



The LifeWatch ERIC Biodiversity
& Ecosystem eScience Conference

PLOVDIV, BULGARIA

7-10 JULY 2026

From the Cosmos to the Coast:

Phytoplankton Classification with Little or No Training Data Using Scattering Covariances and Topological Morphology

Presenter: J.-M. Delouis

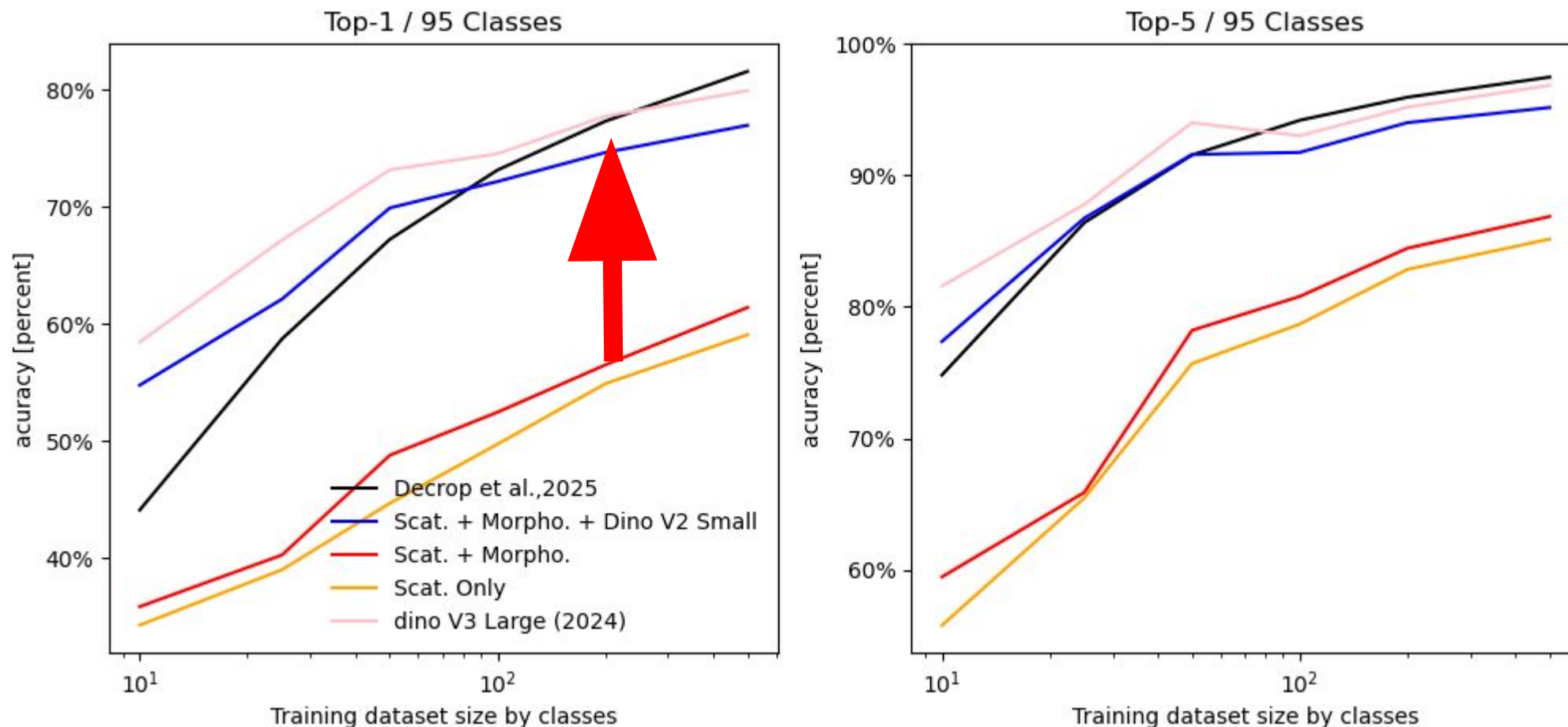
Co-authors: Anne Fouilloux, Tina Odaka, Théo Foulquier, Deyan Samardzhiev

one-image cosmology methods → feature extraction → rare phytoplankton classification



08/07/2026

What We Achieve



Scat. + Morpho.: interpretable & instrument-agnostic, **no training needed**
DINOv2: higher accuracy, **no retraining**
Decrop et al. 2025: reference, task-specific training



The LifeWatch BPNS Phytoplankton Dataset

337k

images

95

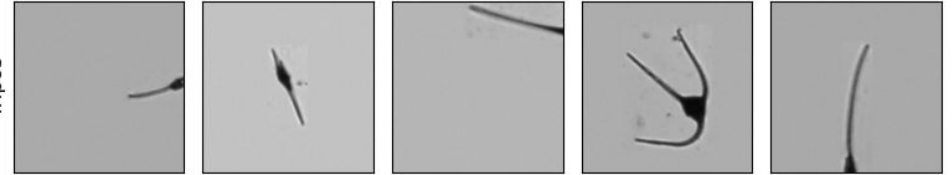
taxa

FlowCam

instrument

- **~337,000 FlowCam images:** Belgian Part of the North Sea (BPNS)
 - **95 phytoplankton taxa**, highly imbalanced
 - rare taxa: < **100 images**
 - dominant taxa: > **10,000 images**
 - **Expert-validated labels**, noisy for morphologically variable groups
 - **Open-world streams:**
 - new life stages and morphotypes appear unpredictably
- **Goal: robust classification with little or no labelled training data**

