ICCV HONOLULU HAWAII
OCT 19-23, 2025

AnimalClue: Recognizing Animals by their Traces

Risa Shinoda [1,3]

Nakamasa Inoue [2,3]

Iro Laina [4]

Christian Rupprecht [4]

Hirokatsu Kataoka [3,4]

1 The University of Osaka

2 Institute of SCIENCE TOKYO

3 AIST

4 VGG University of Oxford

AnimalClue includes footprints, feces, eggs, bones, and feathers, totaling 159,605 bounding boxes from 968 species, and annotations for 22 traits.

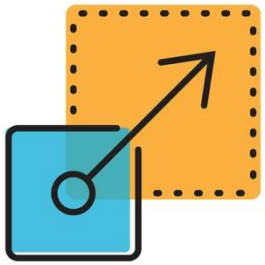
Motivation

Identifying diverse animal species remains difficult,
as most models overlook nocturnal or camouflaged animals.

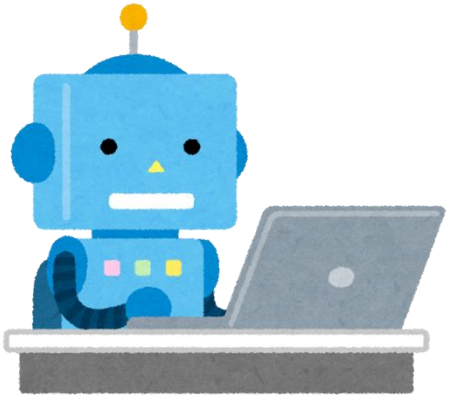
Wildlife monitoring can use indirect evidence to overcome this limitation,
but it requires highly domain-specific knowledge.

However, manual inspection limits scalability, creating a strong need for vision models that can identify species from such traces.

Highly domain-specific knowledge



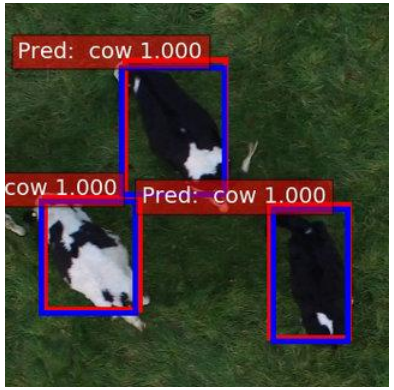
Limits Scalability



Vision-based Automation

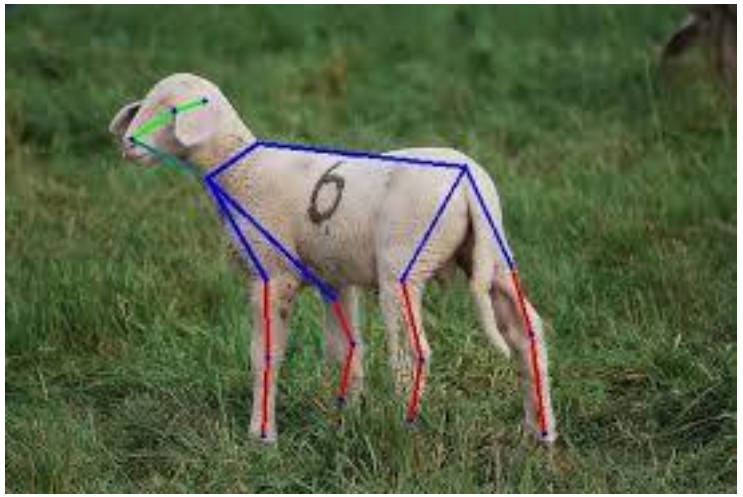
Motivation

Recent advances in computer vision automate key wildlife observation tasks.



Animal Identification
[Andrew+, ICCVW, 2017]

Activity recognition
[Chen+,CVPR,2023]



Pose Estimation
[Cao+,ICCV,2019]

However, indirect evidence identification (e.g., footprints, feces) remains underexplored, despite its importance.

Related Works

- Several studies have tackled trace recognition.
- However, species coverage and bounding boxes are limited, and most tasks focus only on classification, making them unsuitable for real wildlife monitoring scenarios.
- Our AnimalClue dataset includes five major types of traces with segmentation masks and 22 traits.

