

実体顕微鏡画像解析による底生有孔虫種の同定

Species identification of benthic foraminifer by the microscopic image processing

加藤佑奈¹, 阿部 理², 横井大作³, 菊井重和¹, 森本真紀¹

Yuna KATO¹, Osamu ABE², Daisaku YOKOI³, Shigekazu KIKUI¹, Maki MORIMOTO¹

[キーワード Keyword] 底生有孔虫, 実体顕微鏡, 画像解析, 客観的同定, 深層学習

[所属 Institution] ¹岐阜大学教育学部 (Faculty of Education, Gifu University), ²名古屋大学大学院環境学研究所 (Graduate School of Environmental Studies, Nagoya University), ³愛知県立熱田高等学校 (Aichi Prefectural High School of Atsuta)

[要 旨 Abstract]

海底環境の生物指標として用いられる底生有孔虫の種同定は、高度な専門教育と訓練期間が必要不可欠と考えられ、さらにその結果は作業者の能力に大きく依存する。そこで、本研究では実体顕微鏡画像を効率よく取得する撮影システムおよび画像解析アルゴリズムの構築を行い、伊勢湾・三河湾の5地点から採取された底生有孔虫試料群の客観的同定を試みた。その結果、総種数の少ない浮遊性有孔虫による先行研究と同程度のスコアを得ることができた。

1. はじめに

原生動物の一群である有孔虫のうち、海底に生息する底生有孔虫は汽水～海水、熱帯～極域、岩礁～泥底とあらゆる海洋環境に適応し、無酸素下でも生息可能な種を有している。個体群密度が海底面積100 cm²あたり10²～10⁴個と高いことから、比較的少量の試料で統計的に有意な解析が可能であり、さらに多くの種は数週間～数か月の寿命であることから、個体群密度や群集組成が水質・底質の変化に速やかに応答する特徴を持っている。そのため、海底環境変化のモニタリングに用いられるほか、遺骸が堆積物中に長期間保存されることから過去の環境復元のための指標としても利用されている(横井ほか, 2011; 菊井ほか, 2021; 菊井, 2021)。

底生有孔虫を環境モニターとして利用する研究サイトとして、後背地に大都市圏を有する都市型内湾が多くあげられる。都市側内湾は窒素やリン等の栄養塩の流入負荷による貧酸素底層水塊の発生や、河川・沿岸の改修工事等による底質・堆積速度の変化など、短期間で大きな海底環境の変化をもたらす要因が多いため、大きな環境適応力と環境応答速度をもつ底生有孔虫の特徴を最大に応用することができる。底生有孔虫の生態学的研究は、底砂や底泥の化学分析に基づく地球化学的研究の多くと異なり高額な分析装置を必要としないため、理科教育の観点からも優れた環境モニタリング手法といえる。

その一方、底生有孔虫の個体数や群集組成の基盤データとなる、種の同定作業に関しては改善を考慮すべき課題が多い。一般的に用いられる実体顕微鏡下での拾い出し、同定および計数までほぼすべて人力による作業に依存している上、拾い出しには一定の訓練期間が必要とされる。さらに、同定については、有孔虫研究に関する伝統的な研究室を出た、熟練した科学者による監修が不可欠とされ、学生や分野外の若手研究者などの新規参入者にとって容易に開始するのは困難であり、その結果、国内外を問わず、他分野に比べて若手研究者の数は決して多くはない。幸運にも専門家からの訓練や監修を受け、研究に着手できるようになった場合でも、作業自体は反復的・機械的であり、そこに長い時間を費やすよりはむしろ自動処理を目指すべきである。

有孔虫種の客観的同定は、最初にRanaweena *et al.* (2009) による半自動化から始まった。彼らは特定の有孔虫群集に対して同定作業用のテンプレートを自動作成することにより作業負荷の低減を行った。その後、畳み込みニューラルネットワーク (CNN) を用いた深層学習による客観的同定がいくつか開発されており、浮遊性種を対象とした研究 (Zhong *et al.*, 2017; Hsiang *et al.*, 2019; Mitra *et al.*, 2019) が先行し、底生種を含めた研究が続いている (Marchant *et al.*; 2020; Johansen *et al.*; 2021)。深層学習は人工知能による機械学習の一つで

