

scientific data



OPEN

DATA DESCRIPTOR

Machine learning training data: over 500,000 images of butterflies and moths (Lepidoptera) with species labels

Friederike Barkmann¹, Andreas Lindner^{2,3}, Ronald Würflinger⁴, Helmut Höttinger⁴ & Johannes Rüdiger¹✉

Deep learning models can accelerate the processing of image-based biodiversity data and provide educational value by giving direct feedback to citizen scientists. However, the training of such models requires large amounts of labelled data and not all species are equally suited for identification from images alone. Most butterfly and many moth species (Lepidoptera) which play an important role as biodiversity indicators are well-suited for such approaches. This dataset contains over 540.000 images of 185 butterfly and moth species that occur in Austria. Images were collected by citizen scientists with the application “Schmetterlinge Österreichs” and correct species identification was ensured by an experienced entomologist. The number of images per species ranges from one to nearly 30.000. Such a strong class imbalance is common in datasets of species records. The dataset is larger than other published dataset of butterfly and moth images and offers opportunities for the training and evaluation of machine learning models on the fine-grained classification task of species identification.

Background and Summary

Integration of machine learning and citizen science has a high potential to benefit both fields and can support an efficient and accurate collection of data on biodiversity^{1,2}. Developments in technology such as camera traps and voice recorders as well as engagement of the public in data collection has considerably increased the amount of available biodiversity data^{3,4}. Manual processing of the collected data such as species identification by experts is time-consuming, expensive and limited in the amount of data that can be processed. Machine learning approaches and especially deep learning have advanced considerably in the last years and can significantly reduce the time that is needed for image processing and classification^{5,6}. Deep learning algorithms were already applied to identify species from different taxonomic groups such as plants^{7,8}, insects^{9,10} and vertebrates^{11,12}. In addition, deep learning models can provide direct feedback to users of species identification applications and increase the educational value of such applications². They are already included in species identification apps such as iNaturalist or Flora Incognita and provide users with species classifications with high accuracy^{8,13}. While the application of deep learning algorithms has a high potential to accelerate the processing of biodiversity data, their implementation can be impeded by the large amount of data that is needed for their training^{4,14}. Datasets collected by citizen scientists can be a valuable resource for model training, especially when high data quality can be ensured through stringent quality assuring procedures¹.

Not all species groups are well suited for identification from images. While especially many insect species can be difficult to identify by macroscopic characteristics alone, butterflies can be recognized comparably easily. In addition, they react sensitively to changes in environmental conditions^{15,16}, inhabit a wide range of terrestrial habitats¹⁶ and are representative for many (but not all) groups of terrestrial taxa^{17–19}. They are therefore widely used as biodiversity indicators^{20–22}. Butterflies are perceived positively by the public²³ and are well-suited for observation in citizen science projects. Developing efficient species identification methods for this important indicator group can therefore benefit from the large amounts of data that is collected in such initiatives. Deep

¹Department of Ecology, University of Innsbruck, Innsbruck, Austria. ²Advanced Computing Austria ACA GmbH, Wien, Austria. ³Research Area Scientific Computing, University of Innsbruck, Innsbruck, Austria. ⁴Billa Foundation Blühendes Österreich, Wiener Neustadt, Austria. ✉e-mail: johannes.ruediger@uibk.ac.at

