Wildlife@Home

Travis Desell & Susan Ellis-Felege
Departments of Computer Science & Biology
University of North Dakota
http://csgrid.org/csg/wildlife

What is Wildlife@Home?

- A *citizen science* project that combines both crowd sourcing and volunteer computing.
- Users volunteer their brain power by observing videos and reporting observations.
- Users volunteer their computer power by downloading videos and performing.
- A scientific web portal to robustly analyze and compare results from users, experts and the computer vision techniques.

Between 2012 and now, Dr. Ellis-Felege has gathered over 100,000 hours of avian nesting video from the following species:

1. Sharp-tailed grouse (*Tympanuchus phasianellus*), an important game bird and wildlife health indicator species.
2. Piping plovers (*Charadrius melodus*), a federally listed threatened species.
3. Interior least terns (*Sternula antillarum*), a federally listed endangered species.
4. Blue Winged Teal (*Anas discors*), in collaboration with Ducks Unlimited.

We have recently received over 2 million motion sensor camera images from a new Hudson Bay project, and multiple terabytes of aerial imagery gathered by a Trible UX5 unmanned aerial vehicle.

Sharp-tailed Grouse

Piping Plover

All four current species are ground nesting birds.

Sharp-tailed grouse nest in the dense grass (top left). Nests were monitored in areas of high oil development, moderate oil development and no oil development (protected state land).

Piping plover and interior least tern are shore nesting species (top right). Nests were monitored along the Missouri River in North Dakota.

What's the point?

1. Current cameras that use automated motion detection miss some predators and are not robust enough).

2. Camera footage allows Dr. Ellis-Felege to manage and evaluate studies with large enough sample sizes for statistical significance.

3. Answer biological questions about parental investment and predator-prey interactions for these ground nesting species.

4. Examine the effect of oil development on wildlife in western North Dakota, which is experiencing a boom in fracking.

Most grouse video is sleeping birds and grass blowing in the wind. But occasionally, interesting things happen.

Piping plover and tern video is more interesting, with active bi-parental involvement and less obscuring vegetation.

There are many challenges:

1. Dramatically changing weather conditions
2. Dawn/Day/Dusk/Night lighting conditions
3. Model species (sharp tailed grouse and piping plover) and some predators have cryptic coloration (camouflage).
4. Moving vegetation and insects can cause false negatives.

From all this video, we want to determine:

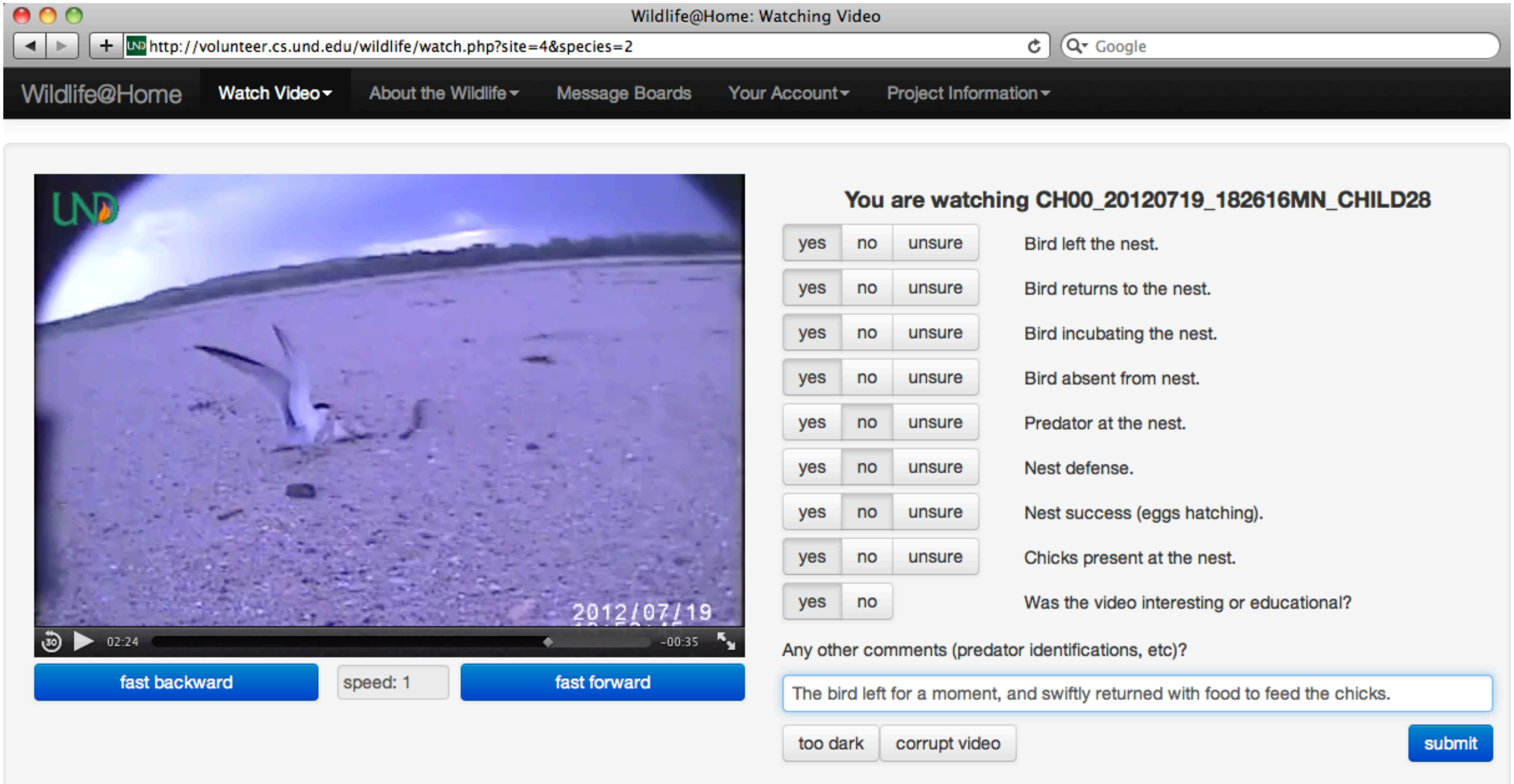
1. Bird Presence
2. Nest Defense
3. Predation Events
4. Nest Success
5. Other events of interest

Analyzing all this video requires both a massive amount of computing power as well as a massive amount of brain power.

Computer vision techniques will need to be run, trained and verified, and updated based on human feedback.