ある。機械学習は、計算機が「学び」を重ねることにより、「違い」を自動で認識することであり、深層学習はその学び方（ルール）を指定することなく、訓練を重ねることにより違いを人道認識する方法である。CNNは深層学習のうち最も成功している方法のひとつであり、特に画像解析に用いられる。画像の認識率を高めるために、訓練に大量の画像データベースが必要であることから、上述の研究はいずれも、公開されているラベル付き画像データベースを用いた事前学習済みモデルを用いて（転移学習法）おおよそ80%程度のスコアが得られている。このうちMitra *et al.* (2019) では、いくつかの学習段階にある被験者（専門家、博士課程大学院生、学部学生など）との浮遊性有孔虫種の同定成績の比較を行い、専門家に次ぐ認識率があると報告している。

約50種しか存在しない浮遊性有孔虫に対して、底生有孔虫は現生種で約10,000種あり、海域・海底環境・緯度・水深によってその群集組成は大きく異なる。例えば貧酸素底層水塊が発生するような閉鎖的な内湾環境などにおける貧酸素度の代表的な指標であるA-E indexを構成する*Ammonia beccarii*と*Elphidium excavatum*は現在の伊勢湾・三河湾においてはほとんど産出しない。きわめて多様な底生有孔虫種の客観的同定を確立するためには様々な海洋環境におけるケーススタディが必要である。

そこで本研究では、わが国の主要な都市型内湾で、夏季に貧酸素底層水塊が広域に拡大する伊勢湾・三河湾域の現生底生有孔虫種の実体顕微鏡画像を取得し、深層学習による客観的同定を試みる。

2. 試料・解析方法

画像を取得する底生有孔虫試料は2011年11月に三重大学生物資源科学部練習船勢水丸のSE1123航海により、伊勢湾2地点、三河湾3地点から採取した（図1）。このうち4地点の底生有孔虫種の拾い出し、同定、群集組成解析は菊井(2021)により実施され、本研究では新たに三河湾1地点（AS-8）を追加し、菊井(2021)と同じ方法で底泥試料の前処理を行った。その後、小分画法を用いて321個体を拾い出し、実体顕微鏡下で種同定を行った。

解析用画像の取得には三眼式実体顕微鏡（Nikon、SMZ745T）を用いた。Ranaweera *et al.* (2009) によれば有孔虫種固有の幾何学的特徴を強調し、種同定の正解率を上げるためには光源の方向が重要となる。Mitra *et al.* (2019) は、顕微鏡にLEDリング照明装置を装着し、光源角度を変化させて一個体について16画像を取得した。本研究でも同様に、実体顕微鏡に4方向独立照射型LEDリング照明装置（アームシステム、LED-R72）を装着し、一個体につき8つの異なる光源で画像を取得した。群集用ホールスライド（60マス）に並べられた有孔虫試料をX-Yステージ（ケニス、XY01）を用いて対象個体が視野中央に配置されるように調節し、直筒部に装着した500万画素の顕微鏡用USBカメラ（SWIFT、EP5R）を用いて画像を取得した（図2）。多数試料から効率よく画像を取得するため、LEDリングとカメラをシングルボードコンピューターRaspberry Pi 400に接続され、Pythonプログラムによって自動制御した。

画像解析による客観的同定は、Mitra *et al.* (2019)の方法を参考にした。彼らは浮遊性有孔虫種の同定をCNN-深層学習で行うため、公開されている大規模画像分類データセット（ILSVRC; Russakovsky *et al.*, 2015）を利用した事前学習済みモデルであるVgg16とReNet50を利用して独自に取得した有孔虫画像から特徴を抽出し同定を行った。本研究ではVgg16のみを用いた。CNNはKerasをフレームワーク、TensorflowをバックエンドエンジンとしたPythonプログラムを用いた。計算にはAMD Ryzen 9 CPU、32GB RAM、NVIDIA GeForce GTX1650 GPU (RAM 4GB)を搭載したWindows 11 PCを用いた。

