniel Danelian, Martin Tetard, Mathias Meunier, Pierre Boulet, et al.. Convolutional neural network application on a new middle Eocene radiolarian dataset. *Marine Micropaleontology*, 2023, 183, pp.102268. 10.1016/j.marmicro.2023.102268 . hal-04207164

HAL Id: hal-04207164

<https://hal.science/hal-04207164v1>

Submitted on 14 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

1 **Convolutional neural network application on a new middle Eocene radiolarian dataset**

2 Veronica Carlsson^{1,2*}, Taniel Danelian¹, Martin Tetard³, Mathias Meunier¹, Pierre Boulet²,
3 Philippe Devienne², Sandra Ventalon⁴

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

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

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

8 ⁴Univ. Lille, CNRS, Univ. Littoral Cote d’Opale, UMR 8187, LOG, F-59000 Lille, France.

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

10 **Abstract**

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

14

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

21

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

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

27

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

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

33 **1 Introduction**

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

43

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

48 between different taxonomic experts. Therefore, artificial intelligence (AI) has been
49 introduced to this field to simplify or automate the work done by micropaleontologists, as for
50 example through automatic image recognition and counting. Several applications of CNNs for
51 automatic image recognition were developed over the last 20 years, since the introduction of
52 SYRACO by Dollfus and Beaufort (1999). Nowadays, CNNs are used on various microfossil
53 groups, such as foraminifera (ex. Mitra et al., 2019, Hsiang et al., 2019; Marchant et al.,
54 2020), coccoliths (ex. Dollfus and Beaufort, 1999; Beaufort and Dollfus, 2004), pollen (ex.
55 Bourel et al., 2020), or even radiolarians (Itaki et al., 2020; Renaudie et al., 2018; Tetard et
56 al., 2020).

57

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

65

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

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

78

79 **2 Materials and methods**

80 **2.1 Core setting and sample preparation**

81 The middle Eocene samples used in this study were collected during ODP Leg 207 from
82 Demerara Rise, situated off the coast of Suriname (Erbacher et al., 2004, Danelian et al.
83 2005). This dataset includes samples from several cores recovered from sites 1258, 1259 and
84 1260. It is noteworthy that the middle Eocene sequence at site 1260 is thick and contains
85 radiolarians of an excellent state of preservation (Danelian et al., 2007). The full sample list
86 used in this study may be found in Supplements 1, Table S1. The sediment samples consist
87 primarily of nannofossil and foraminifera chalk, but also contain abundant and well-preserved
88 siliceous microfossils, composed essentially of radiolarians (Danelian et al. 2007, Meunier
89 and Danelian, 2022), as well as diatoms (Danelian et al. 2007, Renaudie et al. 2010).

90

91 Sediment samples from ODP holes 1258A, 1259A, and 1260A were first processed to
92 disaggregate organic matter and dissolve carbonates and were then sieved through a 45 µm
93 mesh to remove smaller particles. Thereafter, the samples were prepared using a recent
94 random settling protocol described in Tetard et al. (2020). About 1/3 to 1/4 of a microspoon
95 spatula was used from the dried residues, corresponding to approximately 0.2-1.0 g for each
96 sample. Samples were uniformly settled onto 12 mm x 12 mm cover slides using a 3D-printed
97 decanter, as in Tetard et al. (2020), preventing contact between radiolarians and other

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

101

102 Finally, a new set of four samples (see Supplements 1, Table S1) was prepared for manually
103 testing the actual accuracy of the trained CNN. An improved cleaning technique was used,
104 which kept only siliceous particles that are larger than 63 μm , and completely removed all
105 clay, calcite and smaller or broken radiolarians (Sanfilippo et al., 1985, Tetard et al., 2020). In
106 fact, most radiolarians are larger, so using a 63 μm sieve will probably just remove smaller
107 broken pieces or radiolarians rather than the radiolarians themselves. At first, about 2-3 cm of
108 raw sediment sample was placed in a plastic beaker. Thereafter 30 ml of 30 % hydrochloric
109 acid was added and left for two hours and until there was no more reaction. Furthermore, 200
110 ml of distilled water was added to the beaker, which was stirred gently and left to settle for
111 two additional hours. The supernatant was removed and exchanged with 30 ml of 10 %
112 hydrogen peroxide and was left to rest for another two hours. The residue was thereafter
113 washed with a 63 μm sieve into a 100 ml beaker. To remove remaining clay particles, the
114 samples were processed in ultrasonic bath for ten minutes; they were later sieved again with a
115 63 μm mesh and collected afterwards in a filter. Next, the residues were dried in an oven at 50
116 $^{\circ}\text{C}$ and transferred into a glass vial.

