

Identifying Plant Images with Deep Learning

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 492 listed species (274 animals, 218 plants). Expenditures on these conservation plans total more than \$800 million annually.²

Species detection and recognition are still difficult, time-consuming, and expensive challenges 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.³ 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

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 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/aeec.v15i3.2202.

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.⁴

The data will come from Kaggle's plant seedlings classification contest. The data consists of 12 species of 5544 total images pre-split into training and testing datasets (4750 training and 794 testing images). Convolutional neural networks (CNN) will be developed and compared using Keras TensorFlow backend 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. The models will be evaluated with recall and classification accuracy as the primary metrics. Amazon Web Services (AWS) and Google Cloud Platform (GCP) will be utilized for processing power to make these models feasible to train. The application will be a Flask app that incorporates the Python/Keras backend and MongoDB for image data storage.

The goal of this project is to develop a full-stack application that will incorporate a machine-learning algorithm and sorting application to allow for the efficient identification of plant species in wildlife photos for United States government conservation agencies. This application will allow conservationists to simply upload files of images to the app, run the images through the algorithm, sort the files based on model output, and store image data in a MongoDB collection for quantitative analysis. This will allow non-technical researchers and conservationists to easily incorporate machine learning solutions to their problems. The hypothesis is that an automated system and application for observing target plant species, identify weeds and harmful species, and removing non-target images would increase the effectiveness of conservation efforts and save the DoD in resource expenditure and man hours.

⁴ 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 Orjeseck, 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/aeec.v15i3.2202.
- United States. Department of Defense. Natural Resources. By Peter Boice. February 2013. Accessed September 2018. http://www.dodnaturalresources.net/files/terms_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>.

Gantt Chart of Project Progress Timeline