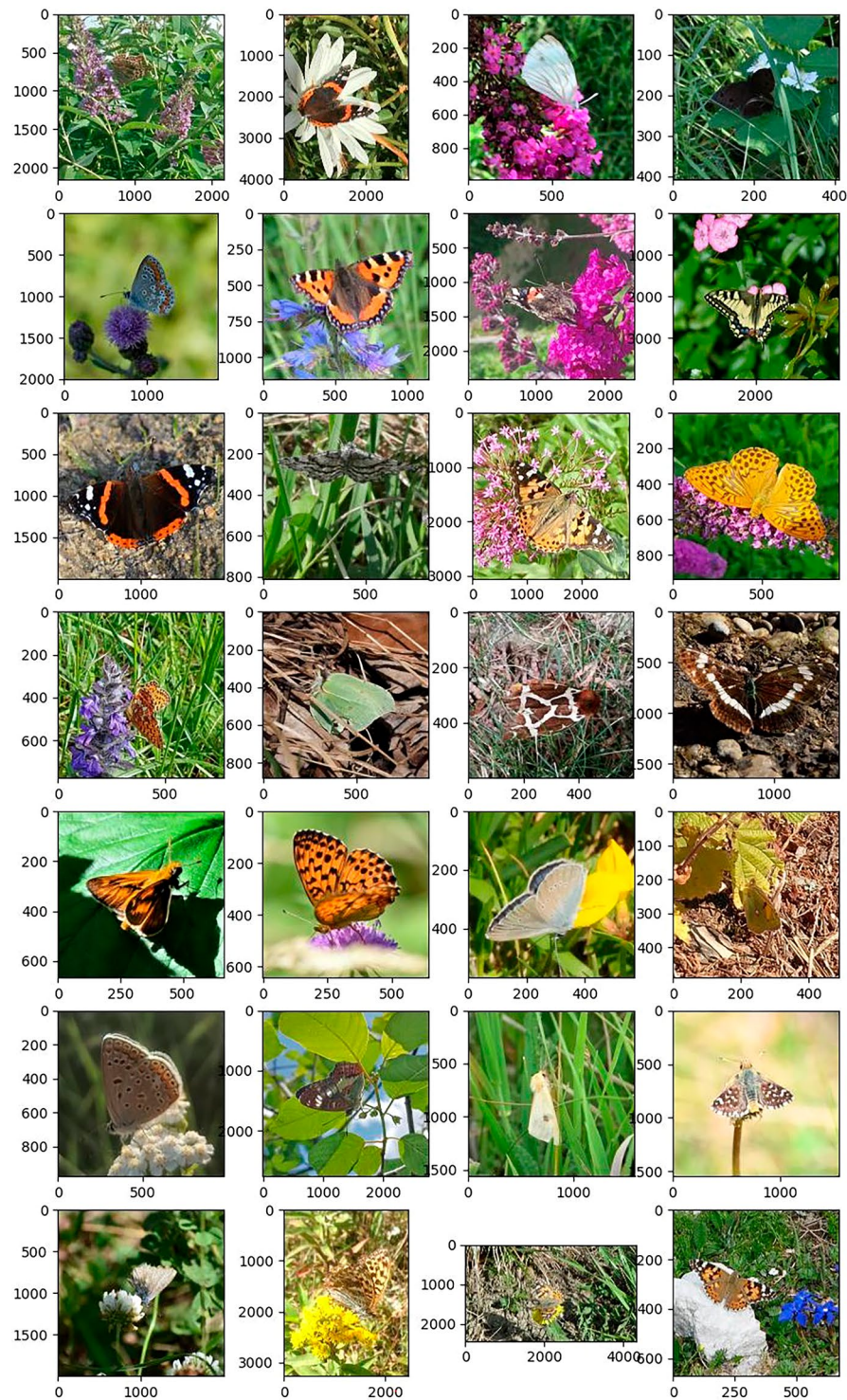


Fig. 1 Random selection of images of the butterfly and moth images dataset collected with the application “Schmetterlinge Österreichs”.

learning models have already been trained to identify butterfly and moth (Macrolepidoptera) species with high accuracy^{10,24–26}. The datasets used in these studies ranged from less than 1000 to over 34,000 images covering ten to 636 species^{10,26}.

The dataset that is presented here, is considerably larger than those used in previous studies. It contains 530,404 images of 185 butterfly and moth species that were recorded in Austria. The dataset was collected by citizen scientists with the application “Schmetterlinge Österreichs” (<https://www.schmetterlingsapp.at/>) of the Billa-Foundation Blühendes Österreich. Correct species identification was ensured by an experienced entomologist. The dataset has a strong class imbalance with the number of images per species ranging from 1 to nearly

