Marine Micropaleontology 158 (2020) 101890

Contents lists available at ScienceDirect

Marine Micropaleontology

journal homepage: www.elsevier.com/locate/marmicro

Research paper

Automated Microfossil Identification and Segmentation using a Deep Learning Approach

L.E. Carvalho[a,b,*], G. Fauth[c], S. Baecker Fauth[c], G. Krahl[c], A.C. Moreira[d], C.P. Fernandes[d], A. von Wangenheim[b]

[a] Graduate Program in Computer Science, Department of Informatics and Statistics, Federal University of Santa Catarina, Florianópolis, SC, CEP 88040-97, Brazil
[b] Universidade Federal de Santa Catarina – UFSC Campus João David Ferreira Lima – Trindade Departamento de Informática e Estatística – Sala 320 Florianópolis/SC – CEP 88040-970, Brazil
[c] itt Fossil, Instituto Tecnológico de Micropaleontologia, Universidade do Vale do Rio dos Sinos, São Leopoldo, RS, Brazil
[d] Graduate Program in Materials Science, Engineering Federal University of Santa Catarina, Florianópolis, SC, Brazil

ARTICLE INFO

Keywords:
3D image Segmentation
Paleontology
Microfossil
Semantic Segmentation
Deep Learning

ABSTRACT

Computational analysis applicability to paleontological images ranges from the study of the evolution of animals, plants and microorganisms to the habitat simulation of living beings from a specific epoch. It can also be applied in several niches, e.g. oil exploration, where several factors can be analyzed in order to minimize costs related to oil extraction. One specific factor is the characterization of the environment to be explored. This analysis can occur in several ways: use of probes, samples extraction, correlation with logs of other drilling wells and so on. During the samples extraction phase, the Computed Tomography (CT) is of extreme importance, since it preserves the sample and makes it available for several analyses. Based on 3D images generated by CT, analyses and simulations can be performed, and processes currently performed manually and exhaustively, can be automated. In this work, we propose and validate a method for fully automated microfossil identification and segmentation. A pipeline is proposed that begins with scanning and ends with the microfossil segmentation process. For the microfossil segmentation, a Deep Learning approach was developed, which resulted in a high rate of correct microfossil segmentation (98% IOU). The validation was performed both through an automated quantitative analysis and visual inspection. The study was performed on a limited dataset, but the results provide evidence that our approach has potential to be generalized to other carbonatic rock substrates. To the extent of the authors' knowledge, this paper presents the first fully annotated MicroCT acquired microfossils dataset made publicly available.

1. Introduction

The applicability of computational image analysis to paleontological data encompasses the possibility of identifying, reconstructing and visualizing microfossils in rock samples not recovered by conventional extraction methodologies. It can also allow microfossils to be taxonomically identified even before physical extraction from the rock sample. In addition, it enables the verification of the microfossil position in a given sedimentary stratum, which can help in taxonomic inference, whereas detailed positional information is lost in the traditional preparation method (Kachovich et al., 2019). Computational analysis of samples can be applied in several niches, e.g. oil exploration, paleoenvironment reconstituition and geologic modeling.

Diagenetic processes can significantly alter the recovery of different microfossils. In this work, the focus was on specimens of foraminifera, widely used in biostratigraphic and paleoenvironmental interpretations. The foraminifera are unicellular organisms, consisting of a carbonate and/or agglutinat carapace, and have a benthic (inhabiting the sediment and water interface) and planktonic (inhabiting the superficial portion of the water column in marine environments) life habit (Boudagher-Fadel, 2015).

In the oil exploration field, there are many factors to be taken into consideration in order to minimize oil prospection costs. One of these factors is the environmental condition, which can be analyzed in multiple ways: use of probes, extraction of samples for petrophysical components evaluation and correlation with logs from other drilling wells.

In the area of samples extraction it is possible to perform different

* Corresponding author.
E-mail address: lcarvalho@incod.ufsc.br (L.E. Carvalho).

https://doi.org/10.1016/j.marmicro.2020.101890
Received 2 July 2019; Received in revised form 14 May 2020; Accepted 18 May 2020
Available online 25 May 2020

analyses on a given sample. For this purpose, the Computed Tomography (CT) plays a central role. More specifically, samples are analyzed with X-ray micro-tomography (MicroCT), which is a radiographic imaging technique that produces 3D images of the material's internal structure with a spatial resolution of around 1 µm (Landis and Keane, 2010). MicroCT is of significance because it preserves the sample and makes it available for different studies. Based on MicroCT generated data volumes, various 3D data analyses and simulations can be performed and several analysis processes can be computationally carried out and automated using state-of-the-art Computer Vision (CV) techniques. These processes are currently performed manually and in a time-consuming manner. One of these processes that can undergo automation through CV is the microfossils identification and extraction in rock samples, which is the focus of this study.

1.1. Objective and strategy

In this work, we propose a CV workflow composed of computational methods that starts with the MicroCT scanning process of a sample and ends with the fully automated identification and extraction of individual microfossils. The main research question we try to answer is: *Is it possible to fully automatically and reliably identify microfossils in carbonatic rock samples?*

The novelty in our work is the use of Deep Learning Convolutional Neural Network (CNN) approaches for the identification and 3D segmentation of microfossils directly in their deposition place. Our approach works directly on MicroCT data gained from carbonate rocks, without the need of any preparation or physical extraction. For this purpose we developed an identification and segmentation strategy that employs a special category of CNN models, namely Semantic Segmentation (SS) neural networks and extends this model in order to be able to process whole 3D MicroCT sample volumes. In order to identify the best model, we extend, train, test and compare a series of different state-of-the-art SS models. To validate our approach we employ a validation strategy where we compare our results to ground truths that were manually generated by experienced micropaleontologists employing state-of-the-art automated image segmentation validation algorithms.

The paper is organized as follows: Section 2 presents a brief overview of computer methods in paleontology, arriving on presently employed methods. Section 3 presents the material and methods employed in our experiments. Section 4 describes the performed experiments and the obtained results. A discussion of the results, conclusions and future works is provided in Section 5.

2. State of the art

Paleontology is a well-established science and its methodological intersection with the computational field started to grow in the 1990's. In the late 1980's, most main paleontology journals still showed an irregular presence of computational methods: some journal issues contained one article describing some computational method application, others presented 2 or 3 articles and very few offered a larger number of them (Tipper, 1991). In the majority of journals and books the insertion of computational methods in the paleontology field still looked uneven.

In the late 1990s, however, with the widespread use of medical CT, a growth in research activities employing tomographic images occurred (Tipper, 1991). This boosted the development of specialized software applications such as: DRISHTI (Limaye, 2012), IMAGEJ (Rueden et al., 2017), AVIZO[1] and so on. These specialized tools helped change how researchers deal with specific problems in several fields, including geology and paleontology, frequently with applications to oil and gas exploration. The applicability of the set of tools and techniques that came to be called Virtual Paleontology (VP) range from animal, plant and microorganisms evolution analysis all the way to virtual reconstruction of a specific extinct environment (Sutton et al., 2014).

On the other side, the application of microfossil study to the oil prospection area had its first appearance in 1890 in Poland (Singh, 2008), but it was in the USA, in 1920, with the use of microfossils to identify the age of samples extracted from drilling rigs, that a bigger advance in the development of the field of Applied Micropaleontology was attained (Molina, 2004).

In the last decade multiple research works contributed to improve the micropaleontology field. The latest efforts aim at the use of VP associated with CNNs in order to identify microfossils (Ge et al., 2017). With this in mind, our research is focusing on pursuing techniques that can identify microfossils on their deposition place, i.e., without the need of previous physical isolation. For this purpose we research some CV fields such as 3D segmentation applied to tomographic image and 3D object recognition, in order to apply them to microfossil identification.

In the next subsections we summarize the results of the systematic literature reviews (SLR) we performed in order to identify state-of-the-art methods and procedures that potentially could be used in microfossil image studies. These reviews followed the approach originally proposed by (Kitchenham, 2004) for SLRs in Computer Sciences, where first we defined a research question: *Is it possible to fully automatically and reliably identify microfossils in carbonatic rock samples?*. This broad question, in order to be more manageable, was split into 2 topics, each of which was explored in depth in a separate SLR:

- Analysis of 3D segmentation methods applied to tomographic images, which could possibly be used to segment microfossils (Carvalho et al., 2018);
- Analysis of methods used for 3D object recognition in a general context, aiming to evaluate which methods could be applied to the microfossils field (Carvalho and von Wangenheim, 2019).

