

Overview of FungiCLEF 2024: Revisiting Fungi Species Recognition Beyond 0-1 Cost

Lukas Picek^{1,2,*}, Milan Šulc³ and Jiří Matas⁴

¹Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Czech Republic

²Inria, LIRMM, University of Montpellier, CNRS, France

²Second Foundation, Czech Republic

³The Center for Machine Perception Dept. of Cybernetics, FEE, Czech Technical University in Prague, Czech Republic

Abstract

The third edition of the fungi recognition challenge, FungiCLEF 2024, organized within LifeCLEF, advances the field of mushroom species identification using computer vision and machine learning. Building on the Danish Fungi 2020 dataset and incorporating new data from the CzechFungi app, FungiCLEF 2024 challenges participants to recognize fungi species from images and metadata, focusing on efficient inference and minimalization of edible and poisonous species confusion. The strict limits on computational complexity ensure that the resulting solutions are practical for use in real-world settings with limited computational resources. The competition attracted seven teams, with five outperforming the provided baseline, which was based on the pre-trained EfficientNet-B1 model. This overview paper provides (i) a comprehensive description of the challenge and provided baseline method, (ii) detailed characteristics of the dataset and task specifications, (iii) an examination of the methods employed by contestants, and (iv) a discussion of the competition outcomes. The results highlight incremental advancements in fungi recognition, showcasing innovative approaches and techniques that push the limits of previous work.

Keywords

LifeCLEF, FungiCLEF, fine-grained visual categorization, metadata, open-set recognition, fungi, species identification, machine learning, computer vision, classification

1. Introduction

Fungi recognition systems based on computer vision and machine learning [1, 2] are transforming the field of mycology, making it easier than ever for researchers, enthusiasts, and professionals to identify mushroom species. Tasks that once required extensive expertise, i.e., studying the existing literature, can now be accomplished in a few seconds. The fungi identification service offered by the Atlas of Danish Fungi [3] exemplifies this advancement: users simply capture an image of their specimen, and the system promptly generates a list of probable species matches. This facilitates efficient manual verification by allowing users to compare their observations with reference photos and species descriptions. Additionally, it encourages users to contribute valuable biodiversity observations, enhancing the overall understanding and documentation of fungal diversity.

Despite the impressive performance of automatic fungi species recognition systems, significant challenges remain due to the complexity and diversity of fungal species. One major challenge is the vast number of fine-grained categories (species) that exist. Many of these species exhibit high visual similarities, making it difficult to distinguish between them even though they may not be genetically related (see Figure 1). This visual similarity can easily lead to misidentification, as the algorithm might not reliably discern subtle differences in color, shape, or texture. Additionally, there is significant intra-class variance within species. The appearance of fungal specimens can vary widely based on several factors, including genotype, age, seasonal conditions, and the local environment. For instance, a mushroom of the same species can look markedly different when it is young compared to when it is mature. Seasonal variations can affect the color and size of fungi, while local environmental conditions

CLEF 2024: Conference and Labs of the Evaluation Forum, September 09–12, 2024, Grenoble, France

*Corresponding author.

✉ picekl@kky.zcu.cz (L. Picek)

ORCID 0000-0002-6041-9722 (L. Picek); 0000-0002-6321-0131 (M. Šulc); 0000-0003-0863-4844 (J. Matas)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



Figure 1: Fungi species intra- and inter-class similarities. Visual similarities and differences across nine selected species and three distinct families (i.e., higher taxonomic ranks). Image taken from [8].

such as humidity, light exposure, and soil type can further influence their appearance [4, 5]. These variations pose a considerable challenge for automatic recognition systems, which must be robust enough to account for such differences to ensure accurate identification. Moreover, the quality and resolution of images submitted for recognition can vary significantly, impacting the system’s ability to accurately classify the observations. Images taken in the wild might suffer from poor lighting, occlusions, or background noise, adding another layer of complexity to the recognition task [6, 7]. These challenges underscore the importance of developing new methods capable of handling the complex details of fungal diversity, pushing the boundaries of current state-of-the-art in fungi species recognition, and fine-grained visual categorization in general.

To allow continual incremental improvements in fungi recognition, we organize an annual research competition – FungiCLEF. The latest edition, which took part of the LifeCLEF 2024 [9] and the FGVC11 workshop at CVPR 2024, builds on the foundations laid by its predecessors [10, 11]. This year’s challenge introduces an updated dataset and continues to emphasize model efficiency by imposing computational and memory constraints. The competition presents participants with fungi recognition scenarios that address real-world applications, including fungi species identification and distinguishing between poisonous and edible mushrooms.

