



Active Learning Meets Foundation Models: Fast Remote Sensing Data Annotation for Object Detection

Marvin Burges, TU Wien*

Philippe Ambrozio Dias, ORNL

Carson Woody, ORNL

Sarah Walters, ORNL

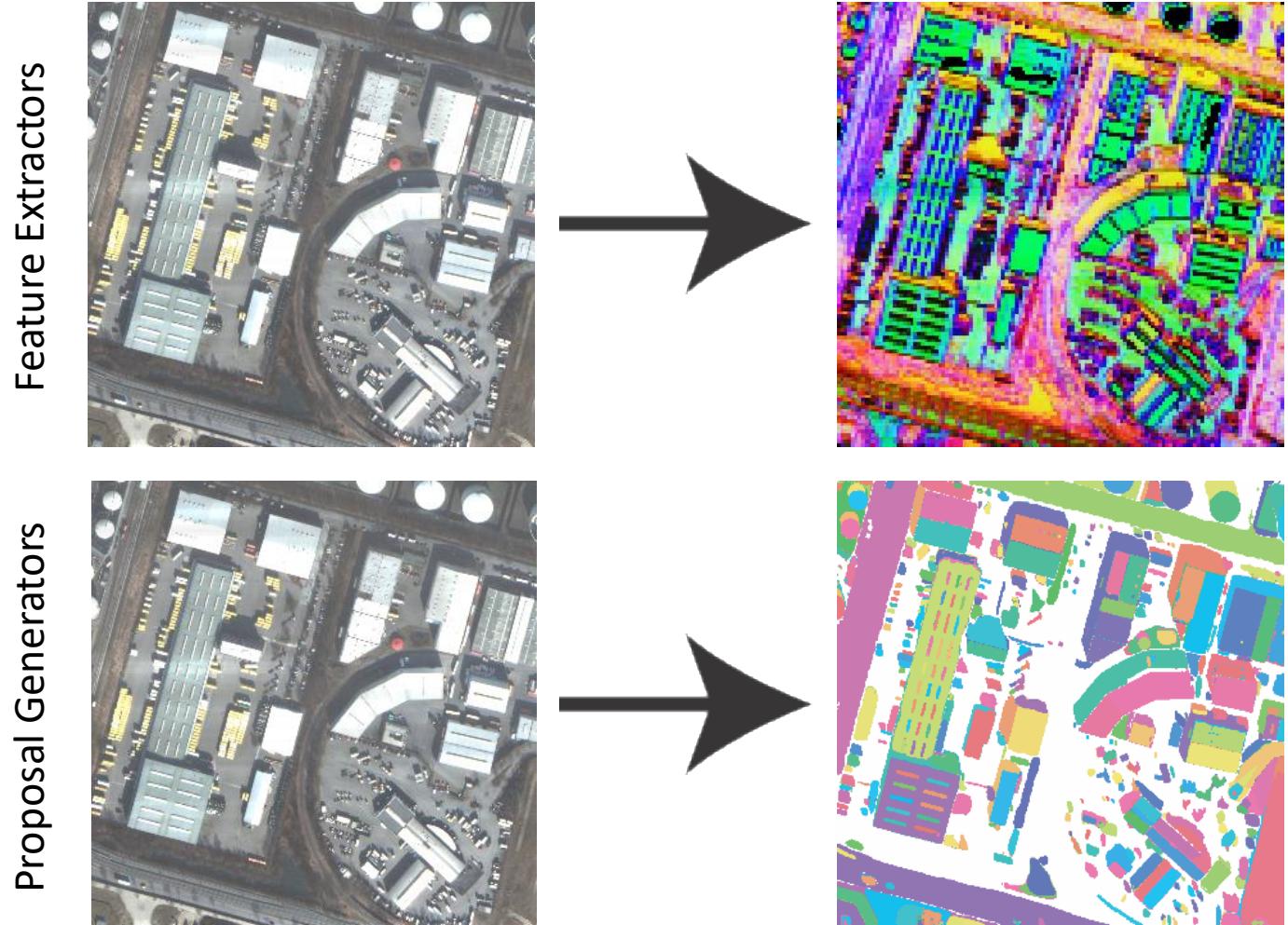
Dalton Lunga, ORNL

Paper ID: 4552

Motivation: Foundation Models are powerful

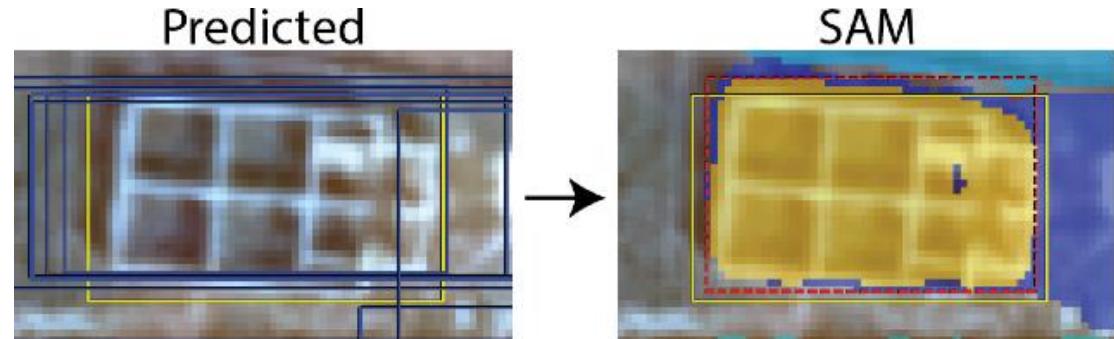
- Feature Extractors (FM-FE):
 - ResNet, DINOV3, ...
 - Strong features
 - But no bounding boxes
- Proposal Generators (FM-PG):
 - Segment Anything
 - Good Masks
 - Weak features
- Labeling new data is expensive

