

Overview of FungiCLEF 2022: Fungi Recognition as an Open Set Classification Problem

Lukáš Pícek¹, Milan Šulc², Jiří Matas³ and Jacob Heilmann-Clausen⁴

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

²Rossum.ai, Czech Republic

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

⁴Center for Macroecology, Evolution and Climate University of Copenhagen, Denmark

Abstract

The main goal of the new LifeCLEF challenge, FungiCLEF 2022: Fungi Recognition as an Open Set Classification Problem, was to provide an evaluation ground for end-to-end fungi species recognition in an open class set scenario. An AI-based fungi species recognition system deployed in the Atlas of Danish Fungi helps mycologists to collect valuable data and allows users to learn about fungi species identification. Advances in fungi recognition from images and metadata will allow continuous improvement of the system deployed in this citizen science project. The training set is based on the Danish Fungi 2020 dataset and contains 295,938 photographs of 1,604 species. For testing, we provided a collection of 59,420 expert-approved observations collected in 2021. The test set includes 1,165 species from the training set and 1,969 unknown species, leading to an open-set recognition problem. This paper provides (i) a description of the challenge task and datasets, (ii) a summary of the evaluation methodology, (iii) a review of the systems submitted by the participating teams, and (iv) a discussion of the challenge results.

Keywords

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

1. Introduction

Automatic recognition of fungi species assists mycologists, citizen scientists and nature enthusiasts in species identification in the wild [1, 2]. Its availability supports the collection of valuable biodiversity data. In practice, species identification typically does not depend solely on the visual observation of the specimen but also on other information available to the observer — such as habitat, substrate, location and time. The main goal for the new FungiCLEF competition was to provide an evaluation ground for automatic methods for fungi recognition in an open class set scenario, i.e. the submitted methods have to handle images of unknown species. Similarly to previous LifeCLEF competitions, The competition was hosted on *Kaggle* primarily to attract machine learning experts to participate and present their ideas. Thanks to rich metadata, precise annotations, and baselines available to all competitors, the challenge provides a benchmark for image recognition with the use of additional information.

CLEF 2022: Conference and Labs of the Evaluation Forum, September 5–8, 2022, Bologna, Italy

✉ picekl@kky.zcu.cz (L. Pícek)

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

0000-0003-4713-6004 (J. Heilmann-Clausen)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).


 CEUR Workshop Proceedings (CEUR-WS.org)



Figure 1: Three fungi specimen observations from the Atlas of Danish Fungi dataset [4]. *Atlas of Danish Fungi*: ©Bedřiška Picková and ©Jan Riis-Hansen and ©Arne Pedersen.

2. Challenge description

The new FungiCLEF 2022 challenge: Fungi Recognition as an Open Set Classification Problem, was organized in conjunction with the Conference and Labs of the Evaluation Forum (CLEF¹) and LifeCLEF² research platform [3], and FGVC9 Workshop³ – The Ninth Workshop on Fine-Grained Visual Categorization organized within the CVPR conference.

The main goal for this challenge was to return the species with the highest likelihood (or "*unknown*") for each given test observation, consisting of a set of images and metadata – the information about habitat, substrate, location, and more is provided for each observation. Photographs of unknown fungi species had to be classified into an "*unknown*" class with label id –1. The baseline procedure to include metadata in the decision problem and baseline pre-trained image classifiers were provided as part of the task description to all participants. Sample observations are visualized in Figure 1. Each row represents one observation.

¹ <http://www.clef-initiative.eu/>

² <http://www.lifeclef.org/>

³ <https://sites.google.com/view/fgvc9/home>

Table 1
FungiCLEF 2022 dataset statistics for each subset.

Subset	Species	Known Species	Unknown Species	Images	Observations
Training	1,604	1,604	0	266,344	×
Validation	1,604	1,604	0	29,594	×
Test	3,134	1,165	1,969	118,676	59,420

2.1. Dataset

The FungiCLEF 2022 dataset is based on data collected through the Atlas of Danish Fungi Web⁴ and mobile (iOS⁵ and Android⁶) applications. The Atlas of Danish Fungi is a citizen science platform with more than 4,000 actively contributing volunteers and with more than 1 million content-checked observations of approximately 8,650 fungi species.

