Bootstrapping Rare Object Detection in High-Resolution Satellite Imagery

Akram Zaytar[1], Caleb Robinson[1], Gilles Hacheme[1], Girmaw A. Tadesse[1], Rahul Dodhia[1], Juan M. Lavista Ferres[1], Lacey F. Hugey[2], Jared A. Stabach[2] , Irene Amoke[3]

[1] *Microsoft AI for Good Research Lab*

[2] *Smithsonian National Zoo & Conservation Biology Institute*

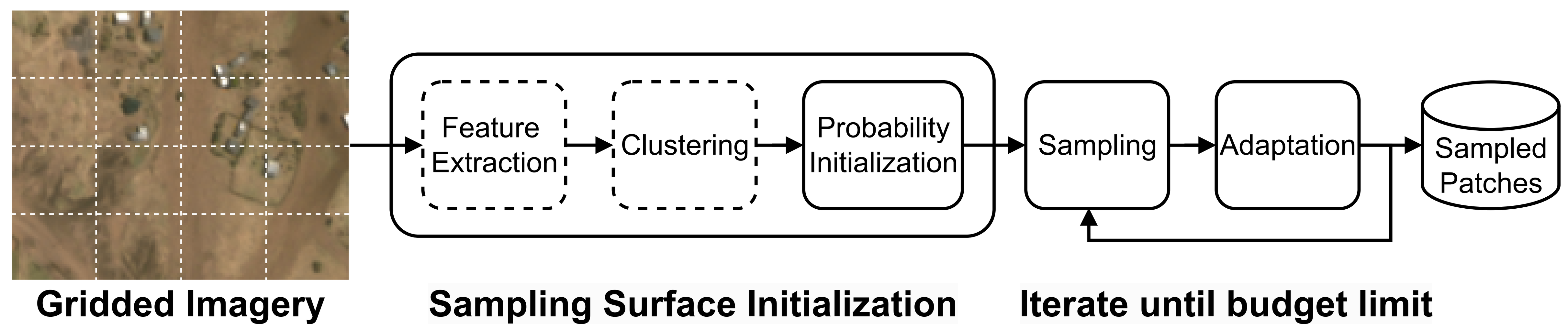
[3] *Kenya Wildlife Trust*

Challenge

- **How can we solve rare object detection tasks in a new domain without labeled data?**
- Problem of “kickstarting”.
- In such problems, we are limited by:
  - Small labeling budget + large amounts of high-resolution satellite imagery
  - No spatial priors
  - Limited computing budget

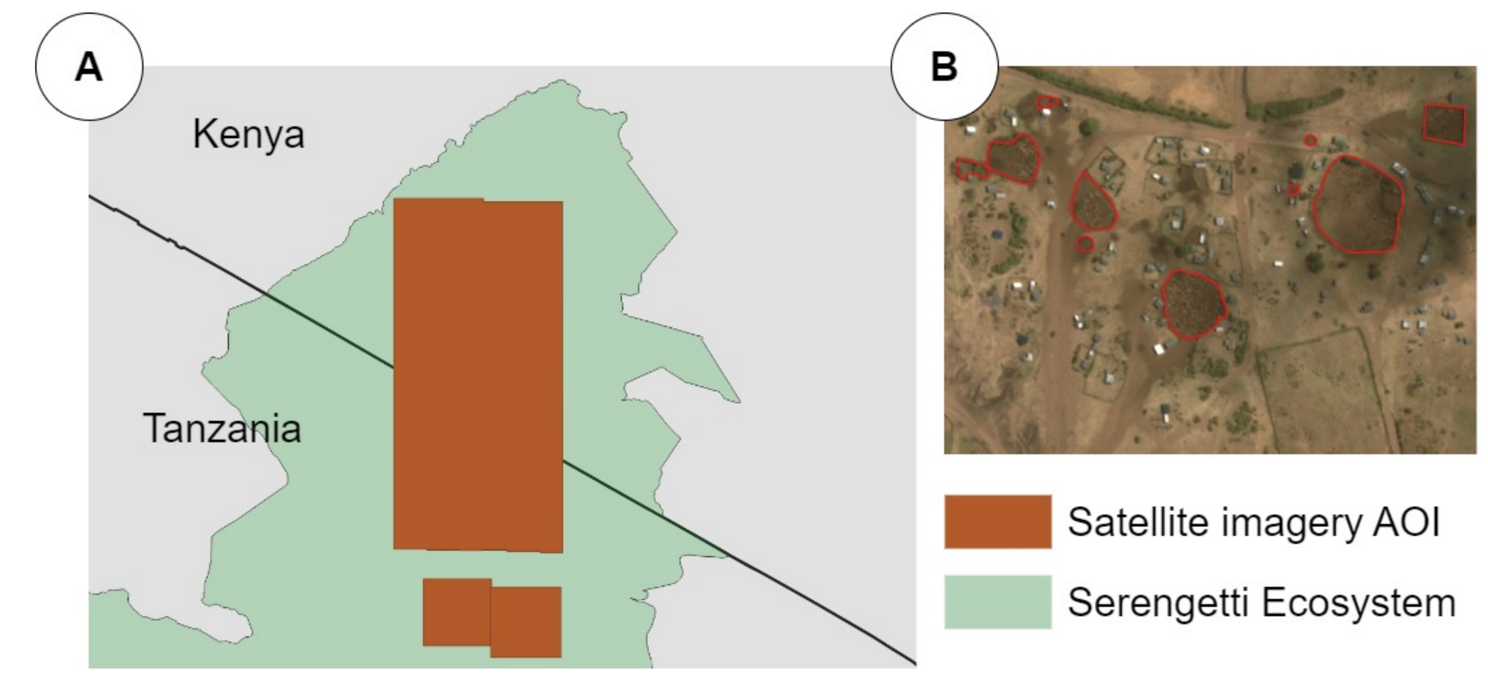
Problem: Given unlabeled imagery and a rare object class of interest, we want to label a core set of representative samples