By promoting innovation in this field, FungiCLEF 2024 aims to bridge further the gap between computer vision capabilities and practical mycological needs, potentially impacting areas ranging from biodiversity research to public health and beyond.

2. Challenge Description

Efficient and scalable species recognition is essential for large-scale initiatives such as citizen science projects [1, 2], which often operate with limited computational resources. Species identification typically relies not just on the visual features of the specimen but also on additional contextual data about habitat, substrate, location, etc. For the FungiCLEF2024 competition, we have developed a benchmark with rich metadata and expert-verified labels for testing the performance of systems that combine visual and contextual information. Given that mushrooms are often foraged for consumption, the competition also addresses scenarios related to misclassification between *edible* and *poisonous* species. This ensures robust outcomes in fungal species recognition, enhancing scientific research and public safety.

To enable use in practical applications, all participants had to submit their models via the HuggingFace evaluation platform and must pass computational limits. Each classification model had to fit limits for prediction time limit (120 minutes) within a given HuggingFace server instance (Nvidia T4 small 4vCPU, 15GB RAM, 16GB VRAM).

Table 1

FungiCLEF 2024 dataset statistics. We list the number of species and images in each subset. The *known* and *unknown* represent how many species were and were not available in the training dataset.

Subset	# of Species	→ known/unknown	# of Images	# of Observations
Training	1,604	1,604 / –	295,938	177,170
Validation	3,299	1,084 / 1,629	91,231	45,021
Test	1,398	749 / 649	41,177	22,412
↳ <i>CzechFungi App</i>	137	94 / 43	393	215
↳ <i>Atlas of Danish Fungi</i>	1,261	721 / 540	40,784	22,197

2.1. Dataset

The FungiCLEF 2024 dataset builds upon the previous editions of FungiCLEF [12, 13], LifeCLEF [14, 15, 16], and the Danish Fungi 2020 dataset [8]. All training data is derived from a citizen science platform – the Atlas of Danish Fungi. Each fungi observation in the provided dataset has undergone expert validation, ensuring high-quality species labels. The dataset features rich observation metadata, i.e., information about habitat, substrate, timestamp, location, etc. Provided subsets (i.e., training, validation, and test) are briefly described below, and their statistics in detail are listed in Table 1.

The training set is based on 177,170 real and expert-verified fungi observations of 1,604 species. The dataset is built exclusively from the Danish Fungi 2020 data by combining the training and public test sets and includes 295,938 images.

The validation set comprises expert-validated observations with species labels collected solely in 2022. This subset includes 3,299 fungi species and contains 45,021 observations with many "unknown" species.

The test set is based on two subsets originating from two applications (Atlas of Danish Fungi and CheckFungi) and two countries, Denmark and the Czech Republic respectively. The CheckFungi dataset is a small subset containing only around 200 submissions and is included primarily as a control set to prevent cheating. The test set was split 80/20 for public and private evaluation, respectively.

2.2. Evaluation Protocol

Same as the previous year [13], the challenge aims at predicting the species given visual observation and metadata and considers scenarios focusing on correct species classification as well as on classifying edible vs poisonous mushrooms. Namely, the goal is to minimize the empirical loss L for decisions $q(x)$ over observations x and true labels y , given a cost function $W(y, q(x))$.

$$L = \sum_i W(k_i, q(x_i)). \quad (1)$$

Different recognition scenarios and their cost function $W(y, q(x))$ are described together with their motivation in the points below:

- **Track 1: Standard classification with "unknown" category** (an open-set scenario). The first metric was the standard classification accuracy, i.e., the average correctness of the predicted class. All species not represented in the training set had to be correctly classified as an "unknown" category. The decision function is simple: each observation is simply represented by an identity matrix, i.e.,

$$W_1(y, q(x)) = \begin{cases} 0 & \text{if } q(x) = y \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

- **Track 2: Cost for confusing edible species for poisonous and vice versa.** Let us have a function d that indicates dangerous (poisonous) species as $d(y) = 1$ if species y is poisonous, and $d(y) = 0$ otherwise. Let us denote c_{PSC} the cost for poisonous species confusion (if a poisonous observation was misclassified as edible) and c_{ESC} the cost for edible species confusion (if an edible observation was misclassified as poisonous).

$$W_2(y, q(x)) = \begin{cases} 0 & \text{if } d(y) = d(q(y)) \\ c_{\text{PSC}} & \text{if } d(y) = 1 \text{ and } d(q(y)) = 0. \\ c_{\text{ESC}} & \text{otherwise} \end{cases} \quad (3)$$

For the benchmark, we set $c_{\text{ESC}} = 1$ and $c_{\text{PSC}} = 100$.