Development set: For training, the competitors were provided with the DanishFungi 2020 (DF20) dataset [4]. DF20 contains 295,938 images — 266,344 for training and 29,594 for validation — belonging to 1,604 species. All training samples passed an expert validation process, guaranteeing high quality labels. Furthermore, rich observation metadata about habitat, substrate, time, location, EXIF etc. are provided.

Test set: The test dataset is constructed from all observations submitted in 2021, for which expert-verified species labels are available. It includes observations collected across all substrate and habitat types. The test set contains 59,420 observations with 118,676 images belonging to 3,134 species: 1,165 known from the training set and 1,969 unknown species covering approximately 30% of the test observations. The test set was further split into public (20%) and private (80%) subsets — a common practice for Kaggle competitions to prevent participants from overfitting to the leaderboard.

2.2. Metadata

The visual data is accompanied by metadata for approximately 99% of the image observations and includes information about attributes related to the environment, place, time and taxonomy. The provided metadata is acquired by citizen scientists and enables research directions on combining visual data with metadata. We include 21 frequently filled-in attributes. The most important attributes are listed and described below.

Substrate: Substrates on which fungi live and fruit are an essential source of information that helps differentiate similarly-looking species. Each species or genus has its preferable substrate, and it is rare to find it on other substrates. We provide one of 32 substrate types for more than 99% of images. We differentiate wood of living trees, dead wood, soil, bark, stone, fruits and others.

⁴<https://svampe.databasen.org/>

⁵<https://apps.apple.com/us/app/atlas-of-danish-fungi/id1467728588>

⁶<https://play.google.com/store/apps/details?id=com.noque.svampeatlas>

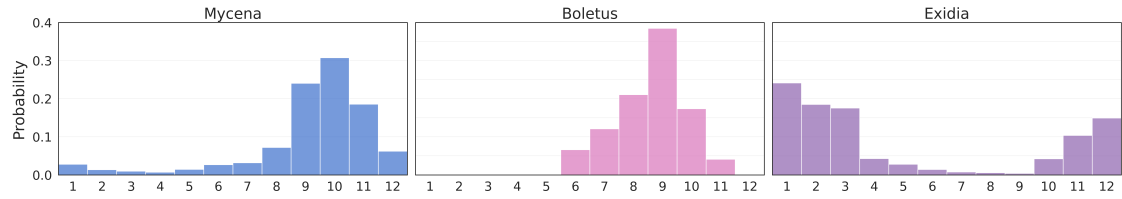


Figure 2: Monthly observations distribution in the FungiCLEF 2022 training dataset. Three genera: *Mycena*, *Boletus*, and *Exidia*. Image taken from [4].

Habitat: While substrate denotes the spots, the habitat indicates the overall environment where fungi grow, which is vital for fungal recognition. We include the information about the habitat for 99.5% of observations.

Location: Fungi are highly location-dependent. We include multi-level location information. Starting from GPS coordinates with included uncertainty, we further extracted information about the country, region and district.

Time-Stamp: Observation time is essential for fungi classification in the wild as fruitbodies' presence depends on seasonality or even the time in a day. Figure 2 shows the monthly observation frequency for three genera.

EXIF data: Since the camera device and its settings affect the resulting image, the image classification models may be biased towards specific device attributes. To allow a deeper study of such phenomena, we include the EXIF data for approximately 84% of images. We included attributes such as White Balance, Color Space, Metering Mode, Aperture, Device, Exposure Time and Shutter Speed.

2.3. Timeline

The competition and data were published in February 2022 through the LifeCLEF, Kaggle, and FGVC challenge pages allowing anyone with research ambitions to register and participate in the competition. The test data were provided jointly with the training data allowing continuous evaluation. Each team could submit up to 2 submissions a day. The deadline for challenge submissions was May 16, setting the competition for roughly three months. Participants submitted CSV files containing the Top1 prediction for each fungi observation. Once the submission phase was closed (mid-May), the participants were allowed to submit post-competition submissions to evaluate any exciting findings.

