Digital Naturalist AI Enabled Biodiversity Tool

Literature survey

Introduction:
A naturalist is someone who studies the patterns of nature, identifies a different kind of flora and fauna in nature. Being able to identify the flora and fauna around us often leads to an interest in protecting wild spaces, and collecting and sharing information about the species we see on our travels is very useful for conservation groups like NCC. When venturing into the woods, field naturalists usually rely on common approaches like always carrying a guidebook around everywhere or seeking help from experienced ornithologists. There should be a handy tool for them to capture, identify and share the beauty to the outside world.

Field naturalists can only use this web app from anywhere to identify the birds, flowers, mammals and other species they see on their hikes, canoe trips and other excursions. In this project, we are creating a web application which uses a deep learning model, trained on different species of birds, flowers and mammals (2 subclasses in each for a quick understanding) and get the prediction of the bird when an image is being given.

Project Requirement:

- ***Software Tools:***
  - Anaconda Navigator
- ***Packages:***
  - Tensorflow
  - Keras
  - Flask

Literature Survey:

[1] How can Big Data and machine learning benefit environment and water management

(by; Alexander Y Sun and Bridget R Scanlon)

Published 1 July 2019 • © 2019 The Author(s). Published by IOP Publishing Ltd

https://iopscience.iop.org/article/10.1088/1748-9326/ab1b7d

A large number of applications on Big Data and ML have already appeared in the EWM literature in recent years. The purposes of this survey are to (1) examine the potential and benefits of data-driven research in EWM, (2) give a synopsis of key concepts and approaches in Big Data and ML, (3) provide a systematic review of current applications,

and finally (4) discuss major issues and challenges, and recommend future research directions. EWM includes a broad range of research topics. Instead of attempting to survey each individual area, this review focuses on areas of nexus in EWM, with an emphasis on elucidating the potential benefits of increased data availability and predictive analytics to improve the EWM research.

[2] A new subset based deep feature learning method for intelligent fault diagnosis of bearing

(by; Zhang Yuyan, Li Xinyu, Gao Liang and Li Peigen)
Published on 15 November 2018
https://www.sciencedirect.com/science/article/abs/pii/S0957417418303324

Intelligent fault diagnosis has attracted considerable attention due to its ability in effectively processing massive data and rapidly providing diagnosis results. However, in the traditional intelligent diagnosis methods of bearing, features are extracted manually. Such a process is not only a grueling and time-consuming work but also greatly affects the diagnosis results. In this study, we propose a new intelligent diagnosis method of bearing, which can learn features automatically. First, a new subset approach is developed and it is helpful to learn the discriminative features from different fault patterns. Second, a subset based deep auto-encoder model is proposed to realize the automatic feature extraction. Additionally, a new self-adaptive fine-tuning operation is designed to ensure the good convergence performance of SBTDA. Finally, to obtain the appropriate configuration, several key parameters are optimized with the particle swarm optimization algorithm. The proposed method is evaluated on three public bearing datasets, and achieves the average testing accuracies of 99.65%, 99.66% and 99.60% respectively. The comparisons with 13 intelligent diagnosis methods demonstrate that SBTDA can obtain higher diagnosis accuracy.

[3] Machine learning for image based species identification

(by; Jana Wäldchen, Patrick Mäder)
First published on 13 August 2018; accepted: 30 July 2018; DOI: 10.1111/2041-210X.1307
https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.13075

Accurate species identification is the basis for all aspects of taxonomic research and is an essential component of workflows in biological research. Biologists are asking for more efficient methods to meet the identification demand. Smart mobile devices, digital cameras as well as the mass digitisation of natural history collections led to an explosion of openly available image data depicting living organisms. This rapid increase in biological image data in combination with modern machine learning methods, such as deep learning, offers tremendous opportunities for automated species identification.

[4] Automated plant species identification - Trends and future directions
(by; Jana Wäldchen, Michael Rzanny, Marco Seeland and Patrick Mäder)
Published on April 5, 2018
https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005993
Current rates of species loss triggered numerous attempts to protect and conserve biodiversity. Species conservation, however, requires species identification skills, a competence obtained through intensive training and experience. Field researchers, land managers, educators, civil servants, and the interested public would greatly benefit from accessible, up-to-date tools automating the process of species identification. Currently, relevant technologies, such as digital cameras, mobile devices, and remote access to databases, are ubiquitously available, accompanied by significant advances in image processing and pattern recognition. The idea of automated species identification is approaching reality. We review the technical status quo on computer vision approaches for plant species identification, highlight the main research challenges to overcome in providing applicable tools, and conclude with a discussion of open and future research thrusts.

[5] A look inside the Pl@ntNet experience
(by; Alexis Joly, Pierre Bonnet, Hervé Goëau, Julien Barbe, Souheil Selmi, Julien Champ, Samuel Dufour-Kowalski, Antoine Affouard, Jennifer Carré, Jean-François Molino, Nozha Boujemaa and Daniel Barthélémy)
Published on 11 April 2015; Issue Date by November 2016; 751–766 (2016).
https://doi.org/10.1007/s00530-015-0462-9
https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005993
Pl@ntNet is an innovative participatory sensing platform relying on image-based plant identification as a means to enlist non-expert contributors and facilitate the production of botanical observation data. One year after the public launch of the mobile application, we carry out a self-critical evaluation of the experience with regard to the requirements of a sustainable and effective ecological surveillance tool. We first demonstrate the attractiveness of the developed multimedia system (with more than 90K end-users) and the nice self-improving capacities of the whole collaborative workflow. We then point out the current limitations of the approach towards producing timely and accurate distribution maps of plants at a very large scale. We discuss in particular two main issues: the bias and the incompleteness of the produced data. We finally open new perspectives and describe upcoming realizations towards bridging these gaps.

[6] A research tool for long-term and continuous analysis of fish assemblage in coral-reefs using underwater camera footage
(by; Bastiaan J.Boom, Jiyin He, Simone Palazzo, Phoenix X.Huang, Cigdem Beyana, Hsiu-Mei Chou, Fang-Pang Lind, Concetto Spampinato and Robert B.Fishera)

First published on September 2014, Pages 83-97; volume23
https://www.sciencedirect.com/science/article/abs/pii/S1574954113001003?via%3Dihub
We present a research tool that supports marine ecologists' research by allowing analysis of long-term and continuous fish monitoring video content. The analysis can be used for instance to discover ecological phenomena such as changes in fish abundance and species composition over time and area. Two characteristics set our system apart from traditional ecological data collecting and processing methods. First, the continuous video recording results in enormous data volumes of monitoring data. Currently around a year of video recordings (containing over the 4 million fish observations) have been processed. Second, different from traditional manual recording and analyzing ecological data, the whole recording, analyzing and presentation of results is automated in this system. On one hand, it saves the effort of manually examining every video, which is infeasible. On the other hand, no automatic video analysis method is perfect, so the user interface provides marine ecologists with multiple options to verify the data. Marine ecologists can examine the underlying videos, check results of automatic video analysis at different certainty levels computed by our system, and compare results generated by multiple versions of automatic video analysis software to verify the data in our system. This research tool enables marine ecologists for the first time to analyze long-term and continuous underwater video records.