Project in Numbers

- 4+ Months of Hardware Development
- 4 Hardware Revisions
- 4000+ Lines of Code
- 3 Computer Vision Algorithms
- 2 different GPU Implementations (RTX 3060, Jetson Nano)
- 40+ Hours of annotations
- 20+ Hours of model training
- 2 Model Architectures
- 3 Trips to Reno, Nevada
- 4 Days of On-Site Installation

Rapid Animal Detection and Driver Warning System to Mitigate Animal Vehicle Collisions Using Artificial Intelligence

Vedant Srinivas, Eastlake High School, Sammamish, Washington, USA

Introduction

Ecological and Economic Impact

As an economy, we continue to be more dependent on our roadways than ever before. According to the Federal Highway Administration, *"over 164,000 miles of highways in the National Highway System form the backbone of our 4-million-mile public road network."*

- Roadways result in habitat fragmentation, which leads to **1 million animal lives** being harmed on the roads of the United States every day (Roadkill Statistics, 2005).
- Road mortality poses a threat to **21 endangered or threatened species** in the United States alone ("Wildlife-Vehicle Collision Reduction Study: Report To Congress," 2008).
- Roadkill has a drastic impact on the environment and poses many ecological risks, including threatening the biodiversity of local ecosystems and potential extinction (Coffin, 2007).
- Drivers braking at high speeds to avoid animal encounters leads to around **200 human fatalities** and tens of thousands of injuries yearly in the US.
- The total cost of damage to vehicles in these collisions reaches roughly **$3.6 billion annually** (Gaskill, 2013).

Figure 1: Aerial view of problem site in Nevada showing habitat fragmentation due to new highway construction along with wild horses spotted in close proximity with vehicles on Highway 439
Image 1: Google Maps, Image 2: https://www.youtube.com/watch?v=GDrbfBvxLa4, Image 3: Finalist

Highway 439 in Nevada (Figure 1), is a classic example of habitat fragmentation due to a new highway construction (HWY 439 opened in 2017). This area is home to several wild horse herds. These herds need to cross the busy highway in order to shuttle between their food and water needs (Figure 1). The crossing of their path (between the feeding grounds and the watering hole), and the new highway (HWY 439) created an animal-vehicle conflict hotspot in Reno (Figure 1, Red Circle).

Between 2007 and 2016 there were over 5000 animal vehicle conflicts on the roads of Nevada. The estimated value of all animal-vehicle collisions in Nevada is over 21 million dollars annually (Cramer & McGinty, 2018).

Reno specifically is where large ungulate collisions are a major problem. Horses are the animal third most often involved in animal-related crashes, with an average of 34.8 fatalities annually (this does not include non-fatal collisions where the animals and/or drivers suffer injuries or property loss). Motorists who hit horses are more than twice as likely to be injured or killed than those that collide with other ungulates, making them the most dangerous animal on the roads of Nevada (Cramer & McGinty, 2018).

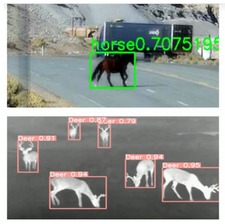
Objectives

The most common and widely implemented solution to avoid animal vehicle conflict is road signage. Animal crossing signs are a passive solution which get ignored over time as a result of highway hypnosis.

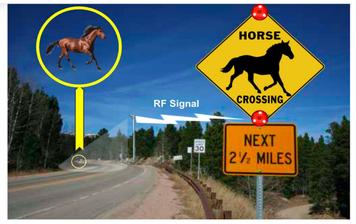
Highway hypnosis is when the brain begins to rely more on predictability, familiarity, and muscle memory to execute the activity of driving rather than real time stimuli. Since animal crossing signs are not as frequently associated with corresponding animal sightings, the brain begins to ignore this stimulus. This leads to worsened reaction times and higher probability of a collision in the event of an animal interaction on the highway. Nevada Department of Transportation (like other DOTs) would like to proactively warn drivers before animals enter busy roadways and endanger themselves and drivers. My objective is to design, engineer and deploy a solution that can,

- in real time, identify and understand animal activity leading up to roads using artificial intelligence, and
- communicate knowledge of active animal presence to oncoming vehicles, giving the drivers time to break away from highway hypnosis, increase focus, and avoid potential collisions.

