Agri-FM+: A Self-Supervised Foundation Model for Agricultural Vision

Md Jaber Al Nahian, Tapotosh Ghosh, Farnaz Sheikhi, Farhad Maleki
University of Calgary, Calgary, AB, Canada

Motivation

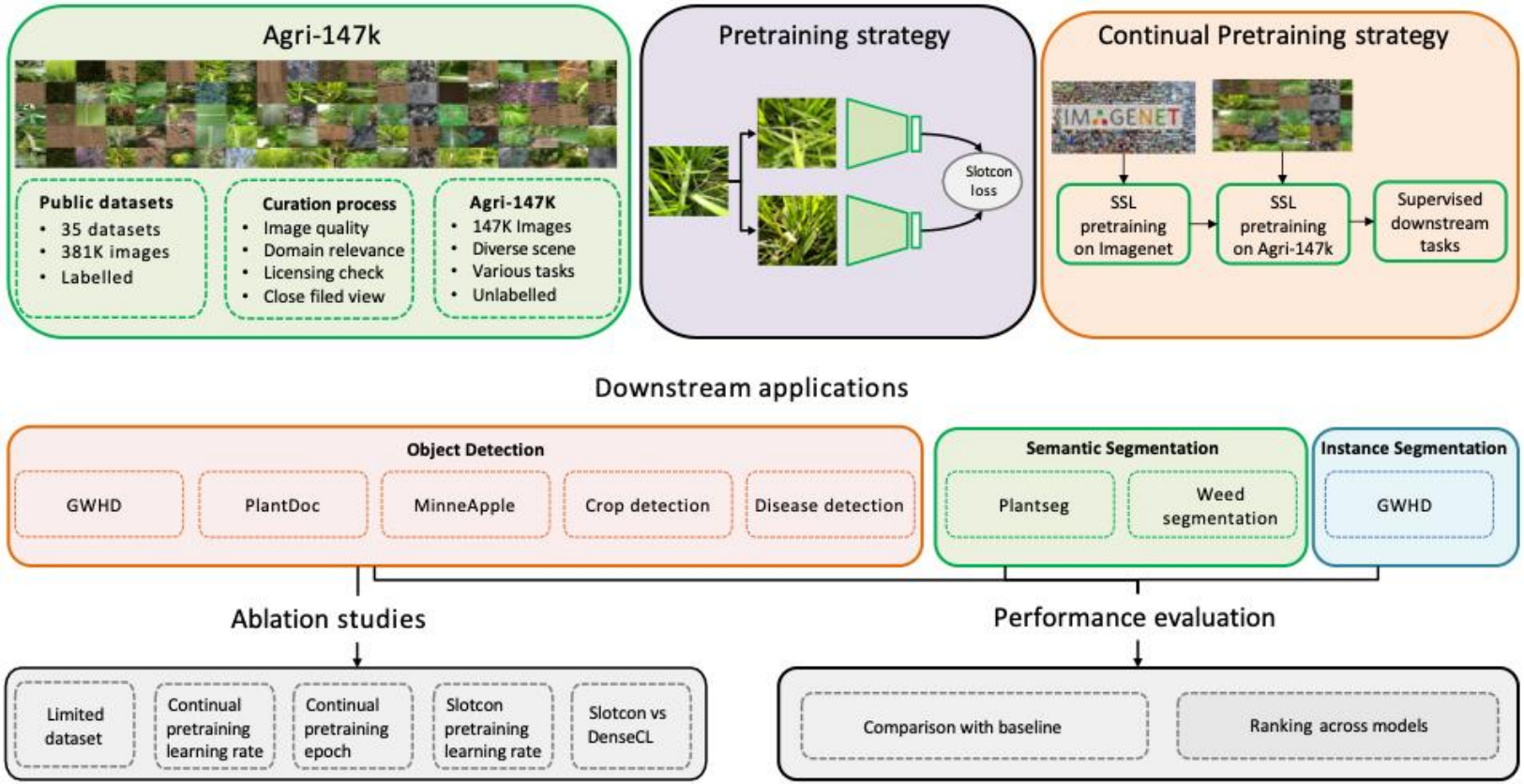
- ✓ Foundation models (e.g., CLIP, DINO, SAM) underperform in agriculture
- ✓ Agricultural tasks require fine-grained, domain-specific understanding
- ✓ Existing FMs lack robustness to diverse field conditions

