PNIE PROCEEDINGS OF THE NATIONAL INSTITUTE OF ECOLOGY OF THE REPUBLIC OF KOREA

Review Article

PNIE 2023;4(4):135-140
https://doi.org/10.22920/PNIE.2023.4.4.135
pISSN 2765-2203, eISSN 2765-2211

Linking Species with the Ecosystem: The Emergence of Big Data and Artificial Intelligence in Ecological Research

Ohseok Kwon[1,2,3]*

[1]Department of Plant Medicine, College of Agriculture and Life Sciences, Kyungpook National University, Daegu, Korea
[2]Department of Plant Protection and Quarantine, Kyungpook National University, Daegu, Korea
[3]Institute of Plant Medicine, Kyungpook National University, Daegu, Korea

ABSTRACT

The intricate relationship between species and their ecosystems has been a focal point of ecological research for decades. With the advent of big data and artificial intelligence, we are now able to explore this relationship with unprecedented depth and precision. This review delves into the transformative role of these technologies in ecological research, emphasizing their potential to enhance our understanding of species-ecosystem linkages.

Keywords: Artificial intelligence, Big data, Ecological research, Ecosystem, Species

Introduction

Ecological research has traditionally relied on field observations and experimental data, often limited by geographical scope and temporal scale. However, the emergence of big data and artificial intelligence (AI) technologies has revolutionized this field, enabling researchers to analyze complex ecological phenomena on a global scale and in real-time.

For last two decades, the realization of AI technology in real life application has become true. When AlphaGo (Silver *et al*., 2016) first came out, people were shocked to watch how fast AI can catch up human intelligence. ChatGPT hit human community hard with fears that AI may take over human intelligence. However