Tripes



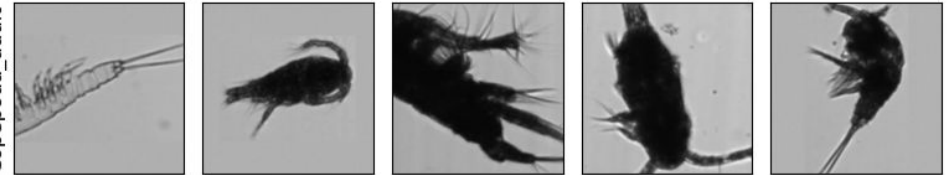
Odontella_rhombus_f._trigona



Pseudo-nitzschia



Copepoda_adult

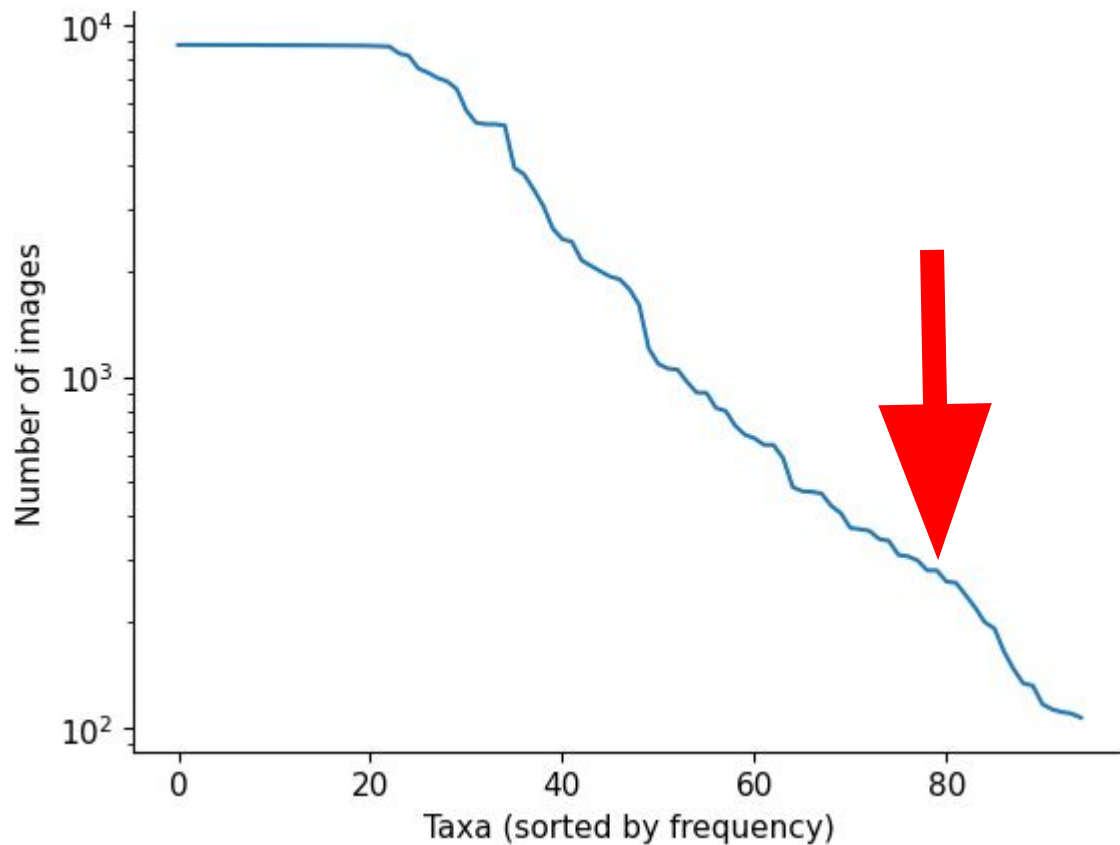


Two Core Challenges in Phytoplankton Identification

Why is phytoplankton identification hard?

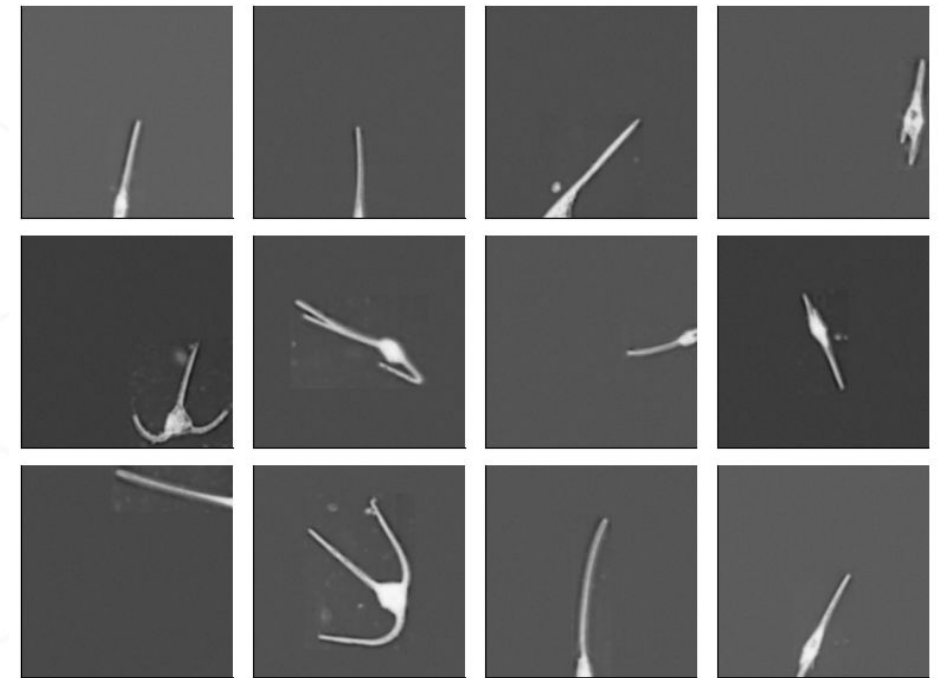
Challenge 1: Extreme Class Imbalance

95 taxa · 337k images · BPNS FlowCam



Challenge 2: Intra-class Variability

Same species · same instrument · different orientations



Tripos -- 12 real FlowCam images

→ same organism, radically different appearance

Why Standard AI Struggles Here

Supervised AI (e.g. CNN, SVM)

Requires large labeled datasets per class

- Rare taxa: only 5-50 images available
- Models overfit and ignore rare classes

Not rotation-invariant by design

- FlowCam images: same organism, any orientation

Black-box: hard to validate

Self-supervised AI (e.g. DINOv2)

- Trained on large dataset without classification
- Classify using embedding (small vector)
- **PROs:**
 - Usable with minimal supervision
 - Very effective thanks to training on very large datasets
 - DINOv2: 142M images
→ 384-dimensional embeddings for any RGB image
- **CONs:**
 - Limited explainability

What can be done with only ONE image

■ Mathematical approach:

- Compute mathematically fixed transform to the image to extract statistical information
- Developed in cosmology, where only **ONE image** is available for training
 - Works for rare taxa: one example suffices
- Requires interpretability :
 - instrument-agnostic