同定結果の成績評価には、次の三種の指標を用いた。

- 1) 正解率 (accuracy) = (正解した試料数) / (全試料数)、
- 2) 種ごと適合率 (precision) = (正解数) / (正解数 + 間違っって予測した試料数)、
- 3) 種ごと再現率 (recall) = (正解数) / (正解数 + 間違っって予測しなかった試料数)。

誤検知が増加すると適合率は下がり、見逃しが増えると再現率は低下する。正解率のみでは評価できない種ごとの誤検知と見逃しを把握するため、これらの指数を用いる。また適合率と再現率から以下のF値（スコア）を求め、種の違いによるバランスのよい総合的な成績評価を行った。

$$F\text{値}(\text{適合率と再現率の調和平均}) = 2 \times \text{適合率} \times \text{再現率} / (\text{適合率} + \text{再現率})$$

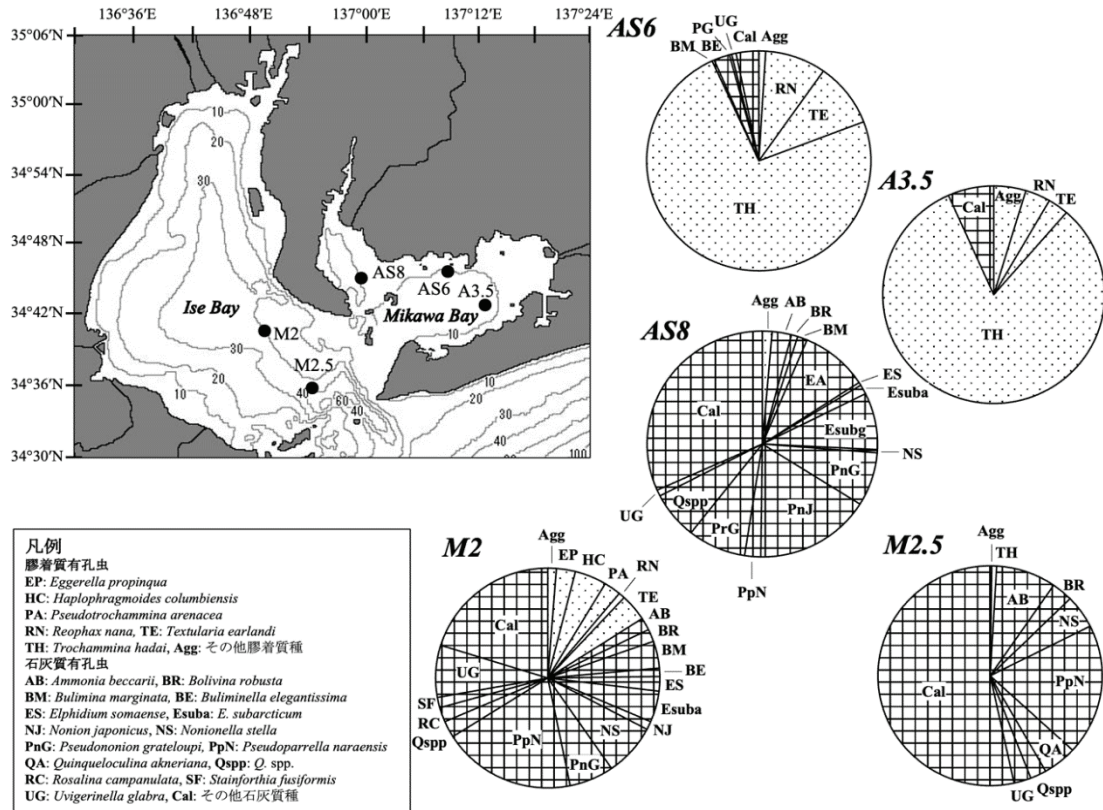


図1. (左上) 伊勢湾・三河湾の海底地形と試料採取地点(黒丸)。等深線は水深を表しており単位はmである。(右側) 各地点の全個体から得られた群集組成。円グラフのドットで示した領域は膠着質の底生有孔虫種、格子で示した領域は石灰質の底生有孔虫であることを意味する。グラフ中の記号は各有孔虫種であり、左下の凡例に学名を記す。ただし、CalおよびAggはそれぞれ相対頻度が3%以下のその他の石灰質種と膠着質種の総和である。