The Solution



Detect animal activity in both day and night



Intelligent signage- is a relevant and time sensitive stimulus for driver warning

Figure 2- Solution Workflow
(Road Sign Image: https://www.amazon.com/Horse-Silhouette-Crossing-Aluminum-Sign/dp/B00R0RBOC4, Background Image: https://trafficorp.com/category/safety-resources/, Horse Image without box: https://pixabay.com/images/search/equine/, Deer Image: https://www.youtube.com/watch?v=VVKAHfoy4VM, Horse with bounding box: Finalist)

My proposed solution uses computer vision (in both day and night time conditions) to identify animal activity in areas leading up to the highways, before the animals even get to the vehicle conflict zones. It uses this early intelligence of animals and their direction of movement to predict if they are a danger to drivers. This information is then used to send a radio frequency signal to devices nearby including RF enabled road signs.

These road signs flash warning lights on receiving the signal and proactively warn the drivers miles ahead of active animal presence in the upcoming area. By converting a passive road sign into an active information channel (Figure 2), the device allows for drivers to reduce speed gradually, and increase attentiveness leading to better outcomes.

System Design

The system consists of four major hardware components (Figure 3, Items 1,3,5 & 6) interleaved with computer vision and AI algorithms running on edge device and long range zigbee protocol for inter device communication. The hardware components include,

- A bispectral camera with thermal video for nighttime detection and optical video for daytime detection
- A GPU system capable of running computer vision algorithms on the edge
- A long range communication infrastructure (XBee emitter and receiver) to activate the warning lights on intelligent road signs.
- A solar power module that can keep the system up and running when off the grid, leading to better scalability of the solution

This coupled with computer vision algorithms that identify the animals and understand their behavior form the backbone of the technology.

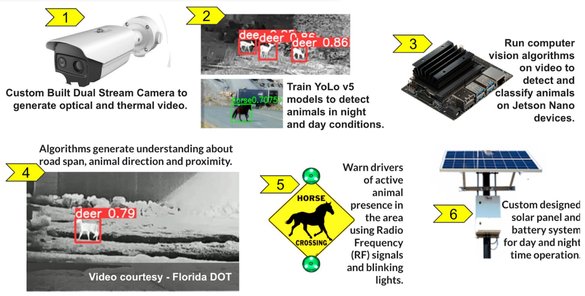


Figure 3- System architecture
Solar Panel Image: https://www.solarelectricsupply.com/side-of-pole-mount-battery-enclosure-solar-system-mapps-760w-24v-530ahr
Jetson Nano Image: https://developer.nvidia.com/embedded/jetson-nano-developer-kit
Camera Image: https://govcomm.us/its-cameras/

Computer Vision

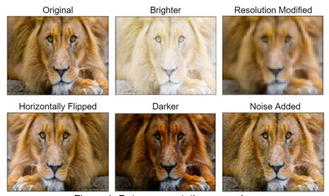


Figure 4- Data augmentation examples
Image Credit: https://www.mygreatlearning.com/

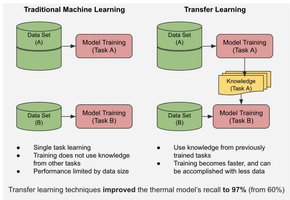


Figure 5- Transfer learning Diagram

Figure 6- Thermal model performance pre and post transfer learning
Image Credit: Florida Department of Transportation

With tools such as AWS AI, Azure ML and Auto ML, training machine learning models with large volumes of data is reasonably straightforward. The challenge presents itself when there is not enough data to train a complete model. In these cases, training with small data sets results in the model overfitting and not generalizing well to a wide range of use cases. Specifically for wild animal detection and classification, the data sets available are limited. This becomes even more challenging when the model is needed for object detection in thermal imagery (not infrared) which is a developing field in wild animal classification. In order to improve model performance using limited data, I used two model training techniques: data augmentation and transfer learning.

Data Augmentation

The optical computer vision model was trained from scratch using 200K images from the COCO dataset (Lin et.al., 2014) with an additional 2000 images of animals, including: horses, cows, deer, bobcats, and panthers. Data augmentation techniques were used to amplify the dataset.