→ Combine both with AL

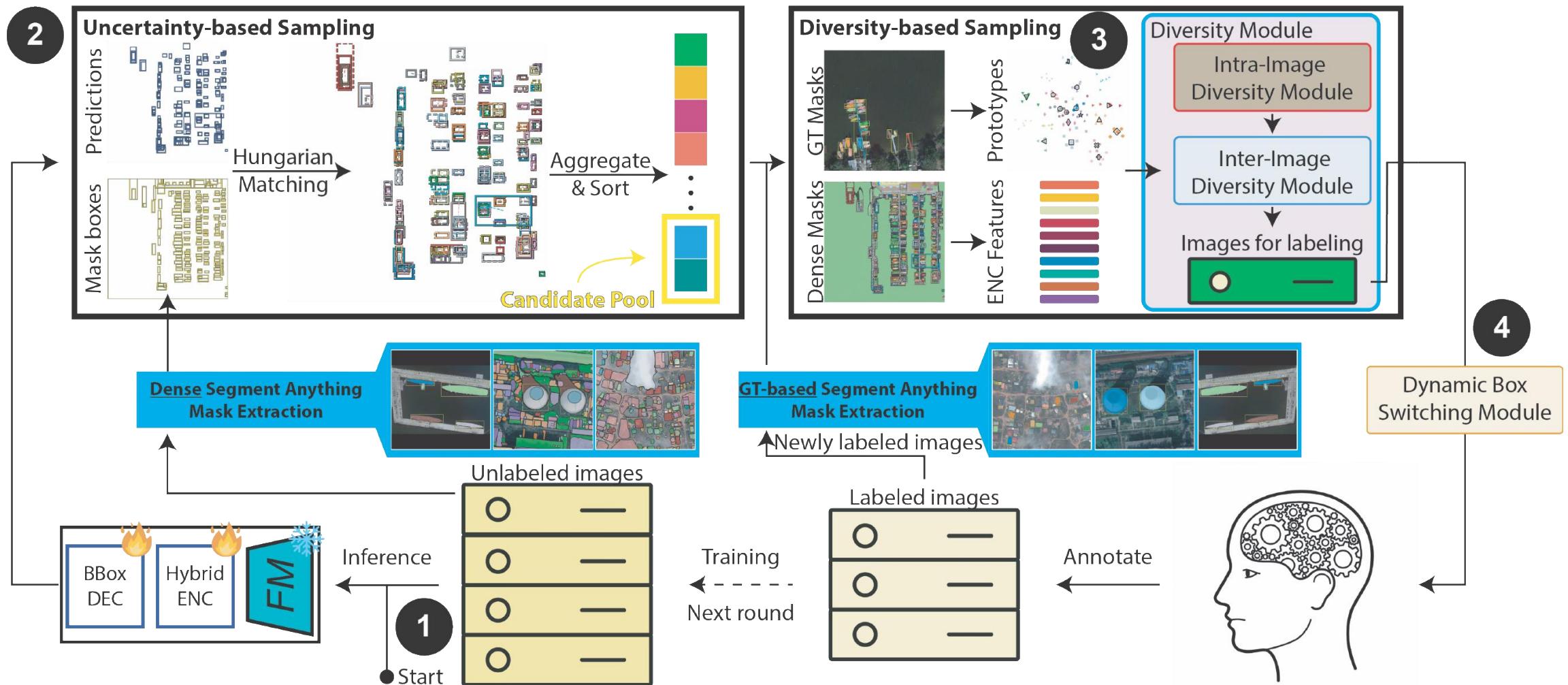


Key Contributions

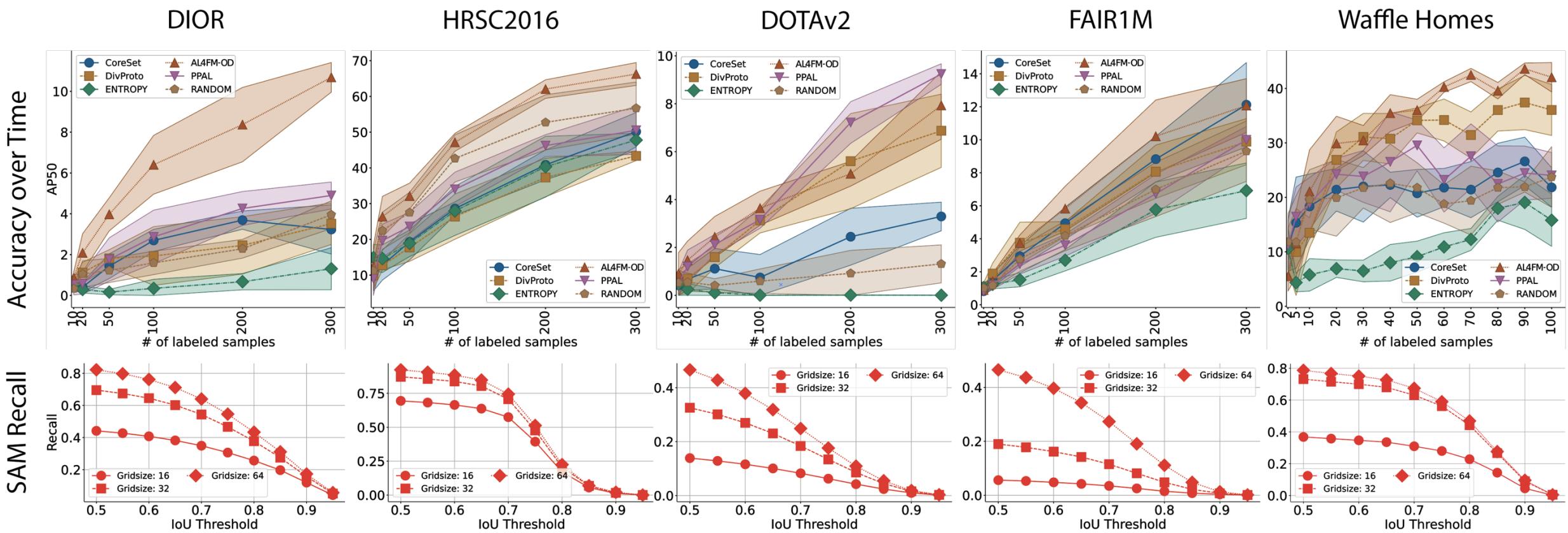
- Combine Foundation Models with AL
 - Reduce annotation effort by leveraging pretrained knowledge
- Dual Source Uncertainty (FM-PG + Detector)
 - Integrate proposals from the FM-PG and the detector predictions
- Mask-Guided Diversity (FM-PG + FM-FE)
 - Ensures diverse and informative samples
- Dynamic Box Switching (FM-PG)
 - Combats cold-start issues