3. 結果と考察

伊勢湾・三河湾の底生有孔虫群集組成の水平分布

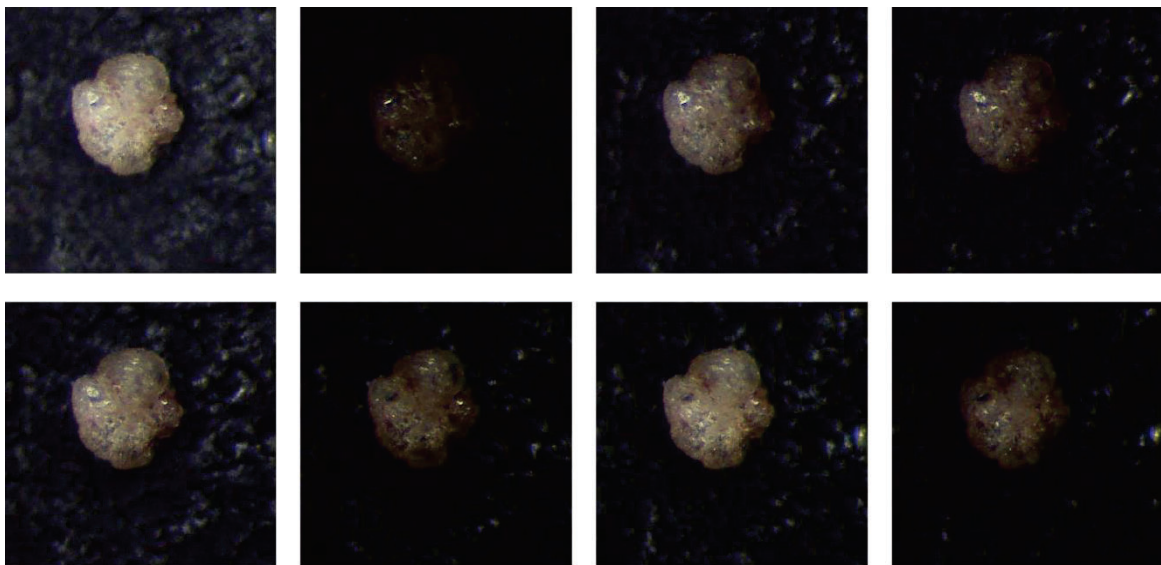
図1に本研究で取り扱った5地点の位置およびその群集組成をまとめた。先述したように、AS8以外については、菊井(2021)と同一である。円グラフから明らかのように、三河湾2地点(AS6、A3.5)と伊勢湾2地点および三河湾1地点(AS8、M2、M2.5)の2グループに大きく分けることができる。前者は*Trochammina hadai*が大きく卓越する低多様性群集でその他を含めて90%近くが膠着質殻から構成されている。後者は石灰質殻がほとんどを占め、高い多様性を持つ群集となっている。また、本研究で調べたAS8地点は、菊井(2021)や横井ほか(2011)で調査した三河湾の典型的な群集構造、すなわち*T. hadai*が大きく卓越する低多様性群集とは明確に異なることが明らかとなった。

取得した実体顕微鏡画像

前項で示した各地点の底生有孔虫群集について、M2地点全358検体、M2.5地点全478検体、AS6地点全332検体、AS8地点全321検体、A3.5地点69検体(一部のみ)の撮影を実施した。その他、人間による種同定作業のために用意した、参照用スライドの底生有孔虫81検体を加えた計1603検体について、各検体8枚ずつの画像を取得したため、総画像数は約12,800枚となった。図3に取得した1検体の8画像の例を示す。本研究では、これらの画像のうち、産出数の多い*T. hadai*、*Cibicides lobatulus*、*Textularia earlandi*、*Uvigerinella glabra*、*Pseudononion japonicum*の5種の判定を行い、それ以外の種については“その他”に含め、上記5種の判定の成績評価に利用した。



