Applicability of Object Detection to Microfossil Research: Implications from Deep Learning Models to Detect Microfossil Fish Teeth and Denticles Using YOLO-v7

K. Mimura[1,2], K. Nakamura[3,2,1], K. Yasukawa[3,2], E. C. Sibert[4], J. Ohta[3,5,1], T. Kitazawa[2], Y. Kato[2,1,6]

[1]Ocean Resources Research Center for Next Generation, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan.

[2]Department of Systems Innovation, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

[3]Frontier Research Center for Energy and Resources, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

[4]Department of Geology & Geophysics, Woods Hole Oceanographic Institution, 266 Woods Hole Rd MS 22, Woods Hole, MA 02543, USA.

[5]Volcanoes and Earth's Interior Research Center, Research Institute for Marine Geodynamics, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka, Kanagawa, 237-0061, Japan.

[6]Submarine Resources Research Center, Research Institute for Marine Resources Utilization, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka, Kanagawa, 237-0061, Japan.

Corresponding author: Kentaro Nakamura (kentaron@sys.t.u-tokyo.ac.jp)

Key Points:

- We trained object detection models under different conditions to detect microfossil fish teeth and denticles from microscopic images.
- The best model can count teeth, denticles and irregularly shaped teeth from samples.
- Object detection may improve the observation efficiency of a wide array of microfossils.

Abstract

Microfossils of fish teeth and denticles, referred to as ichthyoliths, provide critical information for depositional ages, paleo-environments, and marine ecosystems, especially in pelagic realms. However, owing to their small size and rarity, it is time-consuming and difficult to analyze large numbers of ichthyoliths from sediment samples, limiting their use in scientific studies. Here, we propose a method to automatically detect ichthyoliths from microscopic images using a deep learning technique. We applied YOLO-v7, one of the latest object detection architectures, and trained several models under different conditions. The model trained under appropriate conditions with an original dataset achieved an F1 score of 0.87. We then enhanced the dataset efficiently using the pre-trained model. We validated the practical applicability of the model by comparing the number of ichthyoliths detected by the model with those counted manually. This revealed that the best model can predict the number of triangular teeth without manual check, and those of denticles and irregularly shaped teeth with manual check. This object detection method can extend the applicability of deep learning to a wider array of microfossils and has the potential to dramatically increase the spatiotemporal resolution of ichthyolith records for applications across disciplines.

Plain Language Summary

Fossils of fish teeth and denticles, referred to as ichthyoliths, can be used to study the environmental changes of marine conditions throughout Earth's history. However, it is time-consuming and difficult to analyze large numbers of ichthyoliths from sediment samples, limiting their use in scientific studies. Here, we trained several artificial intelligence models to automatically detect ichthyoliths from microscopic images. The best model is suitable for counting the number of fish teeth, denticles, and irregularly shaped teeth fragments with minimal human intervention. We propose that object detection, a deep learning technique used in this study, can be applicable for the study of various microfossils, as well as for increasing the spatiotemporal resolution of ichthyolith records.

1 Introduction

Microfossils such as foraminifers, coccolithophores, radiolaria, and diatoms, have been used to constrain depositional ages and environments of various kinds of seafloor sediments, as well as to provide high-resolution and detailed records of evolutionary processes (Armstrong & Brasier, 2005). Among them, microfossil fish teeth and denticles, referred to as ichthyoliths, are composed of calcium phosphate, which is resistant to dissolution on the deep seafloor (Doyle & Riedel, 1985; Sibert et al., 2017). Therefore, ichthyoliths are observed from almost all types of seafloor sediments, including pelagic clay, where other siliceous and calcareous microfossils are rarely observed. Taking advantage of this, ichthyoliths have provided key constraints for depositional ages (Doyle & Riedel, 1979, 1985; Ohta et al., 2020) and marine environments and/or ecosystems (Britten & Sibert, 2020; Sibert et al., 2014, 2016, 2021) especially in pelagic realms. In addition, ichthyoliths preserve a variety of geochemical systems, including strontium and neodymium isotopes, which can provide additional age constraints on sediments (e.g., Gleason et al., 2002; Ingram, 1992) and insights into deep water circulation patterns and origin of sedimentary components (e.g., Huck et al., 2016; Martin & Haley, 2000; Scher & Martin, 2004; Tanaka et al., 2022; Thomas et al., 2014). Oxygen isotopes in ichthyoliths have also been used to reconstruct changes in ocean temperature (e.g., MacLeod et al., 2018). However, traditional observation methods rely on "handpicking," in which an observer picks fossils

individually under a stereomicroscope (Ohta et al., 2020; Sibert et al., 2017; Tanaka et al., 2022). This process is time-consuming and can only be conducted by a skilled observer, making it difficult to analyze large numbers of ichthyoliths from various sediment samples.

Computer vision technologies are developing rapidly. In particular, image processing using deep learning has been applied to various fields, including earth science (Hoeser & Kuenzer, 2020; Mimura et al., 2023a). Automating previous manual observation processes saves time and provides opportunities for discoveries by increasing the number of fossils that can be observed and processed. The application of deep learning techniques for the classification of foraminifers (Hsiang et al., 2019) and radiolarians (Itaki et al., 2020b; Tetard et al., 2020; Carlsson et al. 2022, 2023), and coccolithophores (Beaufort et al., 2022) is enhancing the resolution in paleoenvironmental studies. These studies detect particles by thresholding and recognize their classes using classification models. However, this method is difficult to directly apply to ichthyoliths because it is sometimes challenging to identify the outline of ichthyoliths by thresholding method (Figure 1). To solve this problem, we have proposed an automated detection of ichthyoliths in microscopic images by combining the object detection model "Mask R-CNN" (He et al., 2017) and image classification model "EfficientNet-V2," both of which are based on deep-learning techniques (Mimura et al., 2022). Although the system showed a good performance, two problems remained. First, due to the scarcity of the learning dataset, the system could only detect triangular teeth, leaving denticles and saw-toothed ichthyoliths undetected (Figure 1). Second, there was a time loss in the combined system, as a well-trained object detection model can distinguish classes without using the classification model.

Recently, we compared the performances of object detection models "Mask R-CNN" and "YOLOv5" (Jocher et al., 2022) in detecting signals of hydrothermal activity in echo sounder images (Mimura et al., 2023a) and showed that the YOLOv5 model achieved much higher performance than that of the Mask R-CNN model. Here, with reference to this, we applied "YOLOv7" (Wang et al., 2022), one of the latest versions of YOLO (You Only Look Once, Redmon et al., 2016), to solve the problem of ichthyolith detection. To overcome the problems associated with the previous system developed by us, we aimed to detect teeth, denticles, and irregular shapes of teeth in a single step.

