Streamlining the study of the Tierra del Fuego forest through the use of deep learning

Leonel Viera[1], Federico González[1], Rosina Soler[2], Lucas Romano[1], Guillermo Feierherd[1]

[1] Instituto de Desarrollo Económico e Innovación, Universidad Nacional de Tierra del Fuego, Antártida e Islas del Atlántico Sur
Hipólito Irigoyen 880, 9410, Ushuaia, Tierra del Fuego, Argentina
{lviera,fgonzalez,lromano,gfeierherd}@untdf.edu.ar
[2] Centro Austral de Investigaciones Científicas, CONICET
Houssay 200, 9410, Ushuaia, Tierra del Fuego, Argentina
rosinas@cadic-conicet.gob.ar

Abstract. Understanding plant-herbivorous relationships allows to optimize the way to manage and protect natural spaces. In this paper the study of this relationship in the ñire forests (Nothofagus antarctica) of the province of Tierra del Fuego (Argentina) is approached. Using trap cameras to monitor such interaction offers the opportunity to quickly collect large amounts of data. However, to take advantage of its potential, a large investment in trained personnel to analyze and filter the images of interest is required. The present work seeks to establish a path to significantly reduce this obstacle using the advances of machine and deep learning in the recognition of objects from images.

Keywords: machine learning, deep learning, computer vision, trap cameras, forests, image recognition, ñire, antarctic nothofagus.

1 Introduction

1.1 Understanding ecosystems with images

To understand the complexity of forest ecosystems, improve their management and at the same time ensure their conservation, it is necessary to have detailed knowledge about their natural dynamics and that of the organisms that live there (number, location and behavior of animals) [1]. Much of the native forests of Southern Patagonia are currently used for livestock (cows and sheep). The direct impact of domestic animals (foraging, trampling), coupled with the natural presence of wild herbivores (guanacos), affect the growth of young trees and the composition of vegetation [2]. Good management of this activity together with monitoring over time could lead to maintaining productivity without degrading the forest environment.

The use of cameras with motion sensors, called trap cameras, in natural habitats significantly speeds up studies on ecology and wildlife conservation over the past two decades [3]. These have become an essential tool for environmentalists, allowing them to study the size and distribution of populations [4] and evaluate habitat used by animals. While they allow thousands of images to be taken [5], effective data extraction is traditionally performed by people (i.e. experts, previously trained volunteers) and it is so slow and expensive that much of the valuable knowledge in these large data repositories remains untapped.

1.2 Machine Learning and Deep Learning

Machine Learning allows computers to solve tasks without being explicitly programmed to solve them [6]. There are several machine learning algorithms, with those classified as supervised being the most interesting for this work. These work by training the algorithm by entering entries associated with their expected results, in order to obtain a function that may classify future entries [6]. For example, when classifying images, the machine trains with many pairs of images and their corresponding labels, where the image is the input and its correct label (for example, "nocturnal") is the output.

Deep learning [7] allows computers to automatically extract multiple levels of abstraction from raw data. Deep convolutional neural networks are a type of deep learning architecture [7] in which each layer of neurons uses convolutional operations to extract information from small regions that overlap between the previous layers [7]. For the classification, the last layer is generally a function with an output between 0 and 1 per class; adding 1 all class exits. These are interpreted as the estimated probability that an image belongs to a certain class, the higher the more certain is the algorithm that the image belongs to that class [8].

Deep learning has changed the way to solve many difficult problems [7], including voice recognition [9], automatic translation of texts [10], image recognition [11], among others. As a counterpart, deep learning requires a large amount of input data (correctly labeled) and high computational resources.

1.3 Combining forest management and deep learning

In the case treated in this article, we will try to automate the identification and counting of animals from the images captured by trap cameras using deep learning techniques, obtaining a more efficient solution to provide data in projects with these conditions [12].

2 Case study

In 2014, a project to assess the impact of native (guanacos) and domestic (cattle) herbivores on ñire forests (Nothofagus antarctica) under livestock production in Tierra del Fuego began. For this, in 4 different forests, fences were installed that prevented the entry of animals and the vegetation was studied inside and outside these fences. At the same time, trap cameras were installed focusing on both the closed zone and the free sectors. Through these cameras it was possible to evaluate the effectiveness of the closures and the use of animals (feed, rest and others), differentiating seasons of the year and times of day.