図2. 実体顕微鏡画像取得システム。左手前の顕微鏡直筒部には500万画素のUSBカメラ、対物レンズ付近に光源角度変化用のLEDリング照明が装着されている。顕微鏡奥の光源調節部とUSBカメラは、中央付近のキーボード型シングルボードコンピュータ（Raspberry Pi 400B）にて制御される。



Trochammina hadai（SE1123航海、AS6地点、試料番号AS6-1-14）

図3. 図2のシステムにより一検体から取得された画像の例。光源の方向と照度を変えることにより、有孔虫の幾何学的特徴を強調するとともに、訓練数を“水増し”させモデルを強化する目的も兼ねている。

客観的同定の結果および成績評価

図4にCNN-深層学習による同定結果の例を示す。前項で得た5種の画像についてはその総数の40%を訓練用、10%を出力チェック用、50%を試験用にランダムに分類し、総数のおよそ30%程度になるようにランダムに選択したその他の画像を加え、訓練および出力チェック画像はモデル微調整プログラム、試験画像を同定プログラムに読み込ませそれぞれ計算を実行した。結果は図のように予測した種名（黒字）がラベルされた画像の配列として出力させた。本研究では手入力で正解ラベル（青字）と誤答ラベル（赤字）を加筆したが、性能評価目的であればこの正誤の入力も自動化させることは可能である（本番ではもちろん使えない）。この例では25の検体数に対して誤答7、正解18となり、正解率は72%となった。*Trochammina hadai*についてはすべて正解し、*Textularia earlandi*は2つとも不正解であった。*Cibicides lobatulus*は3検体配置されていて予測は2検体でそのうち正解は1検体のみとなり、適合率は50%、再現率33%となった。

全375試料の成績を表にまとめた。その他の種を含めた全375試料による正解率は76%となり、先行研究と同程度の数値が得られた。全試料の適合率・再現率・スコアは、5種の数値も重み付け平均値であり、スコアは約80%となり、同じく先行研究に匹敵する数値となった。種ごとの適合率・再現率については、*T. hadai*がいずれも90%を超え、スコアも95%となり、ほぼ客観的同定可能とみなせる水準にあることがわかった。*C. lobatulus*はスコア約70%であったが、適合率が再現率を大きく上回っており、本種を他種と見誤った例がより多かったことがわかる。ただ、これらの2種については、実用レベルと考えられる。

一方、試料の少ない*Pseudonion japonicum*、*T. earlandi*および*Uvigerinella glabra*はスコアが50%前後となり、実用可能とは言い難い結果となった。個別には*P. japonicum*と*T. earlandi*については適合率・再現率いずれも同程度であった一方、*U. glabra*は再現率が大きく上回り、他種を本種と見誤った例がより多かった。本種の結果を調べてみると、*T. earlandi*を*U. glabra*と誤答した例がいくつかあり、画像取得の方法を含め今後改善すべき点である。

また、試料数とスコアが正の関係、すなわち試料数が増加するとスコアが向上する傾向が見られていることから、今後適正な試料数の下限の評価を行う必要がある。

4. まとめ

- 1) 三河湾AS8地点における底生有孔虫の群集組成を調べたところ、三河湾の他地点と大きく異なり、石灰質殻種にほとんど占められた高多様性群集であることがわかった。またこの特徴は伊勢湾口部と類似していた。
- 2) 底生有孔虫種の客観的同定のための、実体顕微鏡画像取得システムを作成した。本システムは光源方向を変化させながら検体ごとに8画像を自動で取得するため、USBカメラと光源をRaspberry Pi上のPythonプログラムで制御している。
- 3) 事前学習済みモデルを用いた転移深層学習による底生有孔虫種の同定アルゴリズムを構築し、伊勢湾・三河湾5地点の底生有孔虫の実体顕微鏡画像による客観的同定を行なった。その結果、全体の正解率とスコアがおおよそ80%となり、先行研究と同程度の数値が得られた。解析した試料がもっとも多い*Trochammina hadai*では95%のスコアが得られた一方、もっとも少ない*Uvigerinella glabra*では46%となった。スコアと試料数に正の関係が認められ、試料数の増加と適正な下限数評価が今後の課題である。

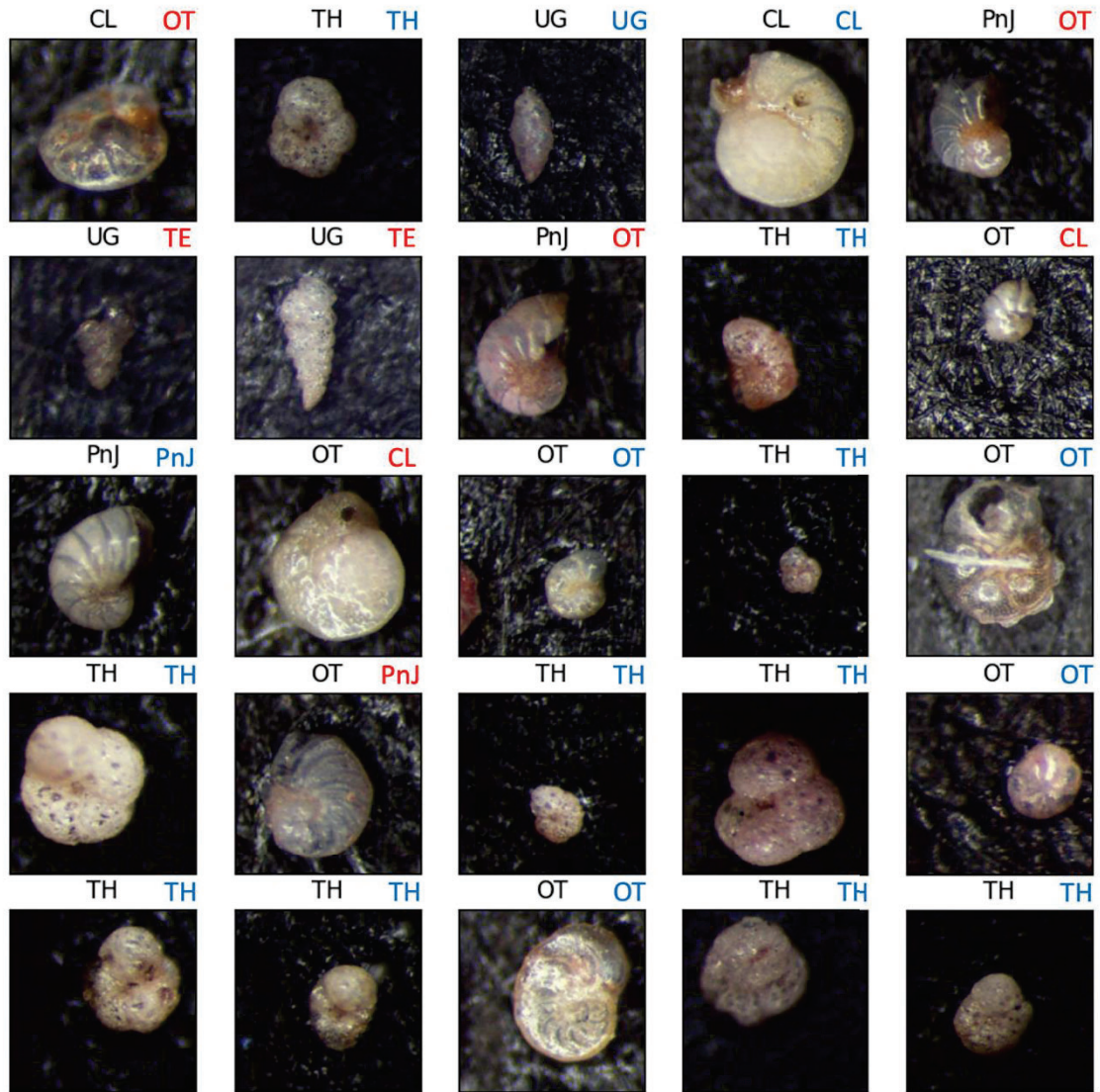


図4. ランダムサンプリングした底生有孔虫画像の同定結果の例。画像上の中央部の黒字は深層学習によって予測された種名で、右端は専門家が同定した種名であり、青字が正解、赤字が不正解を意味する。

