Evaluating Self-Supervised Foundation Models in Holographic Imaging

Silas Dietler [*1] **Yanick Zeder** [2] **Elias Graf** [2] **Kilian Koch** [2] **Andreas Schwendimann** [2]
Tommaso Bendinelli [*1]

Abstract

DINOv2, a large self-supervised computer vision foundation model, has achieved impressive performance on downstream tasks like classification, segmentation, and depth estimation. This success suggests the idea that universal features can be extracted through large-scale pre-training. However, its applicability beyond natural image domains remains relatively unexplored. This study aims to contribute in this direction, by exploring the potential of DINOv2 for a niche but important task: pollen classification based on holographic images. Our findings reveal that features learned by the network in the natural image domain are not informative for this task. However, when DINOv2 is pre-trained on a pollen-specific dataset, it achieves superior performance compared to supervised methods, especially in scenarios with limited data. This superior performance opens doors for new applications such as online few-shot (bio)aerosol particle classification with holographic imaging.

1. Introduction

Self-supervised pre-training has become the dominant paradigm in Natural Language Processing (NLP) for several years now. This approach utilizes vast amounts of data and computational resources to train large-scale networks, commonly referred to as foundation models. These foundation models have a remarkable generalizability, and power applications like ChatGPT and Gemini, which are utilized by millions on a daily basis. (Team et al., 2023; Achiam et al., 2023; Touvron et al., 2023). In computer vision although self-supervised learning has long been explored (Jing & Tian, 2020), only recently it has been scaled to billion parameters (He et al., 2022; Oquab et al., 2023) and shown to follow similar scaling laws observed in NLP: more compute and data yield increasingly better performance.

Of particular interest is DINOv2 (Oquab et al., 2023). Once pre-trained, DINOv2 can extract robust features from images that perform on par with supervised baselines in natural image domains. Additionally, these features, compared to other large-scale image foundation models, can be directly utilized without the need of any network finetuning. Specifically, they can be used together with k-nearest neighbors for tasks such as instance retrieval and classification and with a linear layer for tasks such as segmentation and depth estimation. This significantly simplifies the development of applications built on top of this model.

Despite its recent release, this model has already been used in applications outside the natural image domain, such as radiology Pérez-García et al. (2024) and histopathology Chen et al. (2023). These studies demonstrate that DINOv2 pre-trained on natural images (LVD-142M dataset) perform poorly on domain-specific downstream tasks. However, by repeating the pre-training process with domain-specific data, DINOv2 can achieve performance comparable to, or even exceeding, supervised learning methods on these specialized task. In this study we aim to bring additional evidence for the potential of DINOv2 for domain specific tasks by testing it on pollen recognition from holographic images.

Holographic images, generated using the method described by Berg (2022), find applications in aerosol measurement (Sauvageat et al., 2020) and cloud observation, including the study of particles and ice crystals (Beck et al., 2017). One significant application is the identification of airborne pollen, spores, and microplastics (Sauvageat et al., 2020; Erb et al., 2024; Beres et al., 2023). Pollen monitoring is crucial for managing public health concerns, given the rising number of allergy sufferers and the associated costs, estimated to range between 50 to 150 billion euros in Europe alone (Tummon et al., 2024; Zuberbier et al., 2014). Traditional pollen monitoring relies on the manual method using Volumetric Hirst traps (HIRST, 1952). However, new real-time pollen monitoring systems are emerging to support or replace traditional methods. This study utilizes data measured by one such device, the SwisensPoleno (Sauvageat et al., 2020).

[*]Equal contribution [1]CSEM SA, Alpnach OW, Switzerland [2]Swisens AG, Emmen LU, Switzerland. Correspondence to: Silas Dietler <silas.dietler@csem.ch>, Tommaso Bendinelli <tommaso.bendinelli@csem.ch>.

Published at ICML 2024 Workshop on Foundation Models in the Wild.