The results of these two SLRs will be briefly summarized below. Since a detailed description would exceed the scope of this paper, we refer to the referenced SLRs for more details.

2.1. 3D segmentation applied to tomographic images and 3D object recognition

An initial analysis of image processing methods employed in the fossil identification area showed difficulty in finding any works that explore microfossils. So we generalized our search for methods in other similar areas. We started by performing a systematic literature review on 3D segmentation methods applied to tomographic images (Carvalho et al., 2018). Several works were analyzed which comprehended a vast group of segmentation methods. In our review, it was noticed a tendency on the use of 3D segmentation methods based on models and region growing. However, its use for fossil/microfossil segmentation was not noticed in the literature.

We also analyzed the field of 3D object recognition employing the same SLR methodology (Carvalho and von Wangenheim, 2019). In this SLR for 3D object recognition we could identify two general pipelines. Both pipelines start with the data acquisition, which can basically vary between 3D data (MRI, CT) or 2D data (RGB and RGBD cameras); preprocessing, where methods for artifact removal, image enhancement and image simplification are applied; and data representation, wherein several authors proposed a varying amount of different object representations. Then, it comes the stage where both pipelines differ: In the first pipeline, the data representation stage is used to describe and storage the object representation chosen, which is later used for similarity calculation and object identification; In the second pipeline, the data representation is employed for training a specific recognition architecture, such as a CNN, which is afterwards used for other objects

[1] https://www.fei.com/software/avizo3d/%C2%A0

recognition. Despite having found two general approaches for 3D object recognition, we could not identify, in our review, the application of these approaches directly on microfossil segmentation.

2.2. Deep learning, object recognition and paleontology

The 3D object recognition area has, in the last few years, experienced a growth boosted by the increased availability of new algorithms and models, 3D data and the popularization of a varied palette of 3D sensors. Methods developed in this area find application in a wide range of sectors, from the field of robotics to the security and surveillance domain. The general tendency in this area has been the use of Deep Learning (DL) techniques.

DL is a form of machine learning that enables computers to learn from experience and understand the world in terms of a hierarchy of concepts (Goodfellow et al., 2016). DL employs very deep CNNs, with neural networks that sometimes consist of more than 100 layers, in contrast to the Artificial Neural Networks (ANNs) employed between the 1980's and 2000's, that typically employed only three layers. One key concept here is the Convolutional Layer (CL), a feature extraction structure, first presented in (Lecun et al., 1998), that allows the hierarchical learning and representation of complex knowledge. Because DL CNNs gather knowledge from examples, there is no need for a human computer operator to formally specify all the knowledge that the computer needs. The capacity to represent a hierarchy of concepts in a network dozens of CLs deep allows a DL CNN to learn complicated concepts by building them out of simpler ones; a graph of these hierarchies would be many layers deep (Goodfellow et al., 2016).

One work that employs DL for object recognition is the 3D Object Recognition with Deep Belief Nets approach (Nair and Hinton, 2009), where a network of symmetrically connected neuron-like units, that performs stochastic decisions about whether to be on or off, is presented. In (Socher et al., 2012), a model based on the combination of convolutional and recursive neural networks for the feature learning and classification in RGB-D images is shown. Another DL approach is presented in (Yu et al., 2013), where a robotic vision-based system, which can not only recognize different objects but also estimate their pose through the Max-pooling Convolutional Neural Network (MPCNN) model is introduced. Similarly, the studies presented in (Liang et al., 2014) and (Xia et al., 2015) also shown DL approaches focusing in object recognition. Lastly, in (Xu et al., 2016) it is introduced an efficient 3D object volumetric representation which requires much less memory than a normal volumetric representation. None of these approaches tackles the problem of identification of fossils embedded in rocks. However, these works give an input about methods used in cluttered environments, which is one of the problems that emerged with rock embedded microfossil analysis.

We also performed a search for automatic methods for micro-organisms identification, where we found some research works performed during the past decades (Liu et al., 1994), (Culverhouse et al., 1996), (Beaufort and Dollfus, 2004). The goal in those studies was to reduce time and cost from the identification process and improve the classification reproducibility. Several methods for automatic marine microfossils classification where analyzed: image features (morphology, texture and intensity) provided by image processing methods have been combined with artificial neural network (Hibbett, 2009) (Schulze et al., 2013), statistical (Culverhouse et al., 1996) or rule-based on classifiers (Yu et al., 1996). Other works employed directly the entire image (Francus, 2007) or used a combination of image and morphology (Barbarin, 2014) with CNNs. More recent approaches are presented in the following works: in (Charles, 2011), the author proposes to segment the foreground particles in the image and to identify those containing a single complete elliptic palynomorph. To do so, he employs trained classifiers to distinguish between regions containing a single palynomorph and one containing other materials; in (Bueno et al., 2017), a method to deal with automatic taxa identification based on machine learning methods aiming to automatically classify diatoms is presented; in (Marchant et al., 2019), the authors applied CNNS for classification of down-core foraminifera; in (Mitra et al., 2019), a convolutional neural network is used to identify six species of extant planktic foraminifera and to distinguish those from other taxa; in (Hsiang et al., 2019), a planktonic foraminifera image set is built, using several expert inputs, and then employed a CNN-based image classification (VGG network), comparing their classification approach results against human performed classification. However, most of those works perform a classification analysis without preserving the microfossils deposition place. This is one aspect that our work addresses by performing a SS in the complete volume sample, thus preserving both the deposition place and the microfossil information.

3. Material and methods

This section describes our datasets and the CV approach we developed for fully automated microfossil identification and segmentation in carbonatic rock samples.

3.1. Material

In this work two datasets were employed: a scanned carbonatic rock sample obtained from a drilling rig probe and a set of manually isolated microfossil specimens that were obtained from the same sample. The sample was collected at the Sergipe-Alagoas Basin (Brazil) Holocene sediments (Fig. 1):

- The *carbonatic rock sample* was used as the material for developing our CV approach. The MicroCT scanner used to digitise the sample is a Versa XRM-500 (ZEISS/XRadia) with the following specifications: best resolution (pixel size) 0.7μm, voltage 30–160 kV, power 2–10 W, CCD cameras 2048 × 2048 pixel, optical lenses 0.4×, 4×, 10×, 20× and 40×, a set of 12 filters for beam hardening correction, maximum sample mass capacity 15 kg and sample size limit (diameter / height) 80/300 mm. The sample acquisition parameters were: Spatial resolution 1.08 mm, image size 956x1004x983, no filtering for beam correction hardening, 10× optical lens, 30 kV / 2 W, angular step 0.255 (1600 projections) and exposure time 11 s. Fig. 2 shows the rock sample and an excerpt of one slice of its digitized result.
- The set of *manually isolated microfossil specimens*, extracted from the carbonatic rock sample described above, was used in this work for illustration purposes and as a guide in order to allow us to know how the specimens in the *rock sample* would look like after being properly cleaned and isolated. These microfossils were prepared in the laboratory, following specific precautions so that there were no chemical and/or mechanical changes: (i) the sediment was first immersed in deionized water for approximately 24 h, aiming the chemical disaggregation; (ii) then, it was washed with running water in a 63 μm sieve; (iii) next, the material was dried at 40°C for approximately 48 h. After drying the samples, the main representative microfossils in the sample were selected with the help of a magnifying glass. In this work, the microfossils specimens were stamped with the help of a multidimensional acquisition with the Zeiss Discovery V20 stereoscope (*Z*-stak mode in AxioVision 4.8 software). Fig. 3 presents these microfossils.