- **Track3: A user-focused loss composed of both the classification error and the poisonous/edible confusion.** Assuming the user is interested both in the species classification as well as in low edible to poisonous species confusion (and vice versa), the third cost function simply combines W_1 and W_2 :

$$W_3(y, q(x)) = W_1(y, q(x)) + W_2(y, q(x)). \quad (4)$$

2.3. Baseline

To enable an easier start for all participants and straightforward model evaluation, we provide a weak baseline based on the pre-trained PyTorch EfficientNet-B1 [17] model wrapped into a [HuggingFace repository](#), allowing direct evaluation on the private test set. This repository includes the model weights and inference scripts. In addition to the PyTorch-based submission, we offer an example for submitting an ONNX model. This ONNX model was initially provided for another LifeCLEF [9, 18] competition, SnakeCLEF 2024 [19]. The pre-trained EfficientNet-B1 model used in our baseline was originally published in the Danish Fungi 2020 dataset [8]. This model has demonstrated relatively strong performance in fungi species classification and serves as a robust starting point. Overall, our goal was to offer a comprehensive and accessible starting point for all participants, enabling them to focus on developing novel solutions and improving upon the provided baseline.

2.4. Timeline

The FungiCLEF 2024 competition was launched on March 13, 2024, and was promoted through the LifeCLEF, HuggingFace, and FGVC challenge web pages, inviting participants to register. The competition ran for approximately three months, with the final submission deadline on May 24. Similar to the previous year, the test data remained confidential. Participants were allowed to make up to five submissions per day using the HuggingFace evaluation platform to assess their models on the test set. Two weeks before the deadline, the submission limit was increased to ten per day. After the competition concluded, all participants had the opportunity to submit post-competition entries for further evaluation of their ablations.

2.5. Working Notes

Participants were strongly encouraged to submit both their code and a detailed technical report (Working Notes) to ensure their results can be fully reproduced. All the submitted Working Notes underwent thorough review and were given complex feedback by 2-3 experts with extensive publication records in Computer Vision and Machine Learning. This rigorous review process was designed to guarantee reproducibility and maintain quality standards. The review was single-blind, allowing participants to respond with up to two rebuttals to address any feedback or concerns raised by the reviewers. These working notes provide an in-depth analysis of the techniques employed, including hyperparameter tuning, model ensembling, and loss function selection, offering valuable insights into the development of the method for fungal image classification.

3. Challenge Results

This year, the three tracks of FungiCLEF have three different best-performing submissions by three different teams¹. However, for the official ranking, the Track 3 score was selected. The best-performing submission in Track 1 by *Jack Etheredge* [20] achieved a score of 0.240. The best scores in Track 2 were achieved by team *upupup* [21] with a score of 0.072. Finally, the best score for Track 3, the main competition track, was achieved by team *IES* [22] with a score of 0.362. The official challenge results, in terms of Track 1, Track 2, and Track 3 metrics, are reported in Figure 2. For completeness, we note that a post-competition submission by team *DS@GT* [23] achieved even higher recognition scores, highlighting the competitive and evolving nature of the challenge. This post-competition effort serves as a testament to the ongoing advancements and innovations in the field, extending beyond the official competition period.