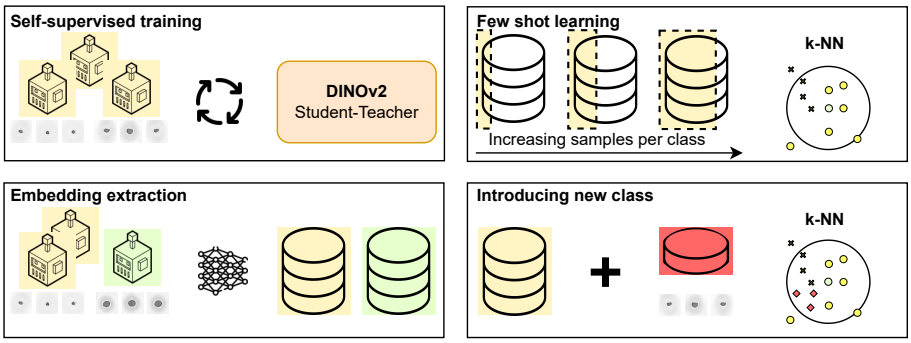

Figure 1. Visualization of the experimental setup utilized in this study to self-supervised train a DINOv2 model on holographic images (Holo ViT) and evaluate its performance across several experiments. **Top left**: Self-supervised training of the DINOv2 student-teacher architecture, where the model learns representations from the holographic training dataset (yellowish) without labeled data. **Bottom left**: Embedding extraction using the trained teacher model to generate embeddings from both the training and test datasets (greenish) which were collected by different instruments. **Top right**: Few-shot learning setup experimenting with the use of subsets of the data with varying sample sizes per class to evaluate k-NN performance on the test set (section 3.1). **Bottom right**: Introduction of new classes in a few-shot learning scenario, showing the addition of new class samples and subsequent k-NN evaluation without retraining the model (section 3.2).

1.1. Contributions

We benchmark DINOv2 against a strong supervised baseline on the task of (bio)aerosol particle classification using holographic images. Our results demonstrate that:

1. A k-NN classifier using DINOv2 features from the LVD-142M dataset outperforms random guessing but lags behind supervised methods.

2. DINOv2 pre-trained on holographic data matches supervised methods on the full dataset and excels in low data regimes with high-quality pre-training data.

3. New classes can be added to DINOv2 classification pipeline without retraining, demonstrating flexibility compared to retrained supervised methods.

2. Preliminary and Experimental Setup

We evaluated DINOv2's performance on our holographic pollen dataset on two tasks: image classification and the integration of new classes after initial self-supervised pretraining. We compared two main configurations: the officially released ViT model pre-trained on the LVD-142M dataset, and a self-supervised ViT model pre-trained on our holographic dataset (Holo ViT). For both, k-Nearest Neighbors (k-NN) served as the final classifier. Additionally, we employed EfficientNet, a well-established supervised learning model, as a strong baseline for comparison. Figure 1 depicts our experimental setup of the self-supervised DINOv2 training and its downstream tasks.

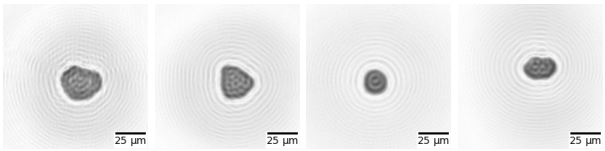

Figure 2. Examples of holographic Pollen images. Classes from left to right: Poaceae, Corylus, Ambrosia, and Betula.

2.1. Holographic Pollen Datasets

Samples from 18 different pollen classes were collected in a laboratory-controlled environment using the SwisensPoleno instrument (Sauvageat et al., 2020). Specifically, the SwisensAtomizer was employed to disperse pollen samples in a clean chamber, where the SwisensPoleno captured each measurement event as a set of two holographic images. These images have a resolution of 200 by 200 pixels with 16-bit grayscale values, with each pixel representing an area of 0.595 by 0.595 micrometers.

After post-processing, we obtained a total of 337'000 images across the 18 classes. During this step, we removed images which were empty or had poor focus. Additionally, we eliminated some images using simple statistical methods, such as comparing the expected shape and size of the particles with the measured ones.

In a similar manner, we collected test sets for the same pollen species using a different instrument, resulting in a total of 57'000 images. Figure 2 illustrates examples of the holographic images generated by the SwisensPoleno. Further details regarding the specific composition of the dataset can be found in the Appendix C.

2.2. DINOv2 Implementation Details

We leveraged the official DINOv2 repository (Oquab et al., 2024) to pre-train various sizes of Vision Transformers (ViTs) on holographic images of pollen, using a self-supervised approach. Modifications to the released standard configuration were implemented as suggested in the original paper. These modifications included replacing the MLP head with a SwiGLU Feed Forward Network (FFN), applying Sinkhorn-Knopp centering, using a patch size of 16 and untangling the heads of DINO and iBOT during training. Additionally, we integrated four registers into our models as described in Darcet et al. (2024).

Model performance was evaluated through a k-nearest neighbors (k-NN) methodology, using teacher model checkpoints to select the configurations that demonstrated optimal validation performance on the holographic dataset.

The computational infrastructure for our self-supervised training consisted of four NVIDIA GeForce RTX 3090 GPUs, each equipped with 24 GB of memory. This setup supported batch sizes of 32, 64, and 128 per GPU, with total batch sizes scaled to 128, 256, and 512. To ensure stable training sessions at these scales, we adjusted the initial learning rate to 0.007, followed by the officially implemented batch size-dependent factorization.

The ViT giant backbone, along with its distilled versions incorporating registers, was directly sourced from the official Github repository associated with the original DINOv2 paper.

2.3. Supervised Benchmark: EfficientNet

To compare the performance of the DINOv2 framework with a supervised learning approach, we employed the pre-trained EfficientNet B0, B4 and B7 architecture (Tan & Le, 2019). The selection of EfficientNet is based on its demonstrated effectiveness in achieving high accuracy with significantly reduced computational requirements making it an ideal benchmark. A classification head, comprising Layer Normalization, Dropout, and a ReLU activation function, was appended to the backbone. Performance metrics were gathered based on the model that demonstrated the highest micro-accuracy on the validation dataset. Throughout the training phase, the entire model was set to trainable, employing an Adam optimizer for adjustments. To augment the robustness of the model, several data augmentation techniques were applied, including zoom, rotation, flipping, Gaussian blur, color jitter, and translation.

2.4. k-NN

For the classification task on top of the ViT CLS token-embeddings, we opted for a $k = 10$ for all the experiments except for those with only two and five samples per class, where the k is set equal to the number of samples. The embeddings were normalized with a normalization layer.

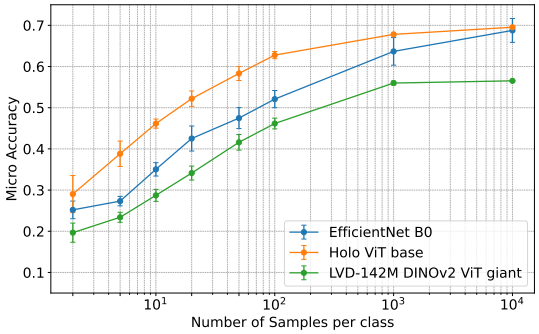

Figure 3. Micro-accuracy of various models in few-shot learning with holographic pollen images. Comparing supervised EfficientNet B0 (blue), official LVD-142 DINOv2 ViT giant (green), and self-supervised Holo ViT base (orange) using k-NN embeddings. The Holo ViT base shows higher accuracy and stability in low-data regimes. In the Appendix B distilled versions of the LVD-142 ViT giant are shown.

3. Experiments and Results

In the following subsections, we report the results of our experiments. To enhance statistical significance, we performed each experiment five times with a different random sample of training data. For each sample, we repeated the k-NN search for DINOv2 and the training of EfficientNet.

3.1. Classification performance

This experiment assessed few-shot learning capabilities across a range of dataset sizes, specifically 2, 5, 10, 20, 50, 100, 1'000, and $min(10'000, \text{total class samples})$. The total samples per class are reported in Table 2 in the appendix. Figure 3 illustrates the results. The Holographic Vision Transformer base (Holo ViT base) model, specifically the one trained on domain-specific data, maintains higher micro-accuracy across various few-shot scenarios compared to EfficientNet B0. Furthermore the error bars indicate that the Holo ViT model not only achieves higher average accuracy but also offer greater stability across different seeds, suggesting robustness in scenarios with scarce data. The Figure 6 in the Appendix B visualizes the performance of the distilled versions of the LVD-142M DINOv2 giant model.