The analyzed sample contains a wide variety of specimens from different species of microfossils. In our analysis, it was considered 14 species among the 4600 specimens found in the sample. The complete dataset, containing the original MicroCT data and the manually segmented images, annotated by specialists, are available at: http://www.lapix.ufsc.br/microfossil-segmentation

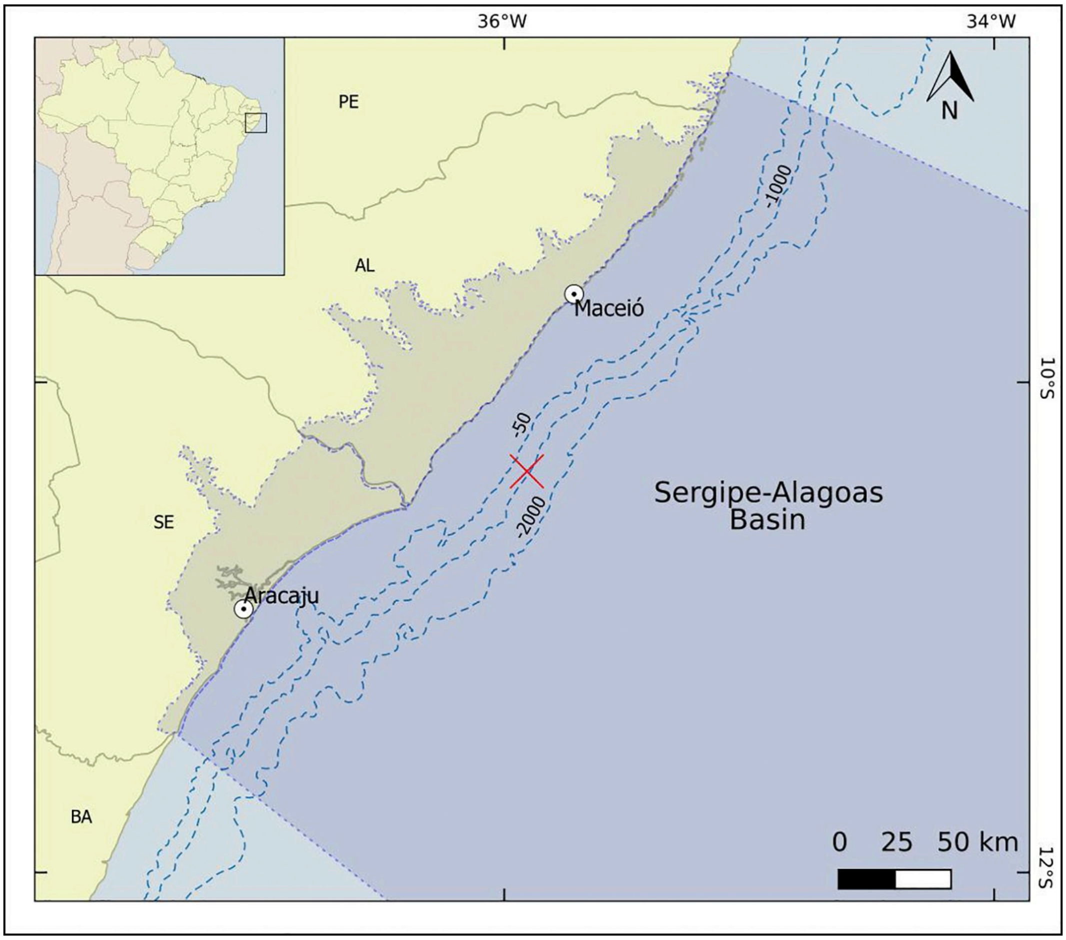

Fig. 1. Sergipe Basin drill section position marked with a red cross. The sample was collected at a depth of approximately 2500 m below the sea floor.The map was generated with the QGIS software (https://qgis.org/en/site/) associated with the geographic coordinates from the place where the samples where collected. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Methods

The CV approach presented here is intended to be embedded into a broader workflow. Fig. 4 presents a general overview of this workflow.

3.2.1. Non-CNN Computer Vision Methods

A prospective search of CV methods for the microfossils segmentation was completed before investigating the use of CNNs. The starting point was to perform a series of experiments using non-CNN, i.e. conventional CV methods for the segmentation of the MicroCT volume.

An extensive list of conventional CV algorithms was analyzed searching for the ones that could potentially generate satisfactory results. The following classical segmentation algorithms were selected: active contours (Kass et al., 1988), simple threshold and threshold with OTSU (Otsu, 1979) - all taking into account the complete tomographic volume. In order to find the best possible parameters for each segmentation algorithm, we performed a broad parameter values search running the algorithms with varied parameter sets. For the active contour algorithm, in order to find the best parameter set, we employed a genetic algorithm to search through possible input parameters. For this purpose we considered 5 input parameters: Number of steps, Sigma, Alpha, Smoothing and Theta. Those 5 parameters, their value variation and the results found are summarized in the Table 1.

The results of these conventional CV algorithms were initially

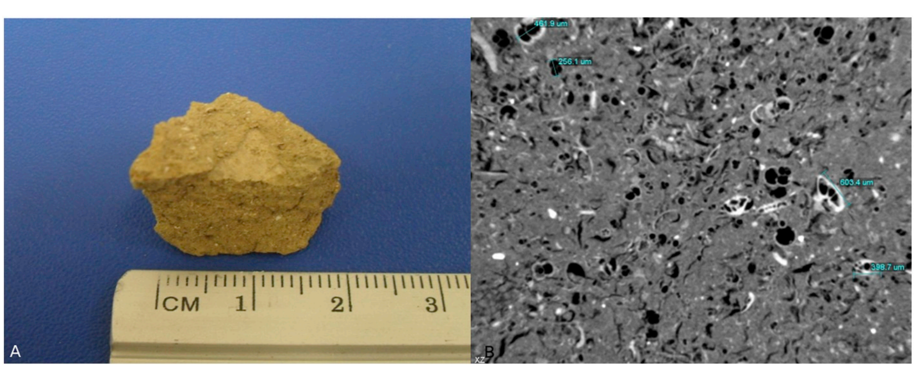

Fig. 2. Analyzed rock sample (A) and one of its microtomography 2D sections (B).

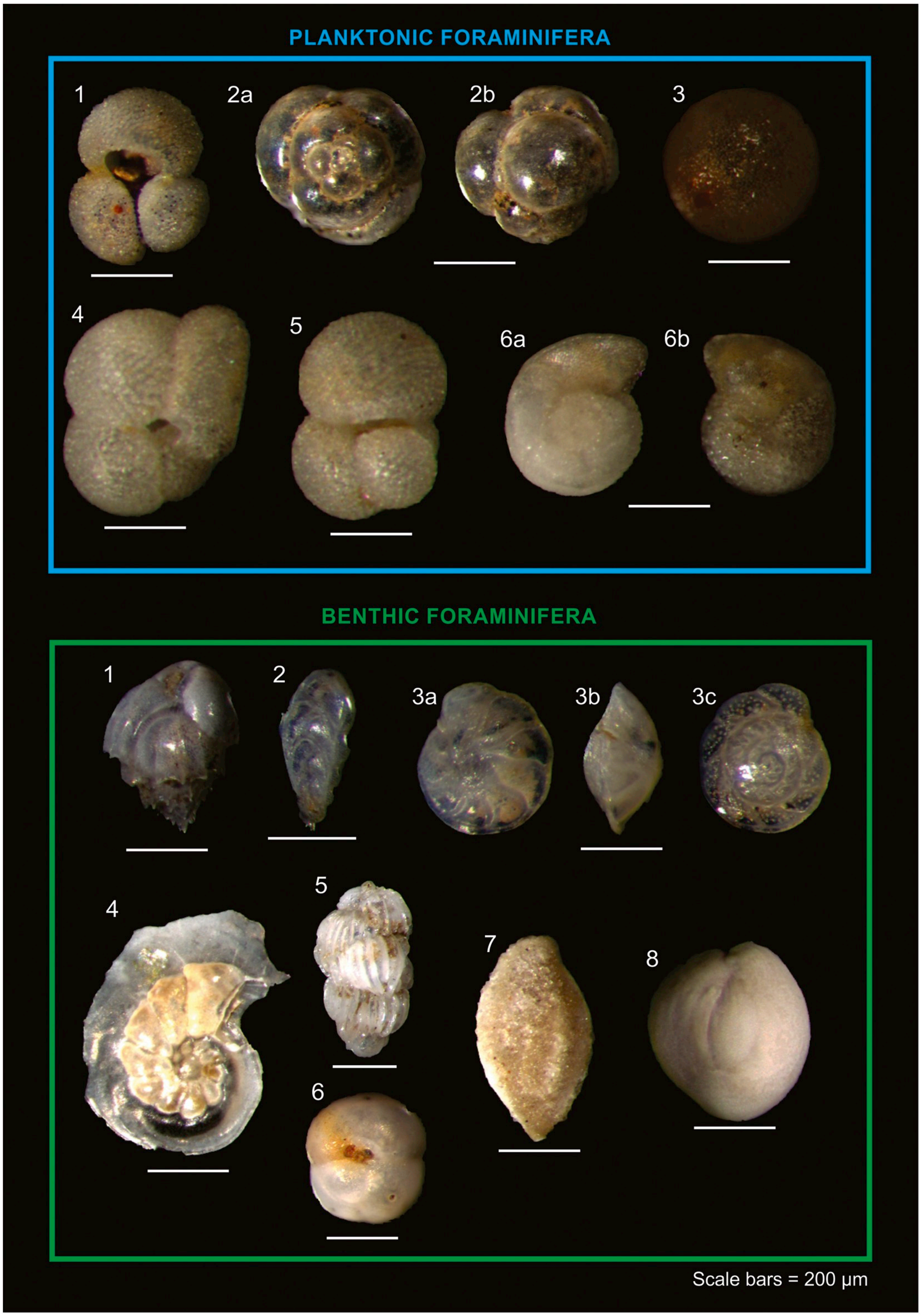

Fig. 3. Analyzed foraminifera taxa. Planktonic Foraminifera: 1) *Globogerinoides ruber*; 2a-b) *Candeina nitida*; 3) *Orbulina universa*; 4) *Trilobatus sacculifer*; 5) *Trilobatus trilobus*; 6a-b) *Globorotalia truncatulinoides*. Benthic foraminifera: 1) *Bulinina*; 2a-c) *Bolivinita*; 3a-c) *Cibicidoides*; 4) *Laticarinina*; 5) *Uvigerina*; 6) *Sphaeroidina*; 7) *Siphonaperta*; 8) *Quinqueloculina*.

analyzed through *visual inspection*. For the conventional CV methods that presented the best results in the visual inspection, we subsequently analyzed its results quantitatively employing the method described in the evaluation metrics section.