Dataset	Track Type	#Species	#Bbox	Task	#Traits
OpenAnimalTracks [42]	Footprint	18	3,579	CLS, DET	0
DFML [30]	Feces	1	1,623	CLS	1
Automated Egg Classification [10]	Egg	2	2,943	CLS	0
Skull2Animal dataset [33]	Bone	4	4,962	CLS	0
FeathersV1 [5]	Feather	595	28,272	CLS	0
AnimalClue (Ours)	All 5 Types	968	159,605	CLS, DET, SEG	22

Dataset Creation

iNaturalist

Research Grade

CC License

Eurasian Lynx

Fieldfare

Red -shouldered Hawk

Image-Label Pairs

Grounding DINO

Segment Anything

Initial Masks

Dataset Creation

Initial Masks

Overlaid texts

Animals in the frame

Distant subjects

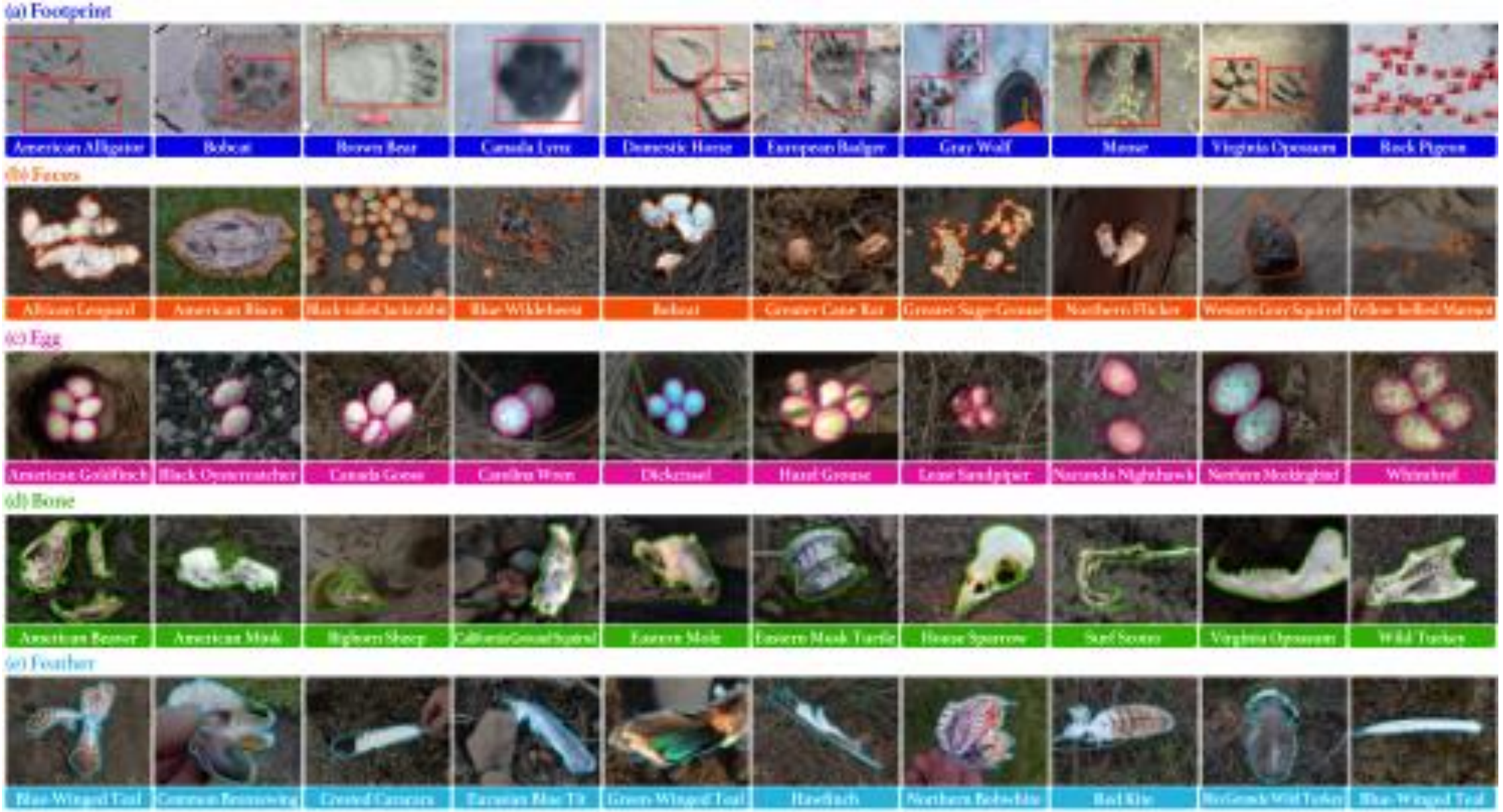
Visible human faces

Remove

Modify

Final Dataset