Furthermore, in the Appendix A, we further evaluate the performance of various model sizes on the complete dataset to compare the capabilities of different training strategies.

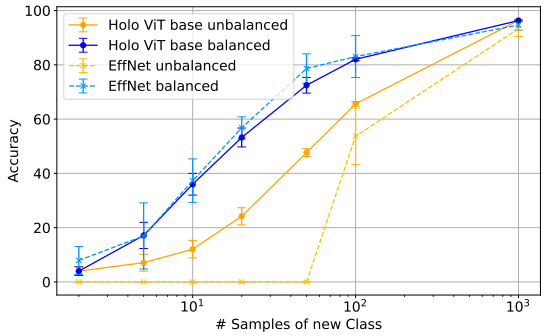

Figure 4. Accuracy of the new class with varying sample sizes, with all other classes kept at 1'000 samples each. Performance of Holo ViT base is compared with (balanced) and without SMOTE (unbalanced), and EfficientNet B0 with unbalanced and upsampled training. Figure 5 visualizes the corresponding precision.

3.2. Introduction of new classes

This experiment assessed the capability to incorporate new, previously unseen classes into the classification system without necessitating retraining, neither self-supervised nor supervised. This was conducted by extracting embeddings using a holographic self-supervised trained DINOv2 ViT base backbone. We obtained embeddings from the 18 classes, each with 1'000 randomly chosen samples and subsequently a new class was added, starting with 2 and incrementally increasing to 1'000 samples. Additionally, we employed the Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al., 2002) to augment the minority class from its original sample size to 1'000.

For comparative analysis, a parallel experiment was conducted using a supervised EfficientNet B0 model. This involved the same data-splits and instead of applying SMOTE to ensure balance in one part of the experiment, we used basic minority upsampling during training time.

Figure 4 illustrates the accuracy of DINOv2 and EfficientNet B0 for the new class. When few samples are available, DINOv2 outperforms EfficientNet B0 without any upsampling technique. This is likely due to optimization challenges faced by EfficientNet B0 with limited data.

Interestingly, with oversampling enabled, both models achieve similar performance. However, DINOv2 offers a clear implementation advantage. EfficientNet needs retraining the entire network for each new class, while DINOv2 only requires updating its vector database. This makes DINOv2 the preferable choice for online applications where efficiency and adaptability are crucial.

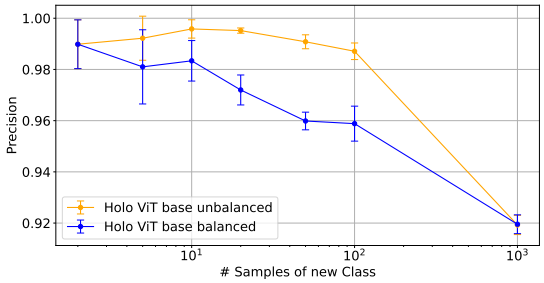

Figure 5. Precision of the new class with varying sample sizes, keeping all other classes constant at 1'000 samples each. Comparing Holo ViT base with (balanced) and without SMOTE (unbalanced). Complements the micro-accuracy results from Figure 4, showing precision stability of Holo ViT base.

4. Conclusion

In this study, we benchmarked the self-supervised learning approach DINOv2 on holographic data against a strong supervised baseline. Our results show that DINOv2, when pretrained on relevant data, outperforms the supervised baseline in the few shot learning regime. Furthermore, DINOv2's ability to adapt to new pollen classes without retraining opens up new practical applications. In the future, we aim to explore how curating the underlying self-supervised pollen dataset affects downstream classification performance and whether the hyperparameter configuration used in the original paper, including augmentations and cropping size, are optimal for the holographic image domain.