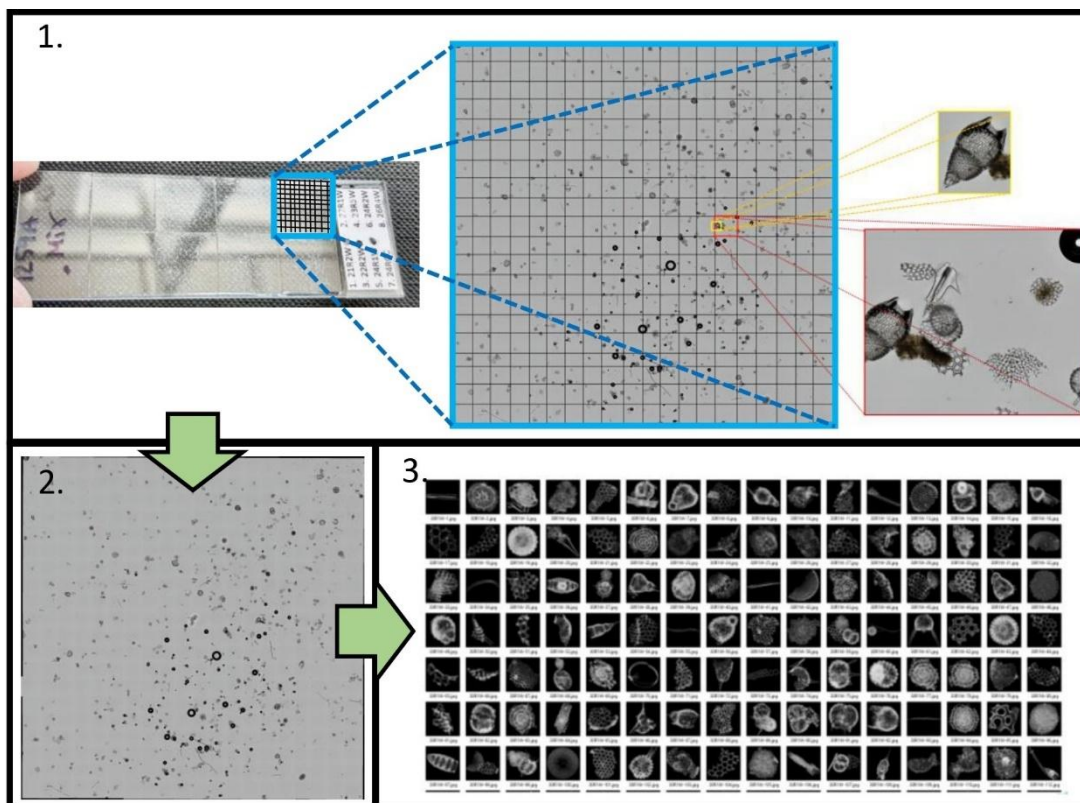
117

118

119 **2.2 Image collection and processing**

120 All samples were photographed using a Nikon Eclipse Ni automatic microscope equipped
121 with a Nikon DS-Ri2 microscope camera and Nikon NIS Element software, using a 20 x
122 objective, allowing a 200 x magnification and 0.36 $\mu\text{m}/\text{pixel}$ resolution. The lowest and

123 highest focal points on the Z-axis were manually determined for each sample since the depth
124 varied depending on the thickness of the glass, coverslip, optical glue, and individual
125 radiolarian specimens. For each field of view (FOV), an image was taken at different focal
126 depths, every 10 μm , depending on the minimum and maximum focal points determined for
127 each sample. The microscope then automatically stacked all images, taken at different focal
128 points for each FOV, in order to create a composite image. The microscope was programmed
129 to capture small images in 20 x 20 (400) FOV, covering about 10 x 10 mm out of the 12 x 12
130 mm available surface for each individual sample. The individual FOV images were
131 subsequently merged into one large mosaic image (see Fig. 1), which has as a result to lose
132 less images of specimens located on the edges of FOVs.
133



134
135 **Figure 1.** 1) For each sample, 20 \times 20 images are automatically photographed in a convolving way and merged into one large
136 “mosaic composed” image, which enables preserving more complete specimens which are not cut in half. 2) The mosaic

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

139

140

141 The composite mosaic pattern images received from the automatic microscope were first
142 converted from RGB to 8-bit grayscale using Adobe Photoshop to decrease its size. The
143 images were then segmented into ImageJ using the ImageJ BioVoxel plugin (Brocher,
144 2022), and a modified script of the Autoradio_Segmenter plugin (Tetard et al., 2020), which
145 enabled each individual particle to become its own individual image, or vignette. For more
146 details, the reader is referred to Tetard et al. (2020).

147

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

150

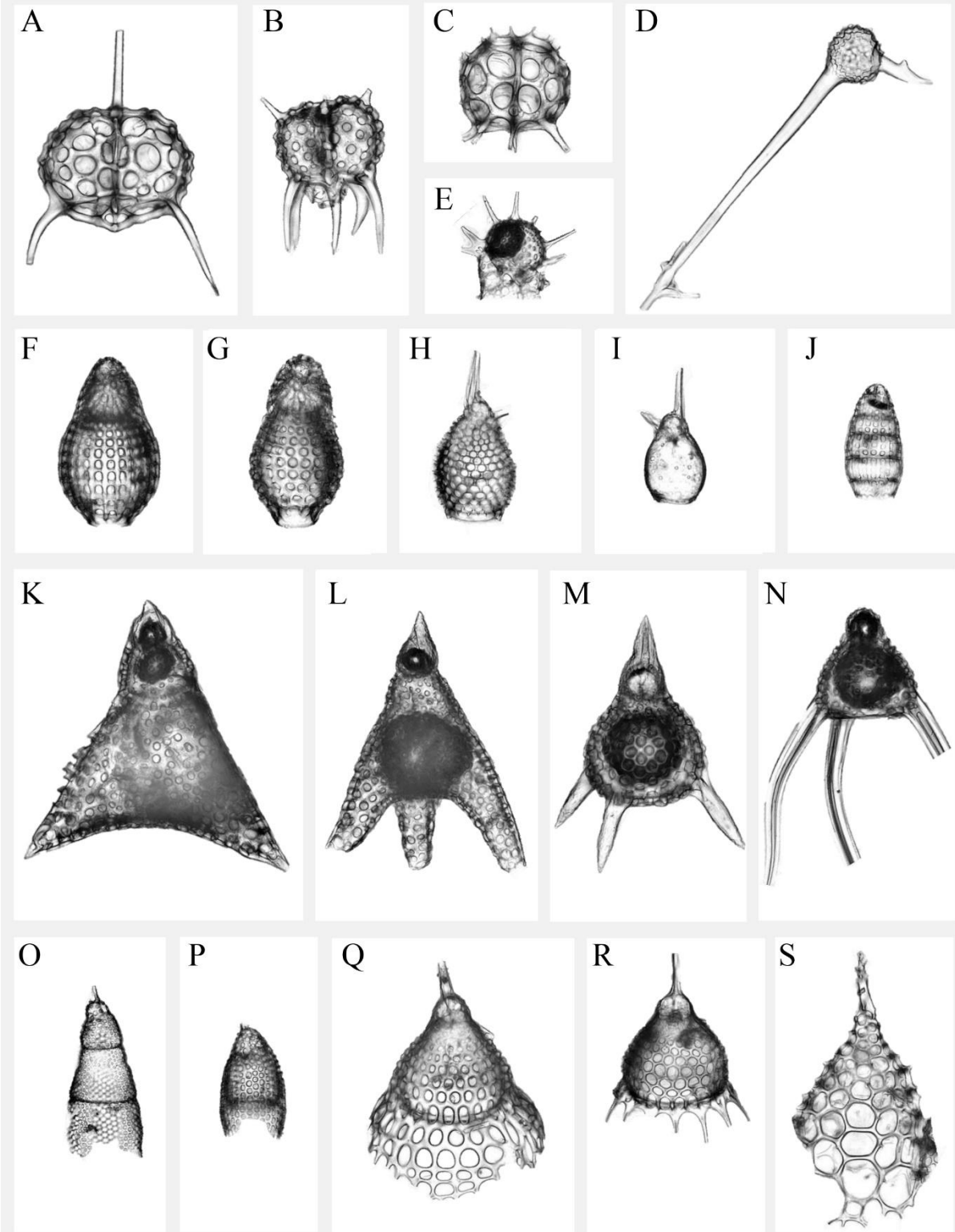
151 **2.3 Taxa selection and dataset**

152 For this study we decided to focus on 39 species (Plates 1 and 2), which are the most common
153 in the Middle Eocene interval of Demerara Rise and most of which are used in biostratigraphy
154 (Riedel and Sanfilippo, 1970, 1978; Sanfilippo and Nigrini, 1998; Meunier and Danelian
155 2022).

156

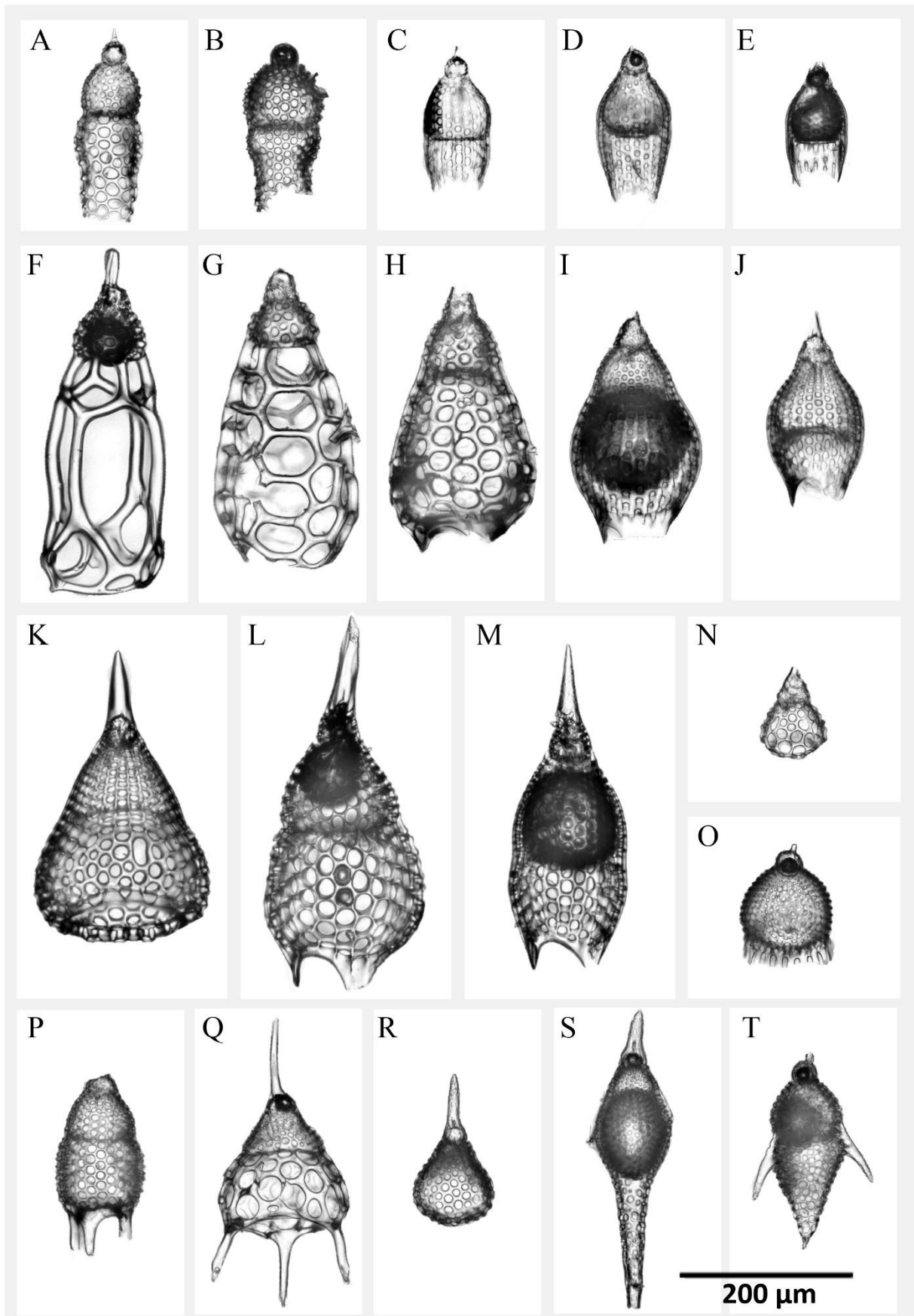
157 **Plate 1.** Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of Demerara Rise; species names
158 are followed by the ODP site and hole, core, section and sampled level from which it comes from. A) *Dendrospyris*
159 *stylophora* (Ehrenberg, 1874) from 1259A-20R-4 W, 53–55 cm, B) *Elaphospyris didiceros* (Ehrenberg, 1874) group from
160 1258A-2R-4 W, 55–57 cm, C) *Liriospyris clathrata* (Ehrenberg, 1874) group from 1259A-20R-4 W, 53–55 cm,
161 D) *Rhabdolithis pipa* Ehrenberg 1854 from 1260A-15R-1 W, 55–57 cm, E) *Lophophaena radians* (Ehrenberg, 1874) group
162 from 1259A-16R-2 W, 55–57 cm, F) *Dictyoprora mongolfieri* (Ehrenberg, 1854) from 1260A-6R-2 W, 55–57 cm,
163 G) *Dictyoprora amphora* (Haeckel, 1887) group from 1258A-2R-4 W, 55–57 cm,