A Tale of Two Interfaces

A Tale of Two Interfaces

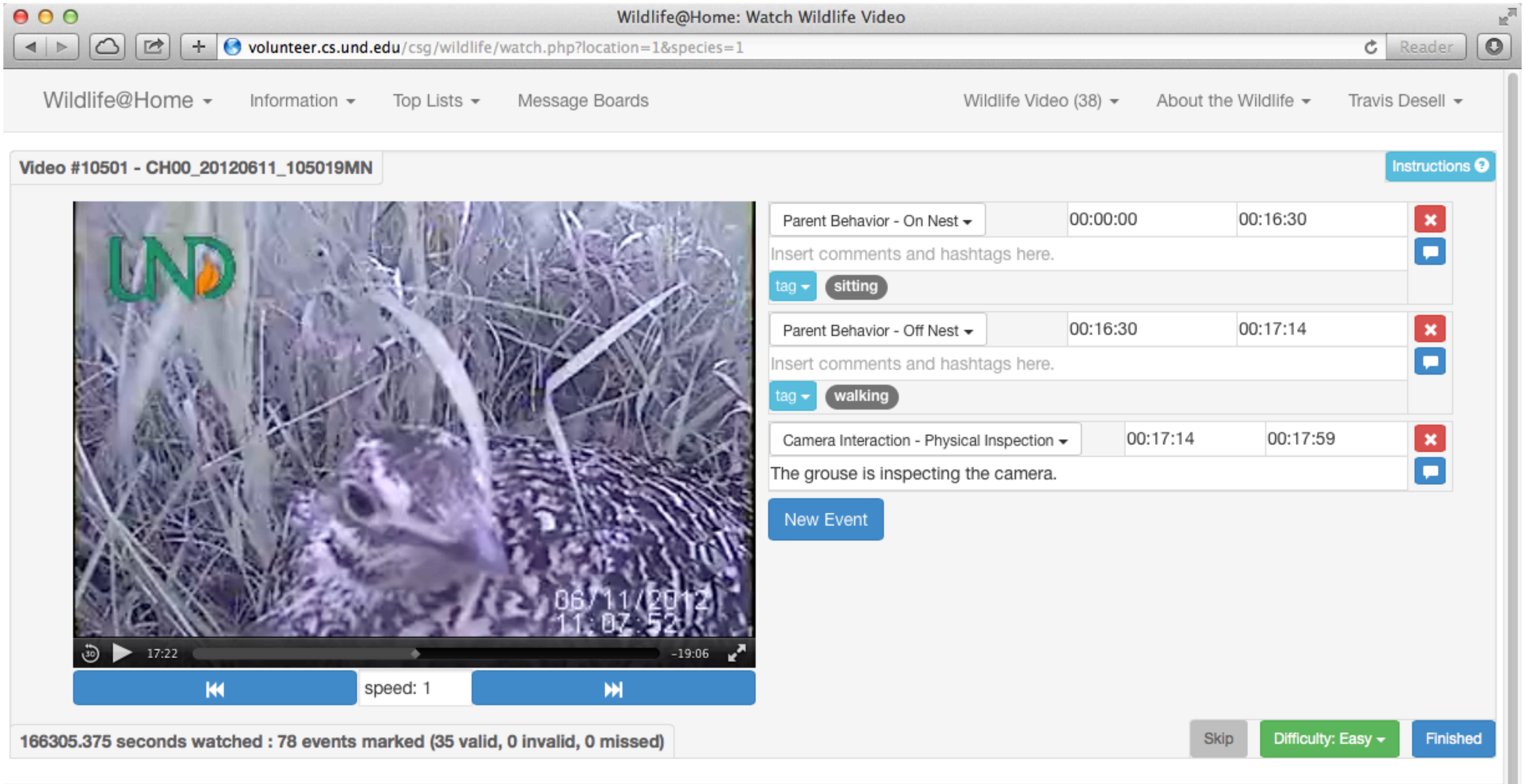


Originally, Wildlife@Home has a simple interface where users could select yes, no or unsure to specify if an event happened at any time during the video.

As we'll see, this simplicity actually had it's costs.

Travis Desell, Robert Bergman, Kyle Goehner, Ronald Marsh, Rebecca VanderClute, and Susan Ellis-Felege. **Wildlife@Home: Combining Crowd Sourcing and Volunteer Computing to Analyze Avian Nesting Video**. *In the 2013 IEEE 9th International Conference on e-Science*. Beijing, China. October 23-25, 2013.

A Tale of Two Interfaces



The interface is significantly more complex, but allows for very accurate specification of when events occur and also a direct comparison to what Dr. Ellis-Felege's experts report.

A Tale of Two Interfaces

Duration (s)	Completed	Observations	Valid	Invalid	Inconvclusive	Valid (%)
< 180	89,645	220,320	206,193	13,129	618	93.58
181 ... 300	8,942	18,715	17,930	649	75	95.80
301 ... 600	6,446	14,022	12,899	1,033	50	91.99
601 ... 1200	3,785	8,396	7,569	744	55	90.15
Total	108,818	261,453	244,591	15,555	798	93.55

Results gathered over 9 months, from August 2013 to April 2014:

- 206 users provided 261,453 observations for 108,818 video segments (~2.4 views to reach a quorum for a video segment)
- 261,453 observations total over 7,411.2 hours of video watched by volunteers. Only 798 were marked inconclusive, and 15,555 marked invalid.
- In the later months of the original interface, video segments were also generated with durations greater than 3 minutes, due to feedback from the users and an interest in seeing how well volunteers would perform on longer video segments. Additional video segments were generated with 5, 10 and 20 minute durations.

A Tale of Two Interfaces

Event Type	Total	TP	TN	FP	FN	Accuracy (%)
Bird Leave/Return	12501	154	8504	287	3556	69
Bird Presence	21230	9407	1338	9270	1215	51
Bird Absence	9540	1092	4680	2173	1595	61
Predator Presence	414	4	393	11	6	96
Nest Defense	33	0	33	0	0	100
Chick Presence	708	12	418	252	26	61

Of the 108,818 video segments marked by volunteers, 25,549 corresponded to videos that were marked by the projects experts.

- True positives (TP) were when a quorum of volunteers marked an event as occurring a video segment, and the times of the video segment overlapped with the time of a similar expert event.
- False positives (FP) were when the marked event did not overlap with the time of a similar expert event.
- True negatives (TN) were when the event was not marked and an expert did not mark the event during that time.
- False negatives (FN) were when the event was not marked and an expert did mark an event during that time.

A Tale of Two Interfaces

Event Type	Total	TP	TN	FP	FN	Accuracy (%)
Bird Leave/Return	12501	154	8504	287	3556	69
Bird Presence	21230	9407	1338	9270	1215	51
Bird Absence	9540	1092	4680	2173	1595	61
Predator Presence	414	4	393	11	6	96
Nest Defense	33	0	33	0	0	100
Chick Presence	708	12	418	252	26	61

Predator presence and nest defense were very accurate, at 96% and 100%.

Bird Leave/Return were fairly accurate at 69%.

Bird absence was not great at 61%.

Bird presence was especially poor at 51% (essentially random guesses).

There were not enough nest success events for comparison.

A Tale of Two Interfaces

5 second buffer

Event	Misses	Type Mismatch	Matches
Parent Behavior - Not In Video	221 (0.23)	23 (0.02)	708 (0.74)
Chick Behavior - In Video	13 (0.93)	0 (0.00)	1 (0.07)
Territorial - Predator	8 (0.53)	1 (0.07)	6 (0.40)
Territorial - Non-Predator Animal	14 (0.93)	0 (0.00)	1 (0.07)
Camera Interaction - Attack	12 (0.57)	9 (0.43)	0 (0.00)
Camera Interaction - Physical Inspection	22 (0.55)	7 (0.18)	11 (0.28)
Camera Interaction - Observation	9 (0.64)	3 (0.21)	2 (0.14)
Error - Video Error	12 (0.09)	7 (0.05)	120 (0.86)
Error - Camera Issue	12 (0.09)	47 (0.34)	78 (0.57)
Parent Behavior - On Nest	484 (0.11)	152 (0.04)	3686 (0.85)
Parent Behavior - Off Nest	315 (0.31)	16 (0.02)	701 (0.68)

10 second buffer

Event	Misses	Type Mismatch	Matches
Parent Behavior - Not In Video	177 (0.19)	26 (0.03)	749 (0.79)
Chick Behavior - In Video	13 (0.93)	0 (0.00)	1 (0.07)
Territorial - Predator	8 (0.53)	1 (0.07)	6 (0.40)
Territorial - Non-Predator Animal	13 (0.87)	1 (0.07)	1 (0.07)
Camera Interaction - Attack	10 (0.48)	11 (0.52)	0 (0.00)
Camera Interaction - Physical Inspection	12 (0.30)	14 (0.35)	14 (0.35)
Camera Interaction - Observation	7 (0.50)	4 (0.29)	3 (0.21)
Error - Video Error	12 (0.09)	7 (0.05)	120 (0.86)
Error - Camera Issue	12 (0.09)	47 (0.34)	78 (0.57)
Parent Behavior - On Nest	409 (0.09)	168 (0.04)	3745 (0.87)
Parent Behavior - Off Nest	253 (0.25)	29 (0.03)	750 (0.73)

We were able to directly compare user observations from the new interface to the expert observations.

Given a buffer time (events matched if the start and end times were within X seconds of each other), we were able to significantly increase user accuracy.