Methods



- **Initialization**: uniform or cluster-based weights of feature vectors (Random CNN, ColorStats, ResNet-18).
- **Clustering Algorithms**: KMeans, Bisecting K-Means, or DBSCAN to create clusters.
- **Sampling Strategies**:
  - **Offline**: Sample based on the static initial probabilities.
  - **Online**: increase the sampling weight of neighbouring samples (e.g., spatially or in feature space).
- **Sampling**: Iterative sampling & re-weighting under a fixed budget.
- **Training**: Post-sampling, use the positive, negative, and annotated samples to train a model.Z

Use Case

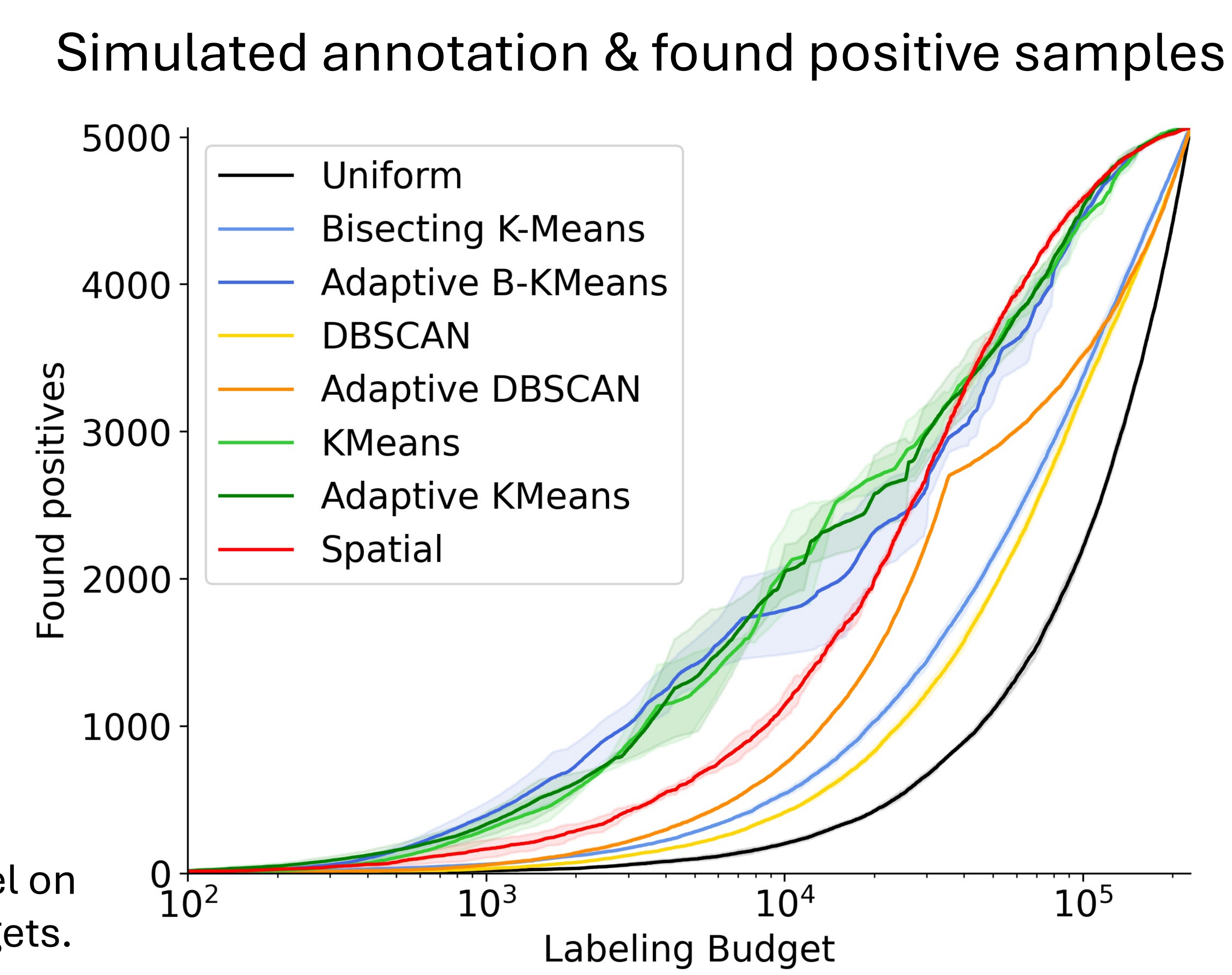


Case study: we want to find “bomas” (cattle enclosers) in high-resolution satellite imagery captured over the Serengeti Mara region of Kenya and Tanzania.

Results

		300 Patch Budget		950 Patch Budget		3,000 Patch Budget	
Method	**Strategy**	**% positive**	**F1**	**% positive**	**F1**	**% positive**	**F1**
Uniform	Random	1.6 ± 0.2	0.06 ± 0.14	1.5 ± 0.3	0.33 ± 0.19	1.8 ± 0.3	0.65 ± 0.04
Spatial	Proximity	**14.0 ± 2.8**	0.15 ± 0.12	**16.5 ± 6.4**	0.42 ± 0.19	**14.0 ± 1.1**	0.76 ± 0.02