3.2.2. CNN-based Segmentation Methods

In the 3D object identification and segmentation field, the most successfully and commonly used SS models in the last years have been the UNET and its variations. The UNET architecture was presented in (Ronneberger et al., 2015), where the authors show its use for medical image segmentation. UNETs provide a general framework that can be

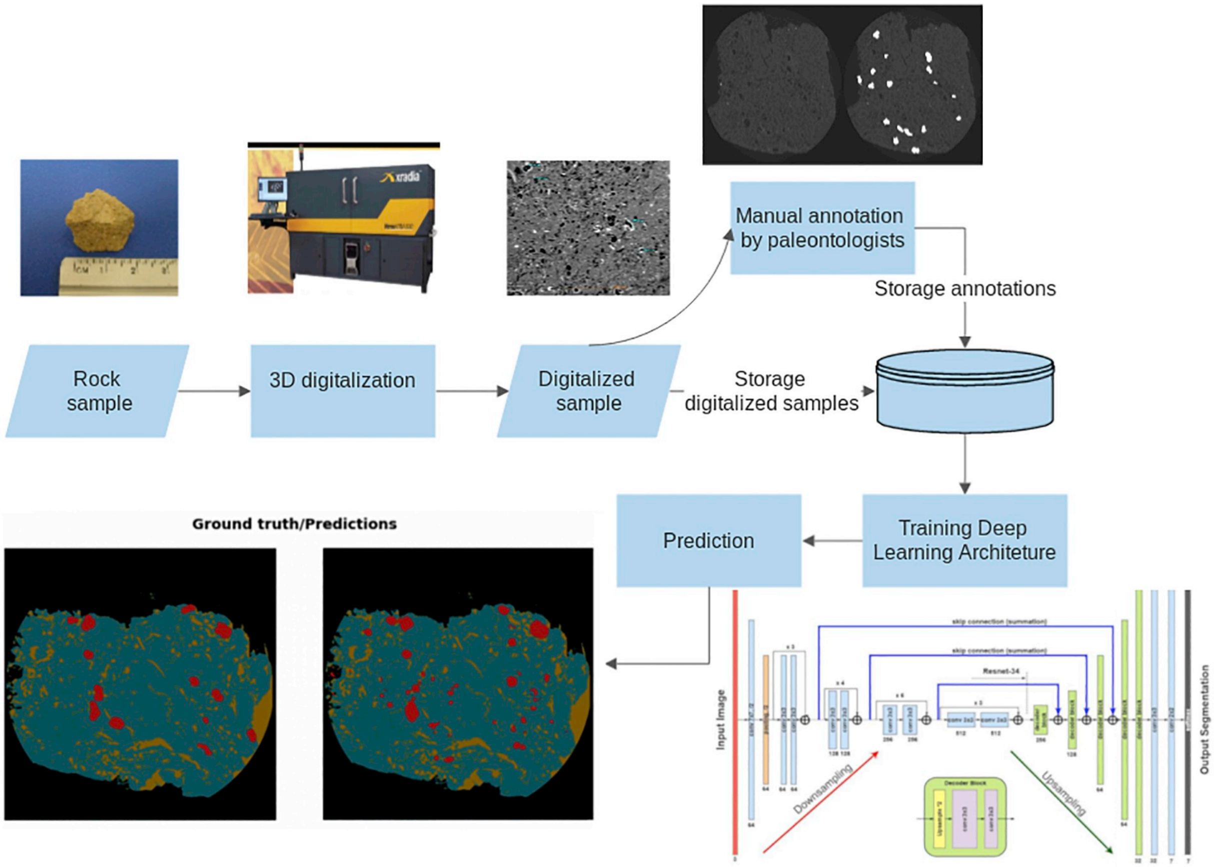

Fig. 4. General workflow. Modified from (Rakhlin et al., 2018).

Table 1
Table summarizing the results of our search for the best active contour algorithm parameters.

Input parameter	Description
Number of steps	For the cases analyzed keeping the same value for the others parameters, the number of steps didn't interfere in the final results. Suggested: keep the number of steps fixed in 100.
Sigma	As sigma increases, it restricts the number of contours. Suggested: vary from 0 to 3.
Alpha	As Alpha increases, the contours are lost, when keeping the other parameter values fixed. However, it can interact with the increase in Sigma, making it possible to use Alpha at higher values. Suggested: Vary between 500 and 1000 according to the sigma parameter used.
Smoothing	With this value set in 1 the borders lose some of their representativeness. From the range 0 to 0.5, no major change was noticed. Suggested: Start from 0.5 increasing with steps of 0.1 until it reaches 1.
Theta	When starting from 0 varying with steps of 10 and going up to 100 it was found that in the initial values, i.e., 0 to 15, the amount of edges is small. On the other hand, if the value is above 40–45, depending on the other parameters, the contours are totally lost. Suggested: Vary with steps of 5 starting at 15 going up to 40.

parameterized with a specific image classification CNN model. The UNET then employs two slightly modified instances of this CNN, an *encoder* and a *decoder*, one for image recognition and another, employed in reverse mode, for the segment mask generation (Badrinarayanan et al., 2017): it uses the encoder to map raw inputs to feature representations and the decoder to take this feature representation as input, process it to make its decision and produce an output. As the UNET produces state-of-the-art semantic segmentation we chose it as our starting point.

We initially employed the UNET model associated with a ResNet34 (He et al., 2016) as our initial structure. We then extended this approach through several state-of-the-art improvements, such as: nearest neighbour interpolation and pixel shuffling (Shi et al., 2016), Leaky Relu (Xu et al., 2015) for activation function, batchnorm (Bjorck et al., 2018) for batch normalization and some efficient ways to set hyper-parameters inspired in (Smith, 2018). For this purpose we employed the fastai[2] implementations of these improvements. Fastai is a high-level CNN framework implemented on top of the lower-level Pytorch[3] CNN framwework.

3.2.2.1. Training speedup. The MicroCT data posed a challenge: the memory requirements imposed by both the dataset size and the UNET architecture would strongly limit the batch size of our training set. In order to overcome this limitation and be able to work initially with larger batch sizes and train the network at a faster pace, we employed a *step-wise progressive improving image resolution training strategy*. For this purpose we performed the training in **three cycles**: we started our transfer learning with the dataset at 1/4 of the original MicroCT image resolution; trained the DL model; re-sized the dataset to 1/2 of the