Our goal: Build a self-supervised, label-efficient foundation model for close-field agricultural vision

Key Contributions

- ✓ Curated **Agri-147K**: 147K high quality, diverse agricultural mages from 35 public datasets
- ✓ Introduced **Agri-FM+:** first agricultural vision foundation model for **close-field** tasks
- ✓ Extensive evaluation: 8 datasets across object detection, semantic and instance segmentation

Methodology



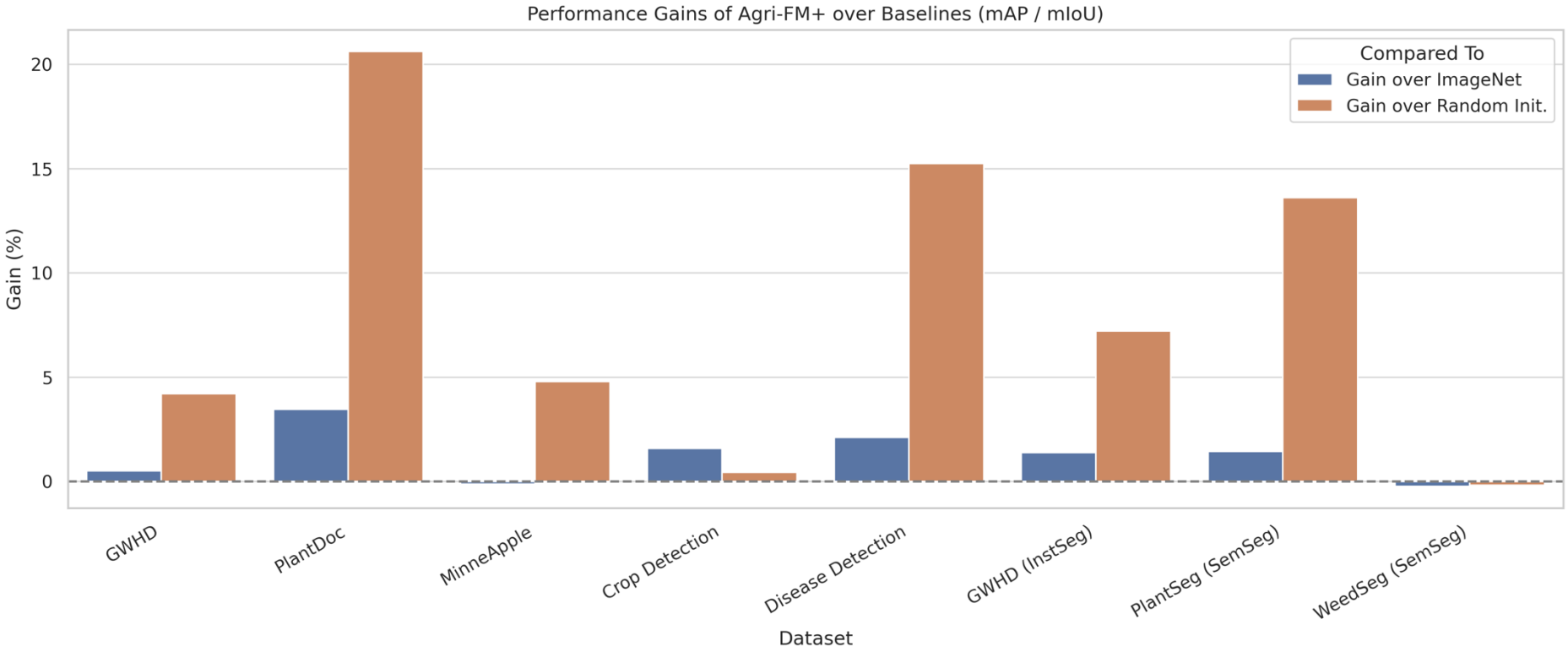
Agri-FM+ Outperform Across All Tasks

Pretrained ResNet-50 Weight	Object Det. (mAP)					Instance Seg. (mAP)	Semantic Seg. ($mIoU$)	
	GWHD	PlantDoc	MinneApple	Crop Det.	Disease Det.	GWHD	PlantSeg	Weed Seg.
Random Init.	44.27	20.99	35.61	32.96	34.57	66.26	16.15	95.13
Supervised ImageNet	47.97	38.15	**40.53**	31.81	47.71	72.09	28.31	95.17
Agri-FM	48.18	38.53	38.21	30.01	47.90	72.41	28.07	**95.87**
Agri-FM+	**48.48**	**41.61**	40.41	**33.41**	**49.83**	**73.48**	**29.76**	94.95

