CrowdPet: Deep learning applied to the detection of dogs in the wild

Victor Capone*, Carlos Figueiredo, Eduardo Valle, Fernanda Andaló

Abstract

Biometry is the statistical study of physical or behavioral characteristics of living beings, mainly applied to the identification of individuals. Visual biometric methods map images to identities, and have been applied almost exclusively to humans. However, animal visual biometry is significantly different than that of humans, and current methods cannot be used interchangeably. The CrowdPet initiative studies and proposes methods to the problem of animal biometric identification, in order to identify and generate analytics related to stray animals. As part of this effort, in this work we present a method for detecting dogs in images, despite changes in pose and illumination, using Deep Learning. Experiments show promising results, in terms of accuracy, using challenging current datasets.

Key words:

animal visual biometrics, deep learning, convolutional neural network

Introduction

Many developing countries have a significant problem with street animal population. Aside the cruel conditions, this situation offers risks to public health, since these animals can be host to diseases and do reproduce at alarming rates. The CrowdPet initiative seeks to apply technological solutions, specifically computer vision and machine learning methods, to approach this problem, by providing a smart mobile app that is able to locate animals lost by they owners and to generate analytics regarding stray dogs. The main goal of this work is to contribute with computational methods for the visual recognition of these animals. More specifically, we study Deep Learning methods to detect animals in the wild, despite changes in scale, pose, and illumination.

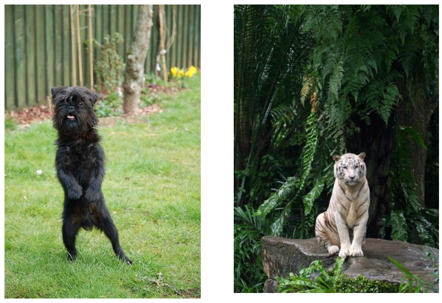

Figure 1. Examples of positive (left) and negative (right) images correctly classified with the VGG16 fine-tuned model. The probabilities of these images containing a dog were 100% and 10%, respectively.

Results and Discussion

We study the use of Convolutional Neural Networks (CNNs) to detect dogs in images acquired in different scenarios. We considered two state-of-the-art CNNs: VGG16[1] and Inception-v3[2], which were originally proposed to object recognition. We use transfer learning (fine tuning) to adapt the CNNS to our binary problem of classifying an image as containing or not a dog.

The fine tuning was made over images of three challenging datasets: *StanfordDogs*[3], *Caltech101*[4], and *Caltech256*[5]. The number of positive (dog) and negative (no dog) images are balanced, and they were split for training (70%) and validation (30%).

The networks were fine-tuned on GPU and the final accuracy was measured in images from the *FlickrDog*[6] and *Kaggle's DogsVsCats* datasets[7]. Table 1 shows the final accuracy for each dataset.

Table 1. Model accuracy for each test dataset.

Model	Dataset	Accuracy
VGG16	FlickrDog	95%
VGG16	DogsVsCats	98%
Inception-v3	FlickrDog	87%
Inception-v3	DogsVsCats	85%

Conclusions

Clearly, the machine learning techniques applied in this work were successful (Figure 1) for the task at hand, obtaining over 90% of accuracy. To be able to tackle the problem of identifying stray animals, research efforts are now focusing on the localization of the animals in the images, using deep learning techniques, such as Faster R-CNN. Posteriorly, the identification itself will be targeted, considering the proposed detection and localization methods.

Acknowledgement

We thank FAPESP, Recod lab, SciPet, and IEEE WIE Unicamp for their support in this work.

[1] Simonyan, K., and Zisserman, A. Very deep convolutional networks for large-scale image recognition. *ICLR*, 2015.
[2] Szegedy, C. et al. Rethinking the inception architecture for computer vision. *CVPR*, 2016.
[3] Khosla, A. et al. Novel dataset for fine-grained image categorization, *CVPR*, 2011.
[4] Fei-Fei L. et al. One-shot learning of object categories. *Pattern Analysis and Machine Intelligence*. **2006,** 28(4), 594-611.
[5] Griffin, G. et al. Caltech-256 Object Category Dataset. *California Institute of Technology*, 2007.
[6] Moreira, T. et al. Where is my puppy? Retrieving lost dogs by facial features. *Multimedia Tools and Applications.* **2017**, 76(14), 15325-15340.
[7] Kaggle's Dogs vs. Cats dataset: www.kaggle.com/c/dogs-vs-cats.