2.4. Evaluation Protocol

The evaluation process consisted of two stages: (i) a public evaluation on the public subset (20%) of the test set, which was available during the whole competition with a limit of two submissions a day, and (ii) a final evaluation on the private test set (80%) after the challenge deadline. The main evaluation metric for the competition was the F_1^m , defined as the mean of class-wise F_1 scores:

$$F_1^m = \frac{1}{N} \sum_{s=1}^N F_{1_s}, \quad (1)$$

where N represents the number of classes – in case of the Kaggle evaluation, $N = 1,165$ (#classes in the test set) – and s is the species index. The F_1 score for each class is calculated as a harmonic mean of the class precision P_S and recall R_S :

$$F_{1_s} = 2 \times \frac{P_s \times R_s}{P_s + R_s}, \quad P_s = \frac{tp_s}{tp_s + fp_s}, \quad R_s = \frac{tp_s}{tp_s + fn_s} \quad (2)$$

In single-label multi-class classification, the True Positives (tp) of a species represents the number of correct Top1 predictions of that species, False Positive (fp) denotes how many times was different species predicted instead of the (tp), and False Negatives (fn) indicates how many images of species s have been wrongly classified.

2.5. Working Notes

All participants with valid submissions were asked to provide a *Working Note* paper – a technical report with information needed to reproduce the results of all submissions. All submitted *Working Notes* were reviewed by 2–3 reviewers. The review process was single-blind and offered up to two rebuttals. The acceptance rate was 75%.

3. Challenge Results

The official challenge results, based on the F_1^m score, are displayed in Figure 3. The best performing team – *xiong* – achieved F_1^m of 80.43% on the private test set and an accuracy of 65.69% on the complete test set. We note that the order would be different in terms of accuracy, as shown in Figure 4, where the best accuracy of 67.08% on the full rest set was achieved by team *GG*, primarily due to a high number of correctly identified out-of-scope observations. In the case of the out-of-scope (OoS) identification performance, i.e. what proportion of out-of-scope observations has been correctly classified as OoS, the best performing team with 44.55% correctly categorized observations was one of the worst-performing teams in terms of F_1^m . As also displayed in Figure 4 participants identified less than 5% OoS observations and only four teams achieved accuracy over 10% on out-of-scope observations. In Figure 5 we have evaluated the species toxicity confusion on the full test set for all the participants, i.e., how often poisonous species are confused for edible ones and vice versa. Interestingly, the more critical confusion where poisonous fungi were misclassified as edible is relatively high even for the best scoring models – 5.70% and 6.63% for team *GG* and team *xiong*, respectively.

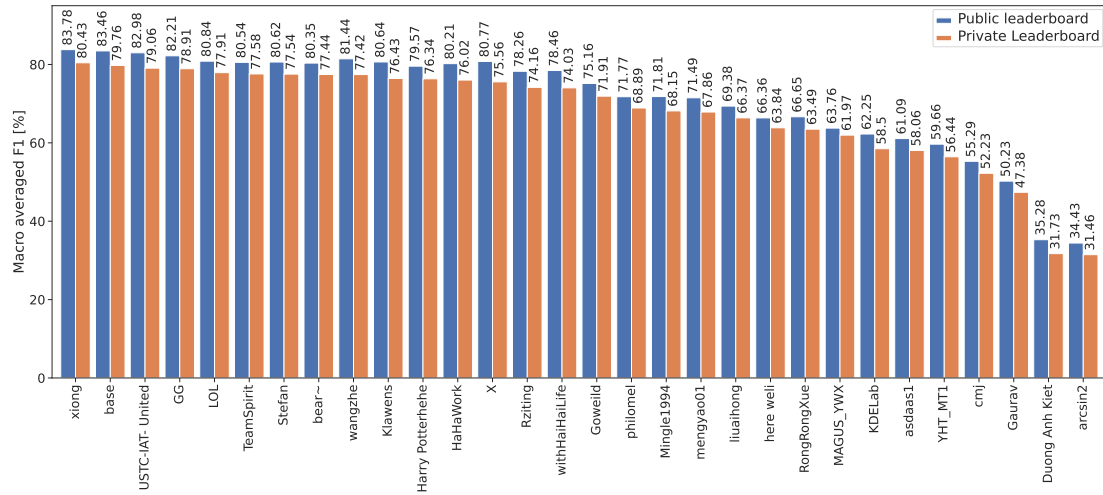


Figure 3: Official FungiCLEF 2022 competition results, sorted by performance on the private set.

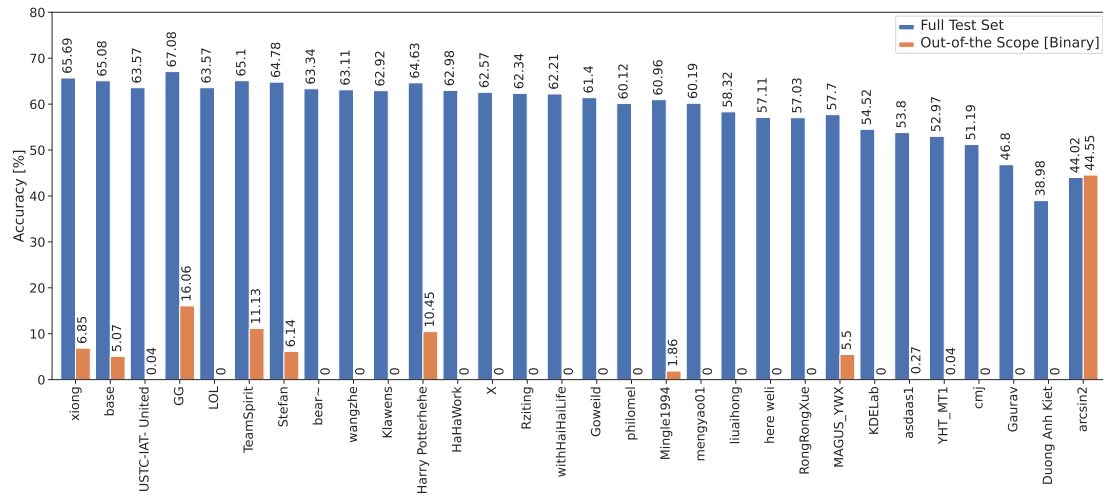


Figure 4: Out-of-scope identification performance on the full test set, i.e. what proportion of out-of-scope observations has been correctly classified as out-of-scope, compared to the accuracy over all observations.

4. Participants and Methods

In total, 38 teams contributed with 701 valid submissions to the challenge evaluation on Kaggle. The results on the public and private test sets (leaderboards) are displayed in Figure 3. Below we summarize the approach of teams with published working notes. More details can be found in the individual working notes of participants [5, 6, 7, 8, 9, 10] which passed the review process, ensuring a sufficient level of reproducibility and quality.

