# The Nature Conservancy: Advancing Computer Vision for Invasive Species Detection

Loralee Ryan, Sam Collier, Jongkyung Kim

## The Nature Conservancy & Key Reserves

The Nature Conservancy (TNC) is a wide-reaching **environmental conservation organization** whose primary mission is to conserve "the lands and waters on which all life depends"

Santa Cruz Island is a TNC-managed island off the coast of California and leading example of successful island restoration. Through 40 years of work, the island is **completely free** of non-native animals and is home to over 1000 species, **12** of which are unique to the island.

The Jack and Laura Dangermond Preserve (JLDP) is a vast and diverse TNC-managed ecosystem in Southern California.

## The Challenge

Invasive-species driven extinctions of birds, amphibians, mammals, and reptiles on islands make up **75% of all globally**.

Remedial actions and long-term effects of invasive species can be ecologically and economically devastating, costing some **$1 trillion** in the last 50 years [1].

The low data availability for invasive species, the similarity between classes of species and poor transferability of models between biomes make for a **hard ML problem**.

## The Solution

Build invasive species detectors using an established network of camera traps. In doing so, push forward the SOTA in camera trap computer vision.

**Main Contribution:** Proof of concept of a camera trap anomaly detection algorithm that learns representations of native species to identify non-native species in new images

**Contribution 2:** Learnings for future applications of CV to camera trap imagery

## Anomaly Detection Algorithm

Two key elements in the algorithm:

**SimCLR:** Learns rich features from images by maximizing agreement between differently augmented views of the same data example via a contrastive loss in the latent space [2]

**OneClassSVM/DBSCAN:** Outlier detection comparing SimCLR representation of a new image with representation of native images

The workflow used in our proof-of-concept is below:

**Data:** 4.5K JLDP images, $X_{tr}, y_{tr}, X_{te}, y_{te}$
$f \leftarrow$ ResNet50;
`# Remove anomalies (bobcats) in train`
$X_{tr}, y_{tr} \leftarrow X_{tr} - X_{bobcat}, y_{tr} - y_{bobcat}$;
`# Train for features & extractor`
$h_{tr}, f_{tr} = SimCLR(X_{tr}, y_{tr}, f)$;
**if** *species_specific* **then**
  **for** $y^i_{tr} \in y_{tr}$ **do**
    `# Train one-class SVM`
    $\text{1CSVM}_i = \text{1CSVM}(h^i_{tr})$ ;
  **end**
**end**

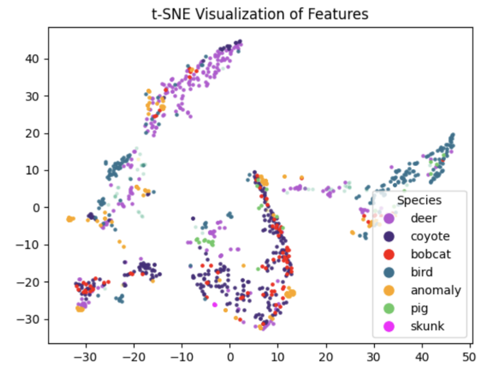
`# To test, run Yolo to crop images`
$X_{crop}, y_{pred} = \text{YOLO}(X_{te})$;
$h_{pred} = f_{tr}(X_{crop})$ ; `# Get SimCLR features`
**if** *species_specific* **then**
  **for** $y^i_{pred} \in y_{pred}$ **do**
    **Result:** $\text{1CSVM}_i(h^i_{pred})$
  **end**
**else**
  **Result:** $\text{DBSCAN}(h_{pred})$
**end**

**Notes:**

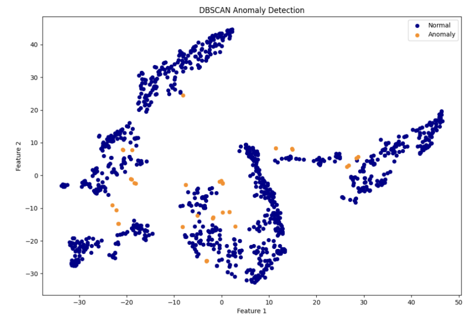
1. We created new anomalous images by swapping crops of bobcats onto different backgrounds
2. We ran both global anomaly detection and species specific anomaly detection
3. t-SNE was used for visualizations

## Initial results

**Result 1:** At species level, model can identify truly anomalous data, but struggles when representations are similar (e.g. bobcat-coyote)




**Result 2:** Model can identify species not present in the train data as anomalous (e.g. skunks)




## Further work

- New models, e.g. isolation forest or GMM's for outlier detection
- Higher quality generated anomalies with GAN's
- Consider low-confidence YOLO predictions as potential anomalies

## References

1. R. D. Zenni, F. Essl, E. García-Berthou, and S. M. McDermott, 'The economic costs of biological invasions around the world', NeoBiota, vol. 67, pp. 1–9, Jul. 2021, doi: 10.3897/neobiota.67.69971.
2. T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, 'A Simple Framework for Contrastive Learning of Visual Representations'. arXiv, Jun. 30, 2020. doi: 10.48550/arXiv.2002.05709.