[2] https://www.fast.ai/
[3] https://pytorch.org/

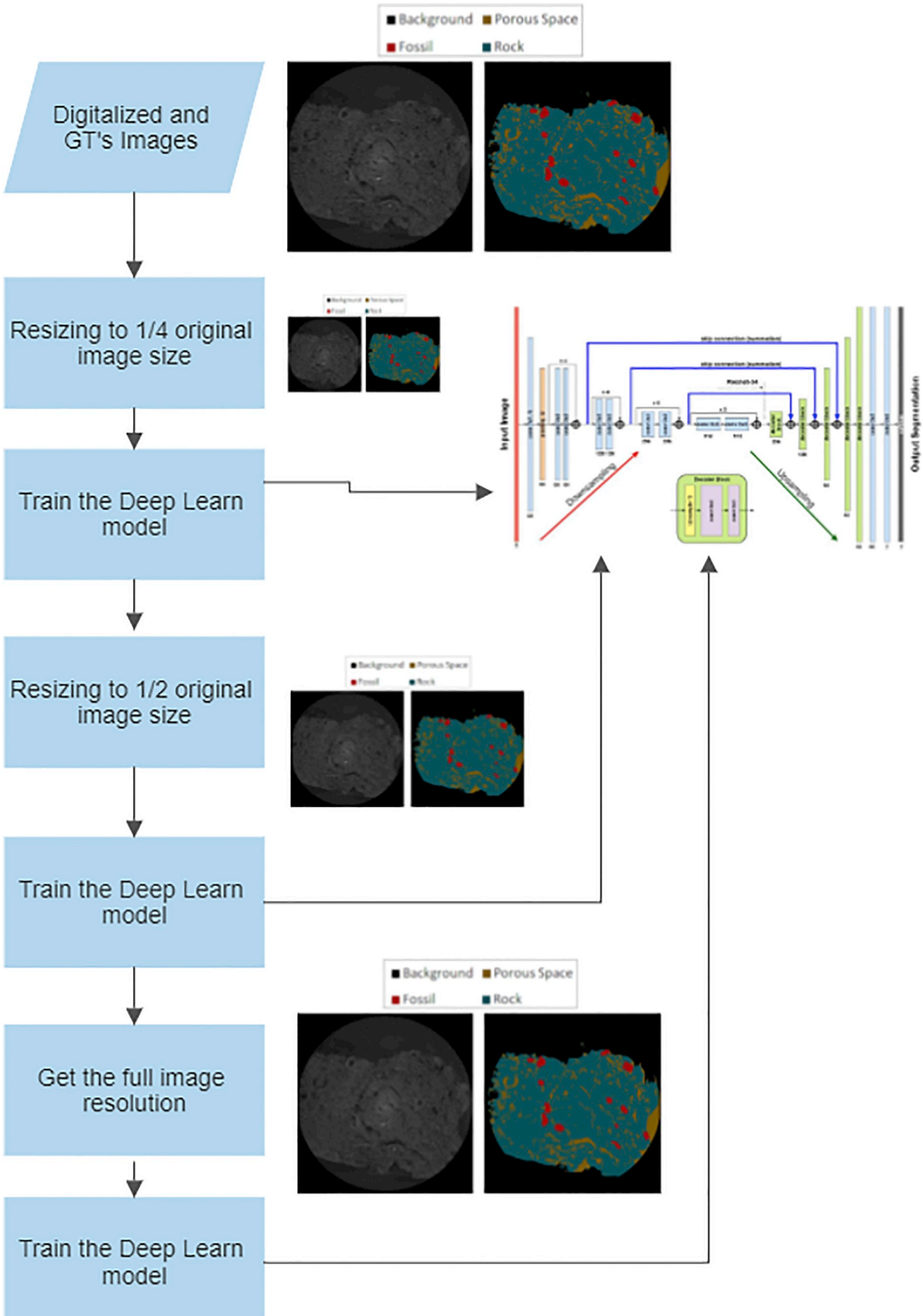

Fig. 5. Progressive image resolution re-sizing approach. Modified from (Rakhlin et al., 2018).

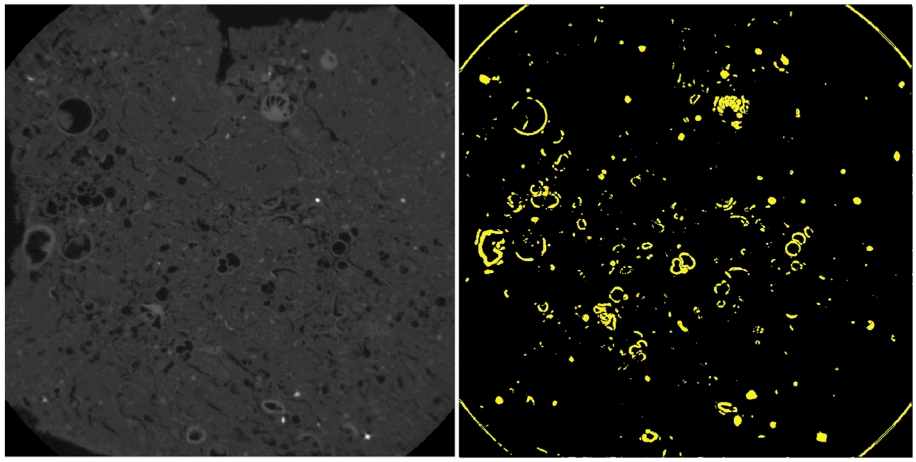

Fig. 6. Best microfossil segmentation that we could obtain using 3D active contours (IOU = 20%).

original resolution; trained again and after that we employed the full CT volume resolution for a final fine tuning training cycle. The outline of this strategy was originally presented in (Howard, 2019). Fig. 5 shows our interpretation of the training workflow based on the progressive image resolution re-sizing approach.

Another strategy for fine-tuning a model is the Differential Learning Rates (DLR) strategy, also presented in (Howard, 2019). This approach for multiple, layer-specific learning rates as the layers get deeper, is justified by the following rationale: When performing transfer learning followed by fine tuning, in the first layers the pre-trained model being adapted will learn generic low level features from the new dataset being used in the transfer learning. These low level features are very probably similar to those of the original dataset, regardless of the image context. Therefore, there is no the necessity of employing high learning rates at these first layers. However, as information gets deeper into the architecture, the feature combinations become more complex and dataset-specific, and are more directly connected to the application context. Accordingly, higher learning rates in deeper layers are desirable in order to allow the network to better adapt to context-specific features.

3.2.3. Evaluation metrics

We evaluated each segmentation method comparing our results to the ground truths generated by micropaleontologists using the Intersection Over Union (IOU) score (Rahman and Wang, 2016), which quantifies similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets. The predicted labels were evaluated against three ground truths generated manually by specialists.

4. Results

This section presents the obtained results of the different algorithms and CNN models we tested.

4.1. Conventional CV algorithms

The best results under the conventional CV algorithms were achieved with the active contours method. For this method we obtained an IOU score of 20%. The obtained active contour segmentation result is shown in Fig. 6.

The obtained results show that conventional CV methods may not be suitable for the task of microfossil segmentation in rock samples.

4.2. CNN-based Semantic Segmentation

For our initial tests with SS CNN models, we started with the following structure: UNET associated with ResNet34 and the binary cross entropy as its loss function, a carbonatic rock sample with several microfossil specimens, scanned with the MicroCT previously described resulting in a total of 1000 slices. We employed an Intel Core i7–7700 CPU3.60GHz, 32GB memory computer and an NVIDIA GeForce GTX 1080 Ti 11GB GPU.

With this initial structure, our first experiment used only the microfossil annotation, performing a binary classification between microfossil or everything else. To improve initial results some strategies such as data augmentation and transfer learning were applied, aiming to minimize the effect of having a small dataset. The following data augmentation techniques were used: brightness and contrast variations, random crop, flip, perspective warp, resize, rotate, symmetric warp and zoom). However, the obtained IOU coefficient, used for the results evaluation, stopped in 40–45%.

Trying to improve the results, we increased the number of classes to four, dividing the class previously named as “everything else” into the following: porous space, rock and background. With this number of classes, the obtained IOU value increased from 40 to 45% to 75–76% and stopped. One problem with this approach is the data balance (Zhu et al., 2018), i.e., the existence in the samples of more annotations from the rock class in comparison with the microfossils. Fig. 7 shows the result obtained after marking and training for the 4 classes setup for a selected slice.

Still using the 4-classes approach, we adjusted the hyper-parameters and applied a few performance-enhancing strategies (Xie et al., 2018), such as the *step-wise progressive improving image resolution training strategy* and the DLR, and explored the batch size in order to obtain a 98% IOU. The microfossil GT and its resulting segmentation with this improved IOU is shown in Fig. 8.

Our experiments resulted in an experimental environment, where we employed the UNET as base model associated with other models in the decoder part (Restnet18, ResNet34, ResNet50, ResNet101), the Cross entropy as loss function and IOU for quality assessment. Table 2 shows the IOU value obtained for each method and Fig. 9 shows the original image, its GT and the prediction results for all the architectures we tested.

After the segmentation we applied the predicted masks, generated by ResNet34, to the original image. The result of this process is the identification of several microfossils. Fig. 10 shows the mask overlap result, the identification of one microfossil specimen (highlighted with the red rectangle) followed by its magnified version and the correlation of this magnified version with the other two versions of the same specimen (physically isolated and digitized with the Versa XRM-500 MicroCT and the Zeiss Discovery V20 stereoscope). Fig. 11 shows the original volume 3D rendered and its generated masks result also 3D rendered.