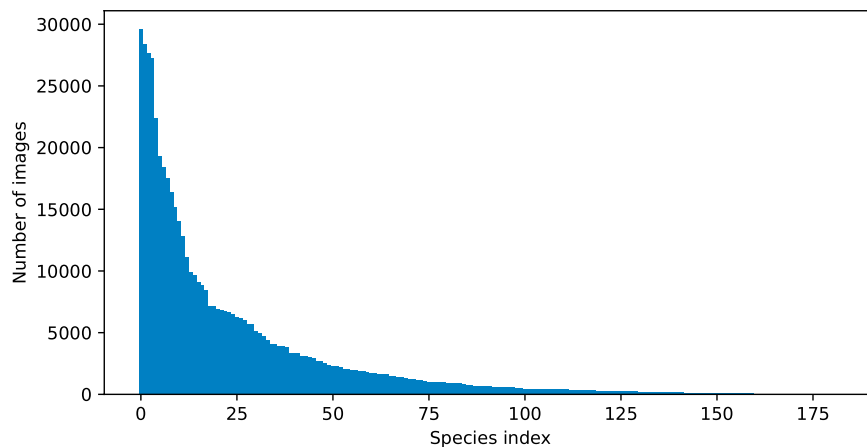


Fig. 2 Distribution of the number of images per species for the dataset with >500,000 images of butterflies and moths that were collected with the application “Schmetterlinge Österreichs” of the Billa Foundation “Blühendes Österreich” between 2016 and 2023 (Figure from Barkmann *et al.* 2025).

30,000. Such an imbalance is common in species records from CS projects and has multiple causes. Some species are more common than others, are easier to detect or more preferably recorded by citizen scientists^{27–29}.

The dataset offers valuable opportunities to train neural networks on the fine-grain classification task of identifying butterfly and moth species and to assess the performance of different neural network architectures and hyperparameter settings. To demonstrate the application of the dataset for the training of deep learning models, it was used to fine-tune a Multi-Axis Vision Transformer model (MaxViT) that was proposed by Tu *et al.*³⁰. A model that was pre-trained on the ImageNet dataset³¹ was used. The dataset has already been used to train a ResNet152 model for a Master’s thesis in which different methods to handle class imbalance and performance for different species were analysed³².

Methods

Butterfly images were taken and uploaded by the users of the application “Schmetterlinge Österreichs” (<https://www.schmetterlingsapp.at/>) of the Billa Foundation “Blühendes Österreich” in Austria between 2016 and 2023. Over 25,000 users were involved in the collection of images. Registered users took images with their smartphones and could directly upload them via a mobile app. The app is also available as a desktop version, which is especially useful for uploading images taken with a camera independent from smartphones. The user who uploaded an observation or other members of the community could propose a species level classification of the images. The species that can be reported include 157 butterfly and 32 moth species. Images of 185 of these species were uploaded. Helmut Höttinger, who is an experienced entomologist and an expert on butterflies and moths, continuously validated the correct classification of all images. Over 11,000 Images that showed eggs, larvae and pupae and images with more than one butterfly species were manually deleted from the dataset. Some images were not detected though (s. Technical Validation).

Data Records

The dataset is available at figshare (<https://doi.org/10.25452/figshare.plus.29135618>)³³.

The whole dataset contains 541,677 images of 185 butterfly and moth species and has a size of 315 GB. Files are organized in a folder structure, with one folder for each species. The size of individual images varies as they were taken with different devices. The mean width of the images is 1887 px (min: 66 px, max 12000 px). The mean height is 1906 px (min: 66 px, max: 8000 px). All images are in the JPEG file format. See Fig. 1 for a random selection of images from the dataset. There are 29,612, images of *Aglais io*, the species that was photographed most frequently, while other species are represented by only one image. For 131 species, there are fewer than 1,000 images, and for 62 species, fewer than 100 (Fig. 2, S Table S1 (see Supplementary information)).

The dataset³³ contains images of 77.6% of the 210 butterfly species (Superfamily Papilionoidea) that occur in Austria, excluding five regionally extinct species³⁴. The moth species that occur in Austria are less well represented as only 32 of the nearly 4000 species³⁵ (of which 1243 can be considered as Macrolepidoptera) can be recorded with the application. The selected moth species are species that can be observed easily and many of them have characteristic morphological features in at least one life stage. In Europe, there are 496 species of butterflies³⁶ and about 8,200 moth species, about 3,000 of which are Macrolepidoptera³⁷.

While some butterfly species such as *Aglais io* and *Vanessa atalanta* have wing patterns that are unique in Austria, species of the genus *Pyrgus* (Fig. 3) or *Erebia* can be highly similar. Other species groups such as the tribus Melitaeini contain species that are highly similar on one side of the wings but mostly have characteristic patterns for species determination on the other side (Fig. 3). Images vary regarding the size of the depicted butterfly or butterflies, the angle at which individuals were photographed and the background of the images (Fig. 4).