⇒ **DIMENSION REDUCTION BASED on INVARIANTS : Rotation, Positions, etc.**

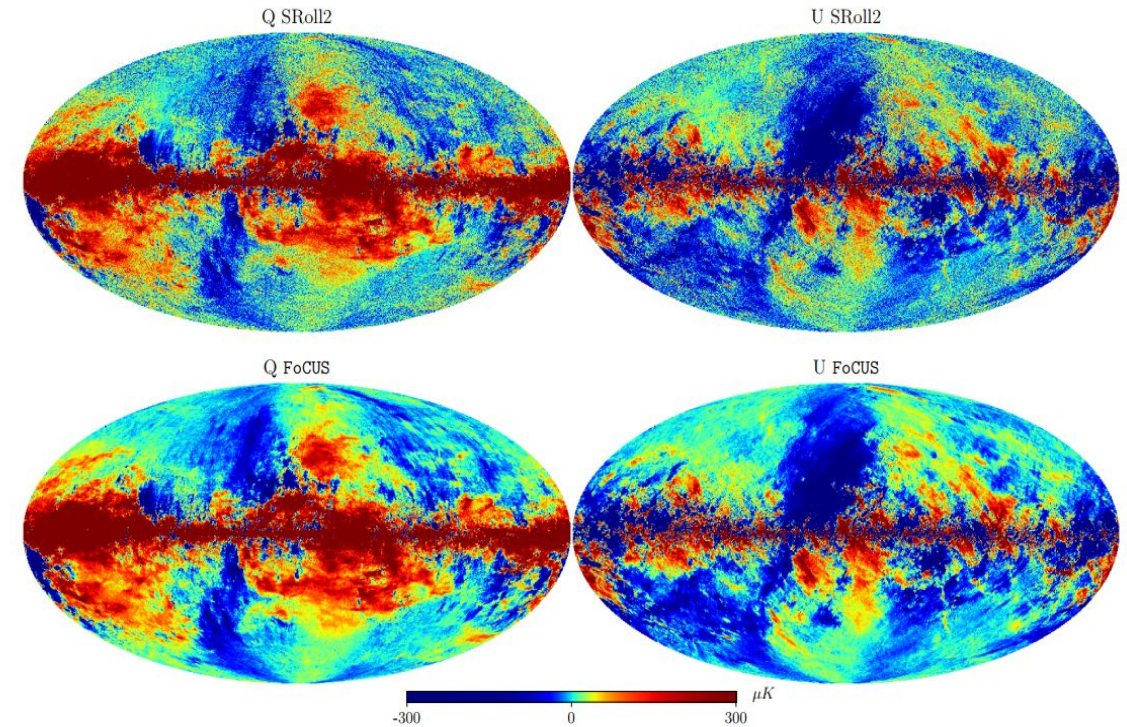
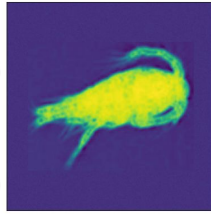


Fig. 6. Stokes Q and U maps at 353 GHz for $N_{\text{side}} = 256$. *Top panels:* input Planck SRoll2 maps, and the *bottom panels* show the corresponding FoCUS maps.

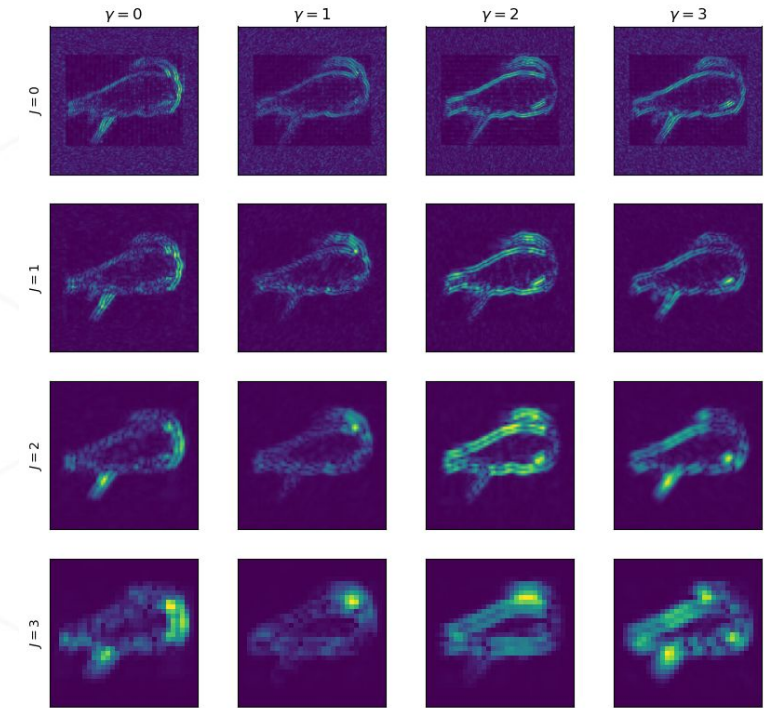
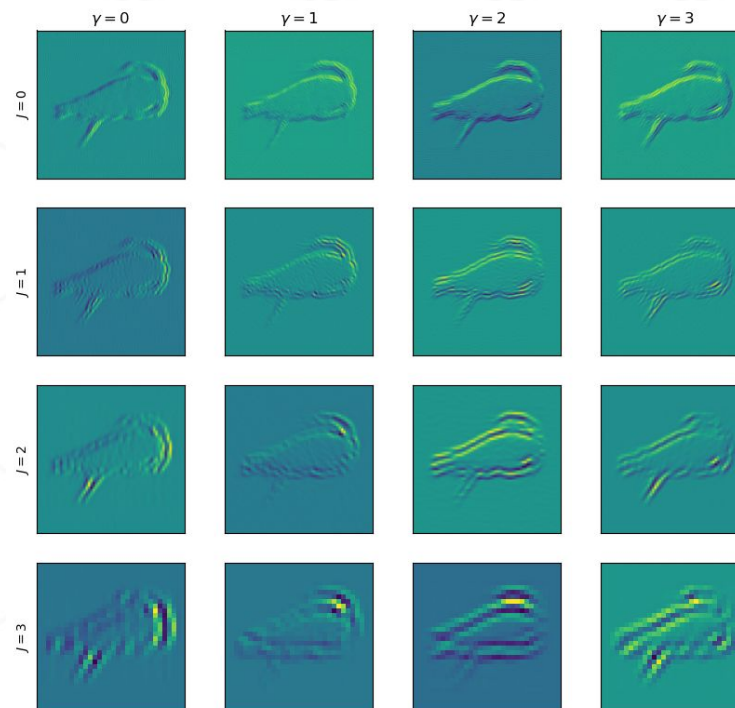
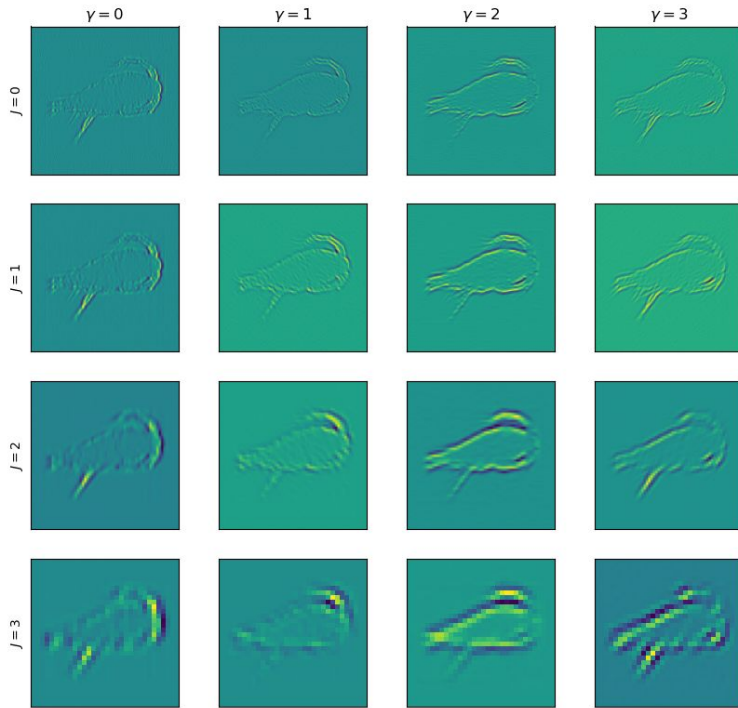
Look for dimension reduction: **Wavelet transform**