The images taken by the cameras were collected from 2015 to 2017, accumulating more than 150,000 photographs. This amount of information produces a high cost of manual image analysis, slows down work and makes it difficult to obtain results.

So far there are several results that show the impact of animals on forest vegetation. For example, both the growth and survival of young trees showed a favorable response to closures. In the closed areas, survival was between 75% and 87%, while outside them it was only 60%. In addition, they achieved greater growth in height. However, this survival varies according to the age of the trees, since the younger (2-3 years) the greater the probability of mortality. Even the smallest plants outside the closures had higher mortality than those of the same age within the closures, as expected. Then, after 5 years the trees were able to survive in the same proportion outside the closures (Fig. 1A).

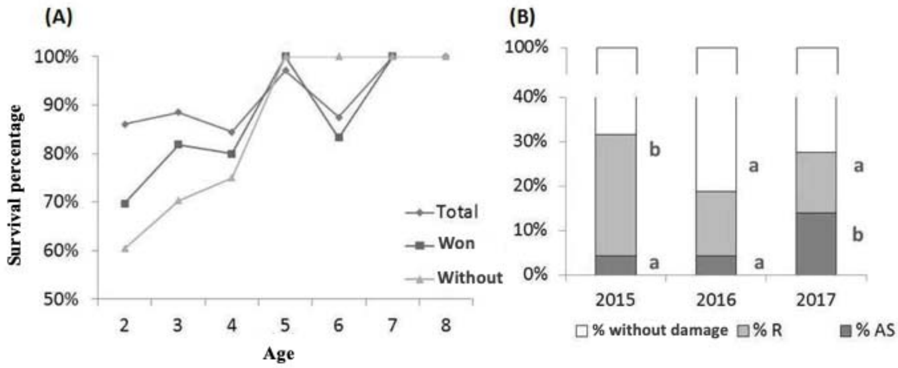

Fig. 1. (A) Survival percentage of the ñire trees inside and outside the closures and according to age. (B) Proportion of damaged trees: eaten by animals (R), desiccation (AS), or without damage. The letters indicate significant statistical differences (p <0.05) between years for each type of damage.

Another example is the identification of two main types of damage: on the one hand branches eaten by animals (called "browsing") and on the other hand drying by climatic factors (frost or drought in spring). The damage by browsing (Fig. 1B) was decreasing over time from the installation of fences, as expected.

These results allow us to understand the effect of animals on the structure and conformation of the forest ecosystem of the province. But it is essential to understand the behavior of the animals (frequency of use of the environment), the time of year with greater use and the possible overlap of presence of domestic and native animals in the forest. Therefore, the use of the entire dataset will increase the accuracy of the conclusions and allow discovering new facets of plant-herbivorous interactions.

2.1 What useful information can be obtained from the images?

For the understanding of plant-herbivore interaction in ñire forests and the generation of outputs similar to those shown in the previous section, it is useful to be able to identify the following image properties:

- Presence or absence of animals (photos with animals and without animals). Only photos with animals are of interest.
- Day or night (photos taken at night and photos taken during the day).
- Classify species in images with animals (for example guanacos, cows, foxes, among others).
- Number of animals per species.
- Use that the animal is making of the forest (for example feeding, resting).

Of the characteristics listed, being able to automatically distinguish images with animals is the task that would most contribute to expedite research, since they represent the minority within the data set; only 0.01% estimated. With the separation made, an expert user could quickly identify the rest of the conditions.

2.2 Dataset status

Within the more than 150,000 thousand images different situations can be observed: day and night photos, with high contrast of light, landscapes with and without snow, animals far away or very close to the camera are some examples. This diversity of conditions makes it necessary to have a large number of classified images, with different conditions, to achieve better accuracy when applying deep learning techniques.

