LifeCLEF 2024 Teaser: Challenges on Species Distribution Prediction and Identification

Chapter · March 2024

DOI: 10.1007/978-3-031-56072-9_3

24 authors, including:

Alexis Joly
National Institute for Research in Computer Science and Control
196 PUBLICATIONS **4,677** CITATIONS

Lukas Picek
University of West Bohemia
35 PUBLICATIONS **227** CITATIONS

Stefan Kahl
Technische Universität Chemnitz
49 PUBLICATIONS **789** CITATIONS

Hervé Goëau
Cirad - La recherche agronomique pour le développement
110 PUBLICATIONS **2,830** CITATIONS

LifeCLEF 2024 Teaser: Challenges on Species Distribution Prediction and Identification

Alexis Joly[1], Lukáš Picek[8], Stefan Kahl[6,11], Hervé Goëau[2], Vincent Espitalier[2], Christophe Botella[1], Benjamin Deneu[1], Diego Marcos[1], Joaquim Estopinan[1], Cesar Leblanc[1], Théo Larcher[1], Milan Šulc[10], Marek Hrúz[8], Maximilien Servajean[7], Jiří Matas[12], Hervé Glotin[3], Robert Planqué[4], Willem-Pier Vellinga[4], Holger Klinck[6], Tom Denton[9], Andrew M. Durso[13], Ivan Eggel[5], Pierre Bonnet[2], Henning Müller[5]

[1] Inria, LIRMM, Univ Montpellier, CNRS, Montpellier, France
[2] CIRAD, UMR AMAP, Montpellier, Occitanie, France
[3] Univ. Toulon, Aix Marseille Univ., CNRS, LIS, DYNI team, Marseille, France
[4] Xeno-canto Foundation, The Netherlands
[5] Informatics Institute, HES-SO Valais , Sierre, Switzerland
[6] K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, USA
[7] LIRMM, AMIS, Univ Paul Valéry Montpellier, Univ Montpellier, CNRS, France
[8] Department of Cybernetics, FAV, University of West Bohemia, Czechia
[9] Google Research, San Francisco, USA
[10] Second Foundation, Prague, Czech Republic
[11] Chemnitz University of Technology, Chemnitz, Germany
[12] Czech Technical University, Prague, Czechia
[13] Department of Biological Sciences, Florida Gulf Coast University, USA

Abstract. Building accurate knowledge of the identity, the geographic distribution and the evolution of species is essential for the sustainable development of humanity, as well as for biodiversity conservation. However, species identification and inventory is a difficult and costly task, requiring large-scale automated approaches. The LifeCLEF lab has been promoting and evaluating advances in this domain since 2011 through the organization of multi-year challenges. The 2024 edition presented in this article proposes five data-driven challenges as a continuation of this effort: (i) BirdCLEF: bird species recognition in audio soundscapes, (ii)FungiCLEF: fungi recognition beyond 0-1 cost, (iii) GeoLifeCLEF: remote sensing based prediction of species, (iv) PlantCLEF: Multi-species identification in vegetation plot images, and (v) SnakeCLEF: snake recognition in medically important scenarios.

Keywords: biodiversity · machine learning · AI · species · identification · prediction · species distribution model

1 Introduction

Accurately identifying and inventorying species is a difficult task requiring high levels of expertise and costly efforts. Since 1992, this problem has been recognized as one of the major obstacles to the global implementation of the Convention on Biological Diversity [1]. Automated approaches have thus been recognized as one of the most promising approaches since 2004 [8]. Since then, automated species identification has progressed a lot, in particular due to recent advances in deep learning [3,5,9,10,21,23,29,30,31,32]. However, even the best models remain uncertain because of the strong ambiguities between species and the scarcity of data for most species [7,22]. The LifeCLEF lab has been evaluating advances in this domain since 2014 and publishes an annual synthesis of the best methods and their performance [11,12,13,14,15,16,17,18,19,20]. Building on this effort, LifeCLEF 2024 consists of five challenges (PlantCLEF, BirdCLEF, GeoLifeCLEF, SnakeCLEF, FungiCLEF), which we briefly introduce in this paper.