On nest - 51% to 85-87%

Off nest - 69% to 68-73%

Absence - 61% to 74-79%

A Tale of Two Interfaces

5 second buffer

Event	Misses	Type Mismatch	Matches
Parent Behavior - Not In Video	221 (0.23)	23 (0.02)	708 (0.74)
Chick Behavior - In Video	13 (0.93)	0 (0.00)	1 (0.07)
Territorial - Predator	8 (0.53)	1 (0.07)	6 (0.40)
Territorial - Non-Predator Animal	14 (0.93)	0 (0.00)	1 (0.07)
Camera Interaction - Attack	12 (0.57)	9 (0.43)	0 (0.00)
Camera Interaction - Physical Inspection	22 (0.55)	7 (0.18)	11 (0.28)
Camera Interaction - Observation	9 (0.64)	3 (0.21)	2 (0.14)
Error - Video Error	12 (0.09)	7 (0.05)	120 (0.86)
Error - Camera Issue	12 (0.09)	47 (0.34)	78 (0.57)
Parent Behavior - On Nest	484 (0.11)	152 (0.04)	3686 (0.85)
Parent Behavior - Off Nest	315 (0.31)	16 (0.02)	701 (0.68)

10 second buffer

Event	Misses	Type Mismatch	Matches
Parent Behavior - Not In Video	177 (0.19)	26 (0.03)	749 (0.79)
Chick Behavior - In Video	13 (0.93)	0 (0.00)	1 (0.07)
Territorial - Predator	8 (0.53)	1 (0.07)	6 (0.40)
Territorial - Non-Predator Animal	13 (0.87)	1 (0.07)	1 (0.07)
Camera Interaction - Attack	10 (0.48)	11 (0.52)	0 (0.00)
Camera Interaction - Physical Inspection	12 (0.30)	14 (0.35)	14 (0.35)
Camera Interaction - Observation	7 (0.50)	4 (0.29)	3 (0.21)
Error - Video Error	12 (0.09)	7 (0.05)	120 (0.86)
Error - Camera Issue	12 (0.09)	47 (0.34)	78 (0.57)
Parent Behavior - On Nest	409 (0.09)	168 (0.04)	3745 (0.87)
Parent Behavior - Off Nest	253 (0.25)	29 (0.03)	750 (0.73)

Also, we feel that the numbers would be even more accurate as a recent survey of users found that 38% do not consider themselves fluent in English - which could hamper their understanding of use instructions for the more complicated new interface.

A Tale of Two Interfaces

	Easy	Medium	Hard
Misses	2529 (0.15)	145 (0.14)	90 (0.20)
Type Mismatch	1056 (0.06)	57 (0.05)	24 (0.05)
Matches	13774 (0.79)	863 (0.81)	330 (0.74)

We also provided a way for users to specify how challenging it was to mark events in a video.

Interestingly, those with the highest accuracy had medium difficulty (as opposed to easy).

Travis Desell, Kyle Goehner, Alicia Andes, Rebecca Eckroad, and Susan Ellis-Felege. **On the Effectiveness of Crowd Sourcing Avian Nesting Video Analysis at Wildlife@Home**. *In the 2015 International Conference on Computational Science*. Reykjavík, Iceland. 1-3 June, 2015.

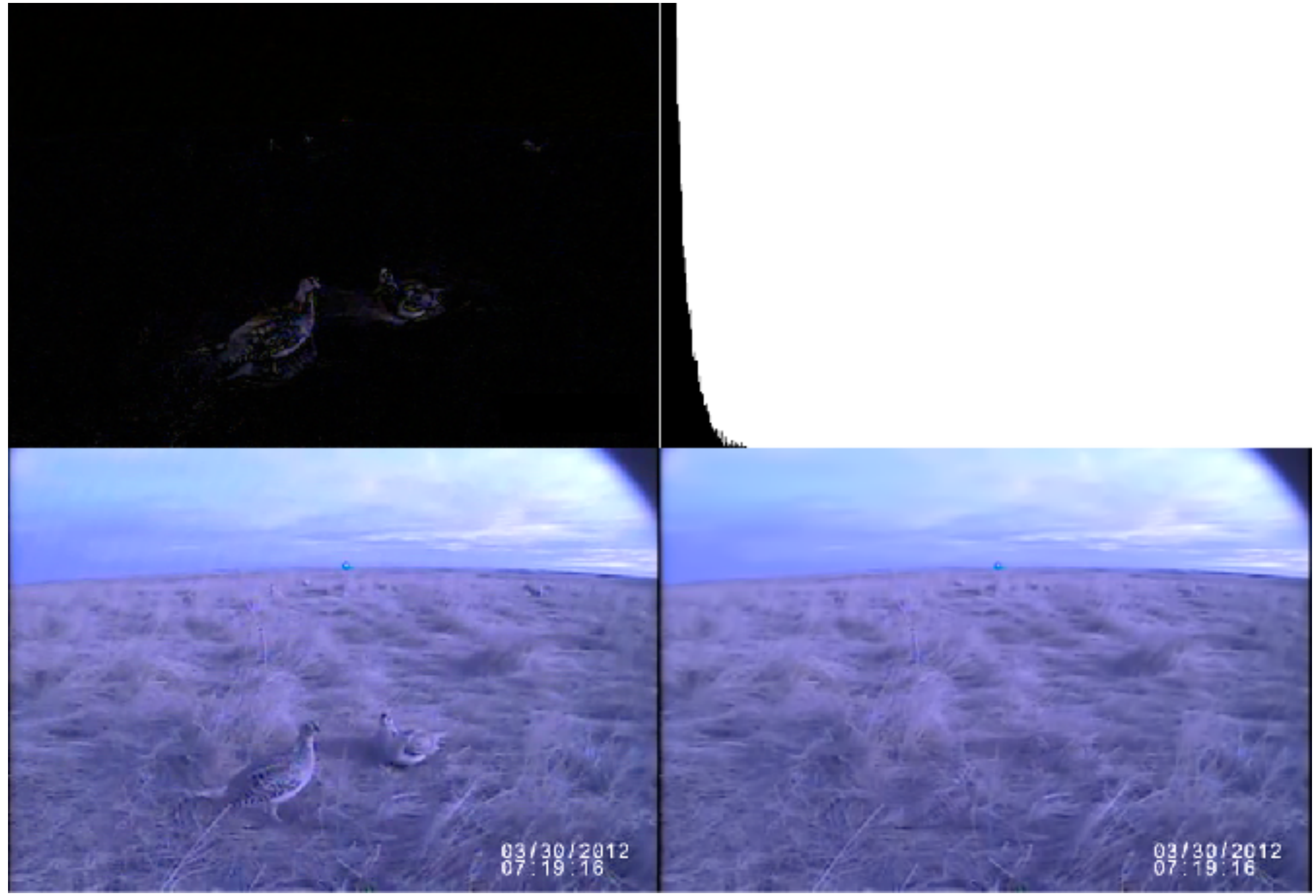
Computer Vision Methods:

Motion Detection
Feature Detection
Background Subtraction
Convolutional Neural Networks

Motion Detection

Initial results gathered using a method called *average window differencing*.

Each frame (lower left) was subtracted from the average of +/- 5 seconds of frames surrounding it (lower right), resulting in a measure of motion (upper left).



Using this, a likelihood of non-noisy motion was for every segment of video.

This was calculated as the average sum of the RGB pixel values in each difference frame divided by the maximum possible difference (3 x width x height x 255).