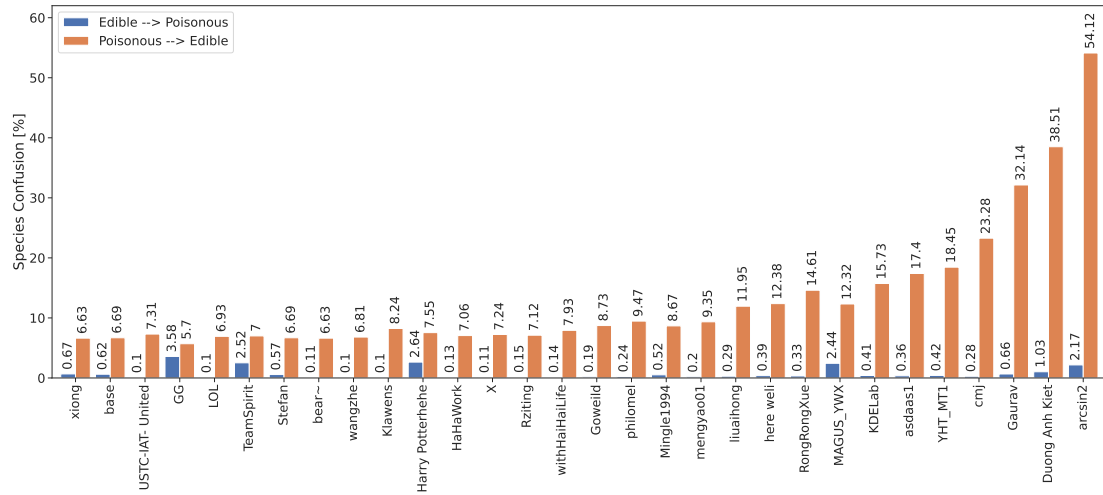


Figure 5: Species toxicity confusion on the full test set: Poisonous \rightarrow Edible denotes poisonous fungi that were misclassified as edible, and Edible \rightarrow Poisonous denotes edible fungi misclassified as poisonous.

xiong [6]: The winning submission by Xiong et al., achieving an impressive F_1^m score of 80.43% on the private test set, used an ensemble of MetaFormer [11] and ConvNext [12] networks. The provided metadata were utilized as inputs to the MetaFormer architecture. To battle the long-tailed distribution of species, the authors used the Seasaw loss [13]. Additional improvements were achieved by test-time augmentation, adding a model trained with pseudo-labels to the ensemble, and adding a thresholding post-processing to deal with out-of-scope observations.

USTC-IAT-United [7]: The submission by Yu et al. used an ensemble of several CNN and Transformer architectures: Metaformer [11], SwinTransformer[14], EfficientNet [15], ViT (Vision Transformer) [16], BEiT [17]. The team scored 3rd with 79.06% of F_1^m score on the private test set. In their working notes, the authors explore the impact of different data augmentation techniques, model architectures, loss functions, and attention mechanisms on the classification performance.

GG [8]: Shen et al. introduced a novel architecture CoLKANet based on VAN (Visual Attention Network) [18] and CoAtNet [19]. It is a combination of large kernel attention and vision transformer. The proposed CoLKANet outperforms Swin [14] and VOLO [20] models in terms of F_1^m by 2.3 and 1.9 percentage points, respectively. ConvNeXt [12] performed similarly to the proposed CoLKANet architecture. Furthermore, the team used techniques such as Label Aware Smoothing [21], Pseudo labelling for tail classes and various augmentation techniques. When TrivialAugment [22] was deployed during the middle stage of experimentation, the team observed a rise in F_1^m of around 0.5%. Progressively, Random Erasing [23], CutMix [24] and Mixup [25] were added, which helped with regularization. The final submission score was achieved by an ensemble of five models: $2 \times$ ConvNeXt, VOLO, Swin, and CoLKANet. The novel CoLKANet is an interesting contribution with potential outside this competition’s scope.

TeamSpirit [5]: Fan et al., who scored sixth in the challenge with 77.58% F_1^m score, propose an image classification method called Class-wise Weighted Prototype Classifier. CWPC decouples closed-set training and open-set inference by constructing class centers from the training set features and their prediction scores. A hard classes mining strategy and the LDAM loss [26] were used to cope with the long-tailed distribution of species. This team encoded the metadata using a multilingual BERT model [27] with RoBERTa [28].

Stefan [10]: Wolf and Beyerer refrained from using ensembles of multiple models, and — for the sake of model simplicity — focused on developing a strong single-model submission. The method is based on a Swin Transformer Large backbone [14], a class-balancing training scheme [29], heavy data augmentation [30] and thresholding the softmax scores to cope with out-of-scope observations. The team scored 7th, in the challenge with 77.54% F_1^m score.

SSN [9]: This team experimented with several ResNet [31], ResNeXt [32], and EfficientNet [33] architectures. For their best submission, feature vectors from two selected architectures, EfficientNetB4 and ResNeXt101, were concatenated with a categorical representation of metadata. The resulted features were later used for training the XGBoost Ensemble Classifier [34]. An interesting benefit of the XGBoost algorithm is that the relative importance of the ensemble features is computed; thus, each feature might be observed and studied. With an absolute F_1^m performance of 48.96%, the XGBoost algorithm with two CNN backbones poses a unique approach for the classification, even though performing worst compared to other participants.

5. Conclusions

This paper presents an overview and results of the first edition of the FungiCLEF challenge organized in conjunction with the Conference and Labs of the Evaluation Forum (CLEF⁷), LifeCLEF⁸ research platform [35] and FGVC.

All submissions with working notes were based on modern Convolutional Neural Network (CNN) or transformer-inspired architectures, such as Metaformer [11], Swin Transformer [14], and BEiT [17]. The best performing teams used ensembles of both CNNs and Transformers. The winning team [6] achieved 80.43% accuracy with a combination of ConvNext-large [12] and MetaFormer [11]. The results were often improved by combining predictions belonging to the same observation and by both training-time and test-time data augmentations.

Participants experimented with a number of different training losses to battle the long tail distribution and fine-grained classification with small inter-class differences and large intra-class differences: besides the standard Cross Entropy loss function, we have seen successful applications of the Seesaw loss [13], Focal loss [36], Arcface loss [37], Sub-Center loss [38] and Adaptive Margin [39].

We were happy to see the participants experimented with different use of the provided

⁷ <http://www.clef-initiative.eu/>

⁸ <http://www.lifeclef.org/>

observation metadata, which often lead to improvements in the recognition scores. Besides the probabilistic baseline published with the dataset [4], we have seen hand-crafted encoding of the metadata into feature vectors, as well as encoding of the metadata with a multilingual BERT model [27] and RoBERTa [28]. The metadata were then combined with image features extracted from a CNN or Transformer image classifier, or directly used as an input to Metaformer [11].

The results of participants' comprehensive experiments with model architectures, loss functions and usage of metadata in fine-grained image-classification will help to improve species recognition services that aid researchers, citizen-science communities and nature enthusiasts. As discussed in Section 3, there is still a great space for improvement in the recognition of out-of-scope classes. Our evaluation of classification errors identified that confusion of poisonous mushrooms for edible is much more common than confusion of edible mushrooms for poisonous. This could be critical in applications that may affect the decision to consume a mushroom, and presents an important aspect to address in the future work.

Acknowledgments

LP was supported by the UWB grant, project No. SGS-2022-017. LP was supported by the Technology Agency of the Czech Republic, project No. SS05010008.

References

- [1] M. Šulc, L. Pícek, 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. Pícek, M. Šulc, J. Matas, J. Heilmann-Clausen, T. S. Jeppesen, E. Lind, Automatic fungi recognition: Deep learning meets mycology, *Sensors* 22 (2022) 633.
- [3] A. Joly, H. Goëau, S. Kahl, L. Pícek, T. Lorieul, E. Cole, B. Deneu, M. Servajean, A. Durso, H. Glotin, R. Planqué, W.-P. Vellinga, A. Navine, H. Klinck, T. Denton, I. Eggel, P. Bonnet, M. Šulc, M. Hruz, 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, Springer, 2022.
- [4] L. Pícek, 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.
- [5] G. Fan, C. Zining, W. Weiqiu, S. Yinan, S. Fei, Z. Zhicheng, C. Hong, Does closed-set training generalize to open-set recognition?, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.
- [6] Z. Xiong, Y. Ruan, Y. Hu, Y. Zhang, Y. Zhu, S. Guo, W. Zhu, B. Han, An empirical study for fine-grained fungi recognition with transformer and convnet, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.
- [7] J. Yu, H. Chang, K. Lu, G. Xie, L. Zhang, Z. Cai, S. Du, Z. Wei, Z. Liu, F. Gao, F. Shuang, Bag of tricks and a strong baseline for fgvc, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.