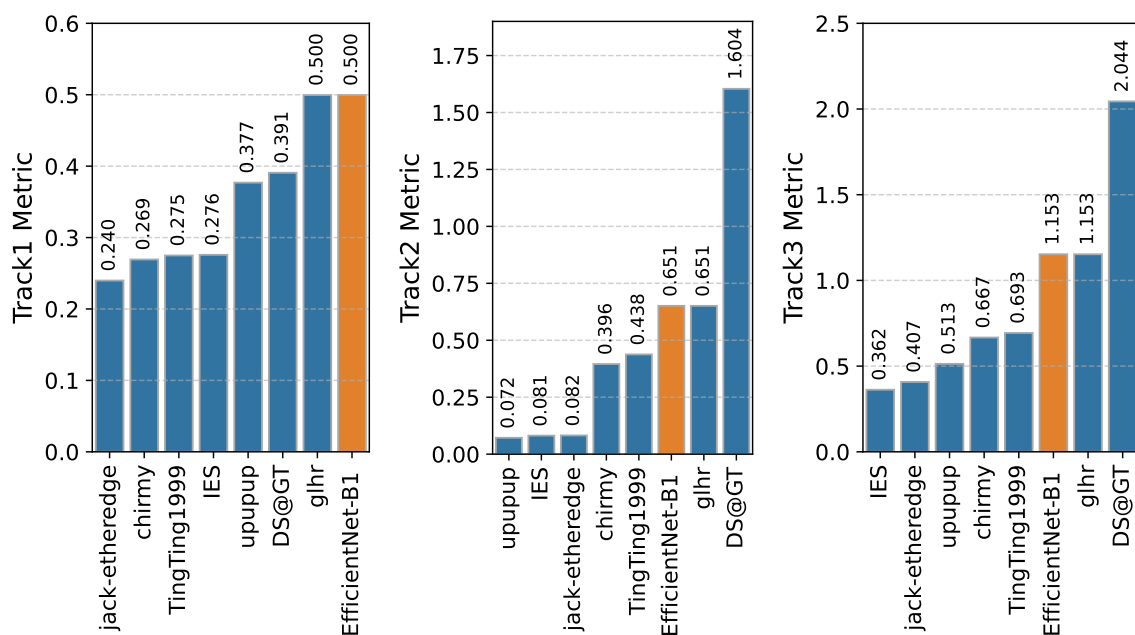


Figure 2: Private Leaderboard of the FungiCLEF 2024 competition. We report performance for all 7 teams. The track metrics are recognition losses, thus **lower is better**. Each track could have a different submission; therefore, Track1 and Track2 scores for the same team do not sum up to the Track3 score. The orange color depicts baseline performance.

4. Participants and Methods

This year, a total of seven teams participated in the FungiCLEF 2024 challenge; of these, five outperformed the baseline EfficientNet-B1 in Track 3, and five submitted working notes from which four passed the review process and were accepted for publication. The methodologies varied and included a range of techniques from state-of-the-art neural network architectures to sophisticated strategies to incorporate the metadata. Details of the best methods and systems used are synthesized below and further developed in participants’ working notes [20, 21, 22, 23].

Team **IES** [22] (Top1) utilized a Swin Transformer V2 Base [24] architecture as a feature extractor and used a similar approach for meta-data integration as Ren et al. [25] from the previous edition of FungiCLEF [13]. Besides, they introduced (i) a poisonous re-ranking that prevents predicting an edible species when there is a significant chance of the sample being poisonous, and (ii) a genus loss that improves the feature space’s regularization.

¹Each team usually had different "best" submissions for each track.

Jack Etheredge [20] (Top2) combined visual information with metadata using MetaFormer-0 and MetaFormer-2 [26] and further improved the ensemble by a vision-only CAFormer-S18 [27], and proposed a novel application of openGAN [28] for open-set recognition of fine-grained images utilizing WGAN-GP [29].

Team **upupup** [21] (Top3) used Dynamic MLP [30] for the fusion of image features and metadata, identifying unknown classes using an entropy-based approach, training with a marginal expected loss for recognizing poisonous mushrooms while maintaining accuracy.

Team **DS@GT** [23] (Top8²) utilized DINOv2 visual embeddings [31] (namely the *dinov2-large* model with register for final submission) which were combined with metadata in a classifier head. The model was trained with a composite loss function, consisting of the Seesaw loss [32] and a binary cross entropy loss for poisonous species classification.

5. Conclusions and Discussion

This paper presents an overview and results evaluation of the third edition of the FungiCLEF challenge, organized in conjunction with the CLEF LifeCLEF lab, and CVPR-FGVC11 – The Tenth Workshop on Fine-Grained Visual Categorization held within the CVPR conference. This challenge continues to push the boundaries of fine-grained visual categorization by bringing together diverse methodologies and innovative approaches from leading research teams worldwide.