164 H) *Rhopalosyringium? auriculaleporis* (Clark and Campbell, 1942) from 1260A-14R-6 W, 55–57 cm,
165 I) *Rhopalosyringium? biauratum* (Ehrenberg, 1874) from 1260A-12R-3 W, 55–57 cm, J) *Dictyomitra parva* (Kim, 1992)
166 from 1258A-2R-4 W, 55–57 cm, K) *Lithochytris vespertilio* Ehrenberg, 1874 from 1260A-10R-5 W, 55–57 cm,
167 L) *Sethochytris triconiscus* (Haeckel 1887) from 1259A-20R-4 W, 53–55 cm, M) *Lychnocanium babylonis* (Clark and
168 Campbell 1942) group from 1258A-3R-3 W, 56–58 cm, N) *Lychnocanoma bajunensis* (Renz, 1984) from 1258A-2R-4 W,
169 55–57 cm, O) *Stichopterygium microporum* (Ehrenberg, 1874) from 1260A-8R-4 W, 54–56 cm, P) *Carpocanopsis*
170 *ornata* (Ehrenberg, 1874) group from 1260A-6R-5 W, 55–57 cm, Q) *Cycladophora spatiosa* (Ehrenberg, 1874) group from
171 1259A-20R-1 W, 55–57 cm, R) *Anthocyrtilis mespilus* (Ehrenberg, 1847) group from 1259A-16R-2 W, 55–57 cm,
172 S) *Zealithapium mitra* (Ehrenberg, 1874) from 1258A-2R-3 W, 55–57 cm.



200 μ m

174 **Plate 2.** Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of Demerara Rise; species names
175 are followed by the ODP site and hole, core, section and sampled level from which it comes from. A) *Lophocyrtis*
176 *alauda* (Ehrenberg, 1874) from 1260A-15R-1 W, 55–57 cm, B) *Aphetocyrtis zamenhofi* Meunier and Danelian, 2023 from
177 1259A-26R-5 W, 54–56 cm, C) *Theocyrtis scolopax* (Ehrenberg, 1874) from 1260A-15R-3 W, 55–57 cm, D) *Phormocyrtis*
178 *embolum* (Ehrenberg, 1874) from 1258A-2R-4 W, 55–57 cm, E) *Phormocyrtis lazari* Meunier and Danelian, 2023 from
179 1260A-8R-6 W, 54–56 cm, F) *Podocyrtis (Lampterium) goetheana* (Haeckel, 1887) from 1259A-18R-1 W, 53–55 cm,
180 G) *Podocyrtis (Lampterium) chalara* Riedel and Sanfilippo, 1970 from 1260A-6R-CC, 63–177 cm, H) *Podocyrtis*
181 *(Lampterium) mitra* Ehrenberg group, 1854 from 1260A-9R-1 W, 55–57 cm, I) *Podocyrtis (Lampterium) sinuosa* Ehrenberg,
182 1874 from 1259A-15R-1 W, 55–57 cm, J) *Podocyrtis papalis* Ehrenberg, 1847 from 1258A-2R-4 W, 55–57 cm,
183 K) *Podocyrtis (Podocyrtoges) ampla* Ehrenberg, 1874 from 1260A-10R-5 W, 55–57 cm, L) *Podocyrtis (Podocyrtoges)*
184 *phyxis* Sanfilippo and Riedel, 1973 from 1259A-16R-1 W, 55–57 cm, M) *Podocyrtis (Podocyrtoges) diamesa* Riedel and
185 Sanfilippo, 1970 from 1259A-26R-3 W, N) *Podocyrtis (Lampterium) puellasinensis* Ehrenberg, 1874 from 1259A-20R-4 W,
186 53–55 cm, O) *Calocyclus hispida* (Ehrenberg, 1874) from 1260A-6R-4 W, 55–57 cm, P) *Thyrsocyrtis (Thyrsocyrtis)*
187 *rhizodon* Ehrenberg, 1874 from 1260A-6R-CC, 63–177 cm, Q) *Thyrsocyrtis (Pentalocorys) triacantha* (Ehrenberg, 1874)
188 from 1260A-8R-6 W, 54–56 cm, R) *Eusyngium lagena* (Ehrenberg, 1874) from 1259A-25R-2 W, 55–57 cm,
189 S) *Eusyngium fistuligerum* (Ehrenberg, 1874) group from 1259 to 18R-1 W, 53–55 cm, T) *Rhopalocanium*
190 *ornatum* (Ehrenberg, 1874) from 1259A-22R-1 W, 55–57 cm.



