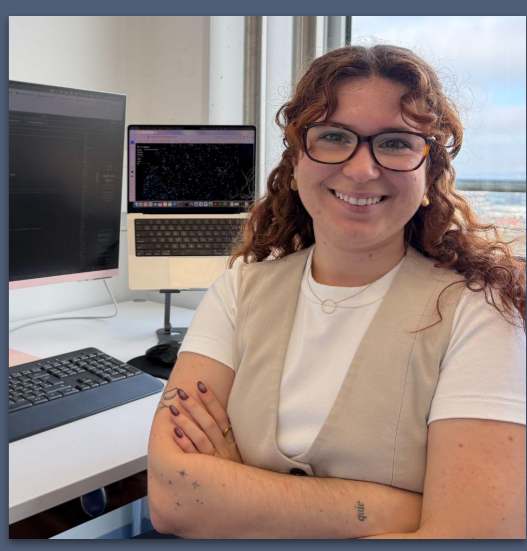


Are Vision Foundation Models foundational for electron microscopy image segmentation?



Caterina Fuster-Barceló and Virginie Uhlmann
 caterina.fusterbarcelo@mls.uzh.ch
 Department of Molecular Life Sciences, UZH

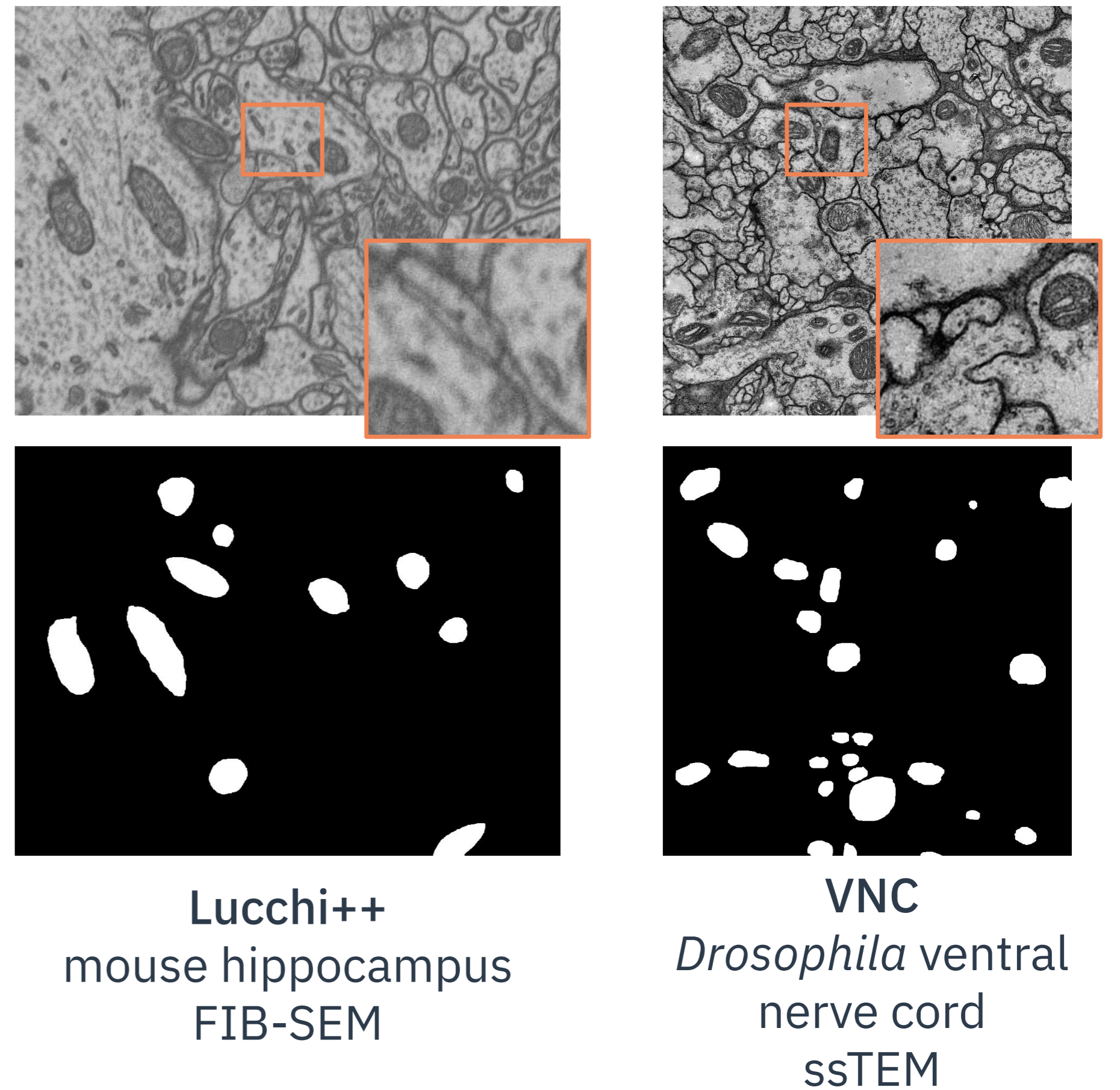
Are FMs FMing?



Key Questions

1. Can Vision Foundation Models (VFMs) generalize beyond their **pre-training distribution and natural images** to be reused to process Electron Microscopy (EM) image data?
2. Can Parameter-Efficient Fine-Tuning (PEFT) techniques such as Low-Ranked Adaptation LoRA) **improve generalization** across EM datasets?
3. Are VFMs with and without PEFT **competitive** against **task-specific** models and other **end-to-end** trained alternatives in practice?

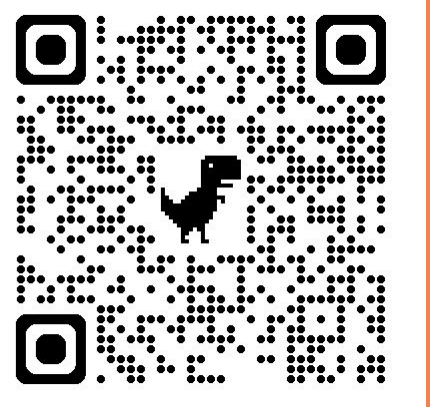
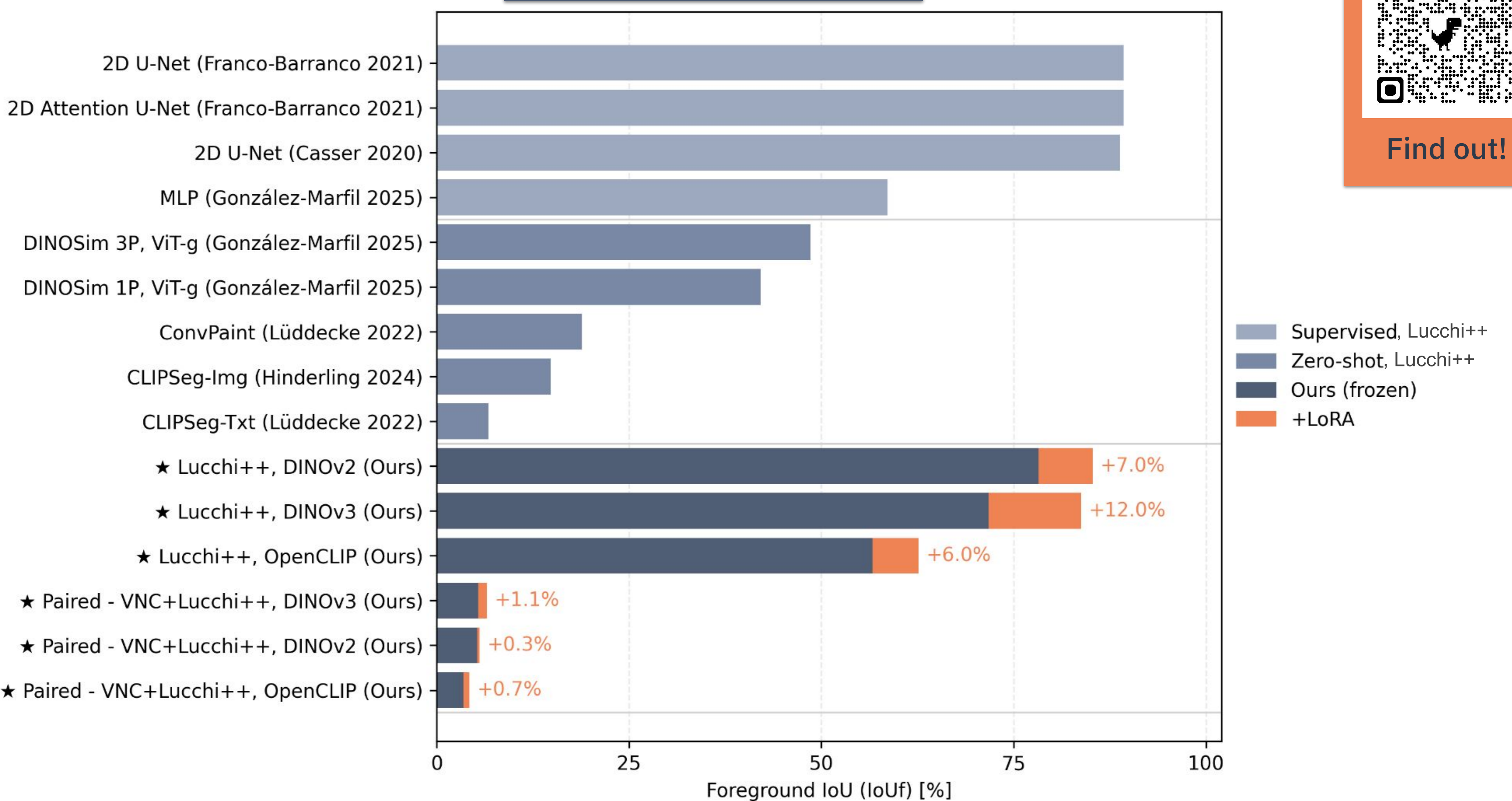
Electron Microscopy Datasets



Models Evaluated

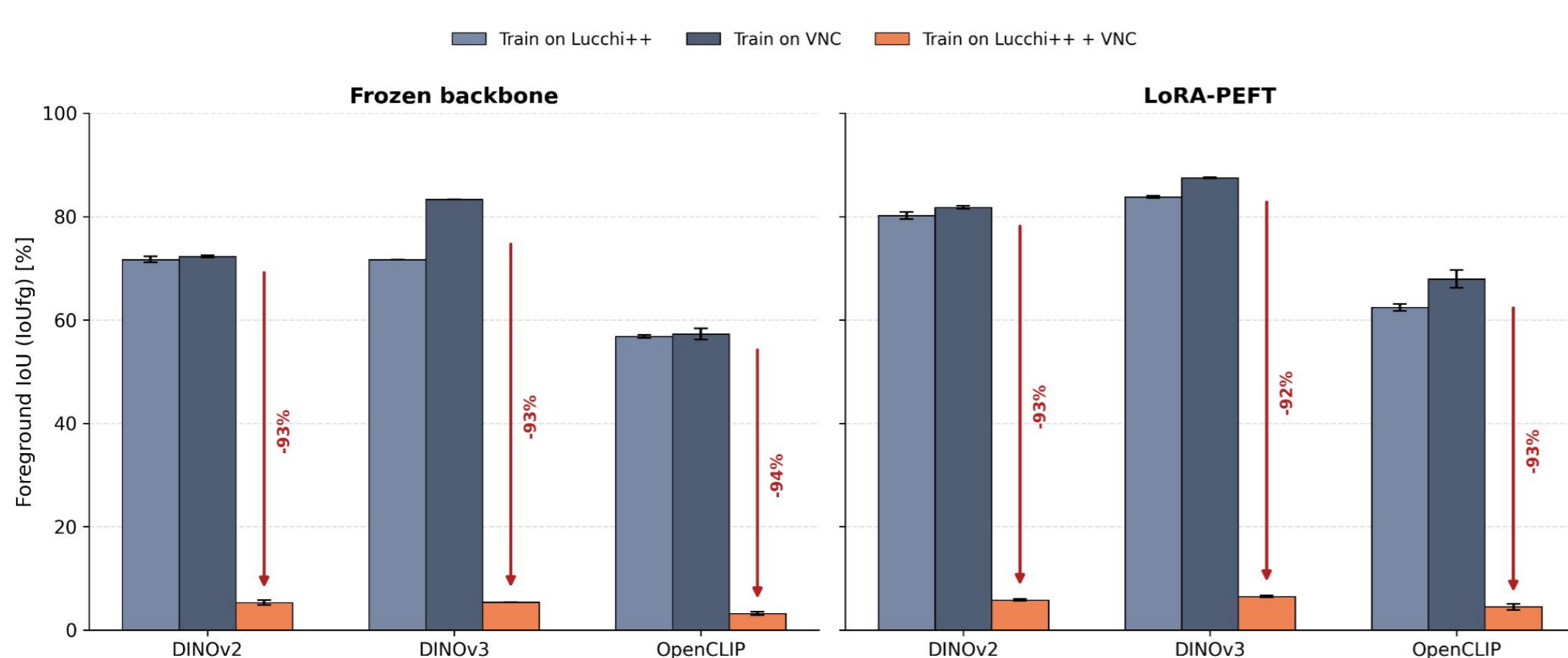
| Model Name | Model Type | Training Dataset |
|-----------------|---------------------|-----------------------|
| DINOv2 | Discriminative ViT | ImageNet-21K |
| DINOv3 | Adv. Discriminative | Large General Dataset |
| OpenCLIP | Contrastive | LAION-2B |

Comparative Results



Find out!

Domain analysis (PCA)



Conclusions

- VFMs adapted with PEFT on a **single target dataset** **perform competitively** with models trained end-to-end.
- Attempting to train a model on **combined EM** data results in a **performance drop** that PEFT cannot fix.
- **Data distributions** in the VFMs' latent spaces remain **strongly separable** even after PEFT, suggesting that **domain alignment tools** for natural images are **insufficient** to account for EM image variability.