ABMI Science in PROGRESS

ABMI Science in Progress presents new work from the ABMI's Science Centre. Results may be subject to change.

Automated Filtering of False Fires to Reduce Processing Costs for Remote Camera Images

John Simms and Jim Schieck

Introduction

Predators, ungulates, and other mid- and large-sized mammals are often used as indicators of ecological condition during management processes like the Alberta Land-use Framework[1]). To determine large-scale patterns of distribution, abundance, and trend for these species, and to understand the cumulative effects of development on their populations, researchers have made extensive use of remote cameras[2]. When an animal enters the camera's detection zone, a picture is triggered by the animal's heat and movement and the resulting image is stored on the camera (Figures 1 & 2). These images are downloaded to a computer and provide the building-blocks for evaluations of species distribution, habitat use, and population trends over time[3]. Remote cameras operate effectively in a wide variety of environmental conditions, so they can be set up at one time and collected months later.

Not all images collected by remote cameras include animals. Sometimes sunlight, vegetation movement, or other factors result in "false fires" of the camera. These false fires result in an increase, and sometimes a great increase, in the number of images collected. Reviewing false fires adds to the time and cost of processing remote camera data. To reduce

Figure 1. Example set-up of a remote camera to survey wildlife.

Figure 2. A full camera image with a fox and its bounding box (a box drawn around the animal).

---

[1] Alberta Environment and Parks. 2008. Land-use framework. Pub No I/321. Alberta Environment and Parks, Edmonton, AB.
[2] Fisher, J. T., S. Bradbury, B. Anholt, L. Nolan, L. Roy, J. P. Volpe, and M. Wheatley. 2013. Wolverines (*Gulo gulo luscus*) on the Rocky Mountain slopes: natural heterogeneity and landscape alteration as predictors of distribution. Canadian Journal of Zoology **91**:706–716.
[3] Burton, A. C., E. Neilson, D. Moreira, A. Ladle, R. Steenweg, J. T. Fisher, E. Bayne, and S. Boutin. 2015. Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. Journal of Applied Ecology **52**:675–685.

costs, we developed an automated process to identify and "tag" as many false fires as possible, while minimizing the number of images containing animals that were wrongly tagged as false fires. This automated process is expected to save processing costs as more and more remote cameras are deployed by researchers and managers throughout Alberta.

Images Used

We obtained remote camera images from the Alberta Biodiversity Monitoring Institute (ABMI). These images had already been labeled by ABMI staff as either animal or false fire. We added a bounding box around the animals in the images (Figure 2) to facilitate modeling of false fires. All deployments were randomly assigned to either the training or validation data sets. Most false fires occurred during the daylight and thus we focused our modeling on daytime images.

Model Training

The training network consisted of 1,325 camera deployments. For each image that contained an animal we extracted three crops of different sizes, each containing the bounding box. All crops were resized to 256 by 256 pixels (Figure 3). The multiple crops increased the volume of training data, and created images with animals at different pixel scales. For false fire images, three similarly-sized crops were extracted from a random location in the image.

Figure 3. Three crops from Figure 2 with the fox at different pixel scales.

We did not have sufficient images to train a deep neural network from scratch. Instead, we started with CaffeNet[4,5], an already trained network, and fine-tuned it for our use. Features learned in large neural networks like CaffeNet are robust enough to be used as a starting point for classifying other image datasets[6]. We used default settings (aside from the number of output units, which we changed from 1,000 to 2) from CaffeNet with a selected learning rate of 0.001 for the initial sequence of iterations. The learning rate was lowered by a

[4] Jia et al., 2014. Caffe: convolutional architecture for feature embedding. arXiv preprint at Xiv:1408.5093.
[5] Krizhevsky. I.S. and G. Hinton. 2012. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems **25**:1106–1114.
[6] Bengio, Y. 2011. Deep learning of representations for unsupervised and transfer learning. JMLR W&CP: Proc. Unsupervised and Transfer Learning.

factor of 5 after every 5,000 iterations, and the training process continued until a total of 20,000 iterations had been completed.

Model Validation

The validation network consisted of 121 camera deployments. In these deployments, we ignored the bounding boxes around animals. For each image, we created 59 croppings so that if an animal was present anywhere in the image it would be represented in a variety of croppings at different pixel sizes.

