TEAM ID: PNT2022TMID00699

Digital Naturalist - AI Enabled tool for Biodiversity Researchers

Literature Survey

Abstract:

To better understand the complexities of natural ecosystems and better manage and protect them, it would be helpful to have detailed, large-scale knowledge about the number, location, and behaviours of animals in natural ecosystems. Having accurate, detailed, and up-to-date information about the location and behaviour of animals in the wild would improve our ability to study and conserve ecosystems. We investigate the ability to automatically, accurately, and inexpensively collect such data, which could help catalyze the transformation of many fields of ecology, wildlife biology, zoology, conservation biology, and animal behaviour into "big data" sciences. Motion-sensor "camera traps" enable collecting wildlife pictures inexpensively, unobtrusively, and frequently. However, extracting information from these pictures remains an expensive, time-consuming, manual task. We demonstrate that such information can be automatically extracted by deep learning, a cutting-edge type of artificial intelligence. We train deep convolutional neural networks to identify, count, and describe the behaviours of 48 species in the 3.2 million-image Snapshot Serengeti dataset. Our deep neural networks automatically identify animals with >93.8% accuracy, and we expect that number to improve rapidly in years to come.

Introduction:

The ever-growing number of digital sensors in the environment has led to an increase in the amount of digital data being generated. This includes data from satellites, weather stations, data from "internet of things" devices, and data collected by members of the public via smartphone applications, to name but a few. These new sources of data have contributed to the era of "Big Data" characterized by large volumes of data, of numerous types and quality, being generated at an increasing speed. This presents challenges and opportunities across a number of domains, including water management, camera trapping, and acoustic analysis. Automated identification of plants and animals have improved considerably in the last few years. In total, nine deep learning systems implemented by three different research teams were evaluated with regard to nine expert botanists of the French flora. Therefore, we created a small set of plant observations that were identified in the field and revised by experts in order to have a near-perfect golden standard. The main outcome of this work is that the performance of state-of-the-art deep learning models is now close to the most advanced human expertise. This shows that automated plant and animal identification systems are now mature enough for several routine tasks, and can offer very promising tools for autonomous ecological surveillance systems.

Artificial intelligence (AI) techniques have profoundly transformed our ability to extract information from visual data. AI techniques have been applied for a long time in security and industrial domains, for example, in iris recognition or the detection of faulty objects in manufacturing. They were nevertheless only recently made more widely accessible after their use in smartphone apps for face recognition and song identification. Combined with increasing access to cloud-based computation, AI techniques can now automatically analyse hundreds of thousands of visual data every day.

Deep learning models (some of the most advanced AI algorithms) are developed with training datasets that allow them to capture discriminant visual patterns. Their performances are then strongly correlated to the quality and completeness of the datasets on which they are trained. Unbalanced, biased, or otherwise poor-quality training datasets will lead to underperforming algorithms in real conditions. During the learning phases, particular attention must be given to any relevant limitations of the training data, and the gap between these and the test data on which the developed algorithms will be evaluated.

S.NO	YEAR	PAPER NAME	ADVANTAGE	DISADVANTAGE
1.	2018	Plant identification: Experts vs. machines in the era of deep learning: deep learning techniques challenge flora experts.	This is a highly powerful tool for modern botany.	It does not help in identification of fauna.
2.	2017	Plant Species Identification Using Computer Vision Techniques: A Systematic Literature	It describes the applied methods categorized according to the studied plant organ, and the studied features, i.e., shape, texture, color, margin, and vein structure.	It does not describe about fauna.
3.	2020	AI Naturalists Might Hold the Key to Unlocking Biodiversity Data in Social Media Imagery	Non-native garden plants dominated, particularly in the urban setting. The AI classifier performed best when photos were focused on	It does not work efficiently on multiple native species.

			single native species in wild situations but also performed well at higher taxonomic levels (genus and family), even when images substantially deviated from this.	
4.	2012	Automated plant identification using artificial neural networks	The information was extracted automatically from images of specimens and only characteristics of the leaf shape and morphology have been used.	This is less accurate than manual methods and the data contains more noise, and so to avoid this it requires more data

Reference:

1. Plant identification: Experts vs. machines in the era of deep learning: deep learning techniques challenge flora experts. Bonnet P., Goeau H., Hang S.T., Lasseck M., Sulc M., Malécot V., Jauzein P., Melet J.C., You C., Joly A.. 2018.
2. Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review Jana Wäldchen, Patrick Mäder Published 7 January 2017
3. AI Naturalists Might Hold the Key to Unlocking Biodiversity Data in Social Media Imagery Tom A

August, Oliver L Pescott, Alexis Joly, Pierre Bonnet Patterns 2020

4. Automated plant identification using artificial neural networks Jonathan Y. Clark, D. Corney, H. L. Tang Computer Science 2012 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)

Conclusion:

A methodology is presented here to provide a practical way for taxonomists to use neural networks as automated identification tools, by collating results from a population of neural networks.