2 PlantCLEF 2024 Challenge: Multi-species Plant Identification in Vegetation Plot Images

Motivation: Vegetation plot inventories are essential for ecological studies, enabling standardized sampling, biodiversity assessment, long-term monitoring and remote, large-scale surveys. They provide valuable data on ecosystems, biodiversity conservation, and evidence-based environmental decision-making. Plot images are typically one square meter in size, and botanists meticulously identify all the species found there. In addition, they quantify species abundance using indicators such as biomass, qualification factors, and areas occupied in photographs. The integration of AI could significantly improve specialists' efficiency, helping them extend the scope and coverage of ecological studies.

Data collection: The test set will be a compilation of several image datasets of plots in different floristic contexts, such as Pyrenean and Mediterranean floras, all produced by experts. The training set will be composed more conventionally of observations of individual plants, such as those used in previous editions of PlantCLEF. More precisely, it will be a subset of the PlantCLEF2023 data focused on Europe and covering 15k plant species. It will contain about 1 million images with trusted labels (aggregated from the GBIF platform) and as many images with potentially noisy labels aggregated through web scraping (based on Google and Bing search engines).

Task description: The main difficulty of the task lies in the shift between the test data (high-resolution multi-label images of vegetation plots) and the training data (single-label images of individual plants). The task will be evaluated as a multi-label classification task that aims to predict all the plant species on high-resolution plot images. The participants will first have access to the training set, and a few months later, they will be provided with the whole test set. Self-supervised, semi-supervised or unsupervised approaches will be strongly encouraged, and a starter package with pre-trained models will be provided. The used metric will be the mean Average Precision (computed columnwise).

3 FungiCLEF 2024 Challenge: Revisiting Fungi Recognition Beyond 1-0 Cost

Motivation: Automatic recognition of species at scale, such as in popular citizen science projects [27,28], requires efficient prediction on limited resources. In practice, species identification typically depends not solely on the visual observation of the specimen but also on other information available to the observer, e.g., habitat, substrate, location, and time. The challenge aims to provide a major benchmark for combining visual information with side information thanks to rich metadata, precise annotations, and baselines available to all competitors. Since mushrooms are often picked for consumption, the competition also considers different scenarios for misclassifying edible and poisonous mushrooms.

Data collection: This year's challenge directly follows up on FungiCLEF 2023 [25], keeping the same training dataset, i.e. the DanihFungi2020 [26]. The data originates from a citizen science project, the Atlas of Danish Fungi, where all samples went through an expert validation process, guaranteeing high-quality labels. Rich metadata (e.g., habitat, substrate, timestamp, location, EXIF, etc.) are provided for most samples. The training set includes 295,938 training images belonging to 1,604 species observed mostly in Denmark). Validation and test datasets will cover a similar number of fungi observations (collection of images and metadata), images, and species but will reflect different periods and cover all seasons. The validation set contains 60k observations with 120k images of 3k species: 1k known from the training set and 2k unknown species. The private test set from last year will be reused. Besides, we will include a new test set originating from the newly created mobile app – CheckFungi – which includes fungi observations mainly from the Czech Republic. Such a setup will allow direct comparison with last year and measure generalization to different locations.

Task description: The challenge aims to provide a major benchmark for combining visual observations with other observed information thanks to rich metadata, species-accurate annotations, and provided baselines. The goal of the task is to create a classification model that returns a ranked list of predicted species for a set of real fungi species observations and minimizes the danger to human life, i.e., the confusion between poisonous and edible species. The classification model will have to fit limits for memory footprint and a prediction time limit within a given Hugging Face server instance. The FungiCLEF 2024 challenge will use several metrics representing different decision scenarios, where the goal is to minimize the empirical loss L for decisions $q(x)$ over observations x and true labels y, given a cost function $W(y, q(x))$.

$$L = \sum_i W(k_i, q(x_i)) \tag{1}$$

Different recognition scenarios and their cost function $W(y, q(x))$ are described together with their motivation in last year's overview (e.g., [24]).