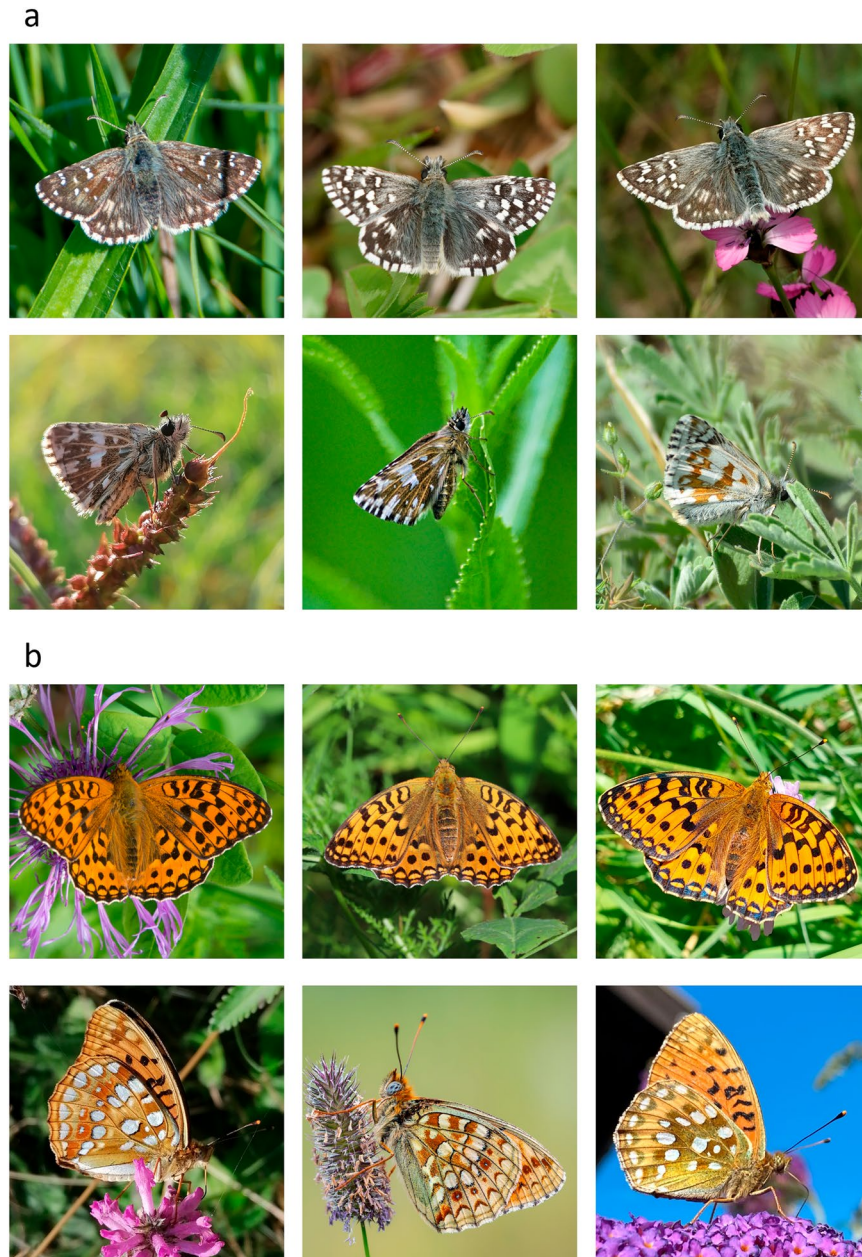


Fig. 3 Examples of similar looking species that can be difficult to identify from images only. **(a)** from left to right: *Pyrgus armoricanus*, *Pyrgus malvae*, *Pyrgus carthami*; **(b)** from left to right: *Fabriciana adippe*, *Fabriciana niobe*, *Speyeria aglaja*.

Technical Validation

Model training. To demonstrate the training of a deep learning model on the dataset³³ and to estimate the number of images that show life stages other than adults or depict more than one species, a deep learning model was fine-tuned using the dataset. Its performance was assessed and misclassified images evaluated.

For model training, only images of species with at least 50 records were used to allow for a reasonable partition of the data and evaluation on species level. This dataset contained 529,835 images of 162 species, 31 of which were moth species. 10% of the images were selected as test data with a stratified approach that ensured that the species were represented proportionally to their number of images in the whole dataset. The remaining images were divided in 80% training and 20% validation data, again using a stratified approach.