At the first step of validation, all croppings for each image were evaluated based on the final network model created during model training, and each crop was assigned a probability of actually being an animal ("potential-animal"). The maximum value for potential-animal was determined across the 59 croppings for the image, and then compared to the user-defined threshold value of 0.6. If the maximum was less than or equal to the threshold value, the entire

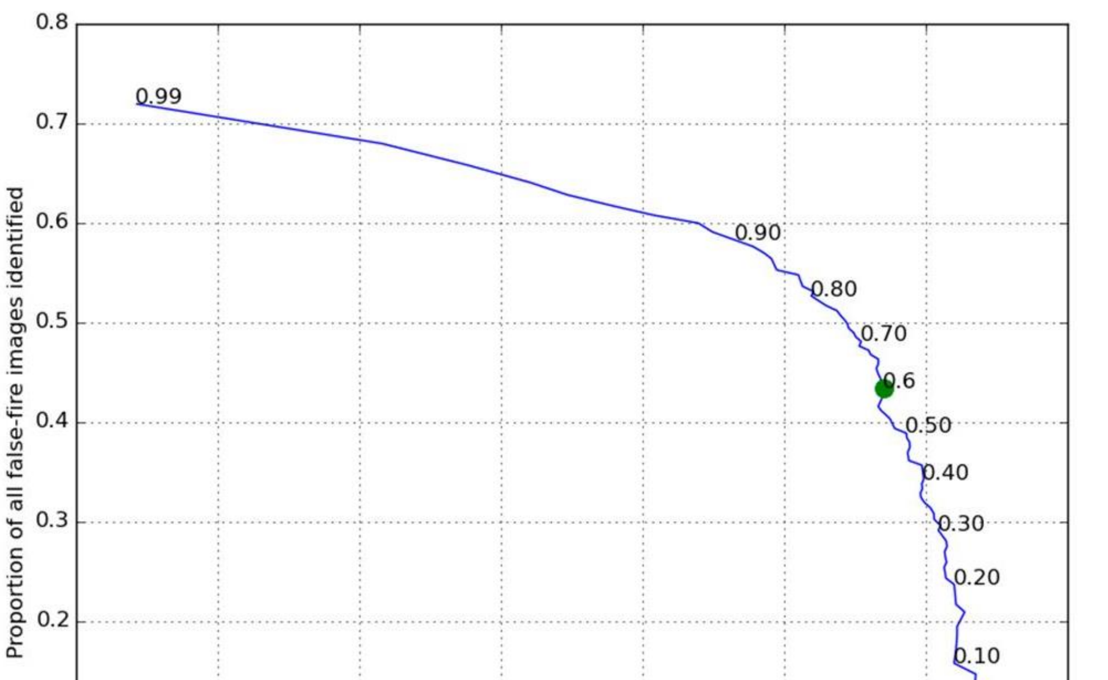

Figure 4. The proportion of false fires that were correctly classified in relation to the accuracy of the model for threshold values ranging from 0.1 to 0.99. The green dot identifies the threshold value of 0.6 that was used in our analyses.

image was classified as a false fire. Using a higher threshold would result in more of the false fires being correctly classified, but would also result in more animal images being incorrectly classified as false fires (Figure 4). There was little value in reducing the threshold below 0.6

because the percentage of false fire images that were correctly classified dropped quickly with little gain in accuracy.

At the second step in the classification, if an image labeled as a false fire was within a sequence of images, and was less than two images and six seconds from a different image labeled as a potential-animal, we adjusted the label from false fire to potential-animal.

From the 121 cameras included in the validation, 34,456 of the 79,451 false-fire images (43.6%) were correctly classified as false-fires by the model. In addition, a very low percentage (0.2%; 80 of 34,536) of the predicted false-fires were incorrect.

Management Implications

- Significant processing costs can be saved by implementing automatic detection of false fires when reviewing and tagging images collected by remote cameras.
- A deep neural network, with adjustments made based on whether an image occurred within a sequence, was used to automatically identify approximately 50% of the camera false fires with a very high degree of accuracy.
- Automated filtering of images will continue to improve over time as better network models are developed.
- Although testing is incomplete, we anticipate that deep neural networks can also be used to automatically tag species in images from remote cameras.