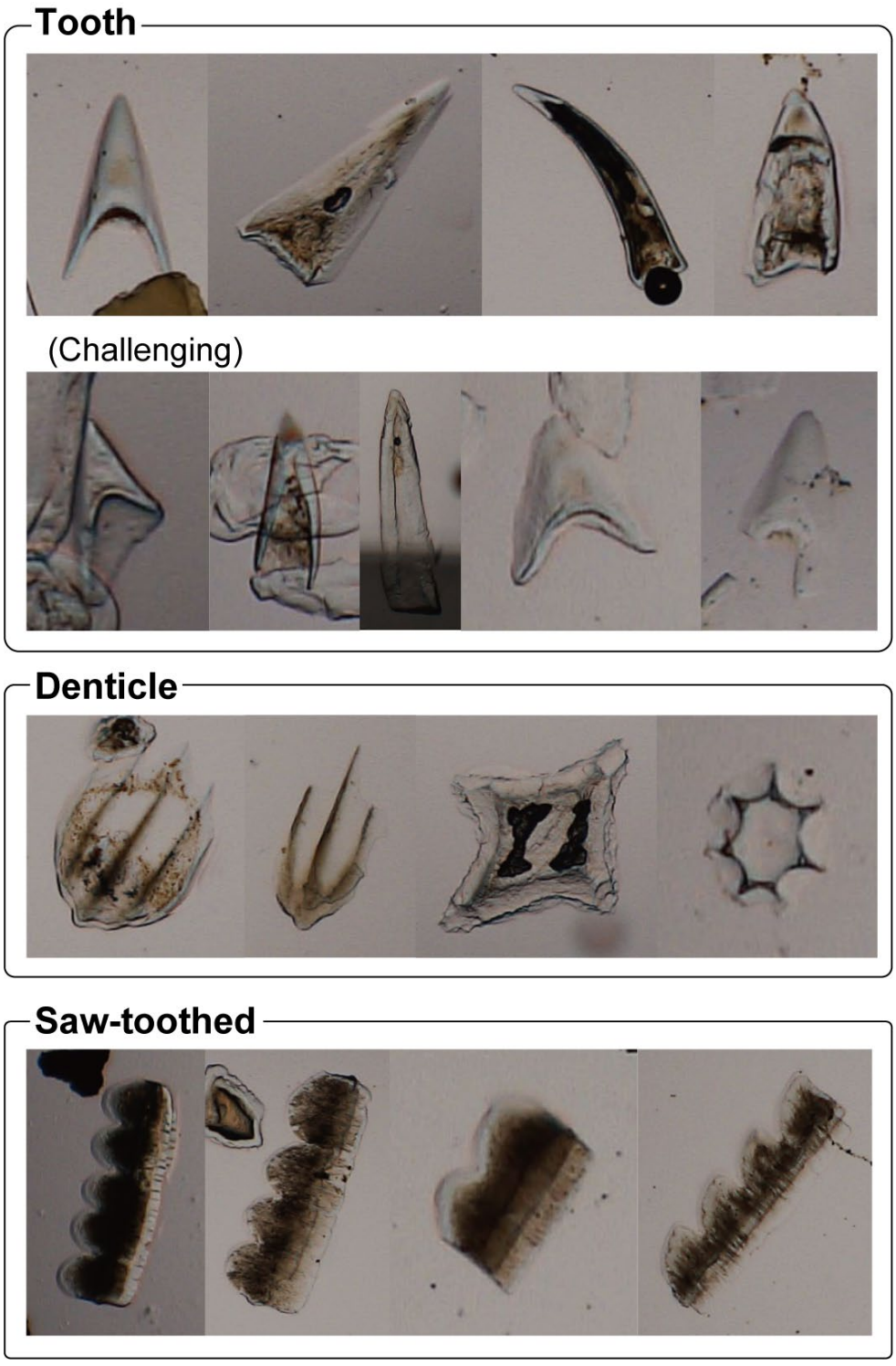

Figure 1. Examples of ichthyolith images categorized into three classes used in this study. Images of teeth considered challenging to detect under the thresholding-based method but can be detected using object detection models are also shown.

2 Materials and Methods

2.1 Sample description

We used pelagic clay samples obtained from the Deep Sea Drilling Project (DSDP) Site 576, Ocean Drilling Program (ODP) Site 1149, Integrated Ocean Drilling Program (IODP) Sites U1366 and U1370, and piston cores KR13-02 PC04 and MR14-E02 PC11. All cores were recovered from the Pacific Ocean at water depths of more than 5,000 m (Table S1). We aimed to cover a variety of depositional ages from the late Cretaceous to the present using DSDP/ODP/IODP samples and to enhance the number of irregular teeth called *Rectangular saw-toothed* (Figure 1) by collecting images from specific horizons of the two piston cores.

2.2 Slide preparation and imaging

Glass slides were prepared from the samples as described by Mimura et al. (2022). Approximately 3−10 g of the sample was mixed well with deionized water and sieved using a 62-μm mesh. Larger particles were collected in a centrifuge, mixed with sodium polytungstate with a specific gravity of approximately 2.8 g/cm^3, and centrifuged at 1,000–1,500 rpm to collect heavier particles, which were proposed by Sibert et al. (2017). The collected particles were washed with deionized water, moved onto glass slides using a pipette, dried at 40 °C, and sealed with a cover glass using a light-curing adhesive.

Imaging of glass slides was also performed as described previously (Mimura et al., 2022). Using a digital microscope RX-100 (Hirox Co., Ltd.), the whole part of the observation realm (~24×36 mm) was divided into ~1,000 squares (~1.15 × 1.15 mm / 1,200 × 1,200 pixels). The z-stack images were automatically acquired using motorized x, y, and z stages. To capture as many ichthyoliths as possible in a complete form, each image overlaps with adjoining images by 20%.

2.3 Generation of datasets

Out of more than 1 million (M) images of the microscopic field of view, 12,219 were selected for “original” datasets. The locations and classes of the ichthyoliths within the images were annotated manually. Ichthyoliths were classified into three classes (Figure 1): triangular tooth (class name: “tooth”), denticle (“denticle”), and forms similar to *Rectangular saw-toothed* (“saw-toothed”).

Two datasets were generated from these images and annotations. The dataset “original_selected” comprised 6,945 images with ichthyoliths, and the dataset “original_all” comprised 6,945 images with ichthyoliths and 5274 images without ichthyoliths (Mimura et al., 2023b). The datasets contained 7,705 triangular teeth, 533 denticles, and 103 saw-toothed shapes. The images and corresponding annotation files were randomly split into three subsets: 80% for training, 10% for validation, and 10% for testing. We note here that images in each subset are the same between the two datasets, except for the image that does not contain ichthyoliths. This enabled us to conduct performance tests on the same dataset (i.e., models trained on the training subset of dataset original_selected can be tested by the testing subset of the dataset original_all).

2.4 Tuning of hyperparameters

We conducted hyperparameter tuning by training the "YOLOv7" model under different initial learning rates ("lr0" in YOLOv7's parameter file) and the final one-cycle learning rates ("lrf"). A stochastic gradient descent algorithm with a momentum fixed at 0.937 was applied for training. The image size was fixed at 640 × 640 pixels and the batch size at 8. The models were trained on a local Windows PC with a single graphic board with 16 GB of memory (GeForce RTXTM 3080 Ti, NVIDIA Inc.).

2.5 Training conditions

YOLOv7 provides several models with various numbers of trainable parameters. In this study, we compared five models, "YOLOv7-tiny," "YOLOv7," "YOLOv7-X," "YOLOv7-W6," and "YOLOv7-E6," each having 6.2M, 36.9M, 71.3M, 70.4M, and 97.2M parameters, respectively. Training of YOLOv7-tiny and YOLOv7 models was conducted on the local Windows PC, while training of the higher models was conducted on the cloud computing platform "Google Colaboratory" (Carneiro et al., 2018). The image size was set to 640 × 640 pixels. However, we also trained YOLOv7-W6 models with a larger image size set to 1,280 × 1,280 pixels, as Wang et al. (2022) proposed for larger models. In all training cases, the batch size was fixed at 8. The models were trained on either the local Windows PC, a local Linux PC with two graphic boards having 24 GB memory (GeForce RTXTM 3090 Ti, NVIDIA Inc.), or Google Colaboratory (see Table 2).

Following YOLOv7's online augmentation method, the images were randomly flipped vertically and/or horizontally, and the colors, scales, and shear of the images were randomly changed every time the training images were loaded.

2.6 Practical test

In the datasets described in Section 2.3, more than half of the images contained at least one ichthyolith, whereas only tens to one hundred ichthyoliths are observed from ~1,000 images in actual observation. We, therefore, conducted a practical test to evaluate the performance of the trained models under more practical conditions.

Three samples at DSDP Site 576, not used in the original datasets described in Section 2.3 or the extended dataset described in Section 3.3, were selected for the practical test. The models detected ichthyoliths from the whole field-of-view images (30,826 in total) taken from 28 slides. Since microscopic images were taken with overlap, duplicated detections were excluded by calculating absolute coordinates in the entire slide (Figure 2). Simultaneously, the slides were manually observed under a polarization microscope. We tested the practical applicability of the trained models by comparing the number of ichthyoliths counted by the models with that observed manually.