191

192

193

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

199

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

209

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

218

219 **2.4 CNN training**

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

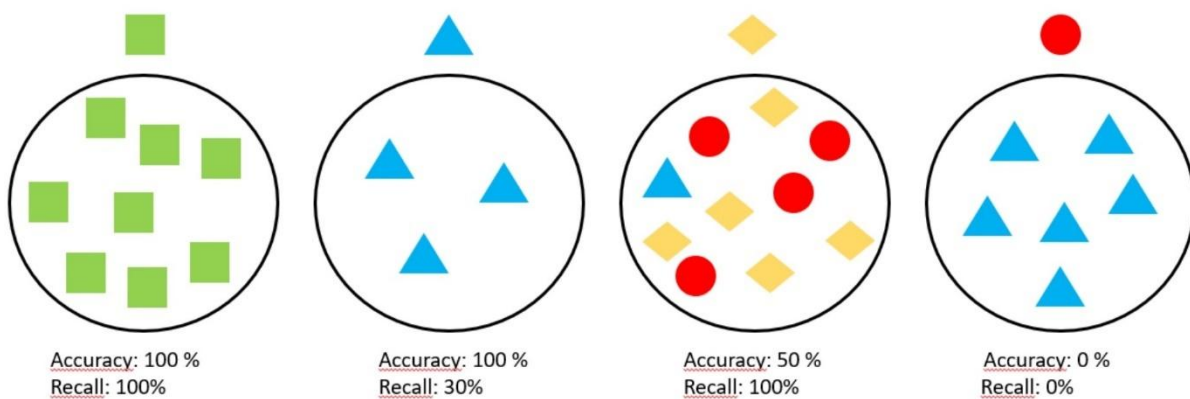
236

237 **2.5 Performance metrics**

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

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

249



250

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

252

253 2.6. Test set to validate the CNN

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

261 We randomly let the ParticleTrieur software pick 200 images selected from four new samples
262 (800 in total), which were unlabeled and contained all particles segmented from composed
263 images, from ODP site 1260, coming from different intervals within those used initially to
264 train the CNN, which were 1260A-6R-3W, 18-20 cm; 1260A-8R-5W, 70-72 cm; 1260A-13R-
265 5W, 66-68 cm; and 1260A-15R-4W, 69-71 cm. In ParticleTrieur we let the trained CNN
266 identify all of them, using a threshold value of 0.5. We selected this low threshold value since
267 it is better having more images classified, even if that will give a somewhat lower accuracy,
268 recall, precision to classify more images rather than that they unclassified. We then examined
269 afterwards what was correctly or incorrectly classified.

270

271 **2.7 Application on 39 species**

272 With the same four samples, we then this time entered all segmented particles from the entire
273 mosaic composed images covering most parts of the coverslips, resulting in a total of 15 932
274 images, which were automatically classified with the CNN. Here we focused on the targeted
275 39 classes representing the selected well-known species. We applied different threshold (1-
276 0.5) values for the pre-trained network and checked how the CNN could recognize these 39
277 most common nassellarian species. The aim of this test was to try future potential
278 applications.

279

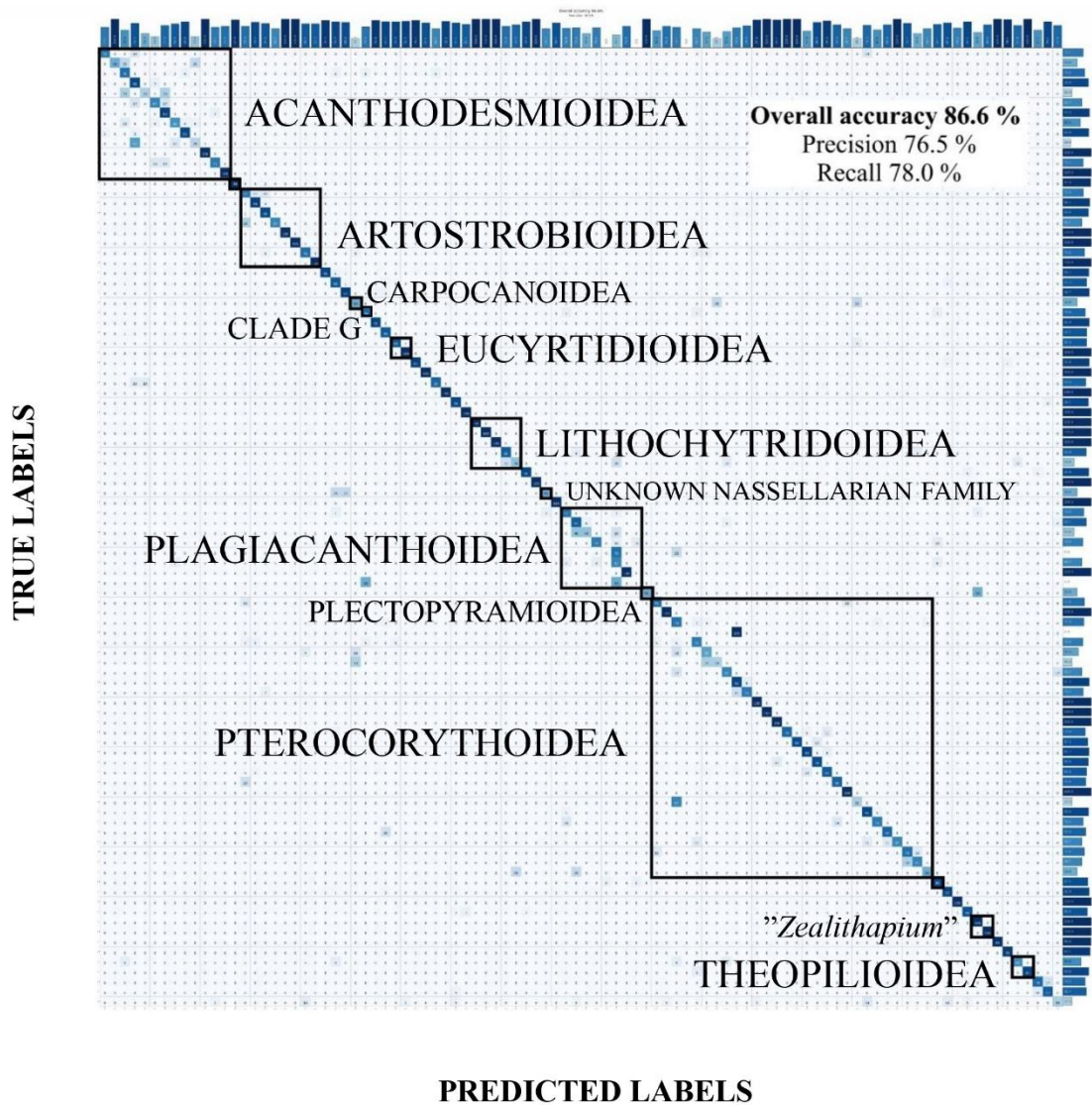
280

281 **3 Results**

282 **3.1 Training of the initial dataset**

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

285 of specimens classified as a class and also belonging to that class, divided by all specimens
 286 classified to that class. Our CNN obtained 78 % of recall, which calculates the ability to
 287 detect all correct classifications, as it corresponds to the number of specimens in a class that
 288 were correctly classified divided by the total number of specimens in that class. The training
 289 and validation iterations are given in Supplements 2 figure S1, which confirms that the data is
 290 neither overfit, nor underfit.
 291



292
 293 **Figure 3.** Simplified confusion matrix, showing the classification between different classes, with a focus on nassellarian
 294 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
 295 shows the predicted classes while the top shows the precision value.