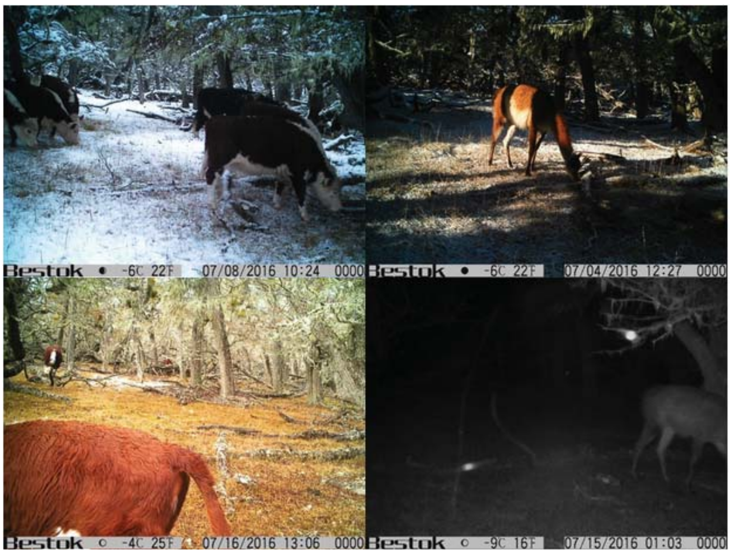

Fig. 2. Examples of the diversity of conditions within the data set. Cows in winter (left above). Guanaco by day with high light contrast (right above). Cows in daylight with a partially photographed animal (left below). Guanaco partially photographed at night (right below).

On the other hand, the trap cameras had different configurations of sensitivity to changes in the environment, which influenced the number of images taken by each of them. Thus, for the same period, there are cameras that have 300 photos while others have 3000, generating a tendency to replicate what happened in a particular forest.

The trap cameras used did not have the possibility of generating metadata that could contribute to improving a rapid classification (for example, identifying day and night with the creation date-time).

Finally, there is only a subset of 1,500 images analyzed by experts, of which 108 are classified as “with animals” and the rest as “without animals”. A quick visual evaluation would allow to differentiate the day and night photos.

3 Workplan

After evaluating the state of the data set, it was decided to work on two main aspects: the classification of images (in order to improve the results of the model to be generated) and establish a set of steps to follow to frame the next advances to be made.

3.1 Sorting images, a social experiment

In order to reduce the unclassified data set and obtain a more complete training set [6] for the deep learning model, a web application was designed and implemented that seeks to get volunteers willing to contribute to the project. The idea is simple: the website shows the user random images that must be classified as "with animals" or "without animals" and, at the same time, "by day" or "by night".

Fig. 3. Mobile version of the image classification screen.

To avoid malicious users or involuntary mistakes, each image must have a minimum of 3 responses and at the same time 90% of them must match the chosen criteria. For example, if three people classify an image as "without animals" and "at night", it is automatically classified with that criterion and will not be shown again to future users. On the contrary, if anyone had differed in the criteria, the image will continue to be shown to future users until reaching a 90% match in their rankings.

Due to this condition, it was necessary to adjust the algorithm that selects random images from the total set since, if they were really random, it would be difficult to get the same image to obtain (at least) three responses. Considering this, the following criteria were designed:

- All images are referenced within a table that has a random order.
- Of all the images, the algorithm limits its random selection to the first 150 unclassified images. These will be shown to the first users.
- As users match their responses and rate a photo, another will take its place. This will always keep a subset of 150 photos pending classification.
- Additionally, expert agents with registered users within the web application can access to revalidate the images or change the classification (if they consider it relevant).

Fig. 4. Summary screen for expert users (with a test database).

During August 2019, it will be promoted in local media and social networks to get people interested in participating. The main public will be inhabitants of the province of Tierra del Fuego. The web application is currently accessible through the link https://bosques.panalsoft.com.

3.2 Methodology for classification

With the purpose of marking a coherent course for the resolution of the problem raised, a series of steps to be solved were devised, which will be useful to address the problem in simpler and more manageable tasks.

Preprocessing. The original images of the data set have a dimension of 1280 × 960 pixels. This size is large in terms of computational calculation, both for training and for evaluation. In that sense, the de facto standard will be followed by reducing the images to 256 × 256 pixels [7]. Then, we will work on the normalization of colors to unify the tones and facilitate training [13, 14]**.**

Use of known architectures. There are different architectures of deep neural networks, with different number of layers, order, size and purpose [7]. It is intended to use some of the most proven architectures: AlexNet [15], VGG [16], GoogLeNet [17] or ResNet [11].