$W^\lambda I$

$$\lambda = (j, \gamma)$$

Scale Angle



Scattering covariance (2024)

Origin of the method: Scattering Transform (Mallat et al. 2012)

Scattering covariance coefficients S_1, S_2, S_3, S_4

*Slide from
Mousset – ENS-DI*

Coefficients associated to a single scale:

$$\begin{aligned} S_1^\lambda &= \langle |W^\lambda I| \rangle \\ S_2^\lambda &= \langle |W^\lambda I|^2 \rangle \end{aligned}$$

Averages over pixels

Coefficients to probe the coupling between scales:

$$\begin{aligned} S_3^{\lambda_1, \lambda_2} &= \text{Cov} [W^{\lambda_1} I, W^{\lambda_1} |W^{\lambda_2} I|] \\ S_4^{\lambda_1, \lambda_2, \lambda_3} &= \text{Cov} [W^{\lambda_1} |W^{\lambda_3} I|, W^{\lambda_1} |W^{\lambda_2} I|] \end{aligned}$$

For instance, with $j = [1, 8]$ and $\gamma = [1, 5]$

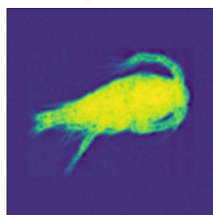
$\Rightarrow \sim 10^3$ coefficients.

$W^\lambda I$ the wavelet transform

$\lambda = (j, \gamma)$

Scale Angle

Scattering coefficients computation

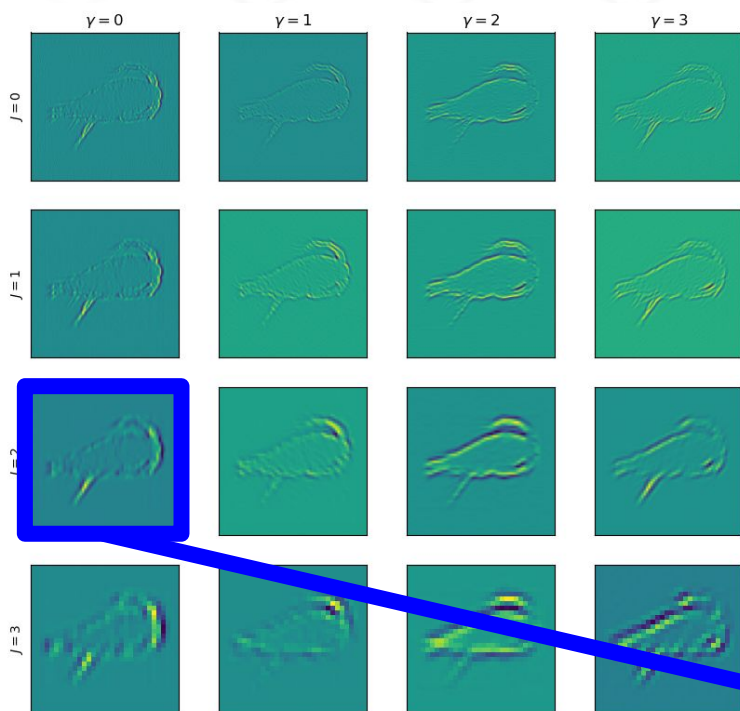


$W^\lambda I$

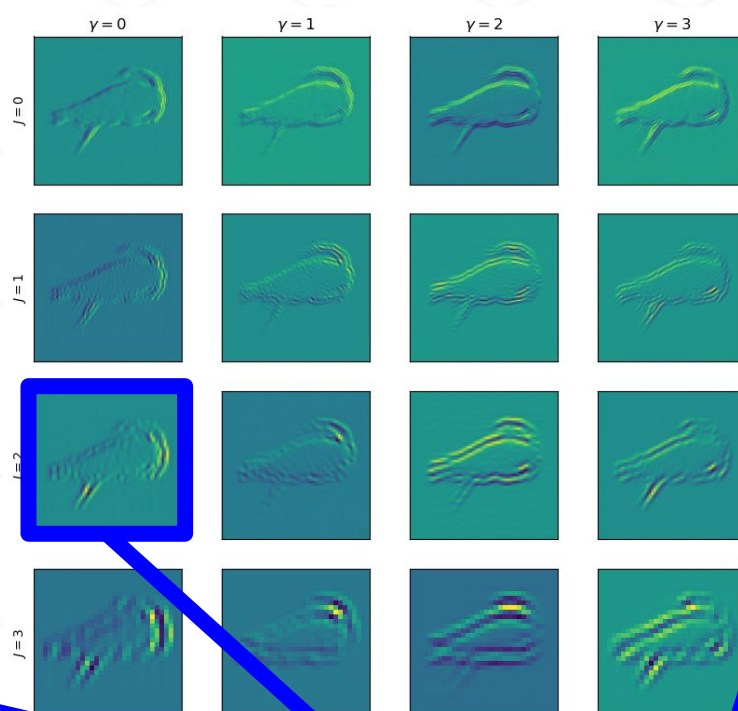
$$S_1^\lambda = \langle |W^\lambda I| \rangle$$