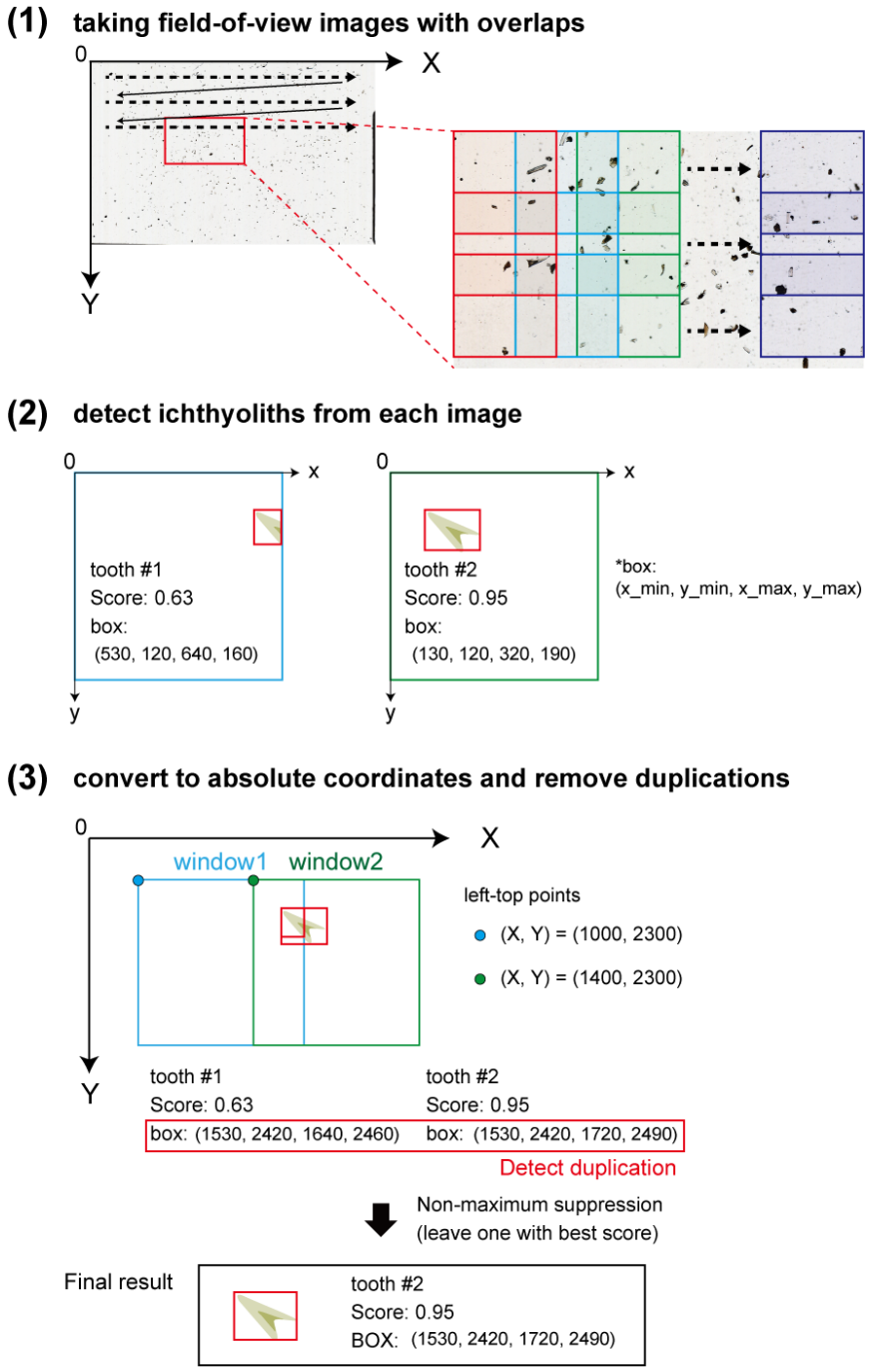

Figure 2. An illustration explaining the algorithm for excluding duplicate detections in this study.

3 Results and Discussion

3.1 Hyperparameter tuning and iteration test

F1 scores of YOLOv7 models trained with different hyperparameters on dataset "original_all" are presented in Table 1. The initial learning rate of 0.0007 and final one-cycle learning rate of 0.05 were the most suitable conditions in this study. Under the same condition, we then conducted and evaluated five training iterations and observed that one standard error (1 SE) of the F1 score was 0.008 (Table S2). When comparing the performance of the models in the following discussion, a difference in F1 scores greater than 2 SE (0.016) was considered significant.

Table 1. F1 scores of the models trained on different hyperparameters of initial learning rate ("lr0") and final one-cycle learning rate ("lrf").

F1 score		lrf		
		0.1	0.05	0.01
lr0	0.001	0.08	0.29	0.75
	0.0007	0.46	**0.82**	0.55
	0.0004	0.68	0.13	0.21

3.2 Comparison of performances under different training conditions

The performance of the models trained on different models and datasets is detailed in #1 to #12 of Table 2. We evaluated the performance of models based on averaged F1 scores of the three classes (macro-F1 score). Comparing the number of parameters (Figure 3a), models with ~70M trainable parameters (YOLOv7-X, YOLOv7-W6) exhibited the highest F1 score, suggesting that these models are suitable for this study. Comparing the image sizes (Figure 3b), we observed that the models trained with the input image size set at 640 exhibited higher F1 scores than those trained with an image size of 1,280. Although the difference in the dataset "selected" is less than 2SE, we suggest that the suitable input image size is 640, as larger input size increases the risk of overfitting (e.g., Sabottke and Spieler, 2020). Finally, comparing the dataset type (Figure 3c), the results exhibited a variety of trends. However, following the discussion above, if we focus on the cases with a number of parameters around 70M and input image size at 640, models trained on the dataset "all" showed higher F1 scores than those trained on the dataset "selected." Thus, we concluded that the suitable training condition in this study is (1) to use models with ~70M parameters (YOLOv7-X or YOLOv7-W6), (2) to set the input image size at 640, and (3) to train on a dataset "all," which is composed of both images containing ichthyoliths and images that do not contain ichthyoliths.

Table 2. Performances of the training with different models and datasets.

Case		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Condition	Dataset	original_selected						original_all						extended_all	
	Environment	Windows	Windows	Colab	Ubuntu	Ubuntu	Ubuntu	Windows	Windows	Colab	Ubuntu	Ubuntu	Ubuntu	Ubuntu	Ubuntu
	Model	YOLOv7-tiny	YOLOv7	YOLOv7-X	YOLOv7-W6	YOLOv7-W6	YOLOv7-E6	YOLOv7-tiny	YOLOv7	YOLOv7-X	YOLOv7-W6	YOLOv7-W6	YOLOv7-E6	YOLOv7-X	YOLOv7-W6
	#Param. (M)	6.2	36.9	71.3	70.4	70.4	97.2	6.2	36.9	71.3	70.4	70.4	97.2	71.3	70.4
	image size	640	640	640	640	1280	640	640	640	640	640	1280	640	640	640
Precision	tooth	0.814	0.885	0.885	0.882	0.778	0.760	0.641	0.868	0.857	0.778	0.791	0.768	0.910	0.931
	denticle	0.671	0.829	0.859	0.803	0.712	0.702	0.672	0.852	0.853	0.756	0.668	0.653	0.832	0.895
	saw-toothed	0.778	0.875	0.799	0.888	0.817	0.714	0.694	0.800	0.900	0.833	0.727	0.833	0.778	0.887
	average	**0.754**	**0.863**	**0.848**	**0.858**	**0.769**	**0.725**	**0.669**	**0.840**	**0.870**	**0.789**	**0.728**	**0.751**	**0.840**	**0.904**
Recall	tooth	0.789	0.832	0.857	0.850	0.907	0.862	0.880	0.859	0.896	0.908	0.916	0.870	0.843	0.813
	denticle	0.755	0.837	0.869	0.776	0.857	0.755	0.837	0.776	0.831	0.918	0.857	0.776	0.878	0.878
	saw-toothed	0.699	0.698	0.800	0.800	0.893	0.500	0.682	0.800	0.899	0.999	0.797	0.500	0.700	0.800
	average	**0.748**	**0.789**	**0.842**	**0.888**	**0.886**	**0.706**	**0.800**	**0.811**	**0.875**	**0.942**	**0.857**	**0.715**	**0.807**	**0.830**
f1 score	tooth	0.801	0.858	0.871	0.866	0.838	0.808	0.742	0.863	0.876	0.838	0.849	0.816	0.875	0.868
	denticle	0.711	0.833	0.864	0.789	0.778	0.728	0.745	0.812	0.842	0.829	0.751	0.709	0.854	0.886
	saw-toothed	0.736	0.777	0.799	0.842	0.853	0.588	0.688	0.800	0.899	0.908	0.760	0.625	0.737	0.841
	average	**0.749**	**0.822**	**0.845**	**0.832**	**0.823**	**0.708**	**0.725**	**0.825**	**0.872**	**0.859**	**0.787**	**0.717**	**0.822**	**0.865**