Acknowledgments

We gratefully acknowledge the contributions of the following individuals and organizations to our research:

- **Funding:** This project is supported by the Chips Joint Undertaking (Grant Agreement No. 101095835) and its members Finland, Germany, Ireland, the Netherlands, Sweden, Switzerland. This work includes top-up funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).

- **Data Providers:** We thank Sophie Erb, Bernard Clot, Benoît Crouzy, and Fiona Tummon from MeteoSwiss (Federal Office of Meteorology and Climatology, Chemin de l'Aérologie, CH-1530 Payerne); Carola Emanuel from Umweltbundesamt Berlin (Wörlitzer Platz 1, DE-06844 Dessau-Roßlau); Julia Burkart from GeoSphere (Hohe Warte 38, AT1190 Wien); Kyu Rang Kim and YoungJong Han from the National Institute of Meteorological Sciences (Research Applications Department, Seogwipo-si, Republic of Korea); and David O'Connor from Dublin City University (DCU Glasnevin Campus, Dublin 9, Ireland) for providing critical datasets for algorithm training.

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

Beck, A., Henneberger, J., Schöpfer, S., Fugal, J., and Lohmann, U. Hologondel: in situ cloud observations on a cable car in the swiss alps using a holographic imager. *Atmospheric Measurement Techniques*, 10(2):459–476, 2017. doi: 10.5194/amt-10-459-2017. URL https://amt.copernicus.org/articles/10/459/2017/.

Beres, N. D., Burkart, J., Graf, E., Zeder, Y., Dailey, L. A., and Weinzierl, B. Merging holography, fluorescence, and machine learning for in situ, continuous characterization and classification of airborne microplastics. *EGUsphere*, 2023:1–31, 2023. doi: 10.5194/egusphere-2023-2853. URL https://egusphere.copernicus.org/preprints/2023/egusphere-2023-2853/.

Berg, M. J. Tutorial: Aerosol characterization with digital in-line holography. *Journal of Aerosol Science*, 165:106023, 2022. ISSN 0021-8502. doi: https://doi.org/10.1016/j.jaerosci.2022.106023. URL https://www.sciencedirect.com/science/article/pii/S0021850222000647.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, June 2002. ISSN 1076-9757. doi: 10.1613/jair.953. URL http://dx.doi.org/10.1613/jair.953.

Chen, R. J., Ding, T., Lu, M. Y., Williamson, D. F., Jaume, G., Chen, B., Zhang, A., Shao, D., Song, A. H., Shaban, M., et al. A general-purpose self-supervised model for computational pathology. *arXiv preprint arXiv:2308.15474*, 2023.

Darcet, T., Oquab, M., Mairal, J., and Bojanowski, P. Vision transformers need registers, 2024.

Erb, S., Graf, E., Zeder, Y., Lionetti, S., Berne, A., Clot, B., Lieberherr, G., Tummon, F., Wullschleger, P., and Crouzy, B. Real-time pollen identification using holographic imaging and fluorescence measurements. *Atmospheric Measurement Techniques*, 17(2):441–451, 2024. doi: 10.5194/amt-17-441-2024. URL https://amt.copernicus.org/articles/17/441/2024/.

He, K., Chen, X., Xie, S., Li, Y., Dollár, P., and Girshick, R. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 16000–16009, 2022.

HIRST, J. M. An automatic volumetric spore trap. *Annals of Applied Biology*, 39(2):257–265, 1952. doi: https://doi.org/10.1111/j.1744-7348.1952.tb00904.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7348.1952.tb00904.x.

Jing, L. and Tian, Y. Self-supervised visual feature learning with deep neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(11): 4037–4058, 2020.

Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., et al. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*, 2023.

Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Assran, M., Ballas, N., Galuba, W., Howes, R., Huang, P.-Y., Li, S.-W., Misra, I., Rabbat, M., Sharma, V., Synnaeve, G., Xu, H., Jegou, H., Mairal, J., Labatut, P., Joulin, A., and Bojanowski, P. Dinov2: Learning robust visual features without supervision, 2024.