$$S_2^\lambda = \langle |W^\lambda I|^2 \rangle$$

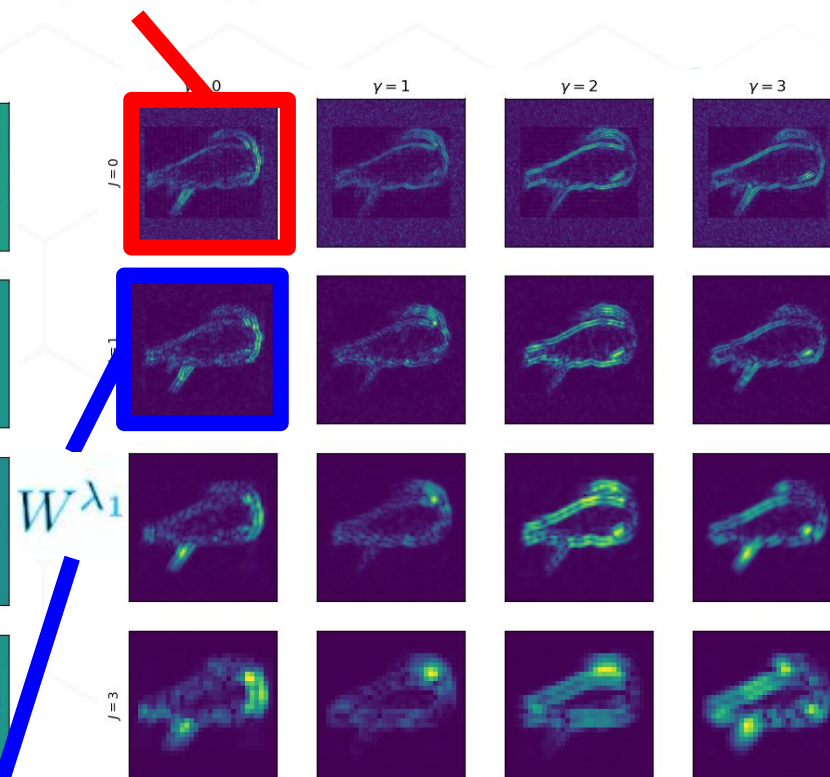
Averages over pixels



Real



Imaginary



Modulus

$$S_3^{\lambda_1, \lambda_2} = \text{Cov} [W^{\lambda_1} I, W^{\lambda_1} |W^{\lambda_2} I|]$$

Angular Rotation Invariance via FFT

Problem: orientation is a nuisance variable

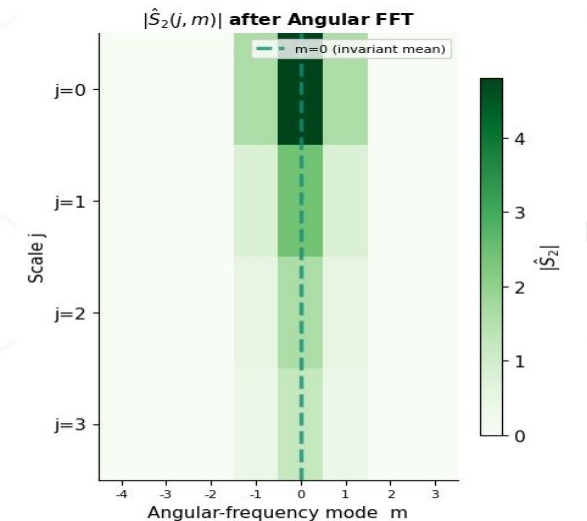
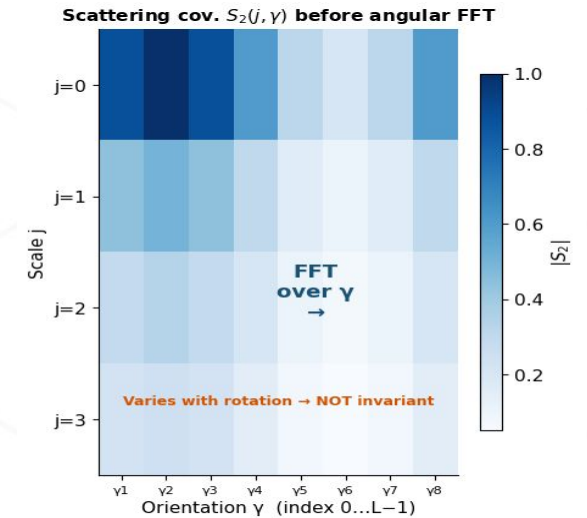
- Scattering covariances are computed for L orientations $\theta_1, \dots, \theta_L$
- Same species photographed at different angles must produce the same descriptor

Solution: Angular FFT

- Apply Fourier transform over the angular index θ
- Keeps only rotationally invariant modes ($m=0$ = mean, $|m|>0$ = oscillation modes)
- Exact rotation invariance by design — not by data augmentation

Two direct benefits:

- Fewer parameters:*
 L orientations $\rightarrow L/2$ independent Fourier modes
- Lower variance:* averaging over orientations = the core scattering idea applied to angles