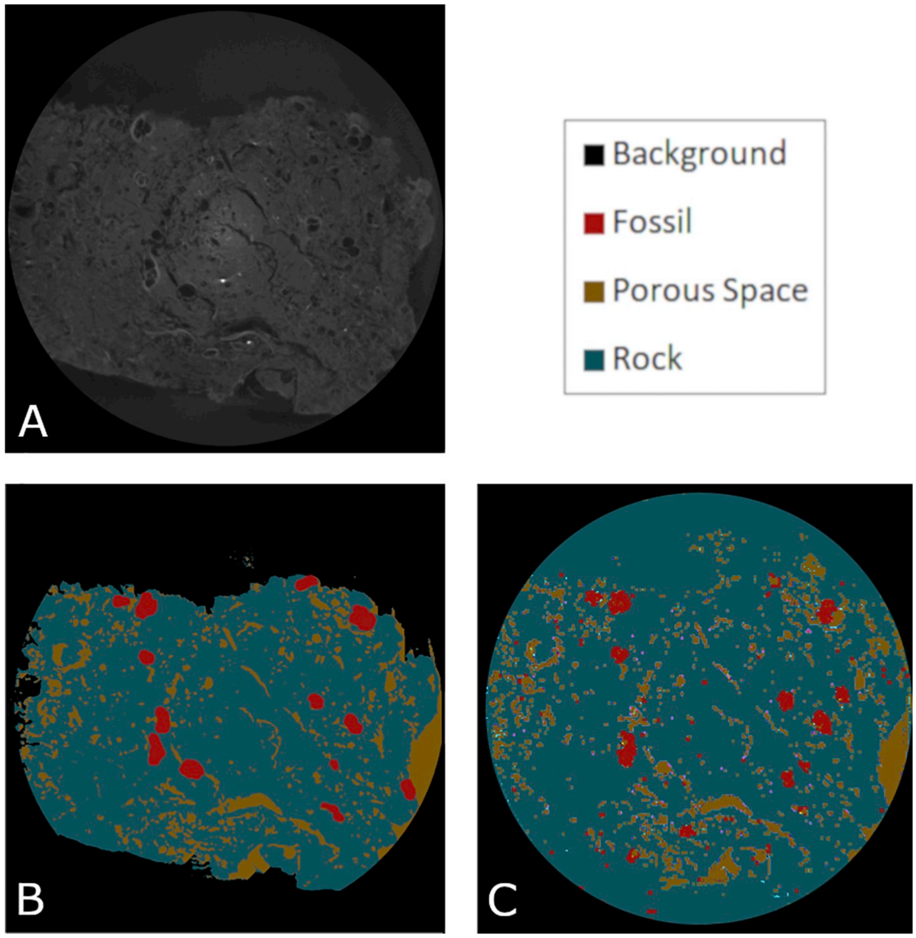

Fig. 7. Obtained microfossil segmentation results with the 4-classes approach. (A) Original digitalized image. (B) Ground Truth manually generated by paleontologists. (C) UNET + ResNet34.

5. Discussion and conclusions

In this paper we present a new non-destructive processing pipeline for the identification of microfossils in carbonate rocks that allows a fully automated segmentation of these fossils without the need of previous physical separation. Furthermore, we developed and validated the CV methods for this identification and segmentation. The validation was quantitatively and automatically performed against a ground truth manually generated by expert micropaleontologists.

An extremely relevant aspect of the developed pipeline for the field of paleontology, more specifically micropaleontology, resides in the non-destructive character of the method. In the micropaleontological study process an essential step is the samples preparation, aiming to separate the microfossils from the other rock and/or sediments. In the traditional laboratory process, the samples are physically disaggregated (ground or milled) and subsequently chemically with addition of reagents (e.g., hydrogen peroxide and acetic acid). Both physical or chemical disaggregation can alter or even destroy microfossils characteristics. In this regard, the imaging method is crucial for the morphological characteristics visualization as reliable as possible, allowing

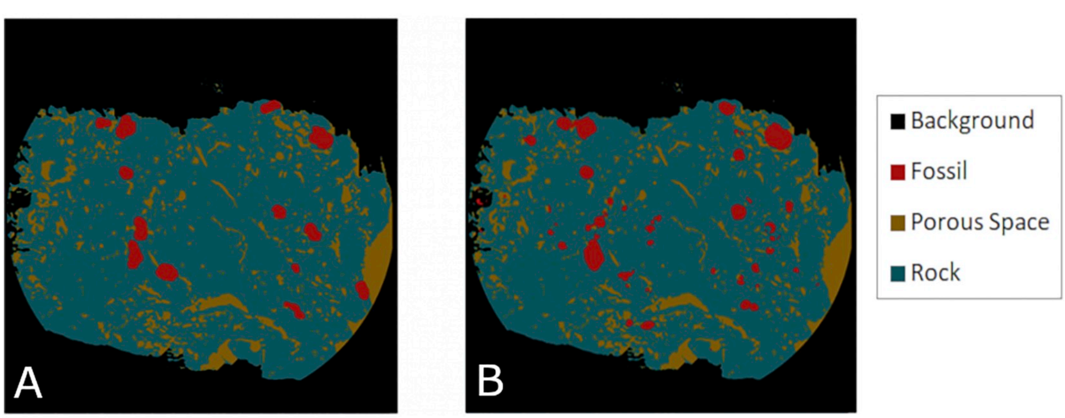

Fig. 8. The ground truth (A) and the obtained microfossil segmentation result (B) with 4 classes, automated hyper-parameters search and additional data augmentation.

Table 2
Segmentation performance in terms of IOU value. Each method was evaluated in a set of 1000 images from annotate microfossil data.

Method	IOU score
Active contours	0.20
UNET + ResNet34 + Only 2 classes (fossil and background)	0.45
UNET + ResNet34	0.76
UNET + ResNet18 + hyper-parameter optimization	0.97
UNET + ResNet101 + hyper-parameter optimization	0.97
UNET + ResNet34 + hyper-parameter optimization	0.98
UNET + ResNet50 + hyper-parameter optimization	0.98

the individuals taxonomic recognition (Kachovich et al., 2019).

Another relevant factor that makes this method meaningful is that it allows the microorganisms' preservation analysis throughout geological time, as well as aspects of fossilization, preservation and even position in which the microfossils are deposited (preserved) in the rocks. It should be emphasized that studies with the taphonomic approach are fundamental for paleoenvironmental conditions and/or diagenetic alteration processes reconstitution over geological time. Also, the use of this tool is strongly indicated in cases where it is extremely difficult to recover microfossils along specific sections and/or intervals where the material (rock) is very compact and even when it presents incipient diagenetic alteration. The microfossils identification is strategic for the exploration of petroleum due to the use in biostratigraphy, which refers to the use of microfossils from different groups to perform the temporal characterization of sedimentary rock strata, fundamental for the petroleum industry and academic science.

A few observations are noted from the obtained results: (i) the importance of employing appropriate hyper-parameters such as learning rate, weight decay, momentum and batch size (the hyper-parameters optimization improved the results by 20%). (ii) a network architecture growth does not imply in better results. It is possible to observe that the ResNet34 shows the same results than the ResNet50 and a better result when compared with a ResNet101. However, at this point we have a hardware limitation: both, ResNet50 and ResNet101, could not run with the full image resolution on the 11 GB GeForce 1080 TI, even with a small batch size. (iii) Analyzing the resulting images and comparing them visually against their ground truth (Fig. 9), we still notice some small errors. However, we understand that this can be be mitigated by adding more training samples, together with GTs from experts, to the training set when applying this pre-trained network to other, new, samples. Also, there are always cutting edge developments that could be tried aiming to reduce even more these small errors. Fig. 10 shows the isolated microfossil digitalized and its correlated identification into the sample.

We understand that this process of microfossil identification without the need of physically isolation has the potential to allow the paleontologist to analyze specific aspects of a sample such as the microfossils deposition. This is important for some applications in the oil and gas industry. It also has the potential to improve the paleontologist's work, because instead of losing time to physically isolate the microfossil he receives the microfossil already identified and can perform other analyses such as class identification and orientation.

5.1. Threats to validity

We employed a dataset that, even if consisted of a very large quantity of images and presented a wide variety of microfossils (more than 4600 specimens), was acquired from a sample obtained from a singular drill core. On the other side, the samples digitization and annotation demand a set of requirements such as: having a MicroCT working and available; the cost of the MicroCT digitisation process; a storage to keep the amount of generated data; and a paleontologist group to analyze and annotate each digitized slice sample. As the workflow we suggest in this paper is new, it was not in place on any of the partners that participated in this work and to obtain more scanned and annotated samples was not possible at this point of our research.

The generalizability of this work could be jeopardized, as we do not have enough data to claim that our approach will be successfully applicable to any carbonatic rock sample. Also, when dealing with deep learning approaches, using small datasets can result in overfitting. On the other hand, our segmentation results were extremely successful and promising. Furthermore, with this singular drill sample we obtained a large variability of species (more than 4600 specimens) which shows the affluence of the analyzed sample. Additionally, from the authors' knowledge, there is not any other publicly available carbonatic rock probe dataset, with or without specialist-annotated microfossils.

In this context, we understand our work as pioneering and pointing

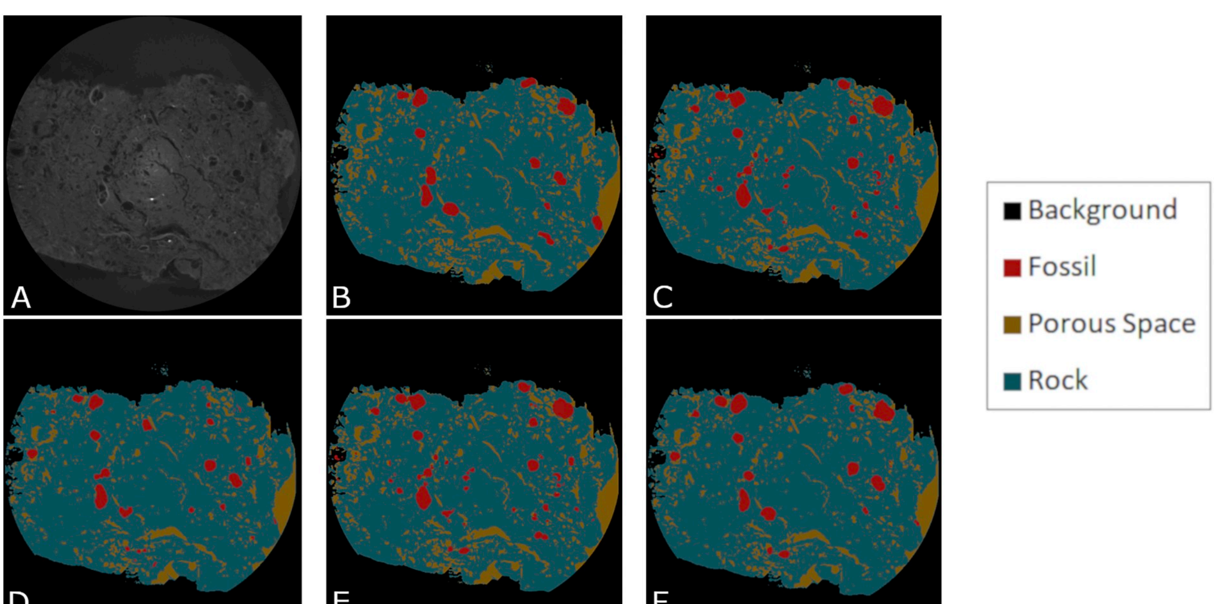

Fig. 9. (A) Original digitalized image. (B) Ground Truth manually generated by paleontologists. (C) UNET + ResNet18 + hyper-parameter optimization. (D) UNET + ResNet101 + hyper-parameter optimization. (E) UNET + ResNet34 + hyper-parameter optimization. (F) UNET + ResNet50 + hyper-parameter optimization.

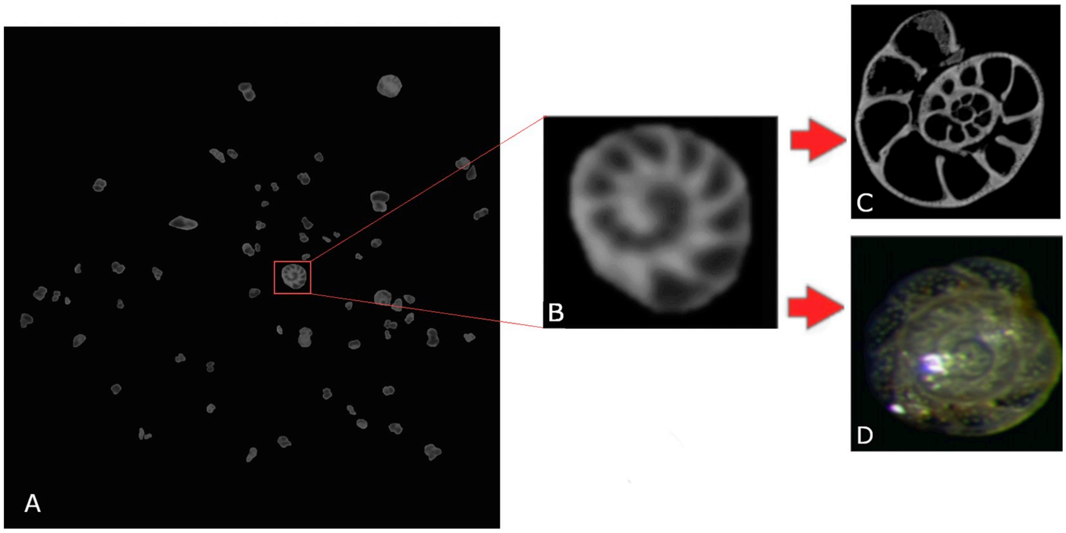

Fig. 10. Result of applying the obtained segmentation mask over the digitized image. (A) contrast-enhanced 2D section image masked from the digitized MicroCT volume with one specific microfossil highlighted in red. (B) Highlighted microfossil extracted and magnified for visualization. (C) Physically isolated microfossil digitized with The Versa XRM-500 MicroCT. (D) *Cibicidoides* multidimensional acquisition with the Zeiss Discovery V20 stereoscope. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

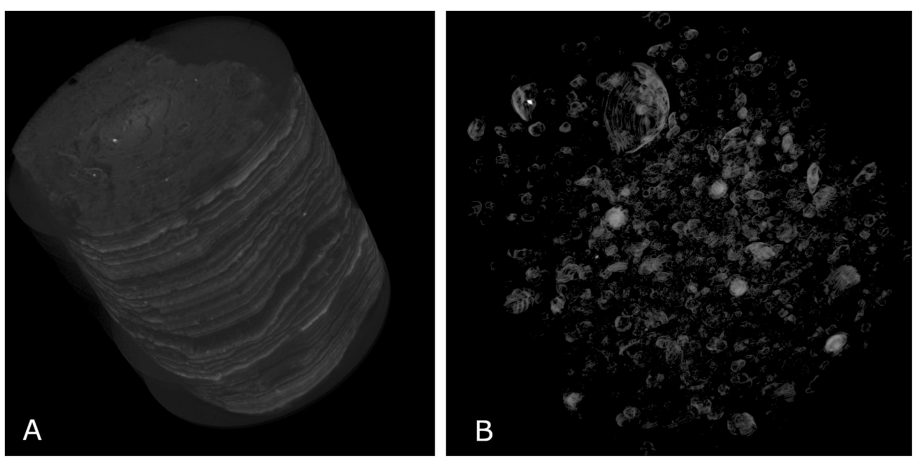

Fig. 11. (A) The original 3D rendered volume and (B) its 3D version of the masking result.

to a promising direction of research that can potentialize both micropaleontological research and associated economical activities, such as oil prospection. Our publicly available fully annotated MicroCT database has also the potential to support research activities to be performed by other groups.

5.2. Conclusions

Summarizing, this work presents the first fully annotated MicroCT-acquired publicly available microfossils dataset additionally with a baseline for microfossil segmentation comparison. Furthermore, it is shown a methodology for microfossil studies through MicroCT-acquired digital models and also a tool for cases where it is extremely difficult to recover microfossils along specific sections.

With the improvement in the available hardware some future work aim to reduce even more the obtained errors by increasing the batch size and image resolution and employ newer deep learning techniques.

Acknowledgments

This study was partially financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and by PETROBRAS through the research project number 902. There are no conflicts of interest. We also would like to acknowledge the reviewers input, which helped to improve the overall quality of this manuscript.

Computer code availability

The Jupyter notebook code and the dataset are available at: http://www.lapix.ufsc.br/microfossil-segmentation.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 2481–2495.

Barbarin, N., 2014. La reconnaissance automatisée des nannofossiles calcaires du Cénozoque. Ph.D. thesis. In: Thèse de doctorat dirigée par Beaufort, Luc Géosciences de l'environnement Aix-Marseille 2014, 2014AIXM4318. http://www.theses.fr/2014AIXM4318/document.