An issue I encountered when training the thermal model was the lack of ground truth data for thermal imagery. I collected videos from existing thermal cameras in high animal activity locations and used them to create a new dataset. In spite of these efforts, I had less than 800 thermal images of wild animals, which was not enough to improve the efficacy of the detection algorithm (compared to over 200K images that were used for the optical model). I annotated these images using the Roboflow platform and further performed data augmentation processes such as reflection, brightness, blurring, and noise to triple the training set size (Figure 4). In spite of this 3X scaling of the data, this data set was still insufficient to train a new model from scratch. To achieve high model performance, I applied the technique of transfer learning (Figure 5) to train on top of the knowledge graph of the optical model (Shorten et al., 2019).

Transfer Learning

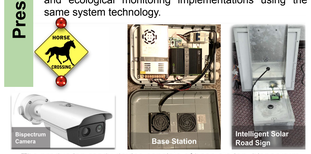
The YOLO V5 base model has three layers: the backbone, which extracts major features from the image; the neck, which normalizes images through three scales; and the head, which classifies the objects of interest inside of the images based on the features extracted by the backbone. I froze the backbone and neck layers and retrained the head layer. I chose to only retrain the head layer because I didn't want certain characteristics of the object, like color, to become limiting factors in the classification of the object, since thermal images will inherently lack color.

Timeline

Present

Growing Demand (Scaling & New Research)

I am currently working on deploying these devices in the Pacific Northwest (I-90 corridor). I'm also working with Oregon and Washington DOT to deploy wildlife tracking and ecological monitoring implementations using the same system technology.



The production ready device consists of base station, bispectrum camera, and intelligent road sign with zigbee wireless transmission

Jan 2022

Camera Replacement (Hardware Revision)

After encountering issues with the infrared cameras getting blinded by headlights and IR reflective paints, I switched to a thermal camera for night vision and found it to be much more reliable.

Oct 2021

Startup Launch (Iyarkai LLC)

After presenting at a conference for ecology and transportation, I received a lot of interest/orders for my device. To scale the manufacturing I started my own company.

Sep 2021

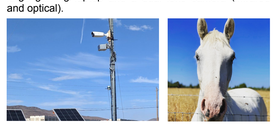
GPU Swap (Hardware Revision)

After encountering issues with the power consumption of the GPU, I redesigned the system onto a smaller compute device and replaced the infrared camera with a thermal camera.

Aug 2021

First Deployment In Reno, Nevada

With the success of my testing in local horse ranches, I did the first deployment of my device in Reno, with a large gaming laptop and a dual lens camera (infrared and optical).



Jun 2021

First Prototype & Data Collection

I built prototypes of my current system to prevent roadkill, which works in both day and nighttime. After devising my final prototype, I did data collection at a farm close to my home, and a horse ranch in Bainbridge Island.

Mar 2021

Ora - First Device for Roadkill Prevention

My first device built to address roadkill was Ora. Ora is an arduino based system that acts as a virtual fence for animals when there are cars on roads. The problem with Ora was that it depended on headlights and only worked during the nighttime. This led me to the development of my latest device.

2016 - 2020

Roadkill Research & Technical Learning

After being affected by roadkill, I became obsessed with solving the problem. I brainstormed ideas to prevent roadkill and sent a letter to president Barack Obama. I was elated to receive a reply, and one quote really resonated with me.

"If you continue to focus on your education and work toward helping your community, there are no limits to what you can accomplish." - Barack Obama

With this motivation fueling me, I spent my time researching the problem, and developing the technical skills necessary to convert my ideas to technology

2017

Elsa - A beautiful dog

I had visited my uncle in Boston and met their dog, Elsa. I bonded a lot with Elsa during the trip. One week after returning home, I got the news that Elsa had been hit by a car.
Living in a suburban community, roadkill is unfortunately something I see often. I've loved animals since a young age and the loss of innocent animal lives continues to be devastating to me.

Hardware Engineering

Power Consumption vs Processing Capacity

The first deployment of the device in July of 2021 was with the Acer Predator Helios 300 Gaming Laptop which contained an NVIDIA RTX 3060 GPU. This GPU gave a high frames per second and had a plethora of CUDA cores (Figure 7), making the model fast and successful. The issue I encountered was with the power consumption of the device. By using 170 Watts, it drained the whole battery of the system in just one night. I then explored the Jetson Nano, which only consumes 20 Watts of power, requiring the creation of more efficient algorithms and models.

Although the Nano had less CUDA cores and was only able to process 2 frames per second, I was able to train lightweight models (YoLo-V5) to perform detection with high level of precision and recall.