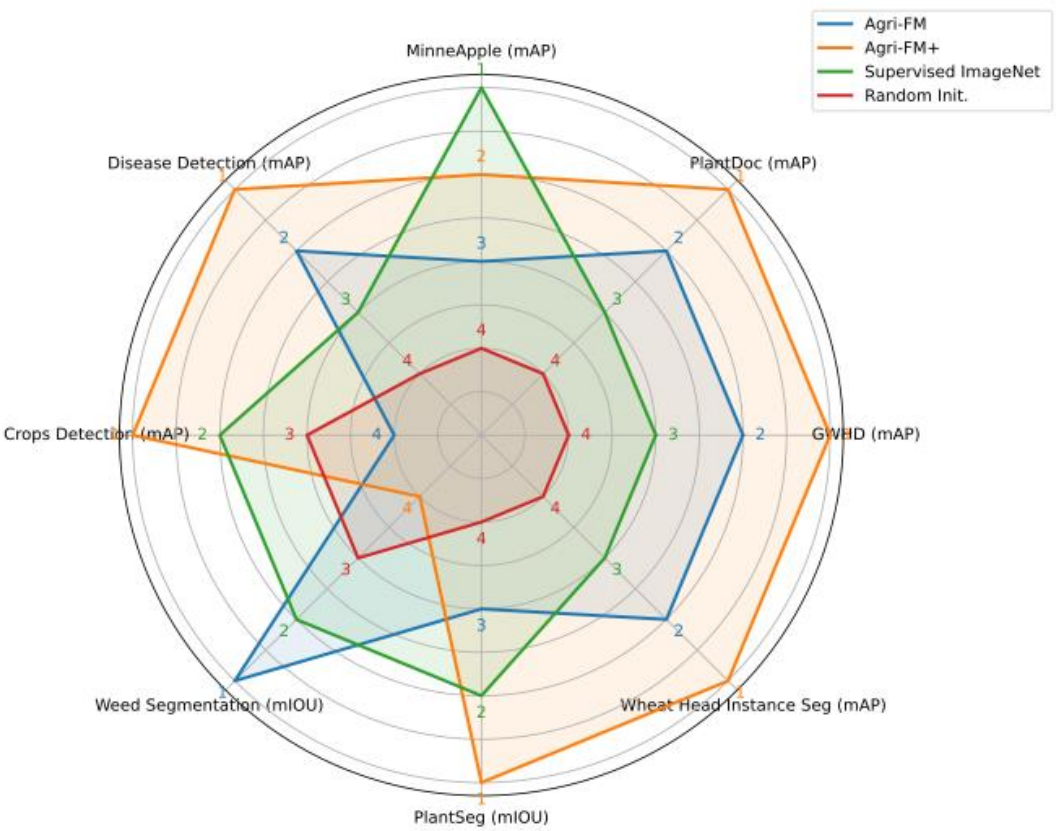
Ablation on Limited Annotation Scenario

Pretrained ResNet-50 Weight	GWHD (mAP)			PlantDoc (mAP)			MinneApple (mAP)			Disease Detection (mAP)		
	5%	10%	20%	5%	10%	20%	5%	10%	20%	5%	10%	20%
Random Init.	34.57	38.32	41.91	3.72	6.19	10.84	19.82	25.81	27.11	15.94	19.13	26.16
Supervised ImageNet	38.33	39.71	41.99	9.17	17.97	22.92	**23.25**	28.55	32.52	22.46	27.29	36.31
Agri-FM	37.92	39.41	41.29	9.16	15.04	19.79	21.44	28.96	32.61	20.62	25.71	34.98
Agri-FM+	**39.24**	**40.41**	**42.44**	**10.73**	**18.09**	**24.22**	22.08	**30.01**	**33.21**	**23.18**	**29.09**	**37.06**

Performance Gains (%) of Agri-FM+ over ImageNet and Random Init.

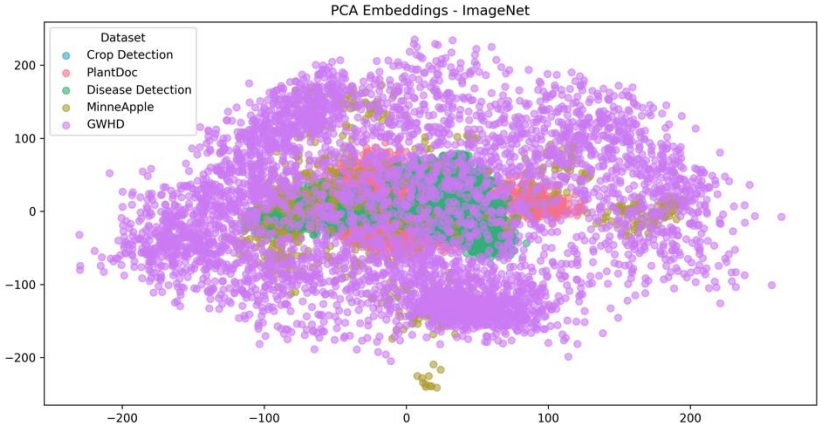