Images were augmented for higher variability of the training data. Images were cropped to up to half of their sizes and the aspect ratio was changed by a value of 0.8 to 1.2. Images were rotated between -50° and 50° , flipped horizontally and vertically with a probability of 30% and distorted with a scale of 0.2 with a probability of 40%. All images were cropped to 224×224 pixels. RGB channels of the images were normalized based on the ImageNet dataset standards with the means 0.485, 0.456, 0.406 and standard deviations 0.229, 0.224, 0.225.



Fig. 4 Examples for the variability of images of the same species. (a) different size of the butterfly in the image, (b) different sides of the wings and angles at which they are photographed, (c) different backgrounds.

Images that were used for model evaluation were only resized and cropped to 224×224 pixels and the same normalization of colour channels as for the training data were applied.

A Multi-Axis Vision Transformer model (MaxViT-T)³⁰ that was pre-trained on the ImageNet dataset³¹ was used. MaxViT models combine elements of convolutional neural networks (CNNs) and Vision Transformers. They outperform other models at image classification of the ImageNet dataset with higher parameter and computing efficiency³⁰.

The model was trained for 300 epochs on 8 Graphics Processing Units (GPUs) with a batch size of 16 images on each GPU. To facilitate longer training, stochastic gradient descent was used as optimizer with a momentum of 0.9. To address the class imbalance of the dataset and ensure better representation of minority classes during training, a weighted loss function was applied. The weights were proportional to the inverse of the number of images in each class.

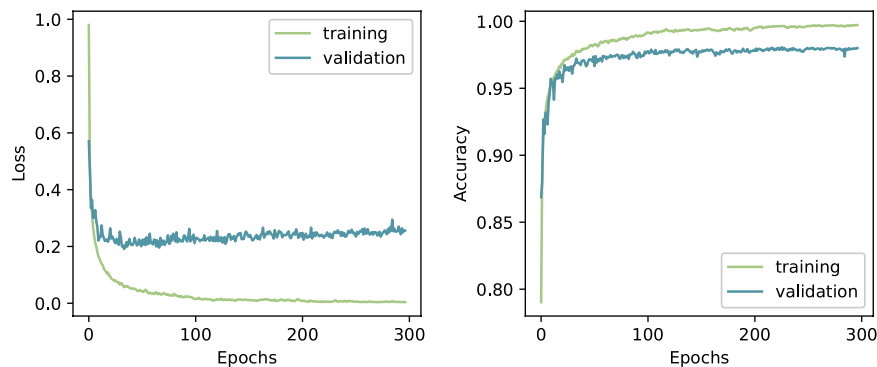


Fig. 5 Accuracy and loss during training of a MaxViT-T model on the butterfly and moth image dataset collected with the application “Schmetterlinge Österreichs”.

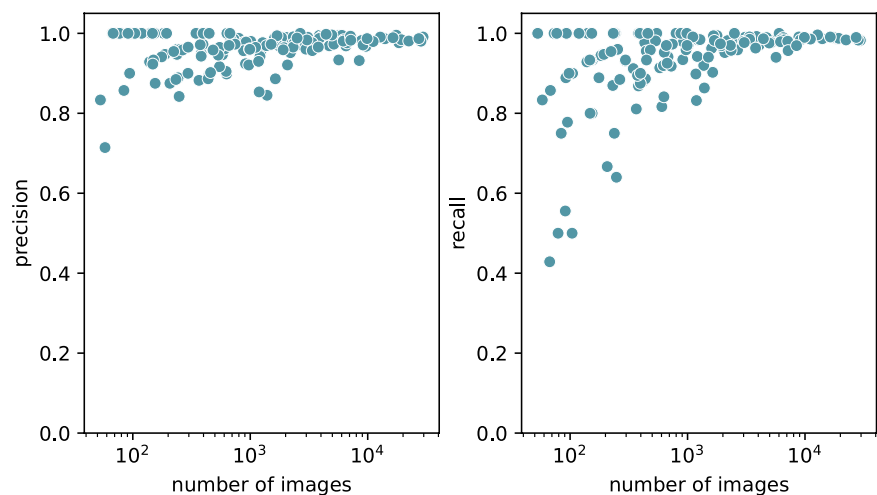


Fig. 6 Precision and recall that were achieved by the MaxViT-T model on test data for the different species ($n = 185$) in the dataset against the number of images of each species.

Model performance on the whole dataset was assessed with the top-1, top-3 and top-5 accuracy. Additionally, precision and recall for each species were calculated. To estimate how many images of eggs, larvae, pupae and multiple species were not detected during data cleaning, all misclassified images were assessed manually.