$m=0$: rotation-invariant ✓ $|m|>0$: orientation modes

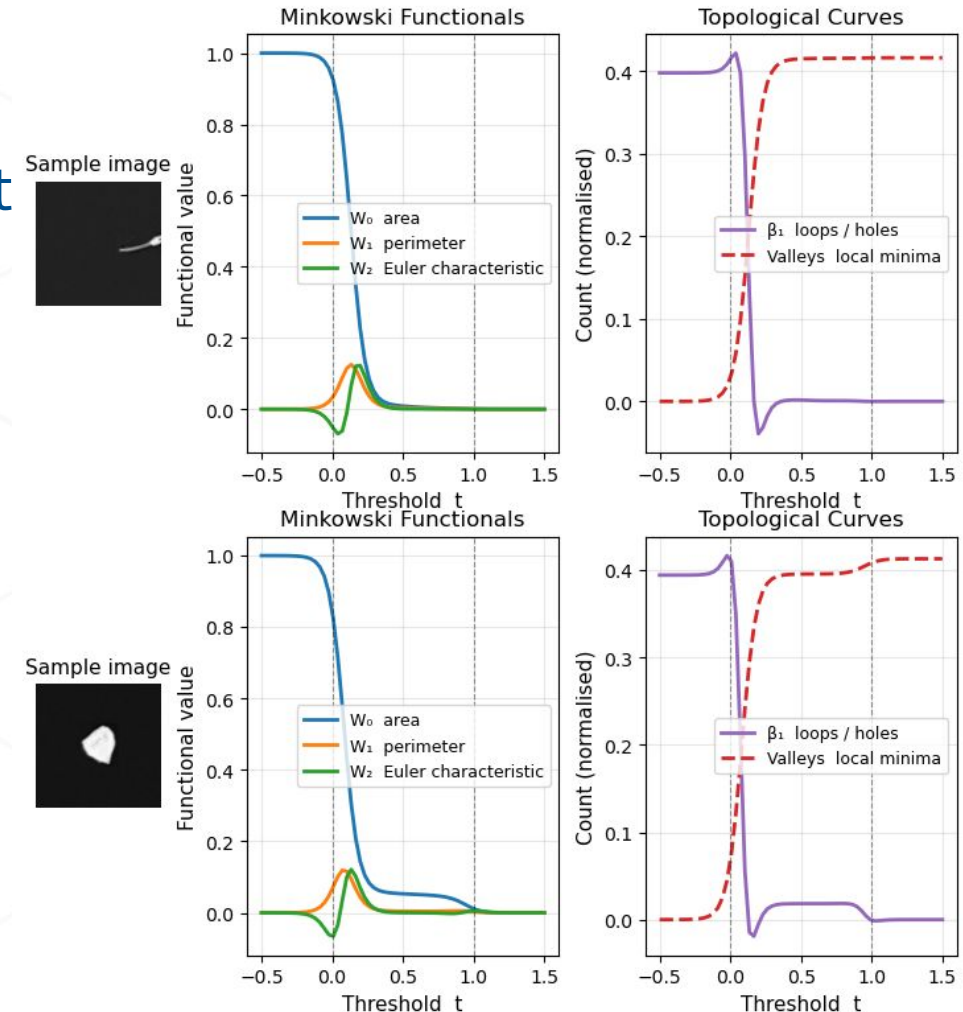
Morphological Functionals

Morphological Functionals as Threshold Curves

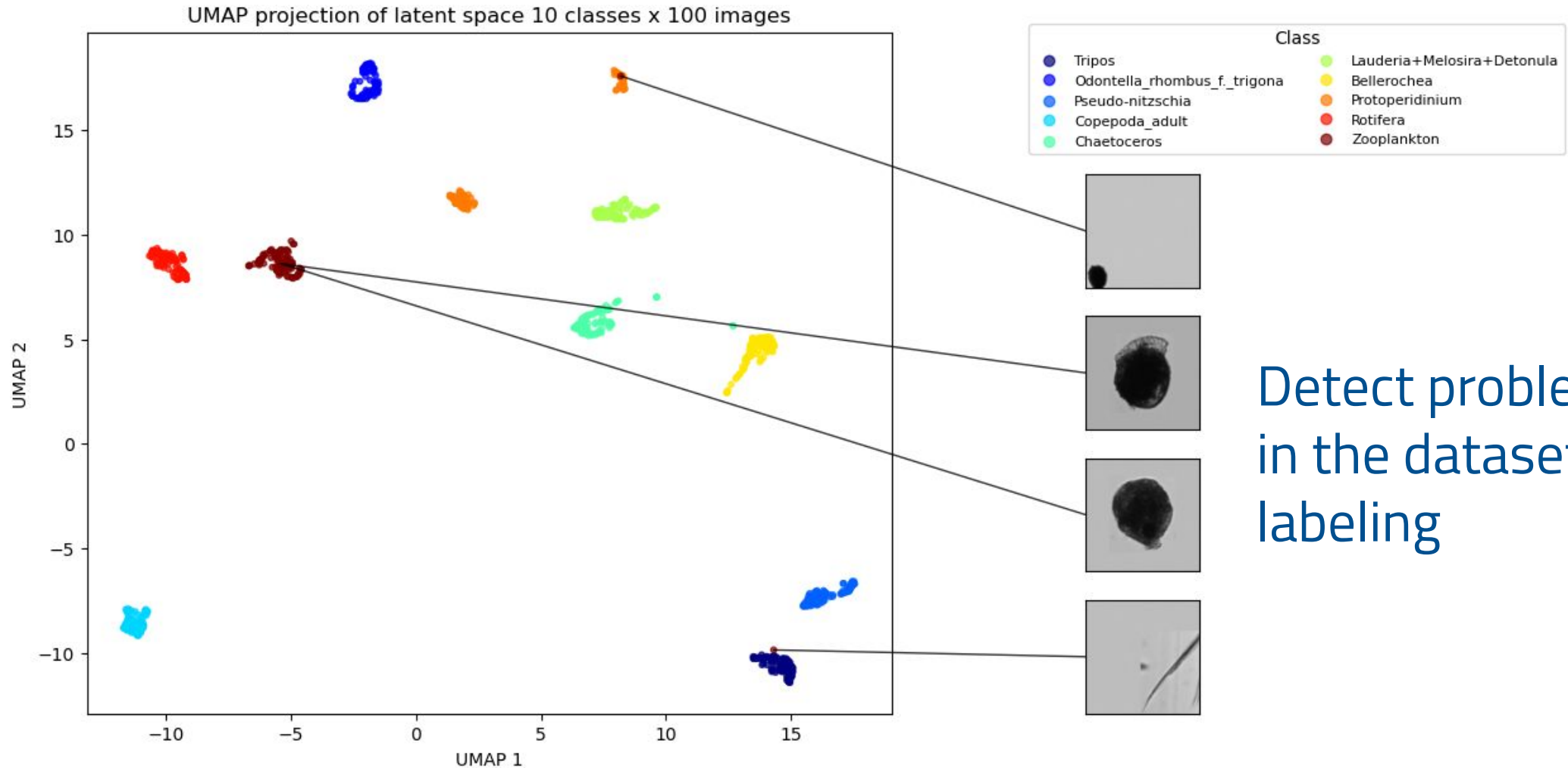
Threshold curves over $t \in [t_{min}, t_{max}]$:

- **Minkowski W_0** : fraction of pixels above t
Area
- **Minkowski W_1** : boundary length of the excursion set
Perimeter
- **Minkowski W_2** : #components #holes
Euler characteristic
- **β_1** : number of independent loops (1st Betti number)
- **Valleys**: count of local intensity minima above t
- All differentiable via sigmoid soft-thresholding.