Pre-training, knowledge transfer. Transfer learning [18] takes advantage of the knowledge obtained to perform a task and applies it to a different and related task. For example, the architectures mentioned above could be pre-trained with the ImageNet data set [19]. Then, it would be re-trained with the set of images of the trap cameras, limiting the specific classifications.

Sort out day and night. This step would involve the first automatic classification, which should be simple due to the high contrast between both situations; saving light-shadow high contrast images. The task has a double purpose, to test the effectiveness of the previous tasks and to introduce the work team in the training of models, adjusting the hyperparameters to achieve more refined results.

Classify presence and absence of animals. This will be the main task and the one that will allow to achieve greater speed in the analysis of the information collected by the trap cameras. In simple words, if achieved correctly, it will be a great filter for experts to focus on the analysis of images with animals. It should be noted that this classification will be carried out in two separate sets provided by the previous task: night and day images. Thus, both the previous classification and this one will be two separate models dedicated exclusively to a particular task.

4 Summary and future work

This work raises the steps to follow to obtain a deep learning model that allows distinguishing (mainly) images with the presence or absence of animals. If a model with acceptable precision is achieved, it would help to speed up the analysis of the Fuegian forest and future projects involving trap cameras. In parallel, a web application was launched that will help classify the original data set, expanding the possibility of achieving a better deep learning model [7].

It is for future work to analyze the classifications achieved with web application and using it to achieve the first models.

Similarly, it would be useful to add new tasks to automate the recognition of the species that appear in the images, the number of animals per species and the use they are making of their environment.

Each of the above tasks should be performed on the sets of nocturnal images with animals and diurnal images with animals. This causes the models to train to be diversified too much. Therefore, it would be interesting to evaluate the possibility of obtaining a model that identifies the absence or presence of animals without requiring a previous filter between night and day.

References

1. Harris G, Thompson R, Childs JL, Sanderson JG: Automatic storage and analysis of camera trap data. Bull Ecol Soc Am (2010)
2. Soler RM, Lencinas MV, Bustamante G, Martínez Pastur G: Atributos de la regeneración natural de ñire (Nothofagus antarctica) en Tierra del Fuego: beneficios y perjuicios que genera el uso silvopastoril. Ecosistemas (2018)
3. O'Connell AF, Nichols JD, Karanth KU: Camera Traps in Animal Ecology: Methods and Analyses. Springer, Tokyo (2010)
4. Silveira L, Jacomo AT, Diniz-Filho JAF: Camera trap, line transect census and track surveys: A comparative evaluation. Biol Conserv (2003)
5. Fegraus EH, et al. Data acquisition and management software for camera trap data: A case study from the team network. Ecol Inform (2011)
6. Mohri M, Rostamizadeh A, Talwalkar A Foundations of Machine Learning. MIT Press, Cambridge, MA (2012)
7. Goodfellow I, Bengio Y, Courville A: Deep Learning. MIT Press, Cambridge, MA (2016)
8. Bridle JS: Probabilistic interpretation of feedforward classification network out-puts, with relationships to statistical pattern recognition. Neurocomputing. Springer, New York (1990)
9. Deng L, Hinton G, Kingsbury B: New types of deep neural network learning for speech recognition and related applications: An overview. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, New York (2013)
10. Sutskever I, Vinyals O, Le QV: Sequence to sequence learning with neural networks. 2014 Advances in Neural Information Processing Systems (NIPS). Neural Information Processing Systems Foundation, La Jolla, CA (2014)
11. He K, Zhang X, Ren S, Sun J: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, New York (2016)
12. Palmer MS, Packer C: Giraffe bed and breakfast: Camera traps reveal Tanzanian yellow-billed oxpeckers roosting on their large mammalian hosts. Afr J Ecol (2018)
13. LeCun YA, Bottou L, Orr GB, Müller KR: Efficient backprop in Neural networks: Tricks of the trade. (2012)
14. Wiesler S, Ney H: A convergence analysis of log-linear training. 2011 Advances in Neural Information Processing Systems (NIPS) (2011)
15. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. 2012 Advances in Neural Information Processing Systems (NIPS) (2012)
16. Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. arXiv preprint (2014)
17. Szegedy C, et al: Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
18. Yosinski J, Clune J, Bengio Y, Lipson H: How transferable are features in deep neural networks? 2014 Advances in Neural Information Processing Systems (NIPS) (2014)
19. Deng J, et al.: Imagenet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2009)