The PyTorch library³⁸ was used for model training and validation. For parallel computing the Distributed Data Parallel (DDP) framework³⁹ and the Accelerate library provided by Hugging Face⁴⁰ were used.

Model training was conducted on the EuroHPC supercomputer LUMI hosted by CSC (Finland) and the LUMI consortium.

Results

The highest validation accuracy of 0.9806 was reached after 225 epochs. The highest training accuracy was 0.9971 (Fig. 5).

On the test dataset, the model achieved an accuracy of 97.87%. Mean recall over all species was 93.54% and mean precision was 96.31%. Precision was $>70\%$ for all species, while recall was $<50\%$ for some of the species which are represented by only few images in the dataset (Fig. 6). See table S1 in the supplements for the number of images, recall and precision for each species.

On the test dataset, 1127 images were not correctly classified by the model. 101 of these showed more than one (mostly two) species, 11 showed eggs, 31 larvae and 7 pupae. These images comprise 0.28% of the test dataset. The number of images showing more than one species is likely higher, as the model can correctly classify such images when identifying the species which the label refers to. Assuming that the dataset contains twice the number of images with more than one species than were detected here, the number of images that do not show adult life stages of only one species is still $<0.5\%$.

Usage Notes

The dataset³³ is highly imbalanced which can negatively affect model performance for minority classes and should be considered when training models on the dataset⁴¹. The dataset does not contain all butterfly and moth species that occur in Austria. Some butterfly species that are difficult to determine to species level from images are not part of the dataset and the moths are represented by only few conspicuous species. The species pairs

Aricia agestis/A. Artaxerxes, *Phengaris alcon/rebeli*, *Colias hyale/alfacariensis* and *Leptidea sinapis/juvernica* are treated as one species each as they cannot be distinguished from images alone. Due to the incomplete coverage the accuracy that can be obtained for an automatic classification of all species in Austria is likely lower than for this dataset. Even though most of such images were manually excluded from the dataset, it still contains few images of eggs, larvae and pupae or images that show more than one species.

Code availability

The scripts that were used for model training and the weights for the MaxViT-T model are available on figshare: <https://doi.org/10.25452/figshare.plus.29135618>.

The scripts that were used for model training are also available on GitHub (https://github.com/FriederikeBarkmann/MaxViT_ButterflyIdentification).

The trained model is also available on HuggingFace⁴².

Received: 9 June 2025; Accepted: 29 July 2025;