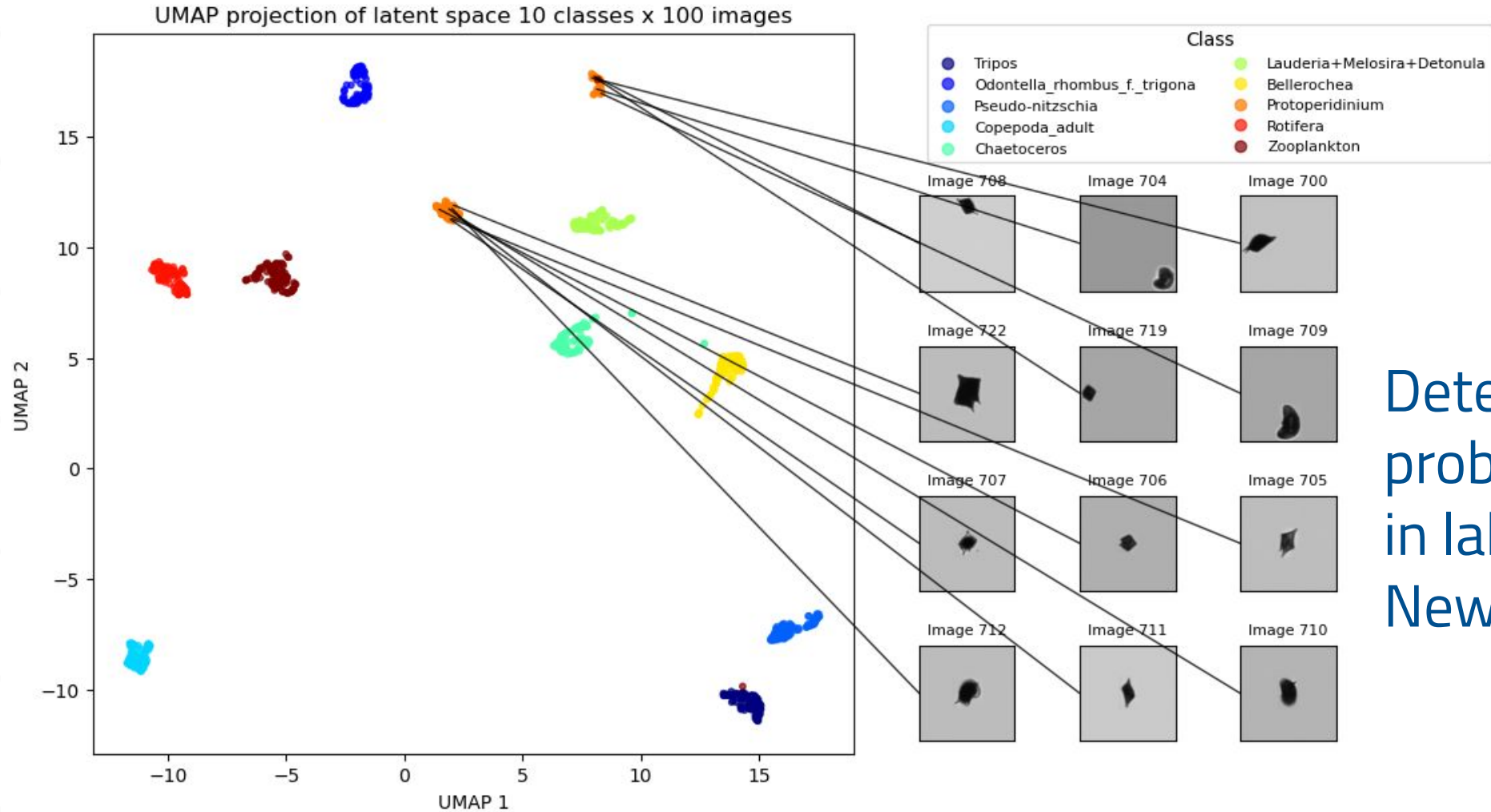
→ Keep only moments to reduce the number of parameters