296

297

298 Some of the classes that are visible in the confusion matrix (Fig. 2 or Fig. S2 in Supplements
299 2 for a more detailed confusion matrix) show a low individual score (diagonal numbers),
300 mainly due to the low number of available specimens (see number next to the label name), or
301 to a high degree of resemblance between closely related or similar looking species. Most of
302 the 39 important species of nassellarians we focused on were classified with a high accuracy
303 (Fig. S2 in Supplements 2).

304

305 **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
<i>Dendrospyris stylophora</i>	35	7	0.71	0.71	0.71
<i>Elaphospyris didiceros</i> group	295	59	0.85	0.90	0.88
<i>Liriospyris clathrata</i> group	77	15	0.78	0.93	0.85
<i>Dictyomitra parva</i>	229	46	0.98	0.98	0.98
<i>Dictyoprora amphora</i> group	124	25	0.75	0.72	0.73
<i>Dictyoprora mongolfieri</i>	271	54	0.89	0.94	0.92
<i>Rhopalosyringium auriculaleporis</i>	104	21	0.84	0.76	0.80
<i>Rhopalosyringium biaurata</i>	24	5	0.83	1.00	0.91
<i>Carpocanopsis ornata</i> group	21	4	0.33	0.50	0.40
<i>Stichopterygium microporum</i>	64	13	0.93	1.00	0.96
<i>Sethochytris triconiscus</i>	22	4	1.00	1.00	1.00
<i>Lithochytris vespertilio</i>	20	4	1.00	1.00	1.00

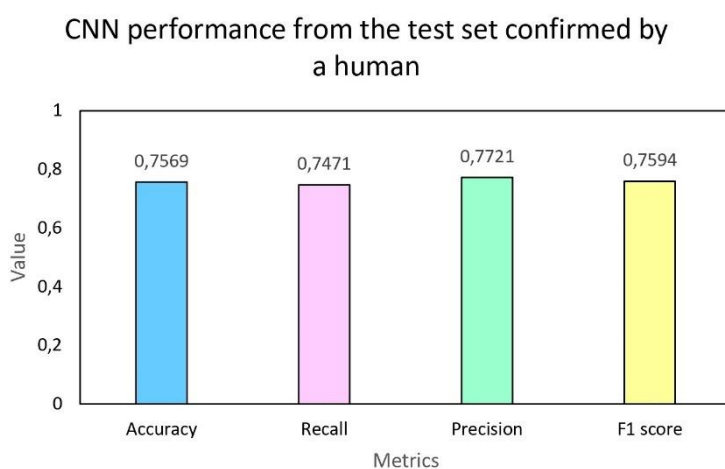
Species	Training (#)	Test (#)	Precision	Recall	F1 score
<i>Lychnocanoma bajunensis</i>	103	21	1.00	1.00	1.00
<i>Lychnocanium babylonis</i> group	56	11	0.75	0.82	0.78
<i>Lophophaena radians</i> group	48	10	0.70	0.70	0.70
<i>Rhabdolithis pipa</i>	28	6	0.75	1.00	0.86
<i>Aphetocyrtis zamenhofi</i>	40	8	0.60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

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

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

318 **3.3 Application on new samples**

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

331

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

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

334

335 **4 Discussion**

336 **4.1 Classification**

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

348

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

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

365

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

369

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

377

378 **4.2 CNN training and new test set score**

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

381 were used to train the CNN. The accuracy of the test is 75.69 %, whereas the training
382 accuracy is 85.6 %. From our labeled dataset, we have purposely removed a lot of “trash”
383 particles, that will say particles which are broken fragments of radiolarians, blurry
384 background particles etc. because if we kept these images, the CNN would rather be
385 overtrained by the thousands of trash images and perform less. Besides that, rare species
386 which consisted of too few species were left untrained by the neural network and therefore the
387 20 % of test did not include that many “trash” images or any rare species, in contrast to the
388 new test set, which were completely randomly selected among any kind of particle that had
389 been segmented. This last test was just to confirm how well the CNN was generally trained.
390 For our last application we tried to just focus on the 39 well known radiolarian species, since
391 it is the radiolarians that are of interest.

392

393 **4.3 Feedback on individual species**

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

400

401 **Table 3.** Training accuracy, recall and number of correctly identified specimens for each one of the 39 targeted species in this
402 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
403 have an identification correctness equal or higher than 0.5. Hyphens “-” correspond to species not found in the samples
404 (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
<i>Dendrospyris stylophora</i>	0.71	0.71	1	4	1	0
<i>Elaphospyris didiceros</i> group	0.85	0.90	16	27	10	48
<i>Liriospyris clathrata</i> group	0.78	0.93	12	2	7	10
<i>Dictyomitra parva</i>	0.98	0.98	–	–	–	108
<i>Dictyoprora mongolferi</i>	0.89	0.94	36	9	38	14
<i>Dictyoprora amphora</i> group	0.75	0.72	1	2	8	68
<i>Rhopalosyringium? auriculaleporis</i>	0.84	0.76	–	2	3	3
<i>Rhopalosyringium? biaurata</i>	0.83	1.00	–	0	1	0
<i>Carpocanopsis ornata</i> group	0.33	0.50	12	0	1	–
<i>Stichopterygium microporum</i>	0.93	1.00	2	3	6	3
<i>Sethochytris triconiscus</i>	1.00	1.00	0	10	–	–
<i>Lithochytris vespertilio</i>	1.00	1.00	–	3	3	6
<i>Lychnocanoma bajunensis</i>	1.00	1.00	46	25	23	10
<i>Lychnocanium babylonis</i> group	0.75	0.82	–	5	14	8
<i>Lophophaena radians</i> group	0.70	0.70	2	0	2	–
<i>Rhabdolithis pipa</i>	0.75	1.00	2	6	1	3
<i>Lophocyrtis alauda</i>	0.75	1.00	–	–	8	9
<i>Aphetocyrtis zamenhofi</i>	0.60	0.75	–	–	0	20
<i>Theocyrtis scolopax</i>	0.00	0.00	–	–	2	1
<i>Calocyclas hispida</i>	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
<i>Phormocyrtis embolum</i>	0.71	0.92	–	–	16	3
<i>Phormocyrtis lazari</i>	0.77	0.71	9	4	–	–
<i>Podocyrtis (Lampterium) chalara</i>	0.98	1.00	7	20	–	–
<i>Podocyrtis goetheana</i> (Lampterium)	1.00	1.00	2	–	–	–
<i>Podocyrtis (Lampterium) mitra</i>	0.97	1.00	–	2	1	–
<i>Podocyrtis (Lampterium) sinuosa</i>	1.00	0.79	–	–	1	2
<i>Podocyrtis (Podocyrtoges) ampla</i>	1.00	0.88	–	–	4	0
<i>Podocyrtis (Podocyrtoges) phyxis</i>	0.89	0.89	–	–	–	0
<i>Podocyrtis (Podocyrtoges) diamesa</i>	0.58	0.92	–	–	–	0
<i>Podocyrtis (Podocyrtis) papalis</i>	0.94	0.83	8	6	1	6
<i>Podocyrtis puellasinensis</i> (Lampterium)	0.80	0.80	0	0	–	–
<i>Thyrsocyrtis rhizodon</i>	1.00	0.67	15	2	4	11
<i>Pentalocorys triacantha</i>	0.78	0.90	12	43	7	4
<i>Eusyringium lagena</i>	0.50	0.67	–	–	3	7
<i>Eusyringium fistuligerum</i> group	0.86	0.75	9	5	2	0
<i>Rhopalocanium ornatum</i> group	0.67	0.67	3	0	6	1
<i>Cycladophora spatiosa</i> group	1.00	0.91	31	20	2	1
<i>Anthocyrtis mespilus</i> group	0.78	0.64	14	3	10	3
<i>Zealithapium mitra</i>	0.54	1.00	7	7	7	1

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

411

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

417

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

422

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

429

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

434

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

438

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

444

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

447

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

452

453 *Lophocyrtis alauda* (Plate 2.A) is well detected in samples coming from 1260A-13R-5W, 66-
454 68 cm and 1260A-15R-4W, 69-71 cm. The somewhat similar looking species, *Aphetocyrtis*

455 *zamenhofi* (Meunier and Danelian 2023) was only found present in 1260A-15R-4W, 69-71
456 cm but other objects or specimens were also mistakenly classified as *A. zamenhofi*, even at
457 samples in which they do not exist. The training accuracies are 60 % for *A. zamenhofi* and 75
458 % for *L. alauda*.

459

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

464

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

472

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

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

483

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

492

493 *Podocyrtis phyxis* (Plate 2.L), an important biostratigraphic index species occurring only
494 within a short interval, was trained in the CNN with an 89 % accuracy and a recall value of 88
495 %; however, it was never classified correctly into its own class in the new test set of four
496 samples. Instead, it was frequently misidentified as *Thyrsocyrtis rhizodon* (Plate 2.P), which
497 is understandable, given that both species have an equal number of segments and consist of a
498 horn and feet (although they differ in size and shape) and are more or less barrel shaped.
499 Unfortunately, the CNN cannot detect size differences because all images are resized to the
500 same dimensions.

501

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

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

509

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

512

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

517

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

523

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

528

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

533

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

546 **5 Conclusions**

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

549

550 We evaluated the performance of our trained Convolutional Neural Network (CNN) on new
551 tests and compared it to human performance, and obtained a testing accuracy of about 75.69
552 %. We thereafter specifically focused on 39 different species which the CNN demonstrated
553 notable success in accurately identifying those species that had been well-trained.

554

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

563

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

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

574

575 **Data availability**

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

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

580 **Acknowledgments**

581 This study was partly funded by the French government through the program
582 “Investissements d’avenir” (I-ULNE SITE/ANR-16-IDEX-0004 ULNE) managed by the
583 National Research Agency. It also received funding from the European Union’s Horizon 2020
584 research and innovation program under the Marie Skłodowska-Curie grant agreement no.
585 847568. It was also supported by UMR 8198 Evo-Eco-Paléo and IRCICA (CNRS and Univ.
586 Lille USR-3380).

