Data Science Capstone
Pre-Proposal
Jackson Crum

Identifying Wildlife Images with Deep Learning

The goal of this project is to develop a machine-learning algorithm and sorting application to allow for the efficient identification of endangered species in wildlife photos for United States government agencies.

The Endangered Species Act of 1973 requires federal agencies to ensure that any action they authorize does not threaten any listed endangered species or result in the damage or destruction of critical habitats of those species. Section 2(a)(2) states that it is "the policy of Congress that all Federal departments and agencies shall seek to conserve endangered species and threatened species and shall utilize their authorities in furtherance of the purposes of this Act."[1] The Department of Defense (DoD) has the highest density of endangered species listed as threatened or endangered of any other federal agency, with 340 out of 420 large military installations requiring active conservation management plans to protect 425 listed species. Expenditures on these conservation plans total more than $800 million annually.[2] Most of these habitat conservation efforts use camera traps to track wildlife. These traps can take millions of images and extracting information from these images is typically done by hand. This method is time-consuming and costly and, as a result, a significant amount of valuable knowledge in the image repositories goes unexamined.[3]

[1] U.S. Congress. Senate. *Endangered Species Act of 1973*. HR 37. 93rd Cong., 1st sess. Introduced in Senate June 12, 1973. https://www.fws.gov/international/pdf/esa.pdf.

[2] United States. Department of Defense. Natural Resources. By Peter Boice. February 2013. Accessed September 2018. http://www.dodnaturalresources.net/files/te__s_fact_sheet_2-21-13.pdf.

[3] Norouzzadeh, Mohammad Sadegh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune. "Automatically Identifying, Counting, and Describing Wild Animals in Camera-trap Images with Deep Learning." *Proceedings of the National Academy of Sciences* 115, no. 25 (November 17, 2017). Accessed September 2018. doi:10.1073/pnas.1719367115.

Animal detection and recognition are still a difficult challenge and researchers have yet to produce a method that provides an efficient solution for all situations. Research conducted by faculty of the Electrical Engineering Department at University of Zilina compared a CNN model to well-known image recognition methods, including PCA, LDA, and SVM models. It was concluded that the CNN outperformed all conventional models and produced a significant increase in accuracy.[4] Another joint university study tested various state-of-the-art neural network models, including AlexNet, GoogLeNet, VGG, and ResNet models, on Serengeti wildlife images. The study showed that deep neural networks performed well on the wildlife dataset, although performance was worse for rarer animals, and saved 99.3% on manual labor. While neural networks have proven successful in the task of image recognition, they rely on large datasets and perform worse on small training datasets.[5]

As the Red-cockaded woodpecker has received more funding than any other species, totally $144.7 million over the past 18 years, woodpeckers will be the primary focus of this project. The ImageNet visual database will be used to obtain labeled images, both of woodpeckers and non-woodpeckers, to populate the training and testing datasets. Convolutional neural networks (CNN) will be developed and compared using Keras and TensorFlow and trained/tested with varying number of layers, parameters, and confidence thresholds. Transfer learning will be employed to test state-of-the-art machine learning models, including VGG-16 and ResNet-50 neural networks. Amazon Web Services (AWS) and Google Cloud Platform (GCP) will be utilized for processing power to make these models feasible to train.

The hypothesis is that an automated system and application for identifying endangered species and removing non-target images would save the DoD in resource expenditure and man hours.

[4] Trnovszky, Tibor, Patrik Kamencay, Richard Orjesek, Miroslav Benco, and Peter Sykora. "Animal Recognition System Based on Convolutional Neural Network." *Advances in Electrical and Electronic Engineering* 15, no. 3 (September 2017). Accessed September 2018. doi:10.15598/aeee.v15i3.2202.

[5] Norouzzadeh, Mohammad Sadegh, et al. "Automatically Identifying, Counting, and Describing Wild Animals in Camera-trap Images with Deep Learning." *Proceedings of the National Academy of Sciences* 115, no. 25 (November 17, 2017).

References

Norouzzadeh, Mohammad Sadegh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S. Palmer, Craig Packer, and Jeff Clune. "Automatically Identifying, Counting, and Describing Wild Animals in Camera-trap Images with Deep Learning." *Proceedings of the National Academy of Sciences* 115, no. 25 (November 17, 2017). Accessed September 2018. doi:10.1073/pnas.1719367115.

Trnovszky, Tibor, Patrik Kamencay, Richard Orjesek, Miroslav Benco, and Peter Sykora. "Animal Recognition System Based on Convolutional Neural Network." *Advances in Electrical and Electronic Engineering* 15, no. 3 (September 2017). Accessed September 2018. doi:10.15598/aeee.v15i3.2202.

United States. Department of Defense. Natural Resources. By Peter Boice. February 2013. Accessed September 2018. http://www.dodnaturalresources.net/files/te__s_fact_sheet_2-21-13.pdf.

U.S. Congress. Senate. *Endangered Species Act of 1973*. HR 37. 93rd Cong., 1st sess. Introduced in Senate June 12, 1973. https://www.fws.gov/international/pdf/esa.pdf.