By introducing an updated dataset, emphasizing model efficiency, and addressing both species recognition and poisonous mushroom identification, FungiCLEF 2024 has fostered innovation and practical solutions for fungi recognition. The diverse approaches employed by the top-performing teams illustrate the evolving landscape of fungi recognition challenges. From sophisticated model architectures to novel techniques for handling unknown species and balancing species classification with poisonous species identification, participants showcased groundbreaking solutions. They built on the findings of previous challenges, particularly in the encoding of metadata, demonstrating continuous advancement in the field.

The strict computational constraints imposed on submissions ensured that the resulting models were not only accurate but also practical for deployment in resource-limited environments and motivated participants to focus on principal improvements rather than training large ensembles of complex models. However, we also observed that enforcing the computational limits through a submission system caused additional technical difficulties with submission to some participants. Future editions shall thrive to further simplify the submission process for the participants.

With the advances in recognition accuracy for well-known species, we propose that future work should focus on the more challenging cases, specifically few-shot classification techniques. This approach, with its potential to push the fungal species recognition forward, would enable more robust identification of rare or newly discovered fungal species with limited training data. As the field progresses, the ultimate goal remains to develop robust, accurate, and accessible fungi recognition systems that can support both expert mycologists and citizen scientists in documenting and understanding fungal biodiversity.

Acknowledgments

LP and JM were supported by the Technology Agency of the Czech Republic, project No. SS05010008 and project No. SS73020004.

²This team encountered some issues while submitting to HuggingFace, but achieved better results than Top1 in their post-competition submissions.

References

- [1] M. Sulc, L. Picek, J. Matas, T. Jeppesen, J. Heilmann-Clausen, Fungi recognition: A practical use case, in: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2020, pp. 2316–2324.
- [2] L. Picek, M. Šulc, J. Matas, J. Heilmann-Clausen, T. S. Jeppesen, E. Lind, Automatic fungi recognition: Deep learning meets mycology, *Sensors* 22 (2022) 633.
- [3] T. G. Frøslev, J. Heilmann-Clausen, C. Lange, T. Læssøe, J. H. Petersen, U. Söchting, T. S. Jeppesen, J. Vesterholt, Danish mycological society, fungal records database (2019).
- [4] D. Taylor, T. Bruns, Community structure of ectomycorrhizal fungi in a pinus muricata forest: minimal overlap between the mature forest and resistant propagule communities, *Molecular Ecology* 8 (1999) 1837–1850.
- [5] L. Boddy, T. H. Jones, Interactions between basidiomycota and invertebrates, in: *British mycological society symposia series*, volume 28, Elsevier, 2008, pp. 155–179.
- [6] M. Sulc, L. Picek, J. Matas, Plant recognition by inception networks with test-time class prior estimation., in: *CLEF (Working Notes)*, 2018.
- [7] K.-H. Liu, M.-H. Yang, S.-T. Huang, C. Lin, Plant species classification based on hyperspectral imaging via a lightweight convolutional neural network model, *Frontiers in Plant Science* 13 (2022) 855660.
- [8] L. Picek, M. Šulc, J. Matas, T. S. Jeppesen, J. Heilmann-Clausen, T. Læssøe, T. Frøslev, Danish fungi 2020-not just another image recognition dataset, in: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 1525–1535.
- [9] A. Joly, L. Picek, S. Kahl, H. Goëau, V. Espitalier, C. Botella, B. Deneu, D. Marcos, J. Estopinan, C. Leblanc, T. Larcher, M. Šulc, M. Hruz, M. Servajean, et al., Overview of LifeCLEF 2024: Challenges on species distribution prediction and identification, in: *International Conference of the Cross-Language Evaluation Forum for European Languages*, Springer, 2024.
- [10] L. Picek, M. Šulc, J. Matas, J. Heilmann-Clausen, Overview of fungiclef 2022: Fungi recognition as an open set classification problem, *CEUR-WS*, 2022.
- [11] L. Picek, M. Šulc, R. Chamidullin, J. Matas, Overview of fungiclef 2023: Fungi recognition beyond 1/0 cost, in: *CLEF 2023-Conference and Labs of the Evaluation Forum*, 2023.
- [12] L. Picek, M. Šulc, J. Heilmann-Clausen, J. Matas, Overview of FungiCLEF 2022: Fungi recognition as an open set classification problem, in: *Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum*, 2022.
- [13] L. Picek, M. Šulc, J. Heilmann-Clausen, J. Matas, Overview of FungiCLEF 2023: Fungi recognition beyond 0-1 cost, in: *Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum*, 2023.
- [14] T. Lorieul, E. Cole, B. Deneu, M. Servajean, A. Durso, I. Bolon, H. Glotin, R. Planqué, R. R. de Castaneda, W.-P. Vellinga, et al., Overview of lifeclef 2021: An evaluation of machine-learning based species identification and species distribution prediction, in: *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 12th International Conference of the CLEF Association*, CLEF 2021, Virtual Event, September 21–24, 2021, *Proceedings*, volume 12880, Springer Nature, 2021, p. 371.
- [15] A. Joly, H. Goëau, S. Kahl, B. Deneu, M. Servajean, E. Cole, L. Picek, R. R. De Castaneda, I. Bolon, A. Durso, 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*, Springer, 2020, pp. 342–363.
- [16] A. Joly, H. Goëau, S. Kahl, L. Picek, C. Botella, D. Marcos, M. Šulc, M. Hruz, T. Lorieul, S. S. Moussi, et al., Lifeclef 2023 teaser: Species identification and prediction challenges, in: *European Conference on Information Retrieval*, Springer, 2023, pp. 568–576.
- [17] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in: *International conference on machine learning*, PMLR, 2019, pp. 6105–6114.
- [18] A. Joly, L. Picek, S. Kahl, H. Goëau, V. Espitalier, C. Botella, B. Deneu, D. Marcos, J. Estopinan,

