# **Paper title:** AI Naturalists Might Hold the Key to Unlocking Biodiversity Data in Social Media Imagery

# **Author:** Tom A.August,Oliver L.Pescott

# **Description:**

1. AI image classifiers can create biodiversity datasets from social media imagery
2. Flickr hosts many images of plants; some can be accurately classified to species by AI
3. Images are spatially aggregated around tourist sites and under-represent native species
4. Images focused on a single, non-horticultural, plant are most reliably identified

**Technique Used**:  Deep Learning And Convolutional Neural Networks (Cnns)

**Advantage:**

### Images Are Spatially Clustered

### The Urban Area Has a Lower Proportion of Well-Classified Images

### The Rural Area Has a Higher Proportion of Images of Naturally Occurring, Native Plants

**Disadvantages:**

A poor match between training and prediction datasets will result in higher error rates, which may not always be associated with low classification scores.

**Paper title:** Machine learning to classify animal species in camera trap

images: Applications in ecology

**Author:** Michael A. Tabak ,Mohammad S. Norouzzadeh,David W. Wolfson

# **Description:**

1. Motion-activated cameras (“camera traps”) are increasingly used in ecological and management studies
2. We trained machine learning models using convolutional neural networks with the ResNet-18 architecture and 3,367,383 images to automatically classify wildlife species from camera trap images obtained from five states across the United States.
3. The trained model classified approximately 2,000 images per minute on a laptop computer with 16 gigabytes of RAM.
4. The use of machine learning to rapidly and accurately classify wildlife in camera trap images can facilitate non-invasive sampling designs in ecological studies by reducing the burden of manually analysing images.

**Technique Used**: Artificial intelligence, deep neural networks

**Advantage:**

1. Deep learning models that automatically classify wildlife
2. This model achieved the highest accuracy (97.6%) to date in using machine learning to classify wildlife in camera trap images
3. The model can also be valuable for researchers studying other species by removing images without any animals from the dataset before beginning anual classification

**Disadvantages:**

1. The burden of classifying images from camera traps has led ecologists to limit the duration and size of camera trap studies
2. The ability to rapidly identify millions of images from camera traps can fundamentally change the way ecologists design and implement wildlife studies

**Paper title:**Overview of ExpertLifeCLEF 2018: how far automated identification systems are from the best experts?

**Author:** Hervé Goëau, Pierre Bonnet, Alexis Joly

# **Description:**

1. Automated identification of plants and animals has improved considerably in the last few years, in particular thanks to the recent advances in deep learning.
2. This paper presents more precisely the resources and assessments of the challenge, summarizes the approaches and systems employed by the participating research groups
3. A picture actually contains only a partial information about the observed plant and it is often not sufficient to determine the right species with certainty
4. Quantifying this intrinsic data uncertainty and comparing it to the performance of the best automated systems is of high interest for both computer scientists and expert naturalists.

**Technique Used**: Deep-learning systems,Automated systems

**Advantage:**

1. The measured performances are very high despite the difficulty of the task,
2. The best results were obtained mostly by systems that were learned on both the trusted and the noisy datasets,
3. The best results were obtained by ensemble classifiers of ConvNets with many data augmentations.

**Disadvantages:**

1. The pictures of plants only contain partial information and that it is often not sufficient to determine the right species with certainty.
2. The observations in the test set do not contain enough information to be identified with confidence when using classical identification keys

**Paper title:** Machine learning for image based species identification

**Author:** Jana Wäldchen, Patrick Mäder

# **Description:**

1. Accurate Species identification is the basis for all aspects of taxomonic research and is an essential component of workflow in biological research.
2. The field of computer vision and Machine Learning resulted in a plethora of papers proposing and comparing methods on automated species

identification.

**Technique Used**: Automated species identification, Computer vision, Convolutional neural network,

**Advantage:**

1. It is fast development and ubiquity of relevant information.
2. Species identification tasks are manifold and were comprehensive.
3. It is widely adopted approach for object detection.

**Disadvantages:**

1. The number of taxonomists and identification experts is drastically decreasing.
2. Modern machine learning approaches only slowly pave their way into the field of species identification,
3. Species are not evenly distributed throughout a larger region as they require more or less specific combinations of biotic and abiotic factors and resources to be present for their development

**Paper title:** Computer Age Statistical Inference:Algorithms, Evidence, And Data Science

**Author:** Bradley Efron, Trevor Hastie

# **Description:**

1. The 21st century has seen a breath-taking expansion of statistical methodology,both in scope and influence.
2. Data Science and Machine learning have become familiar terms in the news .

**Technique Used**: Algorithms, Evidence, and Data Science

**Advantage:**

1. It is a collection of humorous scientification
2. The characters in the fire of poem.
3. The foreshadowing was skillfull.

**Disadvantages:**

* 1. Published first in logic time.
  2. It is the forgotten all about figuring.
  3. It is not expecting to find her spirit.