Published online: 06 August 2025

References

- McClure, E. C. *et al.* Artificial Intelligence Meets Citizen Science to Supercharge Ecological Monitoring. *Patterns (New York, N.Y.)* **1**, 100109, <https://doi.org/10.1016/j.patter.2020.100109> (2020).
- Lotfian, M., Ingensand, J. & Brovelli, M. A. The Partnership of Citizen Science and Machine Learning: Benefits, Risks, and Future Challenges for Engagement, Data Collection, and Data Quality. *Sustainability* **13**, 8087, <https://doi.org/10.3390/su13148087> (2021).
- Chandler, M. *et al.* Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation* **213**, 280–294, <https://doi.org/10.1016/j.biocon.2016.09.004> (2017).
- Besson, M. *et al.* Towards the fully automated monitoring of ecological communities. *Ecology letters* **25**, 2753–2775, <https://doi.org/10.1111/ele.14123> (2022).
- Tuia, D. *et al.* Perspectives in machine learning for wildlife conservation. *Nature communications* **13**, 792, <https://doi.org/10.1038/s41467-022-27980-y> (2022).
- Willi, M. *et al.* Identifying animal species in camera trap images using deep learning and citizen science. *Methods Ecol Evol* **10**, 80–91, <https://doi.org/10.1111/2041-210X.13099> (2019).
- Wang, Z., Cui, J. & Zhu, Y. Review of plant leaf recognition. *Artif Intell Rev* **56**, 4217–4253, <https://doi.org/10.1007/s10462-022-10278-2> (2023).
- Mäder, P. *et al.* The Flora Incognita app – Interactive plant species identification. *Methods Ecol Evol* **12**, 1335–1342, <https://doi.org/10.1111/2041-210X.13611> (2021).
- Hansen, O. L. P. *et al.* Species-level image classification with convolutional neural network enables insect identification from habitus images. *Ecology and evolution* **10**, 737–747, <https://doi.org/10.1002/ece3.5921> (2020).
- Theivaprakasham, H. Identification of Indian butterflies using Deep Convolutional Neural Network. *Journal of Asia-Pacific Entomology* **24**, 329–340, <https://doi.org/10.1016/j.aspen.2020.11.015> (2021).
- Gomez Villa, A., Salazar, A. & Vargas, F. Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological Informatics* **41**, 24–32, <https://doi.org/10.1016/j.ecoinf.2017.07.004> (2017).
- Norouzzadeh, M. S. *et al.* Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E5716–E5725, <https://doi.org/10.1073/pnas.1719367115> (2018).
- Unger, S., Rollins, M., Tietz, A. & Dumais, H. iNaturalist as an engaging tool for identifying organisms in outdoor activities. *Journal of Biological Education* **55**, 537–547, <https://doi.org/10.1080/00219266.2020.1739114> (2021).
- Barta, Z. Deep learning in terrestrial conservation biology. *Biologia futura* **74**, 359–367, <https://doi.org/10.1007/s42977-023-00200-4> (2023).
- Brereton, T., Roy, D. B., Middlebrook, I., Botham, M. & Warren, M. The development of butterfly indicators in the United Kingdom and assessments in 2010. *Journal of Insect Conservation* **15**, 139–151, <https://doi.org/10.1007/s10841-010-9333-z> (2011).
- van Swaay, C., Warren, M. & Lois, G. Biotope Use and Trends of European Butterflies. *J Insect Conserv* **10**, 189–209, <https://doi.org/10.1007/s10841-006-6293-4> (2006).
- Thomas, J. A. Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **360**, 339–357, <https://doi.org/10.1098/rstb.2004.1585> (2005).
- Anderle, M. *et al.* Efficiency of birds as bioindicators for other taxa in mountain farmlands. *Ecological Indicators* **158**, 111569, <https://doi.org/10.1016/j.ecolind.2024.111569> (2024).
- Gerlach, J., Samways, M. & Pryke, J. Terrestrial invertebrates as bioindicators: an overview of available taxonomic groups. *J Insect Conserv* **17**, 831–850, <https://doi.org/10.1007/s10841-013-9565-9> (2013).
- van Swaay, C. *et al.* The EU Butterfly Indicator for Grassland species: 1990–2017: Technical Report. Butterfly Conservation Europe & ABLE/eBMS (www.butterfly-monitoring.net) (2019).
- Roy, D. B., Rothery, P., Brereton, T., Kühn, E. & Settele, J. The design of a systematic survey scheme to monitor butterflies in the United Kingdom (2005).
- Taron, D. & Ries, L. Butterfly Monitoring for Conservation. In *Butterfly conservation in North America. Efforts to help save our charismatic microfauna*, edited by J. C. Daniels, pp. 35–57 (Springer, Dordrecht, 2015).
- Schlegel, J. & Rumpf, R. Attitudes towards potential animal flagship species in nature conservation: A survey among students of different educational institutions. *Journal for Nature Conservation* **18**, 278–290, <https://doi.org/10.1016/j.jnc.2009.12.002> (2010).
- Chang, Q., Qu, H., Wu, P. & Yi, J. Fine-Grained Butterfly and Moth Classification Using Deep Convolutional Neural Networks (2017).
- Nie, L., Wang, K., Fan, X. & Gao, Y. Fine-Grained Butterfly Recognition with Deep Residual Networks: A New Baseline and Benchmark. In *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp. 1–7 (IEEE 2017).
- Mattins, R. F., Sarobin, M. V. R., Aziz, A. A. & Srivarshan, S. Object detection and classification of butterflies using efficient CNN and pre-trained deep convolutional neural networks. *Multimed Tools Appl*, <https://doi.org/10.1007/s11042-023-17563-4> (2023).
- Isaac, N. J. B. *et al.* Distance sampling and the challenge of monitoring butterfly populations. *Methods Ecol Evol* **2**, 585–594, <https://doi.org/10.1111/j.2041-210X.2011.00109.x> (2011).
- Koch, W., Hogeweg, L., Nilsen, E. B., O'Hara, R. B. & Finstad, A. G. Recognizability bias in citizen science photographs. *Royal Society open science* **10**, 221063, <https://doi.org/10.1098/rsos.221063> (2023).
- Arazy, O. & Malkinson, D. A Framework of Observer-Based Biases in Citizen Science Biodiversity Monitoring: Semi-Structuring Unstructured Biodiversity Monitoring Protocols. *Front. Ecol. Evol.* **9**, <https://doi.org/10.3389/fevo.2021.693602> (2021).
- Tu, Z. *et al.* MaxViT: Multi-Axis Vision Transformer (2022).

31. Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255 (IEEE 2009).
32. Barkmann, F. Image-based butterfly species identification with convolutional neural networks. Available at <https://ulb-dok.uibk.ac.at/urn/urn:nbn:at:at-ubi:1-172379> (2025).
33. Friederike Barkmann, Andreas Lindner, Ronald Würflinger, Helmut Höttinger & Johannes Rüdiger. Machine learning training data: over 500,000 images of butterflies and moths (Lepidoptera) with species labels; <https://doi.org/10.25452/figshare.plus.29135618> (2025).
34. Höttinger, H. & Pennerstorfer, J. Rote Liste der Tagmetterlinge Österreichs (Lepidoptera: Papilionoidea & Hesperioidea). In *Rote Listen gefährdeter Tiere Österreichs. Checklisten, Gefährdungsanalysen, Handlungsbedarf. Teil 1: Säugetiere, Vögel, Heuschrecken, Wasserkäfer, Netzflügler, Schnabelfliegen, Tagfalter.*, edited by K. Zulka, pp. 313–354 (Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien, Böhlau, 2005).
35. Huemer, P. Die Schmetterlinge Österreichs (Lepidoptera). *Systematische und faunistische Checkliste* (Tiroler Landesmuseum Ferdinandeum, Innsbruck, 2013).
36. Wiemers, M. *et al.* An updated checklist of the European Butterflies (Lepidoptera, Papilionoidea). *ZooKeys*, 9–45, <https://doi.org/10.3897/zookeys.811.28712> (2018).
37. Karsholt, O. & Razowski, J. *The Lepidoptera of Europe: a distributional checklist* (Brill, 1996).
38. Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems*, edited by H. Wallach, *et al.*, Vol. 32 (Curran Associates, Inc 2019).
39. Li, S. *et al.* PyTorch Distributed: Experiences on Accelerating Data Parallel Training (2020).
40. Gugger, S. *et al.* Accelerate: Training and inference at scale made simple, efficient and adaptable (2022).
41. Buda, M., Maki, A. & Mazurkowski, M. A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural networks: the official journal of the International Neural Network Society* **106**, 249–259, <https://doi.org/10.1016/j.neunet.2018.07.011> (2018).
42. Barkmann, F. & Lindner, A. MaxViT_butterfly_identification; <https://doi.org/10.57967/hf/5986>.

Acknowledgements

We want to thank the about 25.000 users of the application “Schmetterlinge Österreichs” that contributed to collecting this large dataset of butterfly and moth images. We especially want to thank those citizen scientists who contributed the most: Karin Hiebner, Gabriele Kriz, Anna Söllinger, Sissi Lechner, Momcilo Borek, Hansjörg Vogl, Udo Tschernuter, Stefan Greil, Sabina Bergauer, Peter Zych, and Silke Geroldinger. The butterfly app with which the dataset was collected is funded and managed by the Billa Foundation Blühendes Österreich. We especially want to thank Ines Lemberger and Peter Huemer from Blühendes Österreich and Florian Mündler from ApteC. Data preparation and modelling were conducted within the Viel-Falter Butterfly Monitoring (www.viel-falter.at) which has received funding from the Federal Minister of Agriculture and Forestry, Climate and Environmental Protection, Regions and Water Management (BMLUK) and by the project EuroCC Austria which has received funding from the European High Performance Computing Joint Undertaking (JU) and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia under grant agreement No 101101903. We acknowledge the EuroHPC Joint Undertaking for awarding this project access to the EuroHPC supercomputer LUMI, hosted by CSC (Finland) and the LUMI consortium through a EuroHPC Development Access call.

Author contributions

F.B.: conceptualization, data curation, model training and validation, writing original draft. A.L.: model training, adaptation for HPC and data parallelization, writing – reviewing and editing. R.W.: funding acquisition, conceptualization, project administration, resources. H.H.: data curation, validation of species identification. J.R.: conceptualisation, project administration, funding acquisition, writing – reviewing and editing, supervision.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05708-z>.

Correspondence and requests for materials should be addressed to J.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025