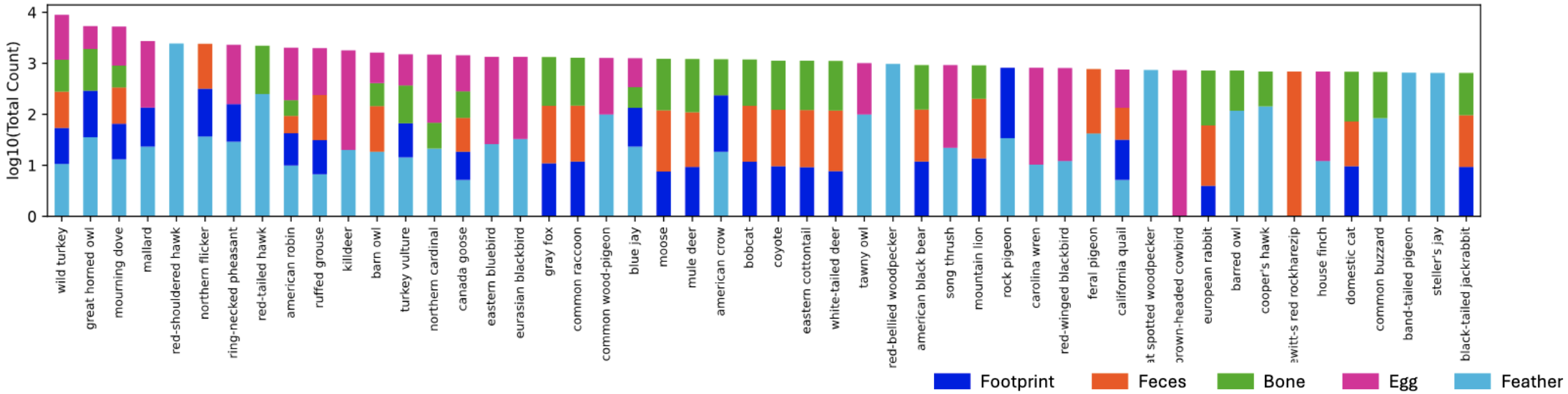
Dataset Examples



Statistics

Our dataset consists of **968 species, 200 families, and 65 orders**. It includes a total of **159,605 bounding boxes** across five trace types:

- **Footprints:** 18,291 boxes from 7,581 images, covering 117 species, 46 families, and 20 orders
- **Feces :** 18,932 boxes from 6,433 images, covering 101 species, 46 families, and 21 orders
- **Bones:** 16,553 boxes from 12,908 images, covering 269 species, 112 families, and 45 orders
- **Eggs:** 29,434 boxes from 9,394 images, covering 283 species, 67 families, and 20 orders
- **Feathers:** 76,395 boxes from 60,491 images, covering 555 species, 89 families, and 30 orders



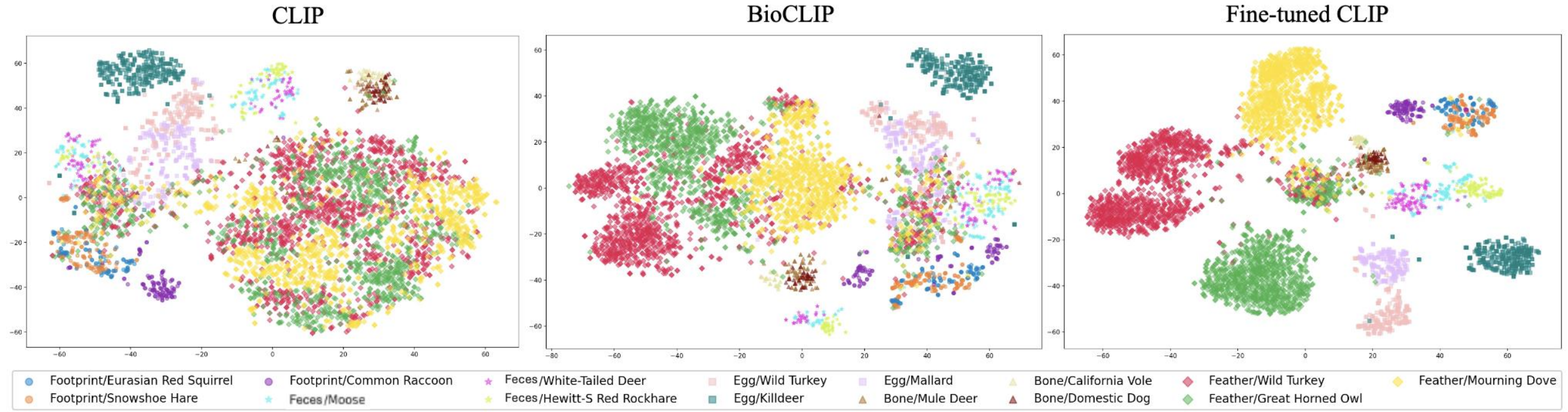
Classification

- Swin-B achieves the highest top-1 accuracy among baselines.
- Feathers show the highest accuracy across taxonomies.
- Eggs follow, consistent with their second-largest number of species.

Model	Species					Family					Order				
	Footprint	Feces	Egg	Bone	Feather	Footprint	Feces	Egg	Bone	Feather	Footprint	Feces	Egg	Bone	Feather
All Categories															
VGG-16 [31]	28.8	29.6	45.2	14.7	56.7	45.6	46.6	61.1	31.0	66.1	62.1	65.1	81.2	54.2	78.7
ResNet-50 [14]	23.7	29.4	41.1	18.3	59.7	41.7	48.6	59.9	33.6	70.1	58.9	64.8	80.8	53.0	81.8
EFNet-B1 [45]	25.9	30.5	41.0	15.0	55.9	42.2	48.4	56.3	29.3	64.4	56.5	61.3	74.5	45.4	77.1
ViT-B [25]	29.2	32.2	46.7	15.0	55.9	47.0	51.8	63.7	28.8	69.9	61.8	69.3	83.1	50.9	81.2
Swin-B [32]	32.3	38.6	49.4	20.5	65.3	49.3	56.8	65.1	37.6	72.4	66.1	70.4	84.0	56.6	77.7
Frequent Categories															
ResNet-50 [14]	24.8	30.6	49.4	26.7	68.5	44.0	51.7	65.5	42.8	74.2	62.9	70.6	83.9	62.7	80.4
Swin-B [32]	33.7	40.3	58.5	25.6	65.3	52.4	60.6	70.0	47.8	72.4	69.2	75.1	86.2	64.4	80.8
Rare Categories															
ResNet50 [14]	7.4	25.0	12.1	1.00	2.16	27.0	25.8	22.4	10.6	16.9	33.3	28.6	63.0	15.4	26.1
Swin-B [32]	14.2	28.1	14.1	4.91	2.52	40.4	22.7	26.3	15.4	22.1	44.4	31.0	66.7	11.5	30.4

Visualization of t-SNE

- Our fine-tuned CLIP model achieves the best feature separation across categories.
- BioCLIP separates feathers and eggs better than CLIP,
  but footprints and eggs remain challenging.
- This suggests that models need to be trained specifically on trace data,
  even if pre-trained on large-scale animal image datasets.



Detection

- RT-DETR shows the best overall performance.
- Performs well on frequent categories, but struggles with rare ones.
- Feathers achieve the best results overall, though minor categories remain challenging likely due to their large category diversity.

Model	Species					Family					Order				
	Footprint	Feces	Egg	Bone	Feather	Footprint	Feces	Egg	Bone	Feather	Footprint	Feces	Egg	Bone	Feather
All Categories															
YOLOv8 [21]	0.10	0.11	0.13	0.08	0.25	0.17	0.16	0.33	0.14	0.19	0.22	0.20	0.50	0.16	0.43
YOLOv11 [20]	0.10	0.12	0.14	0.07	0.25	0.17	0.16	0.36	0.13	0.35	0.24	0.21	0.47	0.16	0.43
Faster-RCNN [39]	0.04	0.06	0.12	0.07	0.08	0.07	0.09	0.17	0.0	0.12	0.08	0.13	0.26	0.10	0.22
DINO [57]	0.08	0.12	0.20	0.07	0.15	0.12	0.17	0.32	0.17	0.23	0.14	0.22	0.52	0.22	0.34
RT-DETR [58]	0.10	0.17	0.04	0.01	0.17	0.21	0.25	0.42	0.17	0.40	0.31	0.28	0.57	0.21	0.50
Frequent Categories															
YOLOv11 [20]	0.17	0.18	0.38	0.11	0.56	0.29	0.26	0.63	0.29	0.70	0.33	0.40	0.73	0.44	0.81
RT-DETR [58]	0.19	0.20	0.21	0.06	0.53	0.33	0.31	0.62	0.39	0.76	0.35	0.46	0.73	0.48	0.84
Rare Categories															
YOLOv11 [20]	0.04	0.05	0.04	0.003	0.14	0.07	0.08	0.07	0.05	0.05	0.10	0.02	0.26	0.007	0.005
RT-DETR [58]	0.006	0.09	0.00	0.00	0.00	0.11	0.21	0.08	0.00	0.00	0.34	0.10	0.38	0.10	0.00

Segmentation

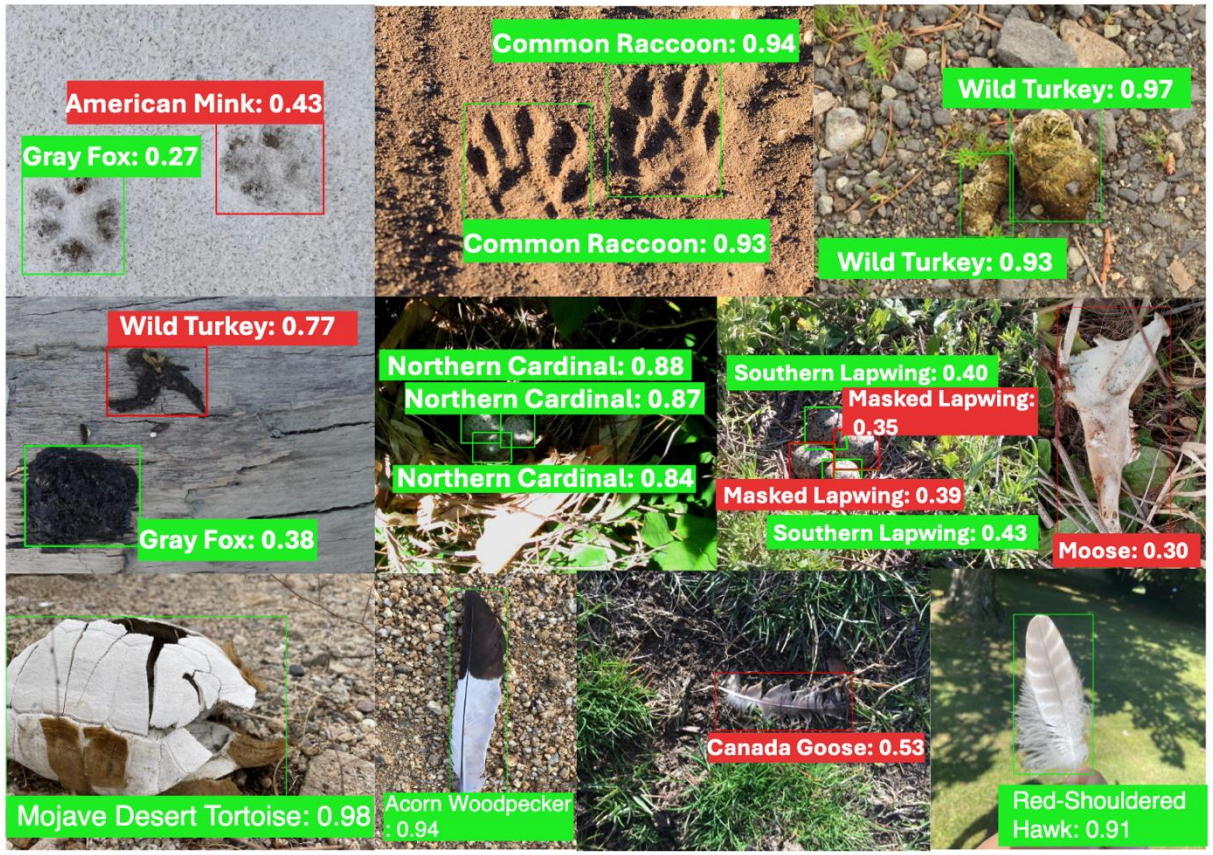
- YOLOv8, YOLOv11, and MaskDINO show competitive results.
- Feathers achieve the highest mAP, even with the largest number of species.

Model	Species				Family				Order			
	Feces	Egg	Bone	Feather	Feces	Egg	Bone	Feather	Feces	Egg	Bone	Feather
All Categories												
YOLOv8 [21]	0.11	0.11	0.07	0.24	0.14	0.29	0.13	0.33	0.20	0.44	0.15	0.41
YOLOv11 [20]	0.11	0.12	0.06	0.24	0.15	0.32	0.13	0.34	0.20	0.45	0.15	0.41
Mask-RCNN [1]	0.08	0.16	0.05	0.08	0.10	0.22	0.08	0.17	0.13	0.35	0.10	0.24
MaskDINO [27]	0.13	0.25	0.07	0.18	0.18	0.32	0.11	0.27	0.23	0.48	0.16	0.37
Frequent Categories												
YOLOv8 [21]	0.14	0.29	0.11	0.48	0.22	0.54	0.29	0.66	0.34	0.65	0.43	0.79
YOLOv11 [20]	0.16	0.33	0.10	0.49	0.24	0.57	0.28	0.68	0.37	0.64	0.45	0.80
Rare Categories												
YOLOv8 [21]	0.07	0.02	0.006	0.12	0.04	0.03	0.06	0.06	0.03	0.26	0.05	0.02
YOLOv11 [20]	0.06	0.03	0.005	0.11	0.07	0.06	0.08	0.05	0.06	0.27	0.05	0.002

Visualization

Object detection (left) and segmentation (right) on AnimalClue using YOLOv11

Faint footprints or unclear traces often lead to species misidentification, reflecting the challenge of real-world images.



Green: Correct Red: Incorrect

Traits Classification

- Feathers achieved the best results.
- For feces, aquatic and predator traits showed high accuracy, likely due to clear morphological differences.
- These results suggest that trait inference depends on the trace type.

	Diet Type	Activity Pattern	Aquatic	Urban	Tropical	Polar	Herding	Predator
Footprint	57.3 / 47.0	69.9 / 65.8	84.4 / 70.2	95.2 / 60.0	91.3 / 63.5	62.2 / 61.8	84.2 / 73.3	80.7 / 94.3
Feces	76.0 / 64.7	72.9 / 72.1	93.1 / 75.1	98.6 / 59.3	82.7 / 70.0	69.1 / 45.6	74.7 / 53.4	88.4 / 85.7
Egg	68.7 / 48.2	95.4 / 73.0	88.4 / 83.4	76.8 / 75.2	90.6 / 69.7	93.8 / 75.4	82.5 / 81.5	96.9 / 73.4
Bone	64.8 / 59.0	60.6 / 60.7	81.5 / 73.3	75.2 / 61.4	72.6 / 64.9	87.5 / 62.7	73.9 / 73.1	76.9 / 72.1
Feather	81.3 / 80.0	89.7 / 63.7	92.1 / 81.6	83.0 / 81.5	91.5 / 69.2	94.7 / 61.9	86.3 / 81.7	91.8 / 90.7

Acc/F1 score

Summary

Project Page

We propose AnimalClue, a dataset for identifying animal species from their traces.

- Covers footprints, feces, eggs, bones, and feathers
- Totaling 159,605 bounding boxes from 968 species, and annotations for 22 traits.
- We evaluate representative vision models and highlight key challenges in trace-based identification.
- We hope this dataset advances research in animal tracking and wildlife conservation.