Beaufort, L., Dollfus, D., 2004. Automatic recognition of coccoliths by dynamical neural networks. Mar. Micropaleontol. 51, 57–73.

Bjorck, N., Gomes, C.P., Selman, B., Weinberger, K.Q., 2018. Understanding batch normalization. Adv. Neural Inf. Proces. Syst. 7694–7705.

Boudagher-Fadel, M.K., 2015. Biostratigraphic and Geological significance of Planktonic Foraminifera. 2 UCL Press, London, UK Revised ed. 306p.

Bueno, G., Deniz, O., Pedraza, A., Ruiz-Santaquiteria, J., Salido, J., Cristobal, G., Borrego-Ramos, M., Blanco, S., 2017. Automated diatom classification (part a): Handcrafted feature approaches. Appl. Sci. 7, 753.

Carvalho, L., von Wangenheim, A., 2019. 3d object recognition and classification: a systematic literature review. Pattern. Anal. Applic. 22, 1243–1292.

Carvalho, L.E., Sobieranski, A.C., von Wangenheim, A., 2018. 3d segmentation algorithms for computerized tomographic imaging: a systematic literature review. J. Digit. Imaging 31, 799–850.

Charles, J.J., 2011. Automatic recognition of complete palynomorphs in digital images. Mach. Vis. Appl. 22, 53–60.

Culverhouse, P.F., Simpson, R., Ellis, R., Lindley, J., Williams, R., Parisini, T., Reguera, B., Bravo, I., Zoppoli, R., Earnshaw, G., et al., 1996. Automatic classification of field-collected dinoflagellates by artificial neural network. Mar. Ecol. Prog. Ser. 139, 281–287.

Francus, P., 2007. Image Analysis, Sediments and Paleoenvironments, 1 ed. Developments in Paleoenvironmental Research, Springer Netherlands 330p.

Ge, Q., Zhong, B., Kanakiya, B., Mitra, R., Marchitto, T., Lobaton, E., 2017. Coarse-to-fine foraminifera image segmentation through 3d and deep features. In: Computational Intelligence (SSCI). 2017 IEEE Symposium Series on, 1–8.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning, 1 ed. MIT Press, Cambridge, MA, USA 800p.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. pp. 770–778.

Hibbett, D., 2009. Automated taxon identification in systematics: theory, approaches and applications. Q. Rev. Biol. 84, 295–296. https://doi.org/10.1086/644681.

Howard, J., 2019. Deep Learning 2019 - Fastai Course. YouTube, Jan. 25 [Video file]. Available: https://youtu.be/XfoYk_Z5AkI [Accessed: May. 14, 2020].

Hsiang, A., Brombacher, A., Rillo, M., Mleneck-Vautravers, M., Conn, S., Lordsmith, S., et al., 2019. Endless forams: 34,000 modern planktonic foraminiferal images for taxonomic training and automated species recognition using convolutional neural networks. Paleoceanography Paleoclimatol. 34, 1157–1177.

Kachovich, S., Sheng, J., Aitchison, J.C., 2019. Adding a new dimension to investigations of early radiolarian evolution. Sci. Rep. 9, 1–10.

Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: active contour models. Int. J. Comput. Vis. 1, 321–331.

Kitchenham, B., 2004. Procedures for performing Systematic Reviews. In: Technical Report. 33. Keele University, Keele, UK, pp. 1–26.

Landis, E.N., Keane, D.T., 2010. X-ray microtomography. Mater. Charact. 61, 1305–1316.

Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. In: Proceedings of the IEEE, pp. 2278–2324.

Liang, D., Weng, K., Wang, C., Liang, G., Chen, H., Wu, X., 2014. A 3d object recognition and pose estimation system using deep learning method. In: 2014 4th IEEE International Conference on Information Science and Technology, pp. 401–404.

Limaye, A., 2012. Drishti: a volume exploration and presentation tool. In: Developments in X-Ray Tomography VIII. 8506. pp. 191–199.

Liu, S., Thonnat, M., Berthod, M., 1994. Automatic classification of planktonic foraminifera by a knowledge-based system. In: Proceedings of the Tenth Conference on Artificial Intelligence for applications, pp. 358–364.

Marchant, R., Tetard, M., Pratiwi, A., de Garidel-Thoron, T., 2019. Classification of down-core foraminifera image sets using convolutional neural networks. bioRxiv In press.

Mitra, R., Marchitto, T., Ge, Q., Zhong, B., Kanakiya, B., Cook, M., Fehrenbacher, J., Ortiz, J., Tripati, A., Lobaton, E., 2019. Automated species-level identification of planktic foraminifera using convolutional neural networks, with comparison to human performance. Mar. Micropaleontol. 147, 16–24.

Molina, E., 2004. Micropaleontología. Concepto, historia y estado actual. In: Micropaleontologia. Prensas Universitárias Zaragoza, 2 ed. Textos Docentes, España, pp. 13–33.

Nair, V., Hinton, G.E., 2009. 3d object recognition with deep belief nets. In: Proceedings of the 22Nd International Conference on Neural Information Processing Systems, pp. 1339–1347.

Otsu, N., 1979. A threshold selection method from gray-level histograms. In: IEEE transactions on systems, man, and cybernetics. 9. pp. 62–66.

Rahman, M.A., Wang, Y., 2016. Optimizing intersection-over-union in deep neural networks for image segmentation. In: International symposium on visual computing, pp. 234–244.

Rakhlin, A., Davydow, A., Nikolenko, S., 2018. Land cover classification from satellite imagery with u-net and lovász-softmax loss. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 257–2574.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. CoRR abs/1505.04597. arXiv:1505.04597.

Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T., Eliceiri, K.W., 2017. Imagej2: Imagej for the next generation of scientific image data. BMC Bioinforma. 18, 529.

Schulze, K., Tillich, U.M., Dandekar, T., Frohme, M., 2013. Planktovision - an automated analysis system for the identification of phytoplankton. BMC Bioinforma. 14, 115.

Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z., 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1874–1883.

Singh, A., 2008. Micropaleontology in petroleum exploration. In: 7th International Conference and Exposition of Petroleum Geophysics, pp. 14–16.

Smith, L.N., 2018. A disciplined approach to neural network hyper-parameters: part 1 - learning rate, batch size, momentum, and weight decay. CoRR abs/1803.09820. arXiv:1803.09820.

Socher, R., Huval, B., Bhat, B., Manning, C.D., Ng, A.Y., 2012. Convolutional-recursive deep learning for 3d object classification. In: Proceedings of the 25th International Conference on Neural Information Processing Systems, pp. 656–664.

Sutton, M., Rahman, I., Garwood, R., 2014. Techniques for Virtual Palaeontology. New Analytical Methods in Earth and Environmental Science, 1 ed. John Wiley Sons Ltd., London 208 p.

Tipper, J.C., 1991. Computer applications in paleontology: Balance in the late 1980s? Comput. Geosci. 17, 1091–1098.

Xia, Y., Zhang, L., Xu, W., Shan, Z., Liu, Y., 2015. Recognizing multi-view objects with occlusions using a deep architecture. Inf. Sci. 320, 333–345.

Xie, J., He, T., Zhang, Z., Zhang, H., Zhang, Z., Li, M., 2018. Bag of tricks for image classification with convolutional neural networks. arXiv preprint arXiv:1812.01187.

Xu, B., Wang, N., Chen, T., Li, M., 2015. Empirical evaluation of rectified activations in convolutional network. CoRR abs/1505.00853.

Xu, X., Dehghani, A., Corrigan, D., Caulfield, S., Moloney, D., 2016. Convolutional neural network for 3d object recognition using volumetric representation. In: 2016 First International Workshop on Sensing, Processing and Learning for Intelligent Machines (SPLINE), pp. 1–5.

Yu, S., Saint-Marc, P., Thonnat, M., Berthod, M., 1996. Feasibility study of automatic identification of planktic foraminifera by computer vision. J. Foraminifer. Res. 26, 113–123.

Yu, J., Weng, K., Liang, G., Xie, G., 2013. A vision-based robotic grasping system using deep learning for 3d object recognition and pose estimation. In: IEEE International Conference on Robotics and Biomimetics. ROBIO 2013, Shenzhen, China, pp. 1175–1180 December 12–14, 2013.

Zhu, W., Huang, Y., Tang, H., Qian, Z., Du, N., Fan, W., Xie, X., 2018. Anatomynet: deep 3d squeeze-and-excitation u-nets for fast and fully automated whole-volume anatomical segmentation. CoRR abs/1808.05238.