4 GeoLifeCLEF 2024 Challenge: Species Presence Prediction Based on Occurrences Data and High-resolution Remote Sensing Images

Motivation: Predicting species presence in an area is vital for ecology and biodiversity conservation. These predictions inform decisions about threatened species, land use planning, protected areas, and eco-friendly agriculture. However, species distribution is influenced by complex local factors that are hard to measure, such as population interactions, landscape connectivity, habitat history, and biases in data collection. Traditional ecological models struggle to account for these factors, leading to coarse-scale resolutions. Additionally, many species are rarely observed due to sampling biases. GeoLifeCLEF aims to evaluate models on an unprecedented scale, covering thousands of species, with a spatial resolution of about 10 meters, and utilizing millions of occurrence data.

Data collection: The 2023 edition of GeoLifeCLEF [6] revealed significant room for dataset improvement. It emphasized the need for more standardized data to enhance predictions in diverse contexts. Therefore, the 2024 edition will introduce new presence-absence data from different parts of Europe thanks to partners from the European Vegetation Archive (EVA) and the network of a large-scale European project on biodiversity monitoring (MAMBO, Horizon EU program). Test sites will be balanced better to represent the diversity of European habitats and regions, and the final test set is expected to comprise several tens of thousands of presence-absence data. Like in 2023, the training data will consist of 5 million non-standardized occurrences from GBIF spanning 38 European countries and over ten thousand plant species. Explanatory variables will include 2023's data with high-resolution remote sensing (e.g., Sentinel-2 RGB-NIR, multi-band Landsat time-series, ASTER elevation raster) and coarser resolution environmental data (e.g., Chelsa climate, SoilGrids, MODIS land use, human footprint, ISRIC soil salinity raster).

Task description: Given a test set of geolocation and year combinations (plot) and given the corresponding high-resolution remote sensing images and environmental covariates, the goal of the task will be to return for each plot the set of species that were inventoried at that location and time by botanical experts over a small area (about $100m^2$). The test set will include only locations for which an exhaustive plant species inventory is available (i.e., in the form or presence/absence data).

5 BirdCLEF 2024 Challenge: Bird Species Identification in Soundscape Recordings

Motivation: Recognizing bird sounds in complex soundscapes is an important sampling tool that often helps reduce the limitations of point counts. In the future, archives of recorded soundscapes will become increasingly valuable as the habitats in which they were recorded will be lost. In the past few years, deep learning approaches have transformed the field of automated soundscape analysis. Yet, when training data is sparse, detection systems struggle to recognize rare species. The goal of this competition is to establish training and test datasets that can serve as real-world applicable evaluation scenarios for endangered habitats and help the scientific community to advance their conservation efforts through automated bird sound recognition.

Data collection: We will build on the experience from previous editions and adjust the overall task to encourage participants to focus on few-shot learning and task-specific model designs. We will select training and test data to suit this demand. As in previous iterations, Xeno-canto will be the primary source for training data, and expertly annotated soundscape recordings will be used for testing. We will focus on bird species for which there is limited training data, but we will also include common species so that participants can train good recognition systems. In search of suitable test data, we will consider different data sources with varying complexity (call density, chorus, signal-to-noise ratio, anthropophony…), and quality (mono and stereo recordings). We also want to focus on specific real-world use cases (e.g., conservation efforts in India) and frame the competition based on the demand of the particular use case. Additionally, we are considering including unlabeled data to encourage self-supervised learning.

Task description: The challenge will be held on Kaggle and the evaluation mode will resemble the 2023 test mode (i.e., hidden test data, code competition). We will use established metrics like F1 score and cmAP, which reflect use cases for which precision is key and also allow organizers to assess system performance independent of fine-tuned confidence thresholds. Participants will be asked to return a list of species for short audio segments extracted from labeled soundscape data. In the past, we used 5-second segments, and we will consider increasing the duration of these context windows to reflect the overall ground truth label distribution better. However, the overall structure of the task will remain unchanged, as it provides a well-established base that has resulted in significant participation in past editions (e.g., 1,397 participants and 21,519 submissions in 2023). Again, we will strive to keep the dataset size reasonably small (<50 GB) and easy to process, and we will also provide introductory code repositories and write-ups to lower the entry-level of the competition.