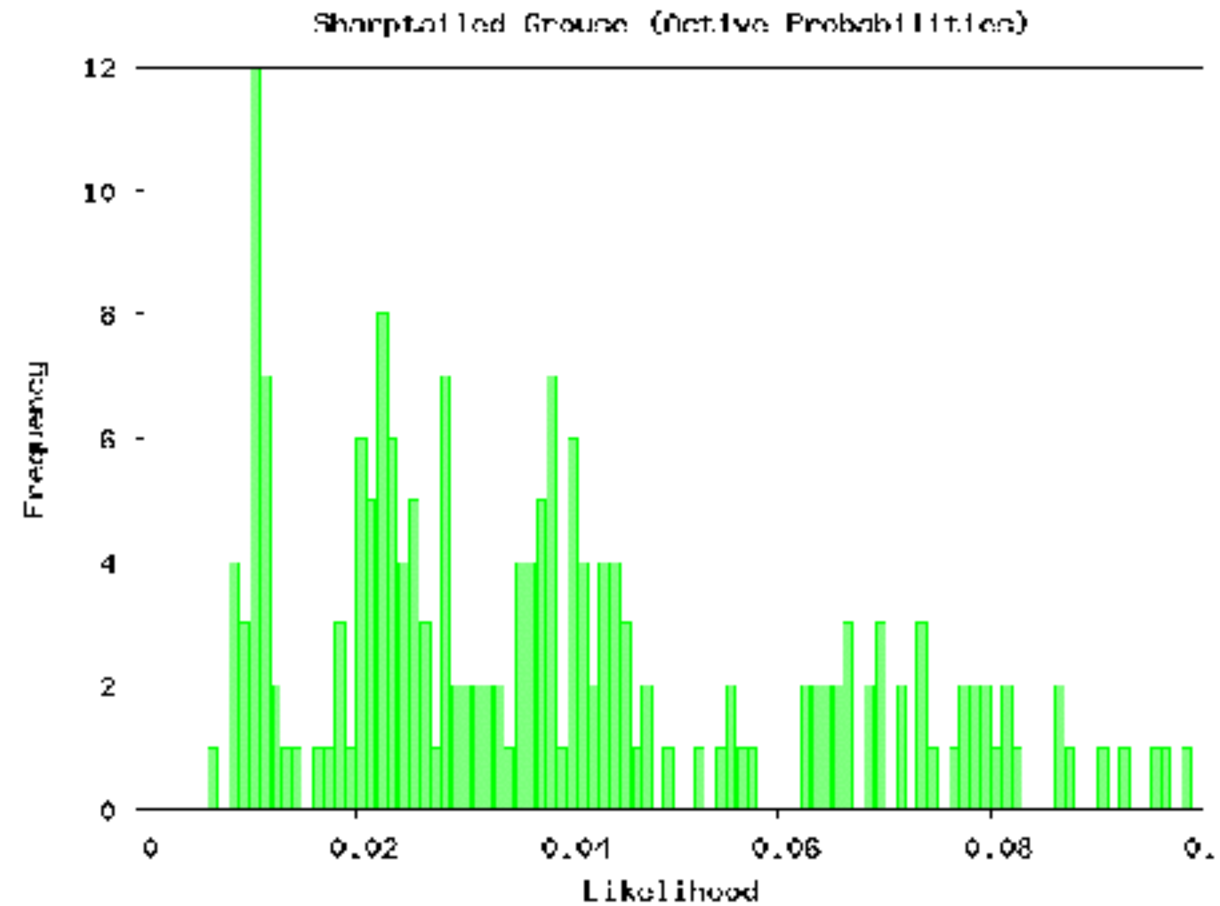
Motion Detection Results

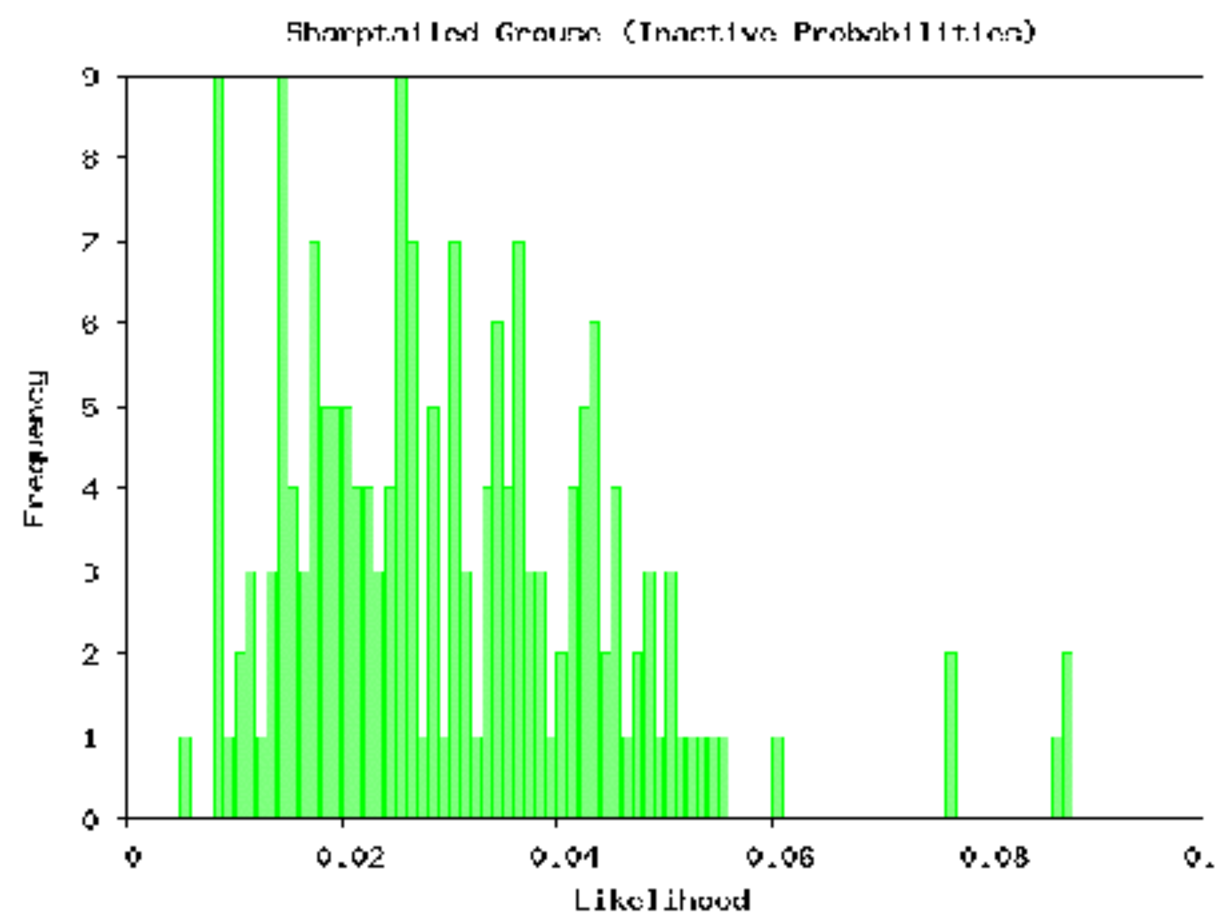
Results for sharp-tailed grouse.

At time of publication:
- 188 videos contained active events (bird return, bird leave, interesting, predator, nest defense)
- 179 contained no active events (bird incubating nest, no bird presence)

Detecting events of interest difficult due to weather, wind and vegetation.

Average and median likelihoods:
- active: 0.039, 0.035
- inactive: 0.030, 0.028





Feature Detection

A feature file was generated by extracting cropped images of birds at their nests in different positions.

Features were extracted using SURF for each image, and then these were merged, by removing any features within a threshold of each other.

This combined feature file was used to calculate a likelihood of a bird being in any segment of video using a bounding rectangle approach.

A rectangle was drawn around all matched features, and the larger the rectangle the less likely there was a strong match to a bird.

Where R_a is the average size of each feature bounding rectangle in each frame of the video segment, and R_f is the size of the frame:

$$\text{likelihood} = 1 - R_a / R_f$$

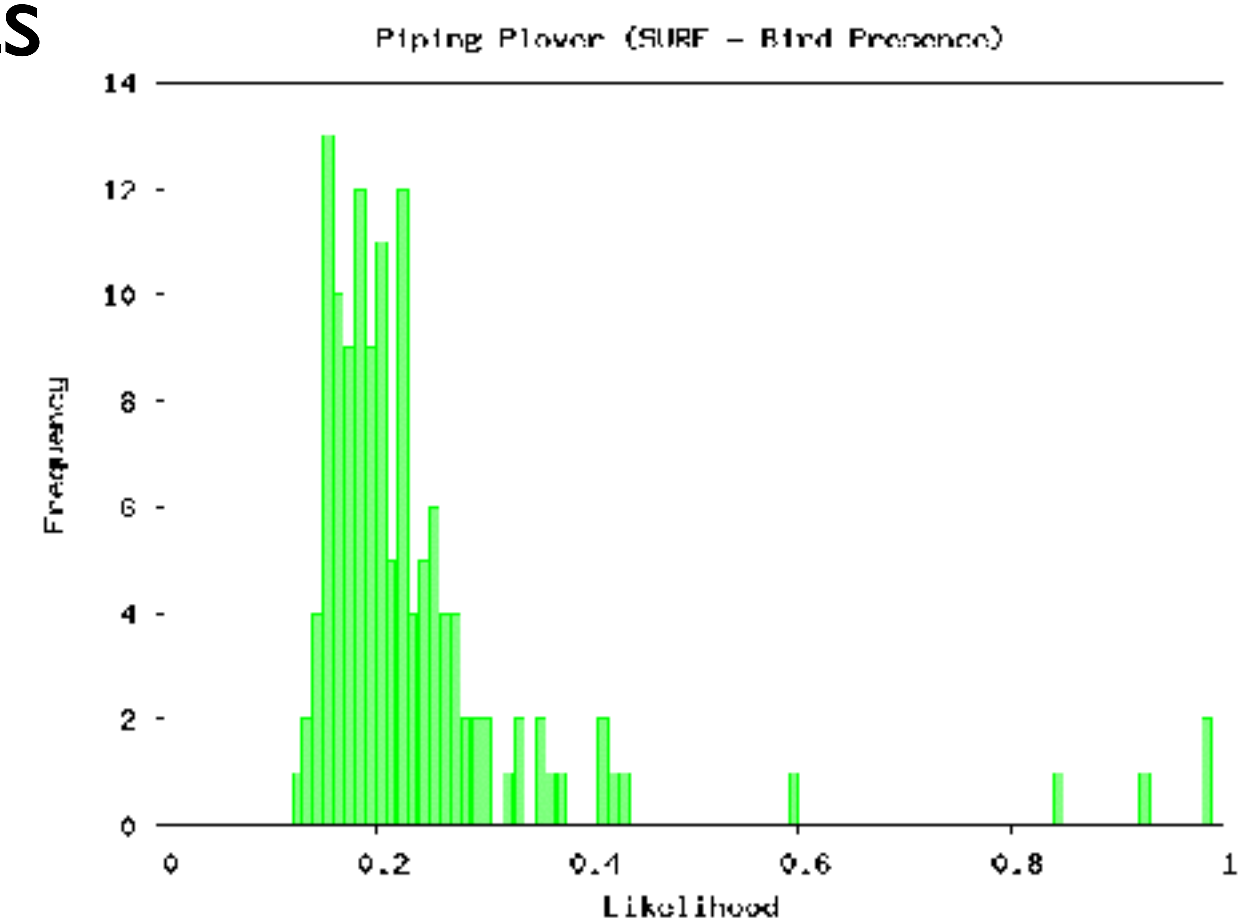
Feature Detection Results

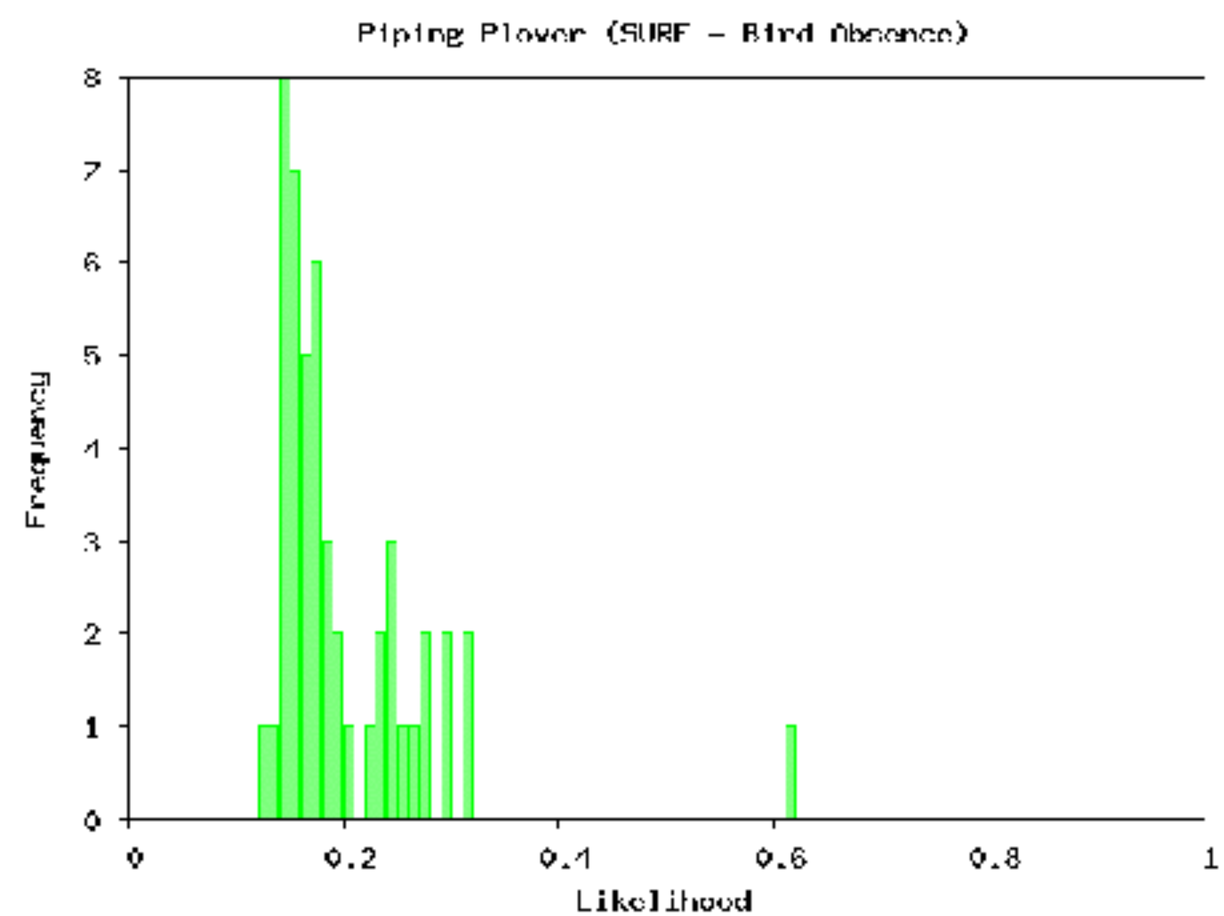
Results for piping plover.

At time of publication:
133 videos contained bird presence
50 contained bird absence

Note: bi-parental investment means not as many videos without a bird at nest.

Average and median likelihoods:
presence: 0.24, 0.21
absence: 0.20, 0.17





Performance Results

At the time of publication, ~70 users had watched over 8400 three minute video segments.
This resulted in ~120 hours of validated observations.

Motion detection was run across the entire video set (~20,000 hours at publication time) and the application processed video at approximately 120 frames per second. At 10 frames per second, this was ~1700 compute hours.

The volunteered hosts processed all videos and returned validated results (meaning each video was analyzed by a volunteer at least twice) in 4-5 days.

Performance Results

SURF feature detection runs much slower (1.7 frames per second).

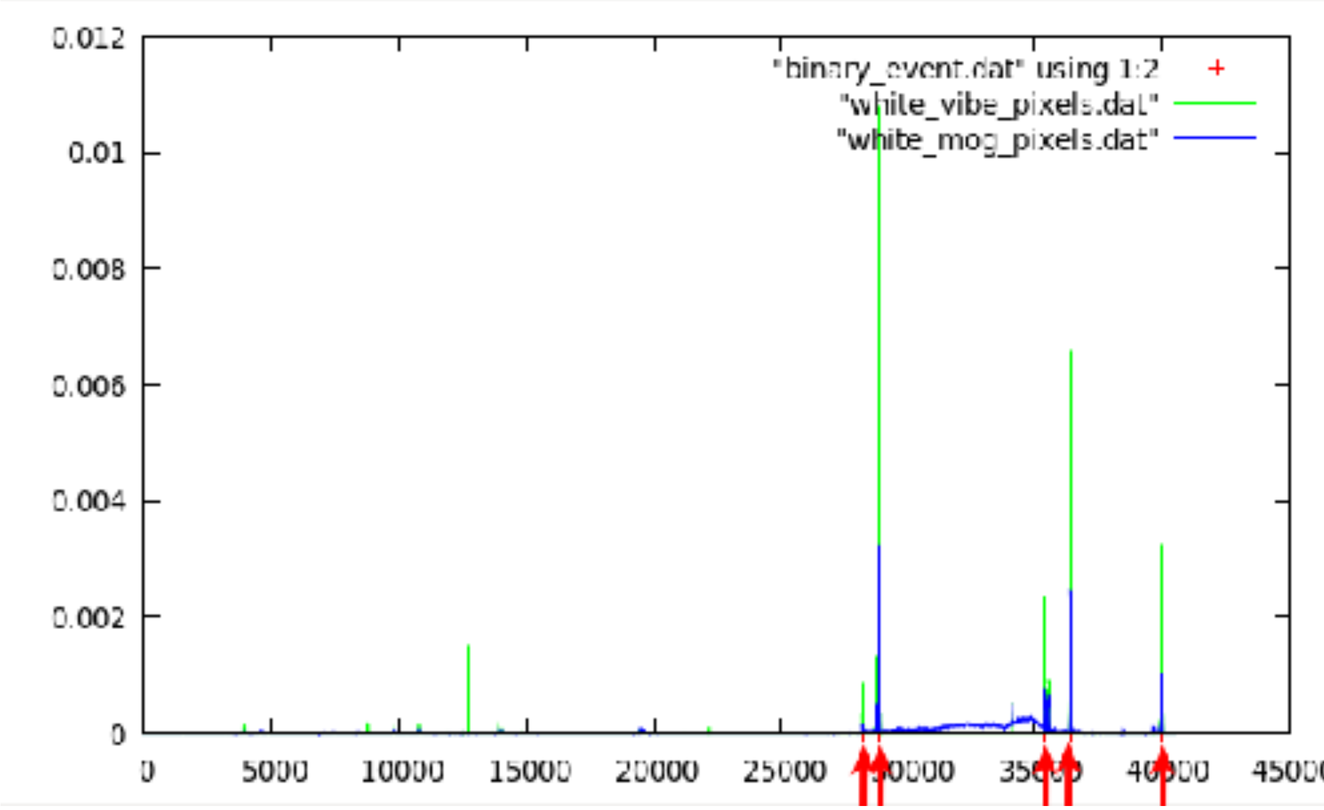
To run this over the piping plover video (682 hours at time of publication), at 10 frames per second or 4000 compute hours results were gathered in under a week.