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

590 **Declaration of interest**

591 None.

592 **References**

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

595

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

600

601 Beaufort, L., and Dollfus, D. 2004. Automatic recognition of coccoliths by dynamical neural
602 networks. *Marine Micropaleontology*, 51, 57-73. doi:10.1016/j.marmicro.2003.09.003.
603

604 Bourel, B., Marchant, R., de Garidel-Thoron, T., Tetard, M., Barboni, D., Gally, Y., and
605 Beaufort, L. 2020. Automated recognition by multiple convolutional neural networks of
606 modern, fossil, intact and damaged pollen grains. *Computers and Geosciences*, 140, 104498.
607 doi:10.1016/j.cageo.2020.104498.
608

609 Brocher, J. 2022. biovoxxel/BioVoxxel-Toolbox: BioVoxxel Toolbox v2.5.3. Zenodo.
610 doi:10.5281/zenodo.5986129.
611

612 Carlsson, V. 2022. Podocyrtils Image Dataset [data set]. Recherche Data Gouv.
613 doi:10.57745/G7CHQL.
614

615 Carlsson, V. 2023. Middle Eocene Radiolarian image dataset from ODP Leg 207 (Demerara
616 Rise) [data set]. Recherche Data Gouv. doi: 10.57745/E9YXW6.
617

618 Carlsson, V., Danelian, T., Boulet, P., Devienne, P., Laforge, A., and Renaudie, J. 2022.
619 Artificial intelligence applied to the classification of eight middle Eocene species of the genus
620 *Podocyrtils* (polycystine radiolaria). *Journal of Micropalaeontology*, 41, 165-182.
621 doi:10.5194/jm-41-165-2022.
622

623 Danelian, T., and Johnson, K.G. 2001. Patterns of biotic change in Middle Jurassic to Early
624 Cretaceous Tethyan radiolaria. *Marine Micropaleontology*, 43, 239-260. doi:10.1016/S0377-
625 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

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

653

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

657

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

663

664 He, K., Zhang, X., Ren, S., and Sun, J. 2016. Deep Residual Learning for Image Recognition.
665 In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition
666 (CVPR), 770-778. doi:10.48550/arXiv.1512.03385

667

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

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

674 H. J., Giles Miller, C., Rios, N., and Hull, P. M. 2019. Endless Forams: >34,000 modern
675 planktonic foraminiferal images for taxonomic training and automated species recognition
676 using convolutional neural networks. *Paleoceanography and Paleoclimatology*, 34, 1157-
677 1177. doi:10.1029/2019PA003612.

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

681

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

686

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

691

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

695

696 Mitra, R., Marchitto, T.M., Ge, Q., Zhong, B., Kanakiya, B., Cook, M.S., Fehrenbacher, J.S.,
697 Ortiz, J.D., Tripathi, 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 polycystine 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

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

766

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

769

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

774

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

777

778 **Plate 1.** Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of
779 Demerara Rise; species names are followed by the ODP site and hole, core, section and
780 sampled level from which it comes from. **A)** *Dendrospyrus stylophora* (Ehrenberg 1874) from
781 1259A-20R-4W, 53-55 cm, **B)** *Elaphospyris didiceros* (Ehrenberg 1874) group from 1258A-
782 2R-4W, 55-57 cm, **C)** *Liriospyris clathrata* (Ehrenberg, 1854) group from 1259A-20R-4W,
783 53-55 cm, **D)** *Rhabdolthis pipa* Ehrenberg 1854 from 1260A-15R-1W, 55-57 cm, **E)**

784 *Lophophaena radians* (Ehrenberg, 1874) group from 1259A-16R-2W, 55-57 cm, **F**)
785 *Dictyoprora mongolfieri* (Ehrenberg 1854) from 1260A-6R-2W, 55-57 cm, **G**) *Dictyoprora*
786 *amphora* (Haeckel, 1887) group from 1258A-2R-4W, 55-57 cm, **H**) *Rhopalosyringium* ?
787 *auriculaleporis* (Clark and Campbell, 1942) from 1260A-14R-6W, 55-57 cm, **I**)
788 *Rhopalosyringium* ? *biauratum* (Ehrenberg, 1874) from 1260A-12R-3W, 55-57 cm, **J**)
789 *Dictyomitra parva* (Kim 1992) from 1258A-2R-4W, 55-57 cm, **K**) *Lithochytris vespertilio*
790 Ehrenberg 1874 from 1260A-10R-5W, 55-57 cm, **L**) *Sethochytris triconiscus* (Haeckel 1887)
791 from 1259A-20R-4W, 53-55 cm, **M**) *Lychnocanium babylonis* (Clark and Campbell 1942)
792 group from 1258A-3R-3W, 56-58 cm, **N**) *Lychnocanoma bajunensis* (Renz 1984) from
793 1258A-2R-4W, 55-57 cm, **O**) *Stichopterygium microporum* (Ehrenberg 1874) from 1260A-
794 8R-4W, 54-56 cm, **P**) *Carpocanopsis ornata* (Ehrenberg, 1874) group from 1260A-6R-5W,
795 55-57 cm, **Q**) *Cycladophora spatiosa* (Ehrenberg 1874) group from 1259A-20R-1W, 55-57
796 cm, **R**) *Anthocyrtis mespilus* (Ehrenberg 1847) group from 1259A-16R-2W, 55-57 cm, **S**)
797 *Zealithapium mitra* (Ehrenberg 1874) from 1258A-2R-3W, 55-57 cm.

798

799 **Plate 2.** Nassellarian radiolarian species commonly occurring in Middle Eocene sediments of
800 Demerara Rise; species names are followed by the ODP site and hole, core, section and
801 sampled level from which it comes from. **A**) *Lophocyrtis alauda* (Ehrenberg, 1874) from
802 1260A-15R-1W, 55-57 cm, **B**) *Aphetocyrtis zamenhofi* Meunier and Danelian, 2023 from
803 1259A-26R-5W, 54-56 cm, **C**) *Theocyrtis scolopax* (Ehrenberg, 1874) from 1260A-15R-3W,
804 55-57 cm, **D**) *Phormocyrtis embolum* (Ehrenberg, 1874) from 1258A-2R-4W, 55-57 cm, **E**)
805 *Phormocyrtis lazari* Meunier and Danelian, 2023 from 1260A-8R-6W, 54-56 cm, **F**)
806 *Podocyrtis (Lampterium) goetheana* (Haeckel, 1887) from 1259A-18R-1W, 53-55 cm, **G**)
807 *Podocyrtis (Lampterium) chalara* Riedel and Sanfilippo 1970 from 1260A-6R-CC, 63-177
808 cm, **H**) *Podocyrtis (Lampterium) mitra* Ehrenberg group, 1854 from 1260A-9R-1W, 55-57

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