- C. Leblanc, et al., Lifeclef 2024 teaser: Challenges on species distribution prediction and identification, in: European Conference on Information Retrieval, Springer, 2024, pp. 19–27.
- [19] L. Picek, M. Hruz, A. M. Durso, Overview of SnakeCLEF 2024: Revisiting snake species identification in medically important scenarios, in: Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum, 2024.
- [20] J. Etheredge, Openwgan-gp for fine-grained open-set fungi classification, in: Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum, 2024.
- [21] B.-F. Tan, Y.-Y. Li, P. Wang, L. Zhao, X.-S. Wei, Say no to the poisonous fungi: An effective strategy for reducing 0-1 cost in fungiclef2024, in: Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum, 2024.
- [22] S. Wolf, P. Thelen, J. Beyerer, Poison-aware open-set fungi classification: Reducing the risk of poisonous confusion, in: Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum, 2024.
- [23] C. Chiu, M. Heil, T. Kim, A. Miyaguchi, Fine-grained classification for poisonous fungi identification with transfer learning, in: Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum, 2024.
- [24] Z. Liu, H. Hu, Y. Lin, Z. Yao, Z. Xie, Y. Wei, J. Ning, Y. Cao, Z. Zhang, L. Dong, et al., Swin transformer v2: Scaling up capacity and resolution, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 12009–12019.
- [25] H. Ren, H. Jiang, W. Luo, M. Meng, T. Zhang, Entropy-guided open-set fine-grained fungi recognition, in: Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum, 2023.
- [26] Q. Diao, Y. Jiang, B. Wen, J. Sun, Z. Yuan, Metaformer: A unified meta framework for fine-grained recognition, arXiv preprint arXiv:2203.02751 (2022).
- [27] W. Yu, C. Si, P. Zhou, M. Luo, Y. Zhou, J. Feng, S. Yan, X. Wang, Metaformer baselines for vision, IEEE Transactions on Pattern Analysis and Machine Intelligence (2023).
- [28] S. Kong, D. Ramanan, Opengan: Open-set recognition via open data generation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 813–822.
- [29] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. C. Courville, Improved training of wasserstein gans, Advances in neural information processing systems 30 (2017).
- [30] L. Yang, X. Li, R. Song, B. Zhao, J. Tao, S. Zhou, J. Liang, J. Yang, Dynamic mlp for fine-grained image classification by leveraging geographical and temporal information, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 10945–10954.
- [31] M. Oquab, T. Darcet, T. Moutakanni, H. Vo, M. Szafraniec, V. Khalidov, P. Fernandez, D. Haziza, F. Massa, A. El-Nouby, et al., Dinov2: Learning robust visual features without supervision, arXiv preprint arXiv:2304.07193 (2023).
- [32] J. Wang, W. Zhang, Y. Zang, Y. Cao, J. Pang, T. Gong, K. Chen, Z. Liu, C. C. Loy, D. Lin, Seesaw loss for long-tailed instance segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 9695–9704.