Travis Desell, Robert Bergman, Kyle Goehner, Ronald Marsh, Rebecca VanderClute, and Susan Ellis-Felege. **Wildlife@Home: Combining Crowd Sourcing and Volunteer Computing to Analyze Avian Nesting Video**. *In the 2013 IEEE 9th International Conference on e-Science.* Beijing, China. October 23-25, 2013.

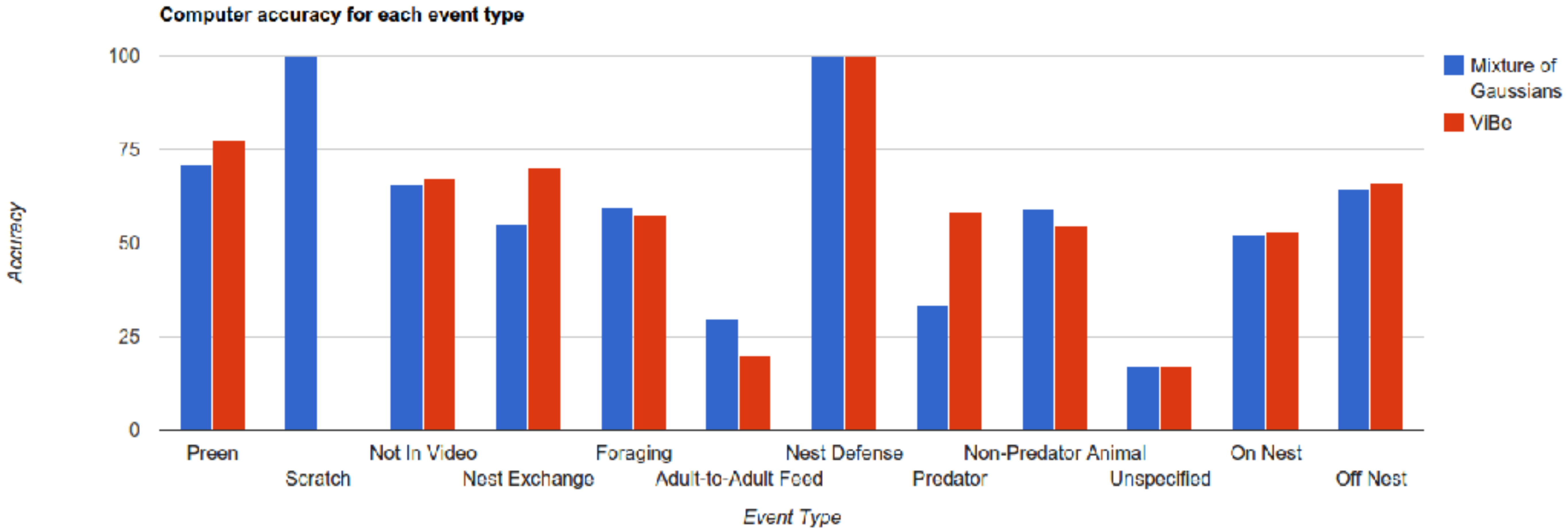
Background Subtraction

Foreground pixels are extracted from an input video file using both the Mixture of Gaussians (MOG) and ViBe algorithms.
Foreground pixels are counted as a percentage of total pixels.
Spikes are classified as an "interesting" event.

- Red arrows indicate scientist classified events (clusters of events).
- Green line indicates pixels marked as foreground with ViBe.
- Blue line indicates pixels marked as foreground with MOG.



Background Subtraction



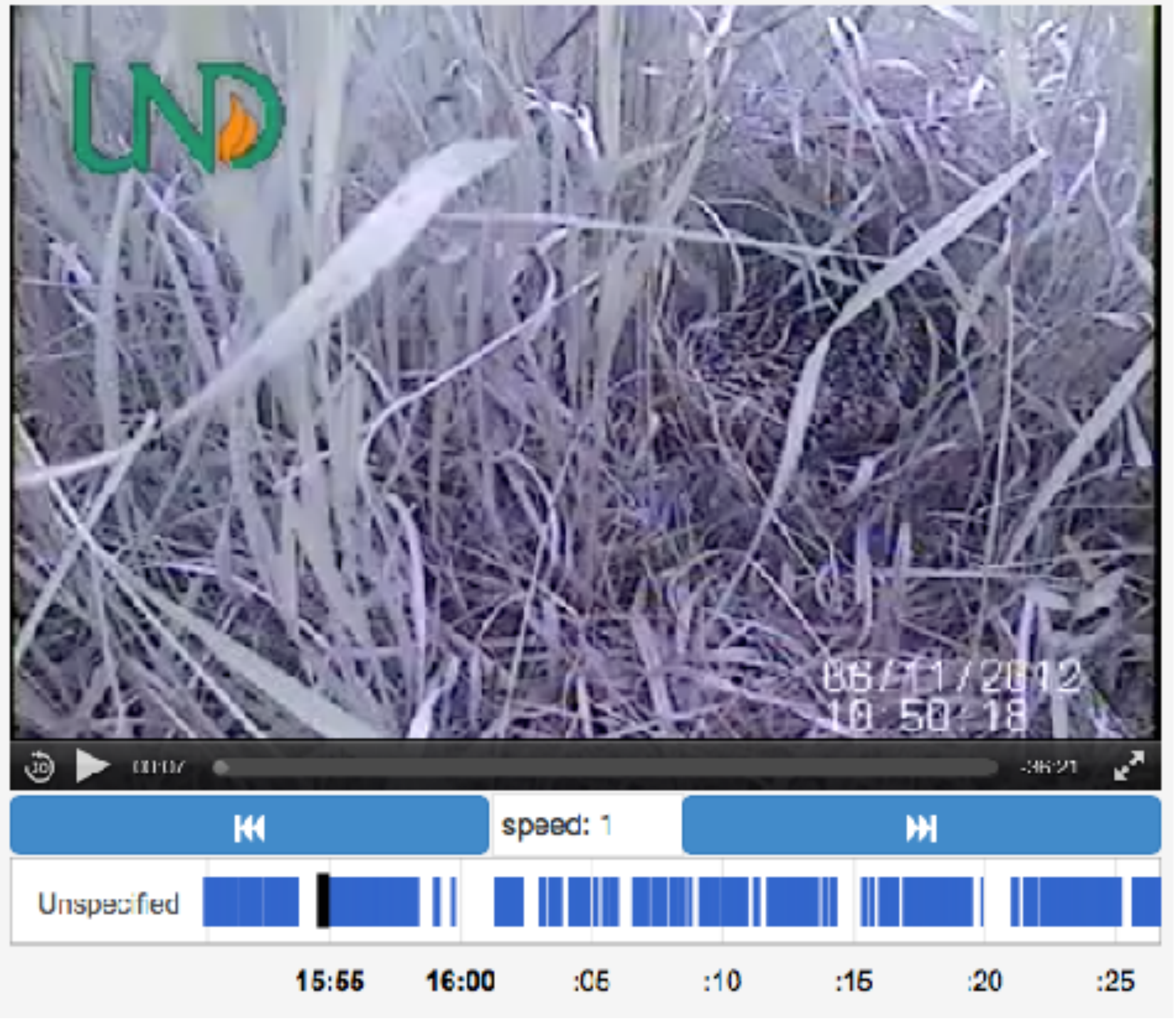
- Accuracy is determined by the number of expert classified events that have a corresponding algorithm spike.
  - 10 seconds in either direction