DBSCAN	Static	3.3 ± 0.7	0.00 ± 0.01	2.8 ± 0.2	0.30 ± 0.16	**4.0 ± 0.01**	0.70 ± 0.03
	Adaptive	**3.3 ± 0.4**	0.02 ± 0.02	**5.2 ± 1.1**	0.31 ± 0.21	**7.1 ± 0.2**	0.55 ± 0.24
KMeans	Static	13.3 ± 3.3	0.30 ± 0.23	28.9 ± 8.2	**0.72 ± 0.04**	29.4 ± 4.2	0.75 ± 0.04
	Adaptive	**28.3 ± 8.6**	0.18 ± 0.07	32.6 ± 7.7	0.42 ± 0.17	28.1 ± 2.2	0.65 ± 0.19
B-KMeans	Static	6.0 ± 0.8	0.28 ± 0.18	6.1 ± 0.2	0.71 ± 0.03	5.6 ± 0.2	**0.77 ± 0.02**
	Adaptive	19.0 ± 1.0	**0.36 ± 0.10**	**39.0 ± 8.6**	0.70 ± 0.13	**33.6 ± 4.8**	0.76 ± 0.01

Table 1. Percentage of positive patches found and performance of downstream U-Net model on the boma detection task for different sampling methods under different patch labeling budgets.



Research Directions

- Investigate how the bootstrapping techniques bridge to subset selection methods for incremental label creation.
- Simulate realistic human annotation scenarios where a visual “window-view” is presented instead of individual patches.
- Extend the proposed methods to other established geospatial datasets and examine the trade-offs.

Paper: