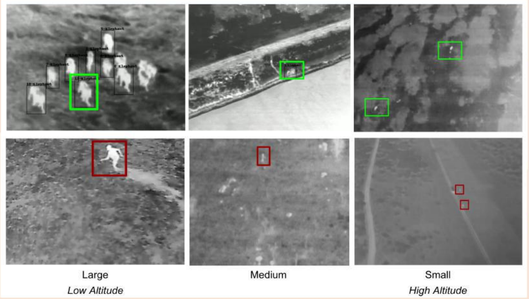
ElSa: A Novel Real-time Wildlife Poacher Detection Solution Leveraging Machine Learning Driven Spatio-temporal Analysis of Nighttime UAV Thermal Infrared Videos

EAEV066

Anika Puri, Horace Greeley High School, Chappaqua, NY, United States

WWF Says African Elephants Will Be Extinct by 2040 If We Don't Act Right Away

To survey a wider area within flying time constraints, UAVs must fly at **heights ~500 ft** (high altitudes)
→ small animal/human sizes



https://sites.google.com/view/elizabethbondi/dataset

Current Solutions

Recent methods leverage computer vision methods:
only study the object's shape in **static images** frames when tested on real-life data:
→ **only 20%** human recognition accuracy
→ **fail to mitigate** high-resolution cameras ($10,000)

Objectives

Software: Develop **a high accuracy** automated computer-vision method for detection of **potential poaching activity**

Hardware Engineering: *Design &* implement **a low-cost engineering design**, built with commodity hardware, for nighttime thermal infrared video capture

Scientific Basis of Research

The **movement patterns** of elephants **differ significantly** from those of humans with respect to speed, turning patterns etc.

→ Previous research fails to leverage this important spatio-temporal nature of *video* data

Methodology/Project Design

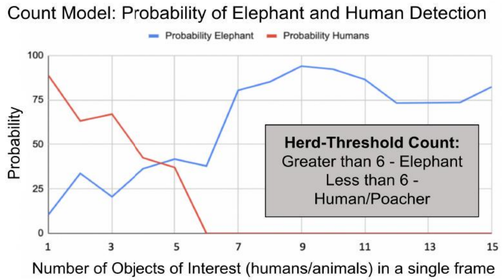
1) **Studied BIRDSAI Dataset:** first large real-life dataset of 172 thermal drone videos collected from 4 national parks in Kenya, South Africa and Uganda
2) **Extracting Spatiotemporal Data:** Measure movement of a moving object (ex. Elephant/human) with respect to a fixed object (ex. Bush, tree)
3) **Outlier Removal and Time Series Smoothing** (Hampel filter & ewm weighted functions)

Training Dataset: 516 total spatiotemporal times series

Machine Learning Model Training: K-Nearest Neighbor (KNN) with Dynamic Time Warping Metric to measure similarity

Herd Count Model Results

of objects of interest can be used **to further improve accuracy**



Results of KNN Spatiotemporal Model

Test Set: 144 spatio-temporal time series

- **36 human patterns**
- **108 elephant patterns**

Method 1: Spatiotemporal Model - Dynamic Time Warping	
Accuracy: 74%	

Method 2: Spatiotemporal Model - Feature Extraction With Herd Count Feature	
Accuracy: 90.9%	

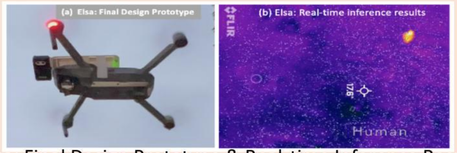
	Actual: Human	Actual: Elephant
Predicted: Human	True Positive 34	False Positive 36
Predicted: Elephant	False Negative 2	True Negative 72

Sensitivity 94.4% Specificity 67%

ElSa is able to detect poachers in nighttime, infrared videos from >500 feet off the ground with a **4X higher accuracy** than the current standard benchmark.

Interpretation & Conclusion

- **First-ever machine learning model** to utilize multi-modal data with **spatio temporal movement patterns** for wildlife conservation.
- When tested using a thermal infrared video dataset collected from four national parks in Africa, my research was able to use movement patterns to detect humans with **over 90% accuracy - a 4X improvement over the existing state-of-the-art methods.**
- Furthermore, my low-cost ($300) engineering design ElSa built with commodity components **mitigates the need for costly high-resolution thermal cameras** ($10,000), **easing the burden** on resource-constrained Parks in Africa and Asia.



Elsa: Final Design Prototype & Real-time Inference Results

Future: Apply this novel methodology to the conservation of other endangered animals