- [8] Y. Shen, X. Sun, Z. Zhu, When large kernel meets vision transformer: A solution for snakeclef & fungiclef, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.
- [9] K. Desingu, A. Bhaskar, M. Palaniappan, E. A. Chodisetty, H. Bharathi, Classification of fungi species: A deep learning based image feature extraction and gradient boosting ensemble approach, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.
- [10] S. Wolf, J. Beyerer, Transformer-based fine-grained fungi classification in an open-set scenario, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.
- [11] 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).
- [12] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, S. Xie, A convnet for the 2020s, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 11976–11986.
- [13] 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.
- [14] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10012–10022.
- [15] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in: International conference on machine learning, PMLR, 2019, pp. 6105–6114.
- [16] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929 (2020).
- [17] H. Bao, L. Dong, F. Wei, Beit: Bert pre-training of image transformers, arXiv preprint arXiv:2106.08254 (2021).
- [18] M.-H. Guo, C.-Z. Lu, Z.-N. Liu, M.-M. Cheng, S.-M. Hu, Visual attention network, arXiv preprint arXiv:2202.09741 (2022).
- [19] Z. Dai, H. Liu, Q. V. Le, M. Tan, Coatnet: Marrying convolution and attention for all data sizes, Advances in Neural Information Processing Systems 34 (2021) 3965–3977.
- [20] L. Yuan, Q. Hou, Z. Jiang, J. Feng, S. Yan, Volo: Vision outlooker for visual recognition, arXiv preprint arXiv:2106.13112 (2021).
- [21] Z. Zhong, J. Cui, S. Liu, J. Jia, Improving calibration for long-tailed recognition, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 16489–16498.
- [22] S. G. Müller, F. Hutter, Trivialaugment: Tuning-free yet state-of-the-art data augmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 774–782.
- [23] Z. Zhong, L. Zheng, G. Kang, S. Li, Y. Yang, Random erasing data augmentation, in: Proceedings of the AAAI conference on artificial intelligence, volume 34, 2020, pp. 13001–13008.
- [24] S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, Y. Yoo, Cutmix: Regularization strategy to train

- strong classifiers with localizable features, in: Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 6023–6032.
- [25] H. Zhang, M. Cisse, Y. N. Dauphin, D. Lopez-Paz, mixup: Beyond empirical risk minimization, arXiv preprint arXiv:1710.09412 (2017).
- [26] K. Cao, C. Wei, A. Gaidon, N. Arechiga, T. Ma, Learning imbalanced datasets with label-distribution-aware margin loss, Advances in neural information processing systems 32 (2019).
- [27] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).
- [28] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach, arXiv preprint arXiv:1907.11692 (2019).
- [29] A. Gupta, P. Dollar, R. Girshick, Lvis: A dataset for large vocabulary instance segmentation, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 5356–5364.
- [30] E. D. Cubuk, B. Zoph, J. Shlens, Q. V. Le, Randaugment: Practical automated data augmentation with a reduced search space, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 2020, pp. 702–703.
- [31] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [32] S. Xie, R. Girshick, P. Dollar, Z. Tu, K. He, Aggregated Residual Transformations for Deep Neural Networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [33] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, in: International Conference on Machine Learning, PMLR, 2019, pp. 6105–6114.
- [34] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, ACM, New York, NY, USA, 2016, pp. 785–794. URL: <http://doi.acm.org/10.1145/2939672.2939785>. doi:10.1145/2939672.2939785.
- [35] A. Joly, H. Goëau, S. Kahl, L. Picek, T. Lorieul, E. Cole, B. Deneu, M. Servajean, R. Ruiz De Castañeda, I. Bolon, H. Glotin, R. Planqué, W.-P. Vellinga, A. Dorso, H. Klinck, T. Denton, I. Eggel, P. Bonnet, H. Müller, Overview of lifeclef 2021: a system-oriented evaluation of automated species identification and species distribution prediction, in: Proceedings of the Twelfth International Conference of the CLEF Association (CLEF 2021), 2021.
- [36] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 2980–2988.
- [37] J. Deng, J. Guo, N. Xue, S. Zafeiriou, Arcface: Additive angular margin loss for deep face recognition, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4690–4699.
- [38] J. Deng, J. Guo, T. Liu, M. Gong, S. Zafeiriou, Sub-center arcface: Boosting face recognition by large-scale noisy web faces, in: European Conference on Computer Vision, Springer, 2020, pp. 741–757.
- [39] H. Liu, X. Zhu, Z. Lei, S. Z. Li, Adaptiveface: Adaptive margin and sampling for face

recognition, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 11947–11956.