Pérez-García, F., Sharma, H., Bond-Taylor, S., Bouzid, K., Salvatelli, V., Ilse, M., Bannur, S., Castro, D. C.,

Schwaighofer, A., Lungren, M. P., et al. Rad-dino: Exploring scalable medical image encoders beyond text supervision. *arXiv preprint arXiv:2401.10815*, 2024.

Sauvageat, E., Zeder, Y., Auderset, K., Calpini, B., Clot, B., Crouzy, B., Konzelmann, T., Lieberherr, G., Tummon, F., and Vasilatou, K. Real-time pollen monitoring using digital holography. *Atmospheric Measurement Techniques*, 13(3):1539–1550, 2020.

Tan, M. and Le, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. *CoRR*, abs/1905.11946, 2019. URL http://arxiv.org/abs/1905.11946.

Team, G., Anil, R., Borgeaud, S., Wu, Y., Alayrac, J.-B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A. M., Hauth, A., et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

Tummon, F., Adams-Groom, B., Antunes, C. M., Bruffaerts, N., Buters, J., Cariñanos, P., Celenk, S., Choël, M., Clot, B., Cristofori, A., Crouzy, B., Damialis, A., Fernández, A. R., González, D. F., Galán, C., Gedda, B., Gehrig, R., Gonzalez-Alonso, M., Gottardini, E., Gros-Daillon, J., Hajkova, L., O'Connor, D., Östensson, P., Oteros, J., Pauling, A., Pérez-Badia, R., Rodinkova, V., Rodríguez-Rajo, F. J., Ribeiro, H., Sauliene, I., Sikoparija, B., Skjøth, C. A., Spanu, A., Sofiev, M., Sozinova, O., Srnec, L., Visez, N., and de Weger, L. A. The role of automatic pollen and fungal spore monitoring across major end-user domains. *Aerobiologia*, 40(1): 57–75, Mar 2024. ISSN 1573-3025. doi: 10.1007/s10453-024-09820-2. URL https://doi.org/10.1007/s10453-024-09820-2.

Zuberbier, T., Lötvall, J., Simoens, S., Subramanian, S. V., and Church, M. K. Economic burden of inadequate management of allergic diseases in the european union: a ga2len review. *Allergy*, 69(10): 1275–1279, 2014. doi: https://doi.org/10.1111/all.12470. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/all.12470.

A. Comparison on Full Dataset

In this section, we present a detailed comparison of the performance of various models on the complete dataset to assess the capabilities of different training strategies. We evaluated the official LVD-142M pre-trained backbones, EfficientNet models, and backbones that underwent self-supervised training on holographic images.

For the EfficientNet models, training was conducted on the entire dataset with the application of minority class upsampling to address class imbalance. Conversely, the official DINOv2 backbones and the Holographic Vision Transformers (Holo ViT) were evaluated on the unbalanced datasets without any upsampling techniques.

Table 1 details the test set micro-accuracies for different models. Notably, DINOv2 models pre-trained on domain-specific data show comparable performance to EfficientNet benchmarks. The holographic self-supervised ViT variants, especially the ViT base with registers, achieve a micro-accuracy that closely matches or even surpasses that of the EfficientNet models. This is significant considering that the EfficientNet models were trained with upsampling techniques to counter class imbalance, whereas the self-supervised methods were directly exposed to this class imbalance without any corrective measures.

Table 1. Comparison of test set micro accuracies for different models on the full holographic dataset. **Holo ViT**: DINOv2 self-supervised models trained on the full holographic dataset. **LVD-142M ViT**: Official DINOv2 backbones pre-trained on the LVD-142M dataset, including distilled versions with fewer parameters. **EfficientNet**: Models pre-trained on ImageNet and fine-tuned on the holographic dataset with minority class upsampling. Micro accuracies for ViT models were obtained using k-NN, while EfficientNet models used a fully supervised approach with a classification head.