Unsupervised Learning Feature Space Structure

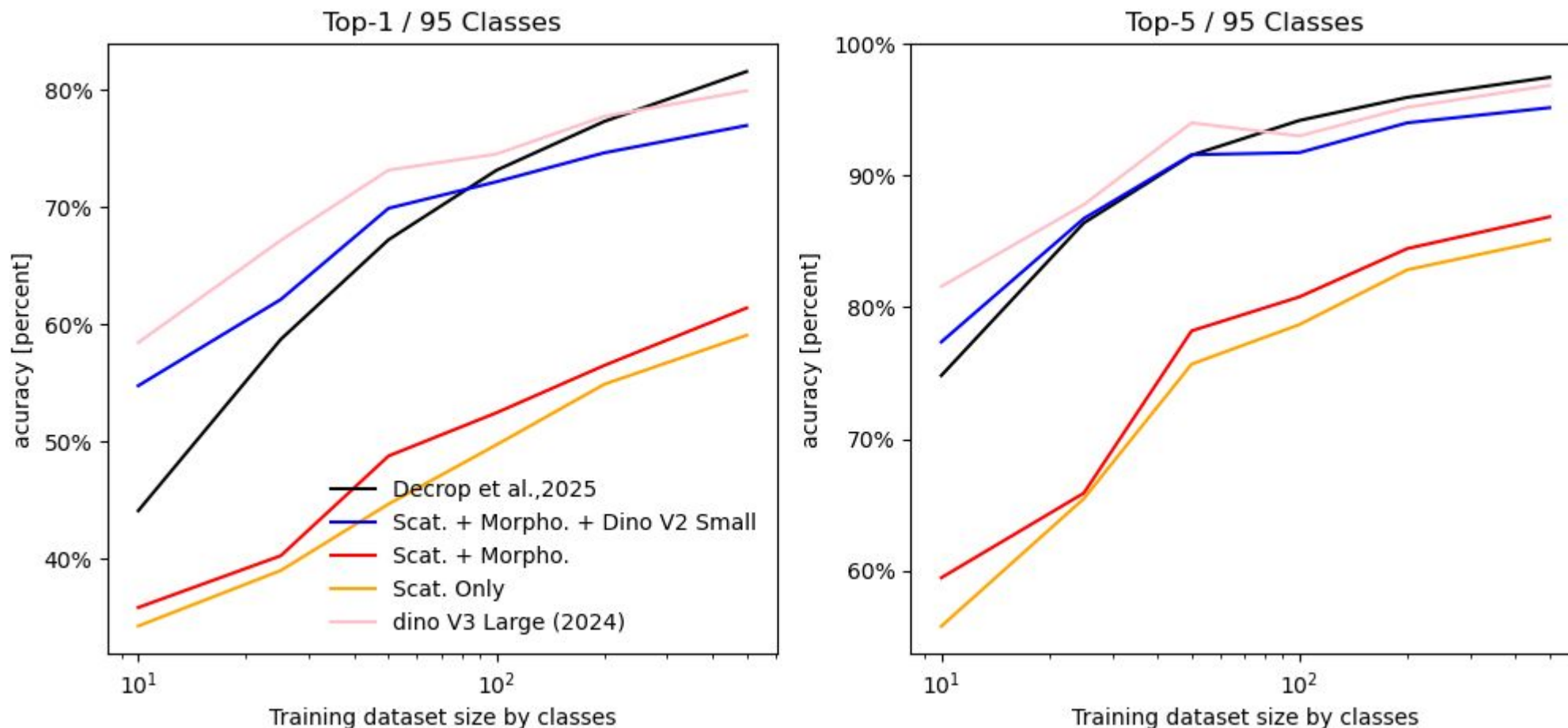


Unsupervised Learning Linear Separability



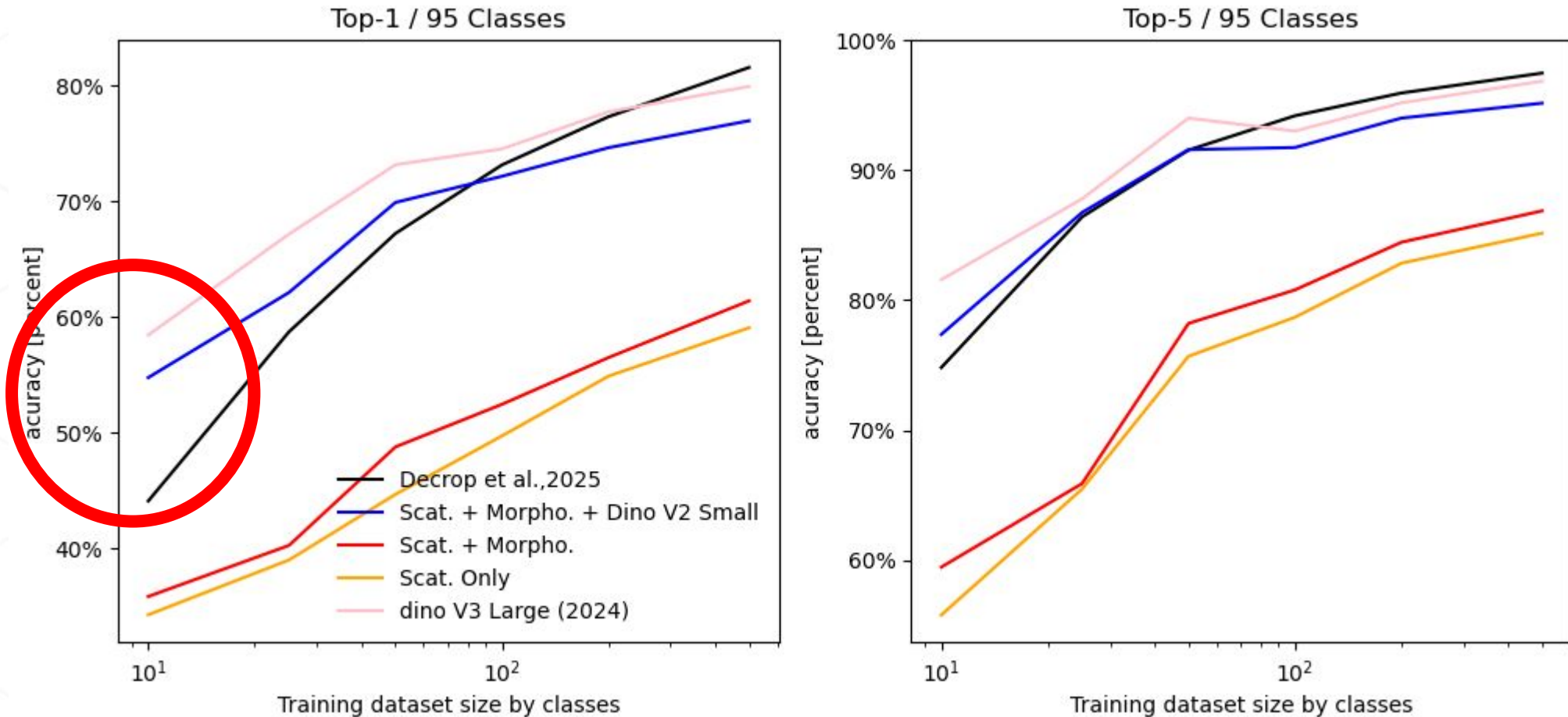
Detect
problems
in labeling:
New taxa?

Supervised Classification vs. Training Set Size



Mathematical transform (Scat.+Morph.) *underperforms* compared to Self-Supervised embedding.

Supervised Classification vs. Training Set Size



SVM+embedding is better for very small datasets: ***Rare taxa***

Conclusion

Mathematical feature extraction:

- Interpretable / explainable AI
- Unsupervised analysis for dataset quality assessment
- Effective with small training sets (rare taxa)
- Adapts to new instruments without retraining

Pre-trained embeddings (DINOv2):

- DINOv2 embeddings offer limited interpretability
- State-of-the-art performance, unsupervised or few-shot (SVM)

⇒ Ongoing research directions:

- ⇒ Reduce Scat. Cov. + Morpho. dimensionality
- ⇒ Add more shape function definition
- ⇒ Learn wavelet shapes to reproduce DINOv2/v3 GRAM matrix (latent space)



The LifeWatch ERIC Biodiversity
& Ecosystem eScience Conference
PLOVDIV, BULGARIA 7-10 JULY 2026

THANK YOU!

Questions?

Jean.marc.delouis@ifremer.fr