- Algorithm accuracy for this video
  - ViBe: 96%
  - MOG: 54%

- Quick lighting changes remain an issue
  - Camera brightness adjustment
  - Overhead shadows created by clouds

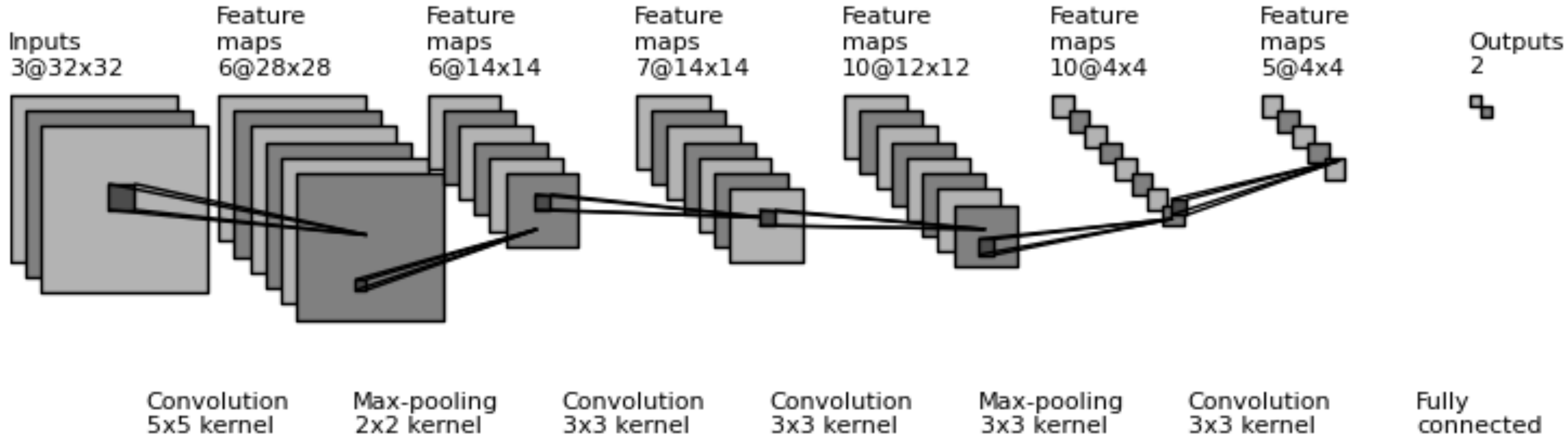
Integrating Results into the Interface

Videos processed by the ViBe algorithm on volunteered computers have the results integrated into the web interface. Regions in blue on the timeline are periods of activity.



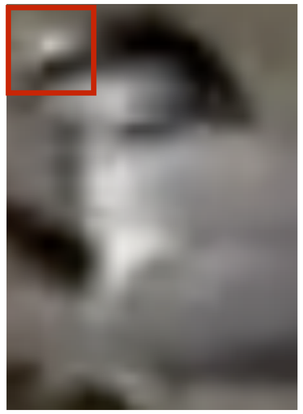
Kyle Goehner, Rebecca Eckroad, Leila Mohsenian, Paul Burr, Nicholas Caswell, Alicia Andes, Susan Ellis-Felege, and Travis Desell. **A Comparison of Background Subtraction Algorithms for Detecting Avian Nesting Events in Uncontrolled Outdoor Video**. *The 11th IEEE International Conference on eScience (eScience 2015)*. Munich, Germany. August 31 - September 4, 2015.

Convolutional Neural Networks



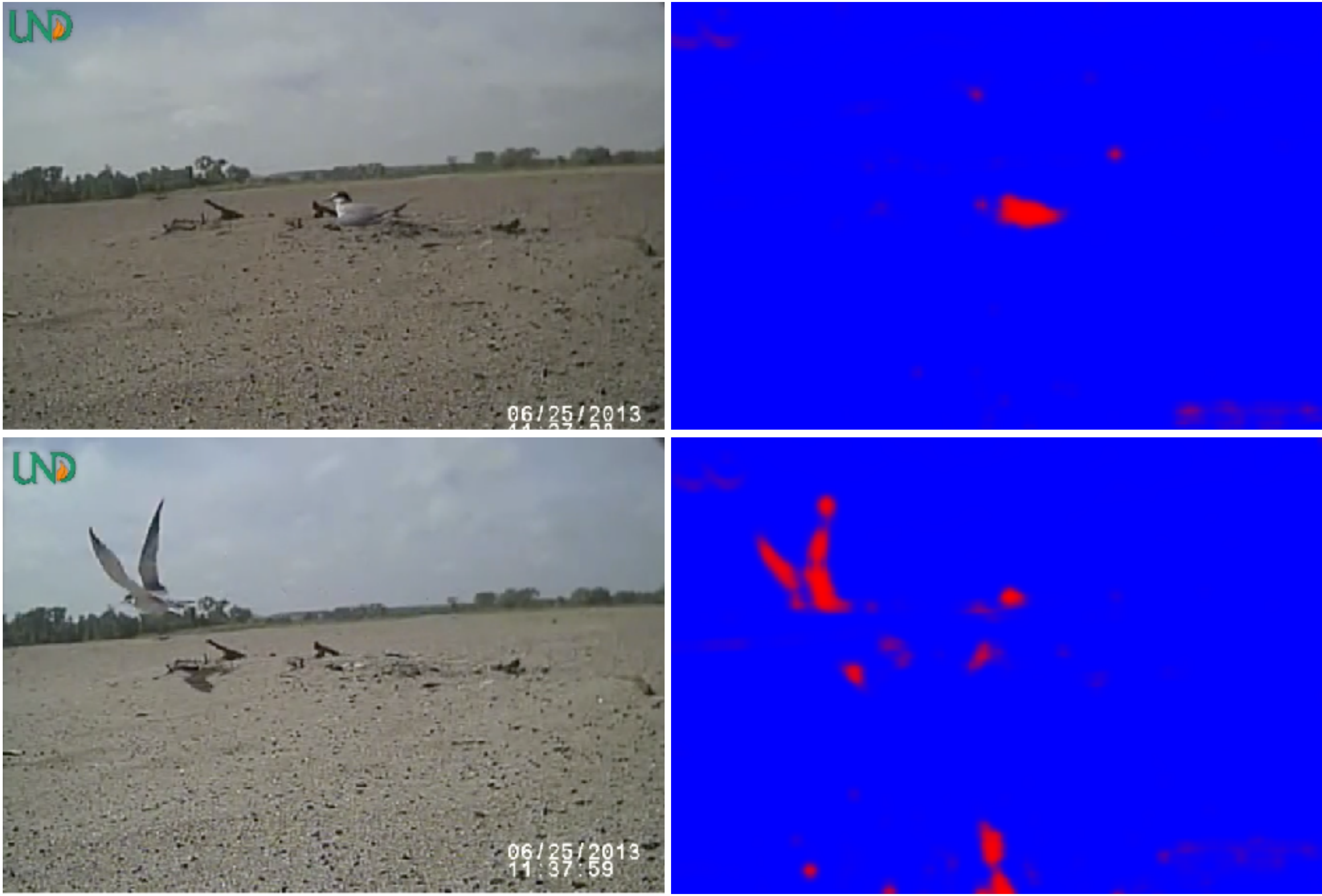
A smaller CNN has been trained with manually extracted imagery. 32x32 (or other sized) training images are generated by striding.