Consistent Top Rankings of Agri-FM+ Across Multiple Vision Tasks

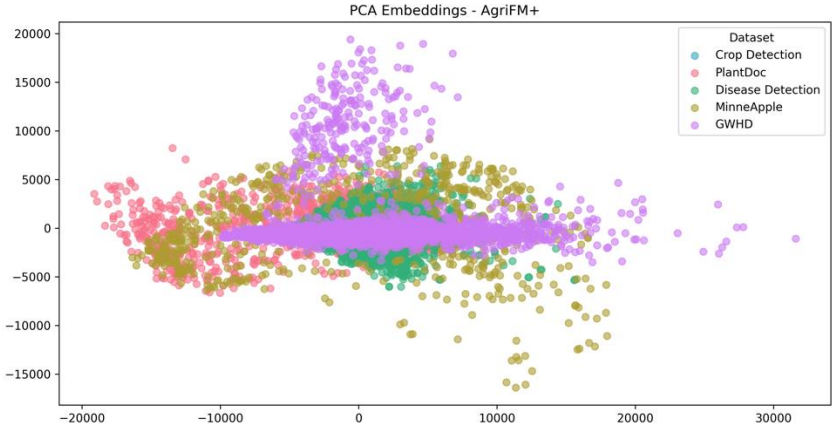


Improved Feature Clustering with Agri-FM+

Feature Representation of ImageNet vs Agri-FM+



ImageNet



Agri-FM+

Results Highlights

- ✓ Benchmarked on **8 datasets** across detection, segmentation, and instance segmentation
- ✓ In fully labeled setup: **+1.27% over ImageNet**, **+8.25% over random init.**
- ✓ In low-label setup (10%): **+1.02% over ImageNet**, **+4.54% over random init**
- ✓ Agri-FM+ ranks **1st in 6 of 8 tasks**

Key Takeaways

- ✓ Agri-FM+ is the first self-supervised foundation model for agricultural vision
- ✓ Demonstrates strong **domain adaptation**, **generalization**, and **label efficiency**
- ✓ Visit our poster to learn more!