Method Overview

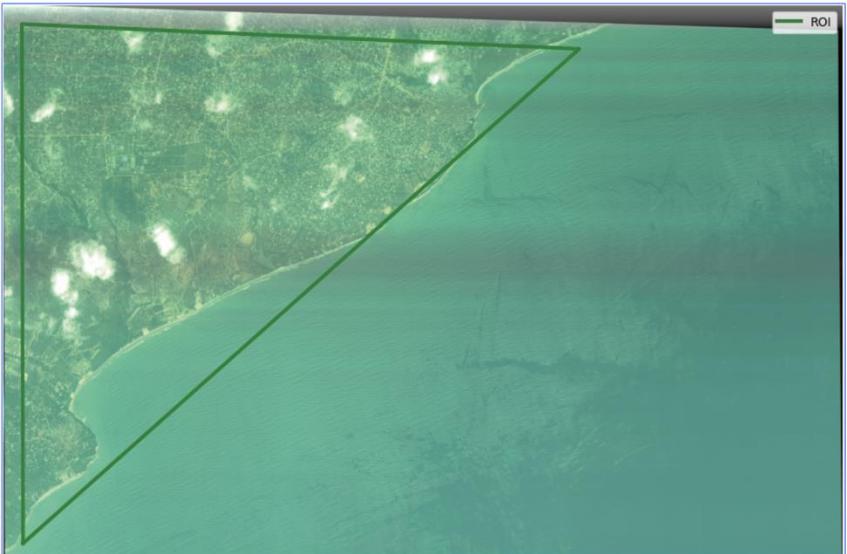


Key Results



User Interface & User Study

Area of Interest



User Interface



User Study Results

Interaction Results

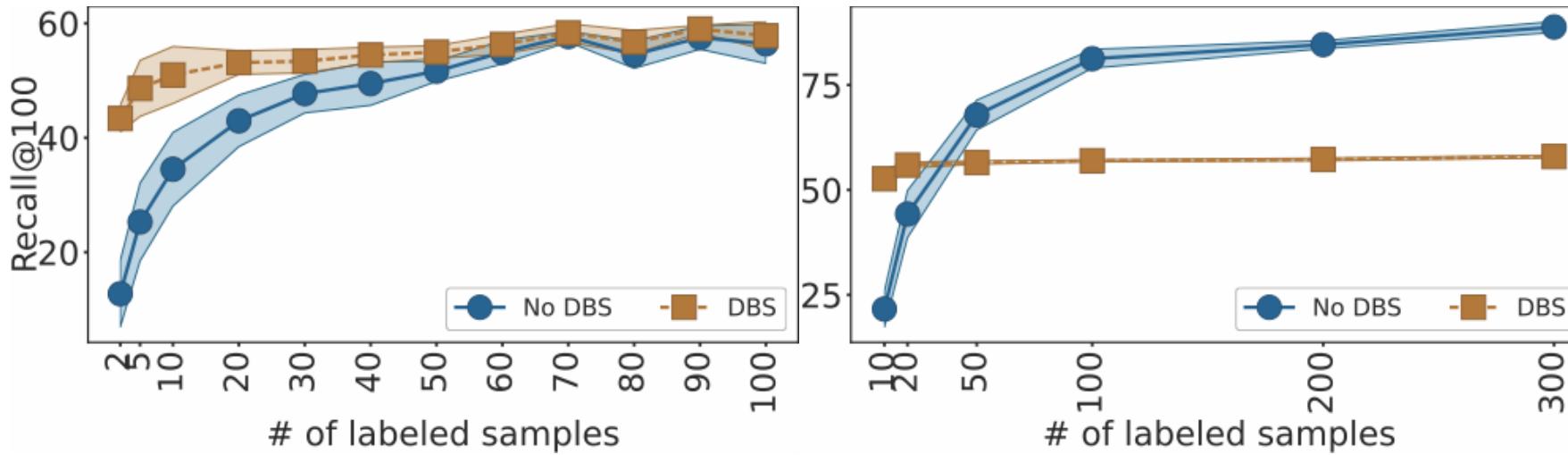
	Random	Ours
Avg. # of image annotated	10.67	11.75
Avg. # of accepted boxes	13.33	51.75
Avg. # of drawn boxes	196	185
Avg. total annotations	209	235