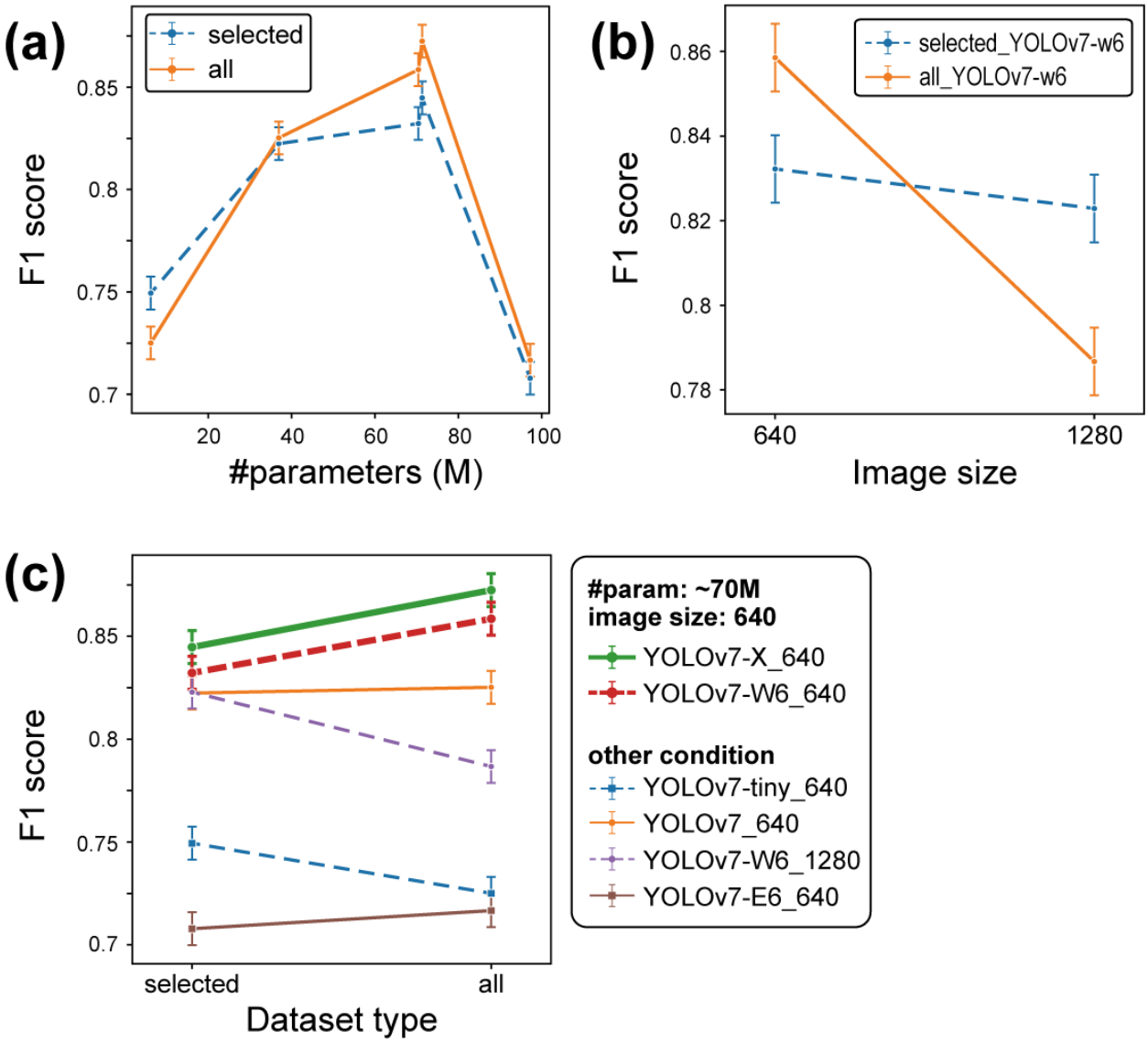

Figure 3. F1 scores compared by training conditions. The x-axis of each graph represents (a) the number of trainable parameters, (b) image sizes, and (c) the type of dataset. Error bars represent ± 1 SE.

3.3 Efficient production of training dataset using detection results

YOLOv7 can output results as text files in the same format as the training labels. Taking advantage of this, we enhanced the sizes of datasets by first predicting a trained model and then checking the result manually. Using the YOLOv7-X model trained on the dataset "all" with an image size of 640 (#9 of Table 2), the existence of ichthyoliths was predicted from ~1,100,000 images generated from the six sites considered in this study. Images from three samples at Site 576 used for the practical test were excluded. We collected 4,463 images in which the model predicted the existence of the class "denticle" or "saw-toothed," which were relatively small compared to the class "tooth. After the manual check of detection results for the 4,463 images, 2,528 images contained ichthyoliths, and 1,935 did not have ichthyoliths; of those containing

ichthyoliths, 1,657 teeth, 1,282 denticles, and 108 saw-toothed ichthyoliths were identified. Of these, the "denticle" was more than twice the number in the original dataset, and the "saw-toothed" was almost the same as the number in the original dataset. As well as the original datasets, images, and annotation information were randomly split into training (80%), validation (10%), and testing (10%) subsets.

The dataset "extended_all" was generated by combining the dataset collected by the above process and the dataset "original_all" (Mimura et al., 2023b). Considering the discussion in Section 3.2, we trained the two models, YOLOv7-X and YOLOv7-W6, on the dataset "extended_all" with an input image size set at 640. The performances of the trained models are shown in #13 and #14 of Table 2.

3.4 Practical test

We conducted a practical test for the four models: YOLOv7-X trained on the datasets "original_all" (#9 of Table 2) and "extended_all" (#13), YOLOv7-w6 trained on the datasets "original_all" (#10) and "extended_all" (#14). The number of ichthyoliths detected by these models and that are manually counted are shown in Table S3. We also calculated the root mean square percentage error (RMSPE), using the following equation:

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}_i - y_i}{y_i}\right)^2} \times 100\ [\%] \ldots (1)$$

where n, $\hat{y}_i$, and y_i indicate the number of samples, the predicted ichthyoliths, and the manually observed ichthyoliths, respectively.

Comparing the models trained on the dataset "original_all" (#9, #10) and "extended_all" (#13, #14), models trained on "extended_all" showed trends closer to y = x for classes tooth and denticle (Figure 4a, b). The high performance of the model trained with the "extended_all" dataset may be attributed to the high variation of false patterns in practical conditions. We realized that models trained on the original dataset confused various triangular particles or patterns with teeth (Figure S1). Since the "extended_all" dataset contains many images that the preliminary model misdetected, the model trained with this dataset is considered to learn false positives efficiently. RMSPEs suggest that using the v7-w6_extended_all model (#14), the number of teeth and denticles from a sample can be estimated with ~7% and ~24% error rates, respectively. On the other hand, RMSPEs for the "saw-toothed" class are > 70%. Furthermore, no clear trend was observed (Figure 4c), indicating that the number of "saw-toothed" cannot be accurately estimated based solely on the model's detection result.

We also manually checked the images detected by models #13 and #14 and removed false positives and duplications that could not be excluded by the algorithm described in Figure 2. After checking model #13's detection, we observed a trend closer to y=x (Figure 4d–f), indicating that combining manual review with model #13 is preferable. Model #13, with manual check, achieved an RMSPE of ~3%, ~9%, and almost no error for counting the number of teeth, denticles, and saw-toothed ichthyoliths, respectively (Table S3).