Compute Device	NVIDIA GeForce RTX 3060	NVIDIA Jetson Nano
Cost	$1400	$99
Power Consumption	170 Watts	20 Watts
CUDA Cores	3584	128
FPS Supported	30 FPS	2 FPS
Remote Desktop GUI	Enabled	Terminal access only
5 Second Recall of Model	99.9%	99.9%

Figure 7- Comparison of RTX 3060 GPU with Jetson Nano

One tradeoff was the remote desktop accessibility. With the laptop, I was able to use Chrome Remote Desktop to get GUI access. This feature is unavailable on the Jetson Nano and I used ngrok.io to SSH (secure shell) into the device, limiting my access to just the terminal.

Infrared vs Thermal Camera

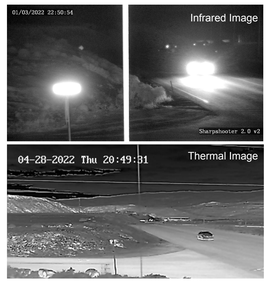


Figure 8- Scene in IR Camera (top) and Thermal Camera (bottom)

Infrared cameras use short wavelength infrared light to illuminate an area of interest. Some of the infrared energy is reflected back to a camera and interpreted to generate an image. Headlights and road signs colored with reflective paint reflect radiation in the short wavelength IR spectrum. The IR camera was consistently blinded by headlights in the nighttime, rendering the image unusable (Figure 8, top).
Thermal imaging systems use mid to long IR wavelengths. Thermal imagers are passive, and only sense differences in heat, working well for detecting warm blooded entities like animals that emit thermal energy. These bodies will create a defined image due to differences in temperature relative to the environment, improving imagery in nighttime conditions (Figure 8, bottom).

Results

The system was benchmarked by collecting baseline data from the current deployment location in Reno, Nevada. From a sample set of 1023 images including horses, trucks, and cars collected from the field, the system achieved a detection recall of 99.5%, with successful classification of 92% of the image. 8% of the images detected as large animals were not classified correctly (horses as dogs or cows). No animals were misclassified as vehicles or vice versa (Figure 9).

	Samples	Percentages
Baseline	1023	
Detection Recall	1018	99.5%
Missed	5	0.5%
True Classification	935	92%
False Classification	83	8%

Figure 9- Data from the current system deployment in Reno, Nevada

Discussion & Future Work

The Eqwis system has been live in Reno, Nevada since September 2021. The system has had a 99% uptime (running continuously) and over 99% recall. The system has maintained its performance across all seasons with no variances in detection recall and classification precision. This paired with the system being significantly cheaper than existing solutions (fences and bridges) makes it highly scalable. To expand to multiple states and geographies, it is important for the system to support a wider range of animals in both day and night conditions. This will entail using GAN based machine learning and image morphing techniques (owing to limited data sets) to train models for other animal classes such as panthers, sheep, coyotes and wolves. In addition, I am developing further applications of the system in animal behavior tracking and ecological monitoring.

References

[1] Redmon, Joseph, et al. *You only look once: Unified, real-time object detection.* Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[2] Shorten, C., Khoshgoftaar, T.M. *A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60 (2019).*
[3] F. Zhuang et al., "A Comprehensive Survey on Transfer Learning," in Proceedings of the IEEE, vol. 109, no. 1, pp. 43-76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.
[4] Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F. *Imagenet: A large-scale hierarchical image database.* In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR 2009), Miami, FL, USA, 20–25 June 2009; pp. 248–255.
[5] Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. *Microsoft COCO: Common objects in context.* In Proceedings of the 13th European Conference on Computer Vision (ECCV 2014), Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2014.
[6] *US VIO Vehicle Registration Statistics*. (2021, January 19). Hedges & Company; https://www.facebook.com/HedgesandCompany/. http://hedgescompany.com/automotive-market-research-statistics/auto-mailing-lists-and-marketing/
[7] *"Wildlife-Vehicle Collision Reduction Study: Report To Congress."* (2008). Federal Highway Administration. http://www.fhwa.dot.gov/publications/research/safety/08034/exec.cfm
[8] Yager, C. E., Cooper, J. M., & Chrysler, S. T. (2012). The Effects of Reading and Writing Text-Based Messages While Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *1*, 2196–2200. https://doi.org/10.1177/1071181312561463