Peer Community In Paleontology

paleo.peercommunityin.org

Open Access

RECOMMENDATION

Performance of machine-learning approaches in identifying ammonoid species based on conch properties

Kenneth De Baets[1]

[1] GeoZentrum Nordbayern, Friedrich-Alexander-Universität Erlangen-Nürnberg – Erlangen, Germany

Cite as: De Baets K (2021). Performance of machine-learning approaches in identifying ammonoid species based on conch properties. *Peer Community in Paleontology*, 100010. DOI: 10.24072/pci.paleo.100010

Published: 1st October 2021

Based on reviews by:
Jérémie Bardin and an anonymous reviewer

Correspondence:
kenneth.debaets@fau.de

A recommendation of

Foxon F (2021). Ammonoid taxonomy with supervised and unsupervised machine learning algorithms. *PaleorXiv* ewkx9, ver. 3, peer-reviewed by PCI Paleo. DOI: 10.31233/osf.io/ewkx9

There are less and less experts on taxonomy of particular groups particularly among early career paleontologists and (paleo)biologists – this also includes ammonoid cephalopods. Techniques cannot replace this taxonomic expertise (Engel et al., 2021) but machine learning approaches can make taxonomy more efficient, reproducible as well as passing it over more sustainable. Initially ammonoid taxonomy was a black box with small differences sometimes sufficient to erect different species as well as really idiosyncratic groupings of superficially similar specimens (see De Baets et al., 2015 for a review). In the meantime, scientists have embraced more quantitative assessments of conch shape and morphology more generally (see Klug et al., 2015 for a more recent review). The approaches still rely on important but time-intensive collection work and seeing through daisy chains of more or less accessible papers and monographs without really knowing how these approaches perform (other than expert opinion). In addition, younger scientists are usually trained by more experienced scientists, but this practice is becoming more and more difficult which makes it difficult to resolve the taxonomic gap. This relates to the fact that less and less experienced researchers with this kind of expertise get employed as well as graduate students or postdocs choosing different research or job avenues after their initial training effectively leading to a leaky pipeline and taxonomic impediment.

Robust taxonomy and stratigraphy is the basis for all other studies we do as paleontologists/paleobiologists so Foxon (2021) represents the first step to use supervised and unsupervised machine-learning approaches and test their efficiency on ammonoid conch properties. This pilot study demonstrates that machine learning approaches can be reasonably accurate (60-70%) in identifying ammonoid species (Foxon, 2021) – at least similar to that in other mollusk taxa (e.g., Klinkenbuß et al., 2020) - and might also be interesting to assist in cases where more traditional methods are not feasible. Novel approaches might even allow to further approve the accuracy as has been demonstrated for other research objects like pollen (Romero et al., 2020). Further applying of machine learning approaches on larger datasets and

additional morphological features (e.g., suture line) are now necessary in order to test and improve the robustness of these approaches for ammonoids as well as test their performance more broadly within paleontology.

References

De Baets K, Bert D, Hoffmann R, Monnet C, Yacobucci M, and Klug C (2015). Ammonoid intraspecific variability. In: *Ammonoid Paleobiology: From anatomy to ecology*. Ed. by Klug C, Korn D, De Baets K, Kruta I, and Mapes R. Vol. 43. Topics in Geobiology. Dordrecht: Springer, pp. 359–426. isbn: 978-94-017-9629-3.

Engel MS, Ceríaco LMP, Daniel GM, Dellapé PM, Löbl I, Marinov M, Reis RE, Young MT, Dubois A, Agarwal I, Lehmann A. P, Alvarado M, Alvarez N, Andreone F, Araujo-Vieira K, Ascher JS, Baêta D, Baldo D, Bandeira SA, Barden P, Barrasso DA, Bendifallah L, Bockmann FA, Böhme W, Borkent A, Brandão CRF, Busack SD, Bybee SM, Channing A, Chatzimanolis S, Christenhusz MJM, Crisci JV, D'elía G, Da Costa LM, Davis SR, De Lucena CAS, Deuve T, Fernandes Elizalde S, Faivovich J, Farooq H, Ferguson AW, Gippoliti S, Gonçalves FMP, Gonzalez VH, Greenbaum E, Hinojosa-Díaz IA, Ineich I, Jiang J, Kahono S, Kury AB, Lucinda PHF, Lynch JD, Malécot V, Marques MP, Marris JWM, Mckellar RC, Mendes LF, Nihei SS, Nishikawa K, Ohler A, Orrico VGD, Ota H, Paiva J, Parrinha D, Pauwels OSG, Pereyra MO, Pestana LB, Pinheiro PDP, Prendini L, Prokop J, Rasmussen C, Rödel MO, Rodrigues MT, Rodríguez SM, Salatnaya H, Sampaio Í, Sánchez-García A, Shebl MA, Santos BS, Solórzano-Kraemer MM, Sousa ACA, Stoev P, Teta P, Trape JF, Dos Santos CVD, Vasudevan K, Vink CJ, Vogel G, Wagner P, Wappler T, Ware JL, Wedmann S, and Zacharie CK (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society* 193, 381–387. doi: 10.1093/zoolinnean/zlab072.

Foxon F (2021). Ammonoid taxonomy with supervised and unsupervised machine learning algorithms. *PaleorXiv* ewkx9, ver. 3, peer-reviewed by PCI Paleo. doi: 10.31233/osf.io/ewkx9.

Klinkenbuß D, Metz O, Reichert J, Hauffe T, Neubauer TA, Wesselingh FP, and Wilke T (2020). Performance of 3D morphological methods in the machine learning assisted classification of closely related fossil bivalve species of the genus *Dreissena*. *Malacologia* 63, 95. doi: 10.4002/040.063.0109.

Klug C, Korn D, Landman NH, Tanabe K, De Baets K, and Naglik C (2015). Ammonoid conchs. In: *Ammonoid Paleobiology: From anatomy to ecology*. Ed. by Klug C, Korn D, De Baets K, Kruta I, and Mapes RH. Vol. 43. Dordrecht: Springer, pp. 3–24. isbn: 978-94-017-9630-9.

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

Appendix

Reviews by Jérémie Bardin and an anonymous reviewer, DOI: 10.24072/pci.paleo.100010.