Final Dataset Results

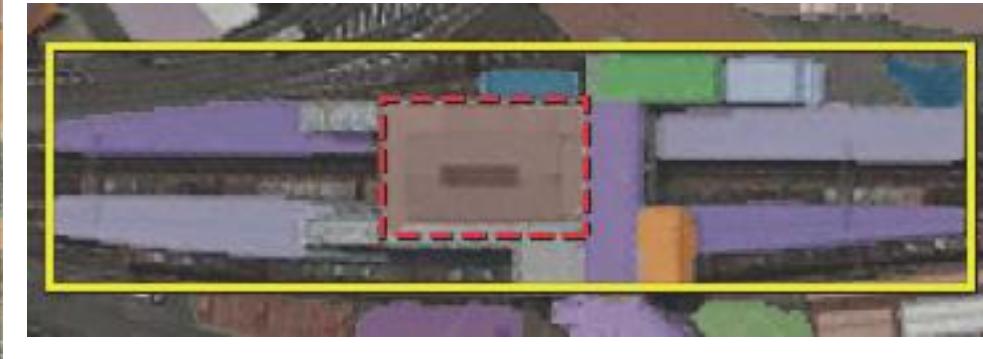
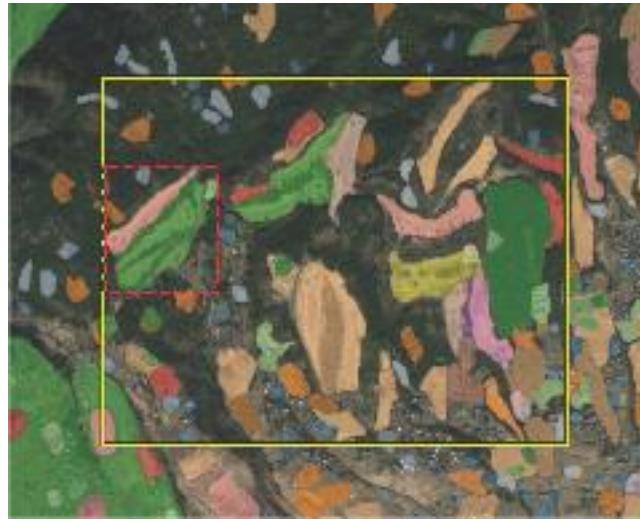
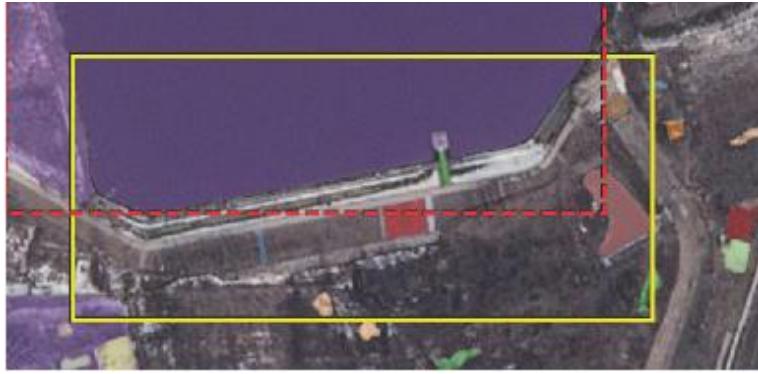
	Random	Ours
AP ₅₀	0.28±0.03	0.34±0.04
AP ₇₅	0.11±0.05	0.17±0.05
AR ₁₀₀	0.35±0.04	0.43±0.03

Limitations

Dynamic Box Switching Module



SAM Segmentation



Conclusion & Future Work

- Foundation models can reduce annotation needs
- Mask-Guided diversity sampling beats box-based sampling
- Dynamic Box Switching handles cold start well
- User study confirms less effort and higher annotation quality

Future Work

Segmentation



Multi-Modal Data



Open-Set Detection



Continuous Learning



GeoAI Group

CVL

GitHub



Thank you

Marvin Burges*
CVL, TU Wien

Paper ID: 4552