表. 同定結果の成績。その他を含めた全375試料の適合率と再現率は、5種の数値の重み付き平均によって求めた。

	試料数	正解数	命名数	命名しなかった数	正解率	適合率	再現率	F1値
全体	375	284		91	0.76	0.80	0.79	0.79
<i>Trochammina hadai</i>	156	145	148	11		0.98	0.93	0.95
<i>Cibicides lobatulus</i>	45	27	35	18		0.77	0.60	0.68
<i>Pseudononion japonicum</i>	28	19	32	9		0.59	0.68	0.63
<i>Textularia earlandi</i>	28	14	30	14		0.47	0.50	0.48
<i>Uvigerinella glabra</i>	14	8	21	6		0.38	0.57	0.46

#画像数は試料数x8

5. 引用文献

- Hsiang, A. Y., Brombacher, A., Rillo, M. C., Mleneck-Vautravers, M. J., Conn, S., Lordsmith, S., Jentzen, A., Henehan, M. J., Metcalfe, B., Fenton, I. S., Wade, B. S., Fox, L., Meilland, J., Davis, C. V., Baranowski, U., Groeneveld, J., Edgar, K. M., Movellan, A., Aze, T., Dowsett, H. J., Miller, C. G., Rios, N., and Hull, P. M. (2019) Endless Forams: >34000 Modern Planktonic Foraminiferal Images for Taxonomic Training and Automated Species Recognition Using Convolutional Neural Networks. *Paleoceanogr. and Paleoclimatol.*, 34, 1157–1177, <https://doi.org/10.1029/2019pa003612>, 2019.
- Johansen, T.H., Sørensen, S.A., Møllersen, K., Godtlielsen, F. (2021) Instance Segmentation of Microscopic Foraminifera. *Appl. Sci.*, 11, 6543, <https://doi.org/10.3390/app11146543>.
- 菊井重和(2021) 伊勢湾・三河湾における底生有孔虫の水平分布, 岐阜大学教育学部卒業論文, 70pp.
- 菊井重和・阿部理・横井大作・森本真紀(2021) 底生有孔虫群集の統計評価法の検討. 岐阜大学教育学部研究報告=自然科学=, 45, 49-55.
- Marchant, R., Tetard, M., Pratiwi, A., Adebayo, M., de Garidel-Thoron, T. (2020) Automated analysis of foraminifera fossil records by image classification using a convolutional neural network. *J. Micropalaeontol.*, 39, 183–202, <https://doi.org/10.5194/jm-39-183-2020>.
- Mitra, R., Marchitto, T. M., Ge, Q., Zhong, B., Kanakiya, B., Cook, M. S., Fehrenbacher, J. S., Ortiz, J. D., Tripathi, A., and Lobaton, E. (2019) Automated species-level identification of planktic foraminifera using convolutional neural networks, with comparison to human performance. *Mar. Micropaleontol.*, 147, 16–24, <https://doi.org/10.1016/j.marmicro.2019.01.005>.
- Ranaweera, K., Harrison, A.P., Bains, S., Joseph, D. (2009) Feasibility of computer-aided identification of foraminiferal tests. *Mar. Micropaleontol.* 72, 66–75.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L. (2015) Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115, 211–252.
- 横井大作・阿部理・井龍康文(2011) 伊勢湾・三河湾における現生底生有孔虫群集の分布. 化石, 89, 15-26.
- Zhong, B., Ge, Q., Kanakiya, B., Marchitto, Mitra, R., Lobaton, E. (2017) A comparative study of image classification algorithms for Foraminifera identification. *2017 IEEE Symposium Series on Computational Intelligence*, 8 pp., DOI: <https://doi.org/10.1109/SSCI.2017.8285164>.