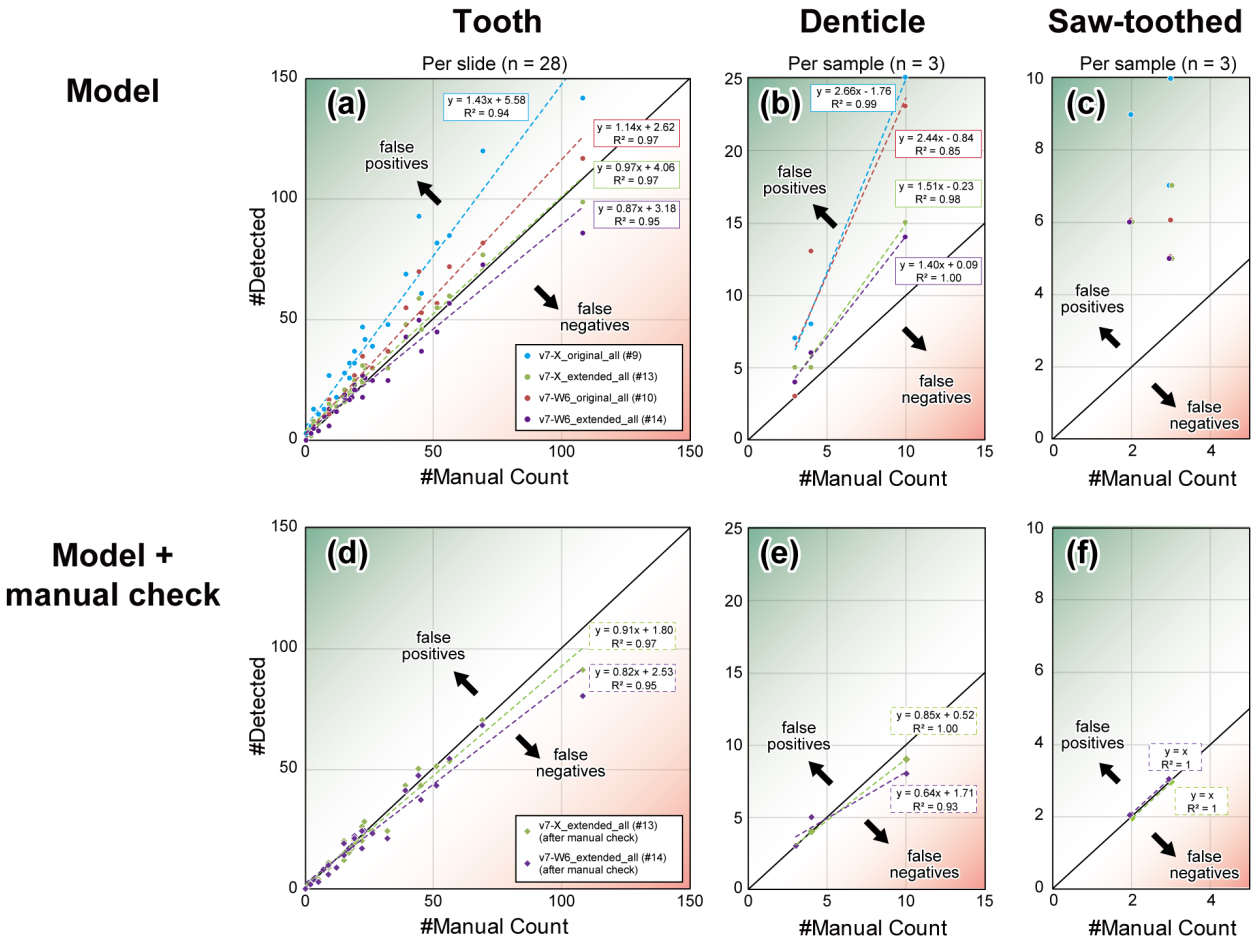

Figure 4. Comparison of the number of ichthyoliths counted manually and those detected by models trained in this study. The black solid lines indicate y = x, which means that the model's detections are identical to manual observations. Plots below and above the y = x line indicate that the model made false negative and false positive errors, respectively. (a–c) Scatter diagram of the number of models' detection and manual count. Regression lines are only indicated in class "tooth" (a), as no clear trend was observed in class "denticle" (b) and class "saw-toothed" (c). (e–f) Comparison between the number of ichthyoliths that a human observer recounted after the best model's detection and the manually counted number. The numbers were compared per slide for teeth, but per sampling horizon for classes denticle and saw-toothed, particles in these slides were contained in only a few slides. Regression lines were obtained using Excel (Microsoft® Excel® for Microsoft 365 MSO, version 2310).

3.5 Advantages of object detection method using YOLO-v7

The application of deep learning to microfossil observations has attracted increasing attention recently (Hsiang et al., 2019; Salonen et al., 2019; Mitra et al., 2019; Romero et al., 2020; Itaki et al., 2020a; Marchant et al., 2020; Tetard et al., 2020; Classon et al., 2022, 2023). A commonly used method in particle detection is to apply rule-based thresholding to detect each particle and subsequently classify them using an image classification model. Although these methods require less work to prepare a dataset, deep learning-based detection has advantages over traditional methods in finding "challenging" particles. While traditional rule-based

thresholding methods struggle to detect particles that overlap, have drastic changes in brightness, or have almost similar brightness to the background (Figure 1) in ichthyolith slides, deep learning-based methods can accurately detect them. Therefore, we propose that object detection would broaden the range of deep learning applications in microfossil studies.

Compared to our previous method (Mimura et al., 2022), which required two steps, object detection by Mask R-CNN and image classification by EfficieneNet-V2, the new method can detect ichthyolith in a single step, which enhances the efficiency of observation. We measured the detection times for processing 10,884 slide images using the two methods on Google Colaboratory. While the previous method required 11,250 s in total, 7,230 s for detection using Mask R-CNN, and 4,020 s for classification using EfficientNet-V2, the new method required only 1,040 s in total process, indicating that the new method is approximately ten times faster than the previous method.

3.6 Implications for biostratigraphic and paleoecological studies using ichthyoliths

We expect the new observation method to make the biostratigraphy of ichthyoliths more precise, advancing progress in paleoceanography and resource geology related to pelagic (red) clay. Pelagic clay covers over one-third of the global ocean (Dutkiewicz et al., 2015) and has huge variation in bulk geochemistry (Dunlea et al., 2015; Mimura et al., 2019). Therefore, pelagic clay is a good recorder of long-term and global/regional environmental changes (Zhou and Kyte, 1992; Kyte et al., 1993; Tanaka et al., 2022; Yasukawa et al., 2023). Moreover, pelagic clay is also attracting attention as a promising resource for rare-earth elements (Kato et al., 2011; Yasukawa et al., 2014; Takaya et al., 2018; Ren et al., 2021). However, the scarcity of microfossils except for ichthyoliths has hampered making precise age models of pelagic clay. Letting machines perform much of the time-consuming observations, substantial amounts of ichthyoliths can be observed, and more accurate age models will be established. This should provide numerous insights into the evolution of pelagic environments from paleoceanographic viewpoints, as well as the ore genesis and potential distributions of the prospective deep-sea mineral resource.

We also expect that this tool will improve our understanding in biological and ecological studies. As a demonstration, we show a downhole variation of denticle/tooth (D/T) ratios at DSDP Site 576 in the western North Pacific Ocean (Table 3, Figure 5), which were generated from the detection results of model #13 combined with manual check. D/T ratio is an index for relative ratios of shark and ray-fined fish, an indicator of marine vertebrate community stability (Sibert et al., 2016). By manual counting in a previous study (Sibert et al., 2016), three stages in the D/T ratios from the late Cretaceous to the present were proposed. Cretaceous ocean (i.e., older than 66 Ma) was characterized by high D/T ratios, reflecting a relatively small number of ray-fined fishes compared to the present ocean. Subsequently, Paleogene ocean (from 66 to ~20 Ma) showed moderate D/T ratios, reflecting the evolution of ray-finned fish after the K/Pg boundary (Sibert and Norris, 2015). Finally, the modern ocean (from ~20 Ma to the present) is characterized by low D/T ratios, which may reflect an extinction event of sharks in the early Miocene (Sibert et al., 2020) and the consequent predominance of ray-finned fish. In the previous study, the trend was clearly exhibited from the South Pacific (DSDP Site 596), but the evidence from the North Pacific (ODP Site 886) was somewhat limited due to the huge hiatus in the Paleogene (Figure 5). Using our deep learning-based image processing method, we found D/T ratios results that were consistent with the previous study from DSDP Site 576 in the North

Pacific site that has continuous Paleogene sedimentation, supporting the pelagic vertebrate community structure proposed in Sibert et al. (2016). While this method is still developing, high throughput data collection provides the opportunity for elucidating the interaction between environmental change and the marine vertebrate community.

Table 3. The total count of ichthyoliths in the three classes was detected by model #13 (YOLOv7-X. trained on the dataset "extended_all" with an image size of 640. Results after the manual check are also provided.

Sample	depth (mbsf)	Age	weight (g)	v7x_extended_all			v7x_extended_all + manual check			Outlier
				tooth	denticle	sawToothed	tooth	denticle	sawToothed	
576B_01_02_77	2.28	P-Q[a]	2.53	2	1	0	2	0	0	1[b]
576B_01_06_52	8.03	P-Q[a]	3.45	16	0	1	13	0	1	
576B_02_03_125	12.46	P-Q[a]	2.99	44	4	2	36	1	0	
576B_02_07_23	17.44	P-Q[a]	6.40	89	3	8	76	1	1	
576B_03_03_125	21.96	Miocene	2.19	12	1	0	9	0	0	1[b]
576B_03_06_81	26.02	Miocene	3.81	123	5	5	101	3	3	
576B_04_01_75	27.96	Oligocene	2.72	38	3	2	31	2	1	1[b]
576B_04_02_25	28.96	Oligocene	3.26	151	5	6	124	4	2	
576B_04_03_25	30.46	Oligocene	3.65	191	9	6	164	9	1	
576B_04_03_75	30.96	Oligocene	3.75	194	7	8	165	4	3	
576B_04_04_75	32.46	Oligocene	5.71	526	15	7	472	9	3	
576B_04_05_75	33.97	Oligocene	4.24	542	32	4	472	24	2	
576B_04_06_21	34.93	Oligocene	4.04	1099	78	17	969	61	3	
576B_04_07_27	36.485	Oligocene	4.84	517	46	3	451	39	1	
576B_05_01_25	36.96	Eocene	3.71	455	44	3	391	32	2	
576B_05_02_75	38.96	Eocene	5.29	497	40	1	417	35	1	
576B_05_03_75	40.47	Eocene	5.26	635	55	4	519	43	4	
576B_05_04_75	41.97	Eocene	4.43	115	13	0	96	12	0	
576B_05_05_25	42.97	Eocene	5.52	294	31	1	223	26	0	
576B_05_06_75	44.97	Eocene	4.63	413	21	0	345	13	0	
576B_06_02_75	48.47	Paleocene	2.85	373	16	0	330	13	0	
576B_06_04_23	50.95	Paleocene	3.89	28	10	0	23	8	0	1[b]
576B_06_04_140	52.11	Cretaceous	3.42	438	52	1	357	47	0	
576B_06_06_140	55.11	Cretaceous	4.16	486	82	0	392	68	0	
576B_07_03_139	60.105	Cretaceous	4.51	444	53	0	308	46	0	
576B_07_07_30	65.01	Cretaceous	2.48	239	21	0	164	18	0	

Note. [a]Pliocene–Quaternary. [b]Too small #tooth compared to uppler/lower horizons.

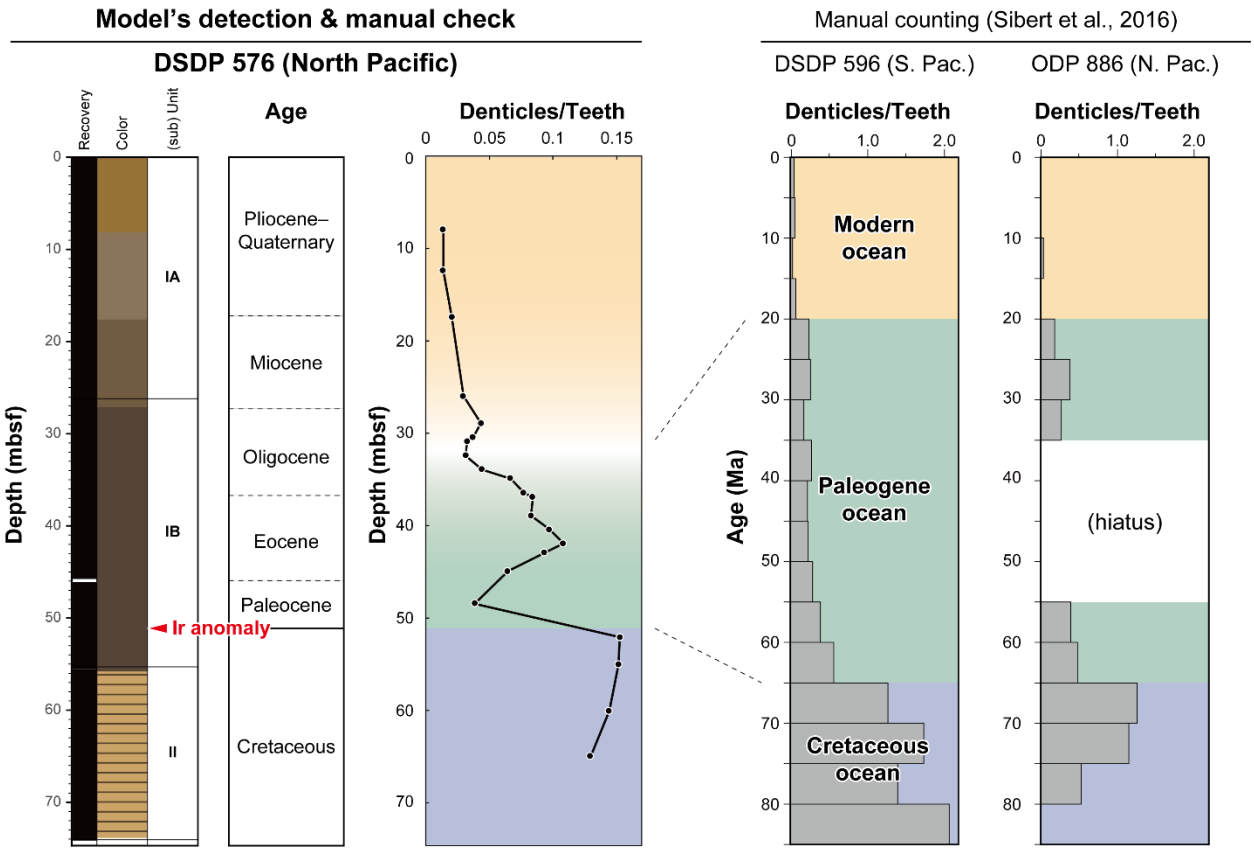

Figure 5. A downhole variation of denticle/tooth ratios at DSDP Site 576, hole 576B, obtained by the detection model proposed in this study. The age model at Site 576 is based on ichthyolith biostratigraphy (Shipboard Scientific Party, 1985) corrected by Ir anomaly (Kyte et al., 1995). The results of three samples with two small numbers of teeth compared to upper and lower horizons were excluded from the plot. D/T ratios obtained by manual counting at DSDP Site 596 (South Pacific) and ODP Site 886 (North Pacific) are also shown.

4 Conclusions

In this study, we proposed a new and efficient method for the observation of ichthyoliths, which is approximately ten times faster than our previous method. Using this method, we expect that studies using ichthyoliths, including biostratigraphy, geochemistry, paleoecology, and the evolution of fishes, will become more precise due to improved sample throughput and identification. Conventional studies on ichthyolith stratigraphy have focused mainly on the presence or absence of each ichthyolith species. In contrast, ratios of the species were hardly considered, possibly due to the enormous amount of manual work required to count the total number of fossils in a discrete sediment sample under a microscope. Since the object detection method is capable of counting the total number of ichthyoliths in a sample, as well as classifying them to a particular type (here, teeth, denticles, or saw-toothed teeth), it can rapidly calculate a ratio of each ichthyolith species within an entire sample slide glass. This tool enables research focusing on quantitative changes in the occurrence of each ichthyolith morphotype, which in turn will provide more accurate depositional ages on pelagic clays, improve geochemical reconstructions, and open the possibilities for high-resolution ecological and evolutionary studies

of fish and sharks at significantly increased spatiotemporal resolution. Finally, while we focused here on ichthyoliths, which are understudied compared to other microfossil groups, the automated deep learning methods presented here can be applied broadly to a wide array of microfossil groups, increasing the throughput of data across many fields of study.

Acknowledgments

The authors thank Prof. Richard Norris at the University of California San Diego for preliminary discussions on ichthyolith analysis. This research used drill core samples collected by the Deep Sea Drilling Project (DSDP), Ocean Drilling Program (ODP), and Integrated Ocean Drilling Program (IODP). This research was funded by the Japan Society for the Promotion of Science (JSPS) KAKENHI grant numbers 20H05658 to YK, 17H01361 to KN, 19J14560, 21K20354, and 23K13192 to KM, and JST FOREST Program (Grant Number JPMJFR227A, Japan), as well as National Science Foundation (NSF) grant number 2403839. The authors declare no conflict of interest.

Data Availability statement

We named a series of program codes “yolov7-slideObservation” and made it available on GitHub (https://github.com/KazuhideMimura/yolov7-slideObservation). The datasets for this study can be accessed at figshare (Mimura et al., 2023b).

References

Armstrong, H. A., & Brasier, M. D. (2005). *Microfossils* (2nd edition). Oxford, UK: Blackwell Publishing. 296 pp. https://doi.org/10.1017/S001675680621238X

Beaufort, L., Bolton, C. T., Sarr, A.-C., Suchéras-Marx, B., Rosenthal, Y., Donnadieu, Y., et al. (2022). Cyclic evolution of phytoplankton forced by changes in tropical seasonality. Nature, *601*(7891), 79–84. https://doi.org/10.1038/s41586-021-04195-7

Britten, G. L., & Sibert, E. C. (2020). Enhanced fish production during a period of extreme global warmth. *Nature Communications*, *11*(1), 5636. https://doi.org/10.1038/s41467-020-19462-w

Carneiro, T., Medeiros Da Nobrega, R. V., Nepomuceno, T., Bian, G., De Albuquerque, V. H. C., & Filho, P. P. R. (2018). Performance analysis of Google colaboratory as a tool for accelerating deep learning applications. *IEEE Access*, *6*, 61677–61685. https://doi.org/10.1109/ACCESS.2018.2874767

Carlsson, V., Danelian, T., Boulet, P., Devienne, P., Laforge, A., & Renaudie, J. (2022). Artificial intelligence applied to the classification of eight middle Eocene species of the genus Podocyrtis (polycystine radiolaria). *Journal of Micropalaeontology, 41*(2), 165–182. https://doi.org/10.5194/jm-41-165-2022

Carlsson, V., Danelian, T., Tetard, M., Meunier, M., Boulet, P., Devienne, P., & Ventalon, S. (2023). Convolutional neural network application on a new middle Eocene radiolarian dataset. *Marine Micropaleontology, 183*, 102268. https://doi.org/10.1016/j.marmicro.2023.102268

Doyle, P. S., & Riedel, W. R. (1985). Cenozoic and late Cretaceous ichthyoliths. In H. M. Bolli, J. B. Saunders, & K. Perch-Nielsen (Eds.), *Plankton Stratigraphy* (pp. 965–995). Cambridge, UK: Cambridge University Press.

Doyle, P. S., & Riedel, W. R. (1979). Ichthyoliths: Present status of taxonomy and stratigraphy of microscopic fish skeletal debris. *Scripps Institution of Oceanography Reference Series* (Report 79–16, pp. 1–231). La Jolla, CA: University of California Scripps Institute of Oceanography.

Dunlea, A. G., Murray, R. W., Sauvage, J., Spivack, A. J., Harris, R. N., & D'Hondt, S. (2015). Dust, volcanic ash, and the evolution of the South Pacific Gyre through the Cenozoic. *Paleoceanography*, *30(8)*, 1078–1099. https://doi.org/10.1002/2015PA002829

406 Dutkiewicz, A., Müller, R. D., O'Callaghan, S., & Jónasson, H. (2015). Census of seafloor
407 sediments in the world's ocean. *Geology*, *43*(9), 795–798. https://doi.org/10.1130/G36883.1

408 Gleason, J. D., Moore, T. C., Rea, D. K., Johnson, T. M., Owen, R. M., Blum, J. D., Hovan, S.
409 A., & Jones, C. E. (2002). Ichthyolith strontium isotope stratigraphy of a Neogene red clay
410 sequence: calibrating eolian dust accumulation rates in the central North Pacific. *Earth and*
411 *Planetary Science Letters*, *202*(3–4), 625–636. https://doi.org/10.1016/S0012-821X(02)00827-0

412 He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2020). Mask r-cnn. In. *IEEE Transactions on*
413 *Pattern Analysis and Machine Intelligence*, IEEE International Conference on Computer Vision
414 (ICCV) 42(2), 386–397

415 Hoeser, T., & Kuenzer, C. (2020). Object detection and image segmentation with deep learning
416 on earth observation data: A review-Part I: Evolution and recent trends. *Remote Sensing*,
417 *12*(10),1667. https://doi.org/10.3390/rs12101667

418 Hsiang, A. Y., Brombacher, A., Rillo, M. C., Mleneck-Vautravers, M. J., Conn, S., Lordsmith,
419 S., et al. (2019). Endless Forams: > 34,000 modern planktonic foraminiferal images for
420 taxonomic training and automated species recognition using convolutional neural networks.
421 *Paleoceanography and Paleoclimatology*, *34*(7), 1157–1177.
422 https://doi.org/10.1029/2019PA003612

423 Huck, C. E., van de Flierdt, T., Jiménez-Espejo, F. J., Bohaty, S. M., Röhl, U., & Hammond, S.
424 J. (2016). Robustness of fossil fish teeth for seawater neodymium isotope reconstructions under
425 variable redox conditions in an ancient shallow marine setting. *Geochemistry, Geophysics,*
426 *Geosystems*, *17*(3), 679–698. https://doi.org/10.1002/2015GC006218

Ingram, B. L. (1992). Ichthyolith strontium isotopic stratigraphy of deep-sea clays: Sites 885 and 886 (North Pacific Transect). *Proceedings of the Ocean Drilling Program: Scientific Results*. https://doi.org/10.2973/odp.proc.sr.145.130.1995

Itaki, T., Taira, Y., Kuwamori, N., Maebayashi, T., Takeshima, S., & Toya, K. (2020a). Automated collection of single species of microfossils using a deep learning–micromanipulator system. *Progress in Earth and Planetary Science*, *7*(19), 1–7. https://doi.org/10.1186/s40645-020-00332-4

Itaki, T., Taira, Y., Kuwamori, N., Saito, H., Ikehara, M., & Hoshino, T. (2020b). Innovative microfossil (radiolarian) analysis using a system for automated image collection and AI-based classification of species. *Scientific Reports*, *10*, 21136. https://doi.org/10.1038/s41598-020-77812-6

Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., Nanocode012, Kwon, Y., et al. (2022). [software] ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation (v7.0). *Zenodo*. https://doi.org/10.5281/zenodo.7347926

Kato, Y., Fujinaga, K., Nakamura, K., Takaya, Y., Kitamura, K., Ohta, J., Toda, R., Nakashima, T., & Iwamori, H. (2011). Deep-sea mud in the Pacific Ocean as a potential resource for rare-earth elements. *Nature Geoscience*, *4*, 535–539. https://doi.org/10.1038/ngeo1185

Kyte, F. T., Leinen, M., Heath, G. R., & Zhou, L. (1993). Cenozoic sedimentation history of the central North Pacific: Inferences from the elemental geochemistry of core LL44-GPC3. *Geochimica et Cosmochimica Acta, 57*(8), 1719–1740. https://doi.org/10.1016/0016-7037(93)90109-A

448 Kyte, F. T., Bostwick, J. A., & Zhou, L. (1995). Identification and characterization of the
449 Cretaceous/Tertiary boundary at ODP Sites 886 and 803 and DSDP Site 576. In I. A. Basov, D.
450 W. Scholl, & J. F. Allan (Eds.) *Proceedings of the ODP*, sci. results, 145 (pp. 427–434). Rea:
451 D.K. TX: College Station. https://doi.org/10.2973/odp.proc.sr.145.132.1995

452 MacLeod, K. G., Quinton, P. C., Sepúlveda, J., & Negra, M. H. (2018). Postimpact earliest
453 Paleogene warming shown by fish debris oxygen isotopes (El Kef, Tunisia). *Science*, *360*(6396),
454 1467–1469. https://doi.org/10.1126/science.aap8525

455 Marchant, R., Tetard, M., Pratiwi, A., Adebayo, M., & de Garidel-Thoron, T. (2020). Automated
456 analysis of foraminifera fossil records by image classification using a convolutional neural
457 network. *Journal of Micropalaeontology*, *39*(2), 183–202. https://doi.org/10.5194/jm-39-183-
458 2020

459 Martin, E.E., & Haley, B.A. (2000). Fossil fish teeth as proxies for seawater Sr and Nd isotopes.
460 *Geochimica et Cosmochimica Acta 64*(5): 835–847. https://doi.org/10.1016/S0016-
461 7037(99)00376-2

462 Mitra, R., Marchitto, T. M., Ge, Q., Zhong, B., Kanakiya, B., Cook, M. S., Fehrenbacher, J. S.,
463 Ortiz, J. D., Tripati, A., & Lobaton, E. (2019). Automated species-level identification of planktic
464 foraminifera using convolutional neural networks, with comparison to human performance.
465 *Marine Micropaleontology, 147*, 16–24. https://doi.org/10.1016/j.marmicro.2019.01.005

466 Mimura, K., Nakamura, K., Yasukawa, K., Machida, S., Ohta, J., Fujinaga, K., & Kato, Y.
467 (2019). Significant impacts of pelagic clay on average chemical composition of subducting
468 sediments: New insights from discovery of extremely rare-earth elements and yttrium-rich mud

at Ocean Drilling Program Site 1149 in the western North Pacific Ocean. *Journal of Asian Earth Sciences*, *186*, 104059. https://doi.org/10.1016/j.jseaes.2019.104059

Mimura, K., Minabe, S., Nakamura, K., Yasukawa, K., Ohta, J., & Kato, Y. (2022). Automated detection of microfossil fish teeth from slide images using combined deep learning models. *Applied Computing and Geosciences*, *16*, 100092. https://doi.org/10.1016/j.acags.2022.100092

Mimura, K., Nakamura, K., Takao, K., Yasukawa, K., & Kato, Y. (2023a). Automated detection of hydrothermal emission signatures from multi-beam echo sounder images using deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 2703–2710. https://doi.org/10.1109/JSTARS.2023.3247467

[Dataset] Mimura, K., Nakamura, K., Yasukawa, K., Sibert, E. C., Ohta, J., Kitazawa, T., &Kato, Y. (2023b) Image datasets for training, validating and testing deep learning models to detect microfossil fish teeth and denticles called ichthyolith using YOLOv7. *figshare*. https://doi.org/10.6084/m9.figshare.22736609.v1

Ohta, J., Yasukawa, K., Nozaki, T., Takaya, Y., Mimura, K., Fujinaga, K., Nakamura, K., Usui, Y., Kimura, J.-I., Chang, Q., & Kato, Y. (2020). Fish proliferation and rare-earth deposition by topographically induced upwelling at the late Eocene cooling event. *Scientific Reports, 10*, 9896. https://doi.org/10.1038/s41598-020-66835-8

Ren, J., Liu, Y., Wang, F., He, G., Deng, X., Wei, Z., & Yao, H. (2021). Mechanism and influencing factors of REY enrichment in deep-sea sediments. *Minerals*, *11*(2), 196. https://doi.org/10.3390/min11020196

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91

Romero, I. C., Kong, S., Fowlkes, C. C., Jaramillo, C., Urban, M. A., Oboh-Ikuenobe, F., D'Apolito, C., & Punyasena, S. W. (2020). Improving the taxonomy of fossil pollen using convolutional neural networks and superresolution microscopy. *Proceedings of the National Academy of Sciences, 117*(45), 28496–28505. https://doi.org/10.1073/pnas.2007324117

Sabottke, C. F., & Spieler, B. M. (2020). The effect of image resolution on deep learning in radiography. *Radiology: Artificial Intelligence, 2(*1), e190015. https://doi.org/10.1148/ryai.2019190015

Salonen, J. S., Korpela, M., Williams, J. W., & Luoto, M. (2019). Machine-learning based reconstructions of primary and secondary climate variables from North American and European fossil pollen data. Scientific Reports, *9*(1), 15805. https://doi.org/10.1038/s41598-019-52293-4

Scher, H. D., & Martin E. E. (2004). Circulation in the Southern Ocean during the Paleogene inferred from neodymium isotopes. *Earth and Planetary Science Letters*, *228*(3–4), 391–405. https://doi.org/10.1016/j.epsl.2004.10.016

Shipboard Scientific Party (1985). Site 576. In *Initial Reports of the Deep Sea Drilling Project*, (Vol. 86, pp. 51–90).

Sibert, E. C., Hull, P. M., & Norris, R. D. (2014). Resilience of Pacific pelagic fish across the Cretaceous/Palaeogene mass extinction. *Nature Geoscience*, *7*(9), 667–670. https://doi.org/10.1038/NGEO2227

510 Sibert, E. C., & Norris, R. D. (2015). New Age of Fishes initiated by the Cretaceous−Paleogene
511 mass extinction. *Proceedings of the National Academy of Sciences*, *112*(28): 8537–8542.
512 https://doi.org/10.1073/pnas.1504985112

513 Sibert, E. C., Norris, R. D., Cuevas, J., & Graves, L. (2016). Eighty-five million years of Pacific
514 Ocean gyre ecosystem structure: Long-term stability marked by punctuated change. *Proceedings*
515 *of the Royal Society B: Biological Sciences*, *283*(1831), 20160189.
516 https://doi.org/10.1098/rspb.2016.0189

517 Sibert, E. C., Cramer, K. L., Hastings, P. A., & Norris, R. D. (2017). Methods for isolation and
518 quantification of microfossil fish teeth and elasmobranch dermal denticles (ichthyoliths) from
519 marine sediments. *Palaeontologia Electronica*, *20*(1), 1–14. https://doi.org/10.26879/677

520 Sibert, E. C., & Rubin, L. D. (2021). An early Miocene extinction in pelagic sharks. *Science*,
521 *372*(6546), 1105–1107. https://doi.org/10.1126/science.aaz3549

522 Takaya, Y., Yasukawa, K., Kawasaki, T., Fujinaga, K., Ohta, J., Usui, Y., et al. (2018). The
523 tremendous potential of deep-sea mud as a source of rare-earth elements. *Scientific Reports*, *8*,
524 5763. https://doi.org/10.1038/s41598-018-23948-5

525 Tanaka, E., Yasukawa, K., Nakamura, K., Ohta, J., Miyazaki, T., Vaglarov, B. S., Machida, S.,
526 Fujinaga, K., Iwamori, H., & Kato, Y. (2022). Secular variations in provenance of sedimentary
527 components in the western North Pacific Ocean constrained by Sr isotopic features of deep-sea
528 sediments. *Geochemistry, Geophysics, Geosystems*, *23*, e2021GC009729. https://doi.
529 org/10.1029/2021GC009729

Tetard, M., Marchant, R., Cortese, G., Gally, Y., de Garidel-Thoron, T., & Beaufort, L. (2020). Technical note: A new automated radiolarian image acquisition, stacking, processing, segmentation and identification workflow. *Climate of the Past*, *16*(6), 2415–2429. https://doi.org/10.5194/cp-16-2415-2020

Thomas, D. J., Korty, J., R., Huber, M., Schubert J. A., & Haines, B. (2014). Nd isotopic structure of the Pacific Ocean 70–30 Ma and numerical evidence for vigorous ocean circulation and ocean heat transport in a greenhouse world. *Paleoceanography*, *29*(5), 454–469. https://doi.org/10.1002/2013PA002535

Wang, C-Y., Bochkovskiy, A., & Liao, H-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696. https://doi.org/10.48550/arXiv.2207.02696

Yasukawa, K., Liu, H., Fujinaga, K., Machida, S., Haraguchi, S., Ishii, T., Nakamura, K., & Kato, Y. (2014). Geochemistry and mineralogy of REY-rich mud in the eastern Indian Ocean. *Journal of Asian Earth Science*, *93*, 25–36. https://doi.org/10.1016/j.jseaes.2014.07.005

Yasukawa, K., Tanaka, E., Miyazaki, T., Vaglarov, B. S., Chang, Q., Nakamura, K., Ohta, J., Fujinaga, K., Iwamori, H., & Kato, Y. (2023). High-Dimensional Chemostratigraphy of Pelagic Clay in the Western North Pacific Ocean Revealed via an Unsupervised Clustering Approach. *Paleoceanography and Paleoclimatology*, *38*(7), e2023PA004644. https://doi.org/10.1029/2023PA004644

549 Zhou, L., & Kyte, F. T. (1992). Sedimentation history of the South Pacific pelagic clay province
550 over the last 85 million years inferred from the geochemistry of Deep Sea Drilling Project Hole
551 596. *Paleoceanography*, *7*(4), 441–465. https://doi.org/10.1029/92PA01063