6 SnakeCLEF 2024 Challenge: Revisiting Snake Species Identification in Medically Important Scenarios

Motivation: Developing a robust system for identifying species of snakes from photographs is an important goal in biodiversity and global health. With over half a million victims of death and disability from venomous snakebite annually, understanding the global distribution of the >4000 species of snakes and differentiating species from images (particularly images of low quality) will significantly improve epidemiology data and treatment outcomes. We learned from previous editions that machines can make accurate predictions (macro averaged F1 of around 90%, and accuracy of around 90%) even in scenarios with long-tailed distributions and 1800 species [4]. However, machines still perform poorly on data from neglected regions. Thus, testing over specific countries (primarily tropical and subtropical) and integrating the medical importance of species is the next step that should provide a more reliable machine prediction.

Data collection: The development dataset from the previous year [24] will be re-used. The dataset covers 1,784 snake species from around the world, with a minimum of three observations (i.e., multiple images of the same specimen) per species. Additionally, country-species and venomous-species mapping will be provided. The evaluation will be carried out on the same datasets as last year in order to allow direct comparison. Additionally, we will enrich the private test dataset with new data from additional neglected regions to allow testing generalization capabilities. The region won't be known to the participants. Since the competition test set is composed of private images with highly restricted licenses from individuals and natural history museums, the dataset will be undisclosed, and participants will not have access to this data.

Task description: The SnakeCLEF challenge aims to be a major benchmark for observation-based snake species identification. The goal of the task is to create a classification model that returns a ranked list of predicted species for each set of images and location (i.e., snake observation) and minimize the danger to human life and the waste of antivenom if a bite from the snake in the image were treated as coming from the top-ranked prediction. The classification model must fit limits for memory footprint and a prediction time limit within a given Hugging Face server instance. Like last year, the 2024 edition will extend the evaluation beyond the 0-1 loss (common in classification) to motivate research in recognition scenarios with uneven costs (e.g., mistaking a venomous snake for a harmless one and vice versa). We will use two custom metrics (further described in last year's overview [24]). The first metric includes the overall classification rate (macro averaged F1) and the venomous species confusion errors. The second metric is a sum of (L) over all test observations.

7 Conclusion

The joint efforts of the LifeCLEF evaluation campaign aim to accomplish several critical goals: (i) promoting non-incremental contribution from participants, (ii) accurately measuring consistent performance gaps, (iii) gradually scaling up the problems, and (iv) enabling the growth of an engaged community. The 2024 edition aligns with this vision, offering challenges enriched with new data, new machine learning tasks (such as the one on vegetation plots), and new evaluation methodologies (such as the code-based submissions through Hugging Face imposing time and memory limits). All information about the timeline and participation in the challenges is provided on the LifeCLEF 2024 web page [2].

Acknowledgements: This work has received funding from the European Union's Horizon research and innovation program under grant agreement No. 101060639 (MAMBO project).

References

1. Convention on Biodiversity. `https://www.cbd.int/`
2. LifeCLEF. `http://www.lifeclef.org/`
3. Banan, A., Nasiri, A., Taheri-Garavand, A.: Deep learning-based appearance features extraction for automated carp species identification. Aquacultural Engineering **89**, 102053 (2020)
4. Bolon, I., Picek, L., Durso, A.M., Alcoba, G., Chappuis, F., Ruiz de Castañeda, R.: An artificial intelligence model to identify snakes from across the world: Opportunities and challenges for global health and herpetology. PLOS Neglected Tropical Diseases **16**(8), e0010647 (2022)
5. Bonnet, P., Goëau, H., Hang, S.T., Lasseck, M., Šulc, M., Malécot, V., Jauzein, P., Melet, J.C., You, C., Joly, A.: Plant identification: experts vs. machines in the era of deep learning. In: Multimedia Tools and Applications for Environmental & Biodiversity Informatics, pp. 131–149. Springer (2018)
6. Botella, C., Deneu, B., Gonzalez, D.M., Servajean, M., Larcher, T., Leblanc, C., Estopinan, J., Bonnet, P., Joly, A.: Overview of geolifeclef 2023: Species composition prediction with high spatial resolution at continental scale using remote sensing. Working Notes of CLEF (2023)
7. Garcin, C., Joly, A., Bonnet, P., Lombardo, J.C., Affouard, A., Chouet, M., Servajean, M., Lorieul, T., Salmon, J.: Pl@ ntnet-300k: a plant image dataset with high label ambiguity and a long-tailed distribution. In: NeurIPS 2021-35th Conference on Neural Information Processing Systems (2021)
8. Gaston, K.J., O'Neill, M.A.: Automated species identification: why not? Philosophical Transactions of the Royal Society of London B: Biological Sciences **359**(1444), 655–667 (2004)
9. Ghazi, M.M., Yanikoglu, B., Aptoula, E.: Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing **235**, 228–235 (2017)
10. Goodwin, A., Padmanabhan, S., Hira, S., Glancey, M., Slinowsky, M., Immidisetti, R., Scavo, L., Brey, J., Sai Sudhakar, B.M.M., Ford, T., et al.: Mosquito species identification using convolutional neural networks with a multitiered ensemble model for novel species detection. Scientific reports **11**(1), 13656 (2021)
11. Joly, A., Botella, C., Picek, L., Kahl, S., Goëau, H., Deneu, B., Marcos, D., Estopinan, J., Leblanc, C., Larcher, T., et al.: Overview of lifeclef 2023: evaluation of ai models for the identification and prediction of birds, plants, snakes and fungi. In: International Conference of the Cross-Language Evaluation Forum for European Languages. pp. 416–439. Springer (2023)
12. Joly, A., Goëau, H., Botella, C., Glotin, H., Bonnet, P., Vellinga, W.P., Müller, H.: Overview of LifeCLEF 2018: a large-scale evaluation of species identification and recommendation algorithms in the era of ai. In: Jones, G.J., Lawless, S., Gonzalo, J., Kelly, L., Goeuriot, L., Mandl, T., Cappellato, L., Ferro, N. (eds.) CLEF: Cross-Language Evaluation Forum for European Languages. Experimental IR Meets Multilinguality, Multimodality, and Interaction, vol. LNCS. Springer, Avigon, France (Sep 2018)
13. Joly, A., Goëau, H., Botella, C., Kahl, S., Servajean, M., Glotin, H., Bonnet, P., Planqué, R., Stöter, F.R., Vellinga, W.P., Müller, H.: Overview of LifeCLEF 2019: Identification of Amazonian Plants, South & North American Birds, and Niche Prediction. In: Crestani, F., Brascher, M., Savoy, J., Rauber, A., Müller, H., Losada, D.E., Bürki, G.H., Bürki, G.H., Cappellato, L., Ferro, N. (eds.)

CLEF 2019 - Conference and Labs of the Evaluation Forum. Experimental IR Meets Multilinguality, Multimodality, and Interaction, vol. LNCS, pp. 387–401. Lugano, Switzerland (Sep 2019). https://doi.org/10.1007/978-3-030-28577-7_29, https://hal.umontpellier.fr/hal-02281455

14. Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.P., Champ, J., Planqué, R., Palazzo, S., Müller, H.: LifeCLEF 2016: Multimedia Life Species Identification Challenges. In: Fuhr, N., Quaresma, P., Gonçalves, T., Larsen, B., Balog, K., Macdonald, C., Cappellato, L., Ferro, N. (eds.) CLEF: Cross-Language Evaluation Forum. Experimental IR Meets Multilinguality, Multimodality, and Interaction, vol. LNCS, pp. 286–310. Springer, Évora, Portugal (Sep 2016). https://doi.org/10.1007/978-3-319-44564-9_26, https://hal.archives-ouvertes.fr/hal-01373781
15. Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.P., Lombardo, J.C., Planque, R., Palazzo, S., Müller, H.: LifeCLEF 2017 Lab Overview: Multimedia Species Identification Challenges. In: Jones, G.J., Lawless, S., Gonzalo, J., Kelly, L., Goeuriot, L., Mandl, T., Cappellato, L., Ferro, N. (eds.) CLEF: Cross-Language Evaluation Forum. Experimental IR Meets Multilinguality, Multimodality, and Interaction, vol. LNCS, pp. 255–274. Springer, Dublin, Ireland (Sep 2017). https://doi.org/10.1007/978-3-319-65813-1_24, https://hal.archives-ouvertes.fr/hal-01629191
16. Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.P., Planque, R., Rauber, A., Fisher, B., Müller, H.: LifeCLEF 2014: Multimedia Life Species Identification Challenges. In: CLEF: Cross-Language Evaluation Forum. Information Access Evaluation. Multilinguality, Multimodality, and Interaction, vol. LNCS, pp. 229–249. Springer International Publishing, Sheffield, United Kingdom (Sep 2014). https://doi.org/10.1007/978-3-319-11382-1_20, https://hal.inria.fr/hal-01075770
17. Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.P., Planqué, R., Rauber, A., Palazzo, S., Fisher, B., et al.: Lifeclef 2015: multimedia life species identification challenges. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction, pp. 462–483. Springer (2015)
18. Joly, A., Goëau, H., Kahl, S., Deneu, B., Servajean, M., Cole, E., Picek, L., De Castaneda, R.R., Bolon, I., Durso, A., et al.: Overview of lifeclef 2020: a system-oriented evaluation of automated species identification and species distribution prediction. In: International Conference of the Cross-Language Evaluation Forum for European Languages. pp. 342–363. Springer (2020)
19. Joly, A., Goëau, H., Kahl, S., Picek, L., Lorieul, T., Cole, E., Deneu, B., Servajean, M., Durso, A., Glotin, H., et al.: Overview of lifeclef 2022: an evaluation of machine-learning based species identification and species distribution prediction. In: International Conference of the Cross-Language Evaluation Forum for European Languages. pp. 257–285. Springer (2022)
20. Joly, A., Goëau, H., Kahl, S., Picek, L., Lorieul, T., Cole, E., Deneu, B., Servajean, M., Durso, A., Bolon, I., et al.: Overview of lifeclef 2021: an evaluation of machine-learning based species identification and species distribution prediction. In: International Conference of the Cross-Language Evaluation Forum for European Languages. pp. 371–393. Springer (2021)
21. Lee, S.H., Chan, C.S., Remagnino, P.: Multi-organ plant classification based on convolutional and recurrent neural networks. IEEE Transactions on Image Processing **27**(9), 4287–4301 (2018)

22. Lorieul, T.: Uncertainty in predictions of deep learning models for fine-grained classification. Ph.D. thesis, Université Montpellier (2020)
23. Norouzzadeh, M.S., Morris, D., Beery, S., Joshi, N., Jojic, N., Clune, J.: A deep active learning system for species identification and counting in camera trap images. Methods in Ecology and Evolution **12**(1), 150–161 (2021)
24. Picek, L., Chamidullin, R., Hruz, M., Durso, A.M.: Overview of fungiclef 2023: Fungi recognition beyond 1/0 cost. In: Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum. CEUR-WS (2023)
25. Picek, L., Šulc, M., Chamidullin, R., Matas, J.: Overview of fungiclef 2023: Fungi recognition beyond 1/0 cost. In: Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum. CEUR-WS (2023)
26. Picek, L., Sulc, M., Matas, J., Heilmann-Clausen, J., Jeppesen, T., Læssøe, T., Frøslev, T.: Danish fungi 2020 - not just another image recognition dataset. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (January 2022)
27. Picek, L., Šulc, M., Matas, J., Heilmann-Clausen, J., Jeppesen, T.S., Lind, E.: Automatic fungi recognition: Deep learning meets mycology. Sensors **22**(2), 633 (2022)
28. Sulc, M., Picek, L., Matas, J., Jeppesen, T., Heilmann-Clausen, J.: Fungi recognition: A practical use case. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 2316–2324 (2020)
29. Van Horn, G., Mac Aodha, O., Song, Y., Cui, Y., Sun, C., Shepard, A., Adam, H., Perona, P., Belongie, S.: The inaturalist species classification and detection dataset. CVPR (2018)
30. Villon, S., Mouillot, D., Chaumont, M., Subsol, G., Claverie, T., Villéger, S.: A new method to control error rates in automated species identification with deep learning algorithms. Scientific reports **10**(1), 1–13 (2020)
31. Wäldchen, J., Mäder, P.: Machine learning for image based species identification. Methods in Ecology and Evolution **9**(11), 2216–2225 (2018)
32. Wäldchen, J., Rzanny, M., Seeland, M., Mäder, P.: Automated plant species identification—trends and future directions. PLoS computational biology **14**(4), e1005993 (2018)