Initial CNN trained on 72,951 training images, and then later further trained on 17,000 "confusers".

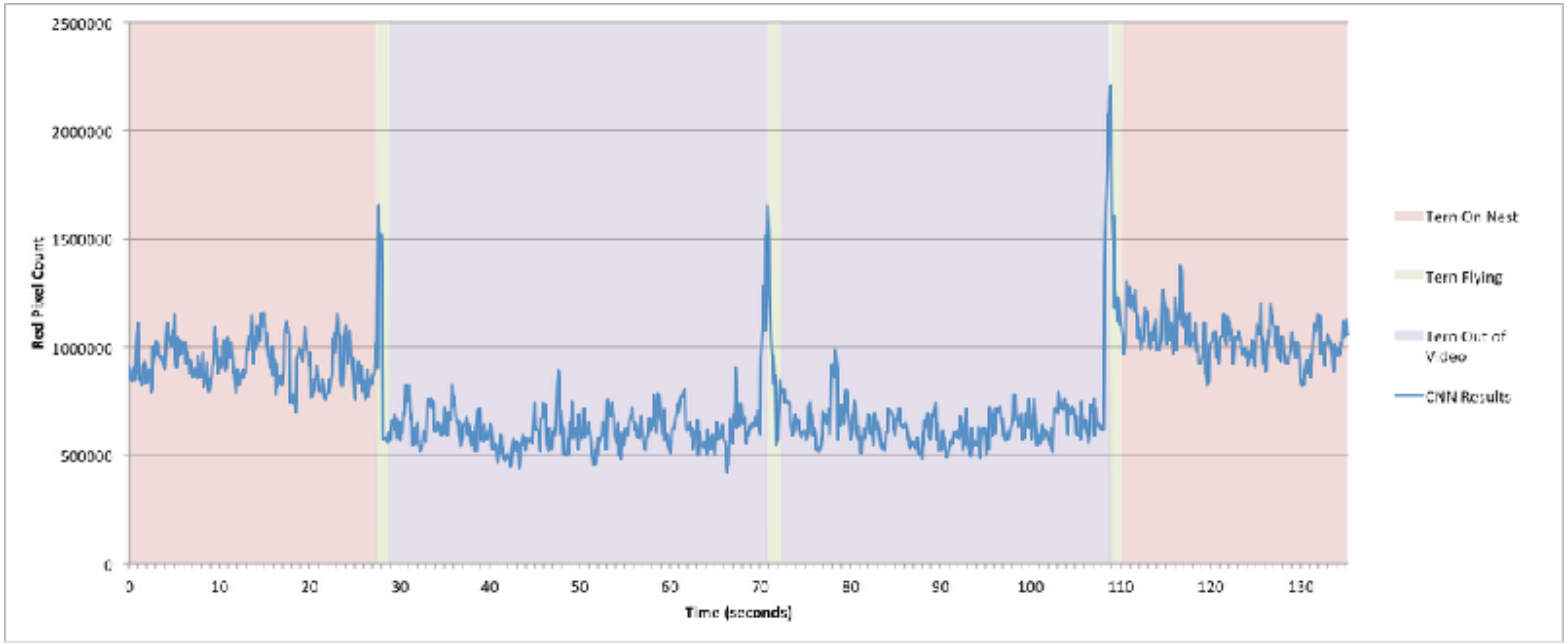


Connor Bowley, Alicia Andes, Susan Ellis-Felege and Travis Desell. **Detecting Wildlife in Uncontrolled Outdoor Video using Convolutional Neural Networks.** *The IEEE 12th International Conference on eScience (eScience 2016)*. Baltimore, MD, USA. October 23-27, 2016.

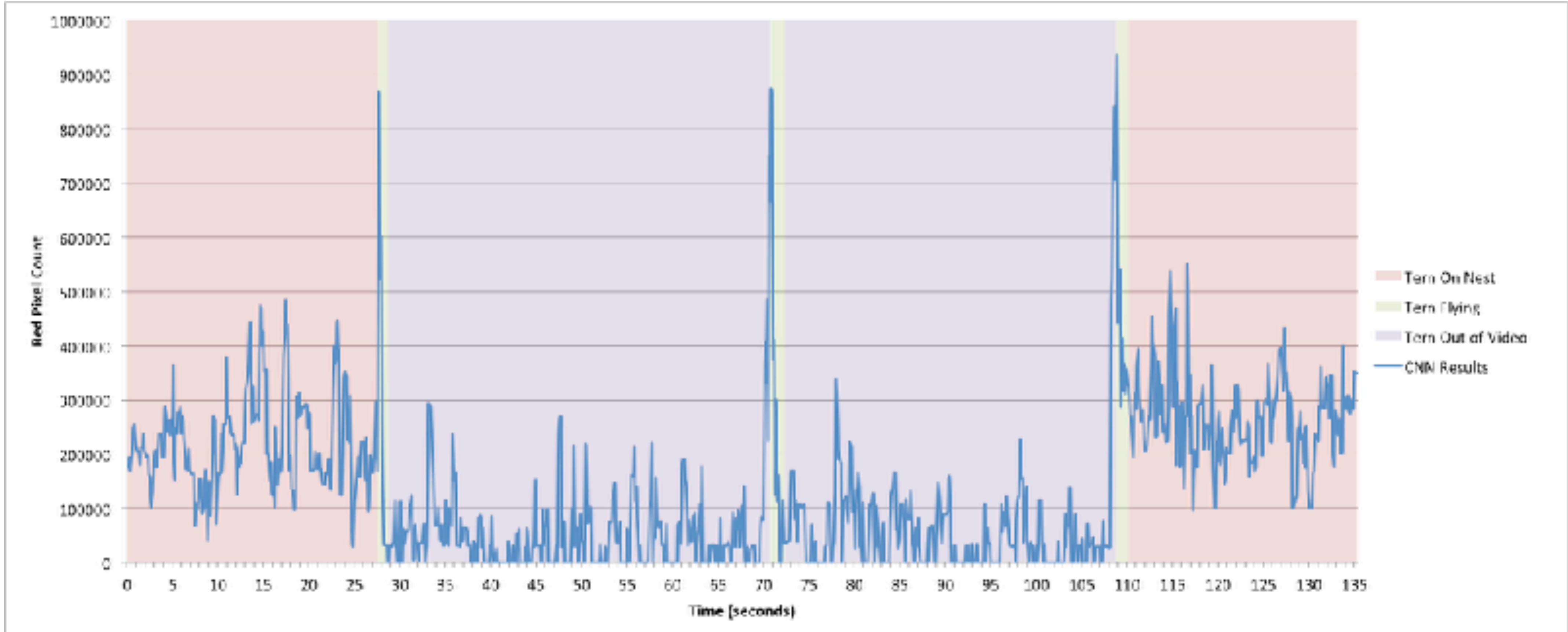
CNN Results



CNN Results



2500000
2000000
1500000
1000000
500000
0
Red Pixel Count
0
10
20
30
40
50
60
70
80
90
100
110
120
130
Time (seconds)
Tern On Nest
Tern Flying
Tern Out of Video
CNN Results



1000000
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
Red Pixel Count
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
Time (seconds)
Tern On Nest
Tern Flying
Tern Out of Video
CNN Results

First Wildlife@Home Data Release

We have made available our first data release of 200+ videos along with the volunteer and expert observations for reproducibility and use by the computer vision community:

http://csgrid.org/csg/wildlife/data_releases.php

Acknowledgements

Wildlife@Home is currently being supported by NSF award no. 1319700 through the Division of Intelligent Information Systems's Information Integration and Informatics (III) program.

Wildlife@Home has been generously supported by a collaborative research award and new faculty SEED grant from UND's Office of Research Development and Compliance. The project's video streaming server is hosted by UND's Computational Research Center and the volunteer computing server is hosted by UND's Scientific Computing Center. DNA@Home is under partial support from a Basic Sciences SEED Grant.

North Dakota Game and Fish has provided financial support for field logistics to collect sharp-tailed grouse videos.

The US Geological Survey has provided financial support for camera equipment, video storage, and field assistance to collect data for the piping plover and interior least tern.

And of course all our volunteers.

Questions?

http://people.cs.und.edu/~tdesell/

http://csgrid.org/csg/wildlife/

tdesell@cs.und.edu