**Paper title:** Fifty years of the Biological Records Centre

**Author:** HELEN E. ROY\*, CHRISTOPHER D. PRESTON , DAVID B.ROY

# **Description:**

1. Phenology is widely seen as one of the clearest ways of documenting ecological responses to climate change.
2. The richness of the invertebrate datasets is one of the most notable features of biological recording in Britain; the invertebrate recording schemes and societies provide unprecedented sources of data on fauna that are otherwise often neglected
3. Technological advances have revolutionized biological recording
4. The possibilities offered by modern computing have allowed the development of analytical techniques that maximize the use of the

largely unstructured datasets accrued through biological recording

**Technique Used**: Biological recording, wildlife observations

**Advantage:**

1. The value of the inspiring contributions made by volunteers meticulously documenting our wildlife to inform conservation and research will undoubtedly ensure an exciting future for biological recording.
2. The molecular revolution is also providing alternative approaches to monitoring biodiversity
3. Biological records have been widely used to predict the changes in species distribution as a consequence of projected climate change

**Disadvantages:**

1. **‘**If we want to maintain the richness of our flora and fauna, we need to

hold on to as much as we can and not just see one species as replaceable by another’.

1. The contributions within this special issue highlight the breadth and value of biological records to advancing knowledge.
2. Most recording schemes and societies focus on the compilation of the records required to develop an atlas documenting the distribution of species.
3. For the less popular groups, this may take many years.

**Paper title:** Avoid Oversimplifications in Machine Learning: Going beyond the Class-Prediction Accuracy

**Author:** Sung Yang Ho, Limsoon Wong,Wilson Wen Bin Goh

# **Description:**

1. The class-prediction accuracy provides a snapshot of learning

Performance

1. The objective is to identify rules and feature sets that allow differentiation of the various class labels.
2. The class-prediction accuracy does not inform on the mechanism and processes undertaken inorder to arrive at the predict.

**Technique Used**: Data science, Class-prediction accuracy

**Advantage:**

1. Class-prediction accuracy provides a quick but superficial way of determining classifier performance.
2. It does not provide explainability in its decision-making process and is not objective, as its value is also affected by class proportions in the validation set.
3. Instead, it needs to be enriched with accompanying evidence and tests that supplement and contextualize the reported accuracy.

**Disadvantages:**

1. High Accuracy Does Not Imply Reproducibility
2. High Accuracy Does Not Imply Meaningfulness
3. High Accuracy Does Not Imply that FeaturesUsed Are Better Than Random

**Paper title:** Opportunities and Risks for Citizen Science in the Age of Artificial Intelligence

**Author:** Luigi Ceccaroni, James Bibby,Erin Roger,Paul Flemons,Katina Michael,Laura Fagan and Jessica L. Oliver

# **Description:**

1. Automated reasoning is an area of computer science and mathematical logic dedicated to understanding different aspects of reasoning.
2. Computer vision and hearing are interdisciplinary fields that explore how computer algorithms and systems can classify and/or identify content and achieve high-level understanding from digital images, videos, or audio

recordings

1. AI can be described as intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals.

**Technique Used**: Artificial Intelligence, speech recognition

**Advantage:**

1. Improving image or audio classification
2. Accelerating the digitization of biodiversity research specimens
3. Training of computer-vision and computer-hearing algorithms using citizen science data.
4. Facilitating sharing the meaning of terms
5. Using social media for collaborative species identification and occurrence

**Disadvantages:**

1. When contributing expertise to develop and train AI, they are not properly and fairly acknowledged, respected, and rewarded;
2. They think that new technologies could be driven more by short-term commercial necessity than longer-term social good
3. They are not comfortable sharing their data because of concerns that their data might be unfairly appropriated (especially for commercial purposes);
4. They are forced (because of ethical considerations) to provide too-frequent re-confirmation of their willingness to share their data openly. (See GDPR(2016) as an example of where good intention can sometimes become burdensome.)

**Paper title:** Rethinking the Inception Architecture for Computer Vision

**Author:** Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe

**Description:**

1. Convolutional networks are at the core of most state-of-the-art computer vision solutions for a wide variety of tasks. Since 2014 very convolutional networks started to become mainstream, yielding substantial gains in various benchmarks.
2. Although increased model size and computational cost tend to translate to immediate quality gains for most tasks (as long as enough labeled data is provided for training), computational efficiency and low parameter count are still enabling factors for various use cases such as mobile vision and big-data scenarios.
3. Here we are exploring ways to scale up networks in ways that aim at utilizing the added computation as efficiently as possible by suitably

factorized convolutions and aggressive regularization.

**Technique Used**: VGGNet, GoogLeNet.

**Advantage:**

1. Spatial Factorization into Asymmetric Convolutions
2. Factorization into smaller convolutions
3. Utility of Auxiliary Classifiers
4. Efficient Grid Size Reduction
5. Model Regularization via Label Smoothing

**Disadvantages:**

1. The combination of lower parameter count and additional regularization with batch-normalized auxiliary classifiers and label-smoothing allows for training high quality networks on relatively modest sized training sets.
2. Still our solution uses much less computation than the best published results based on denser networks
3. The common wisdom is that models employing higher resolution receptive fields tend to result in significantly improved recognition performance.