MODEL	PARAMETERS	MICRO-ACC
OUR MODELS:		
HOLO VIT LARGE REG	300M	72.15%
HOLO VIT BASE REG	80M	**75.12%**
HOLO VIT SMALL REG	22M	72.75%
OFFICIAL DINOV2 BACKBONES:		
LVD-142M VIT GIANT REG	1.1B	58.66%
DISTILLED VIT LARGE REG	300M	**59.89%**
DISTILLED VIT BASE REG	80M	59.47%
DISTILLED VIT SMALL REG	22M	59.87%
EFFICIENTNET BENCHMARK:		
EFFICIENTNET B0	4M	71.89%
EFFICIENTNET B4	18M	72.22%
EFFICIENTNET B7	65M	**74.48%**

B. Few shot learning capabilities of distilled backbones

To evaluate the performance of the smaller, distilled versions of the officially released LVD-142M DINOv2 backbone the same experiment as described in section 3.1 were conducted. The Figure 6 shows that distilled versions were able to represent the features for the holographic images better resulting in almost comparable results to the EfficientNet B0 which was supervised trained on the data. However, the Holo ViT base backbone is still outperforming all official backbones.

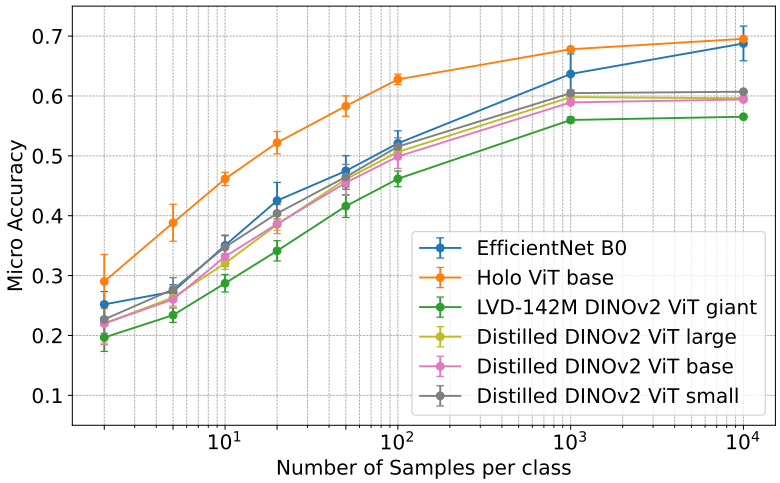

Figure 6. Micro-accuracy of various models as a function of the number of samples per class in few-shot learning scenarios with holographic pollen images. The performance of distilled versions of the official LVD-142M DINOv2 backbone (ViT giant, ViT large, ViT base, ViT small) is compared to the self-supervised Holo ViT base and the supervised EfficientNet B0. The results show that while distilled versions improve upon the original ViT giant, the self-supervised Holo ViT base still outperforms all official backbones.

C. Dataset splits

The Table 2 visualizes the dataset split of the holographic Pollen dataset used in this paper.

Table 2. The distribution of the holographic Pollen dataset across 18 classes. The training and validation sets follow a 80/20 split, while the test set is collected from different sources, ensuring it is not a trivial subset of the training set. This separation ensures robust evaluation of the model's performance.

CLASS	TRAIN	VALIDATION	TEST
ALNUS	32'404	8'016	2'643
AMBROSIA	6'978	1'708	7'896
ARTEMISIA	7'980	1'990	5'590
BETULA	50'848	12'890	5'458
CARPINUS	18'210	4'512	600
CORYLUS	25'542	6'354	4'444
CUPRESSUS	5'758	1'496	405
FAGUS SYLVATICA	15'138	3'766	2'827
FRAXINUS EXCELSIOR	4'126	988	4'837
OSTRYA SP.	2'230	510	2'154
PICEA	10'652	2'564	1'345
PINUS	29'984	7'474	3'139
PLANTAGO LANCEOLATA	4'894	1'220	2'258
POACEAE	60'776	15'312	5'467
POPULUS	33'902	8'654	588
QUERCUS ROBUR	5'362	1'344	2'050
ULMUS	5'302	1'276	4'742
URTICA	17'454	4'312	1'214
TRAIN SIZE	**337'540**		
VALIDATION SIZE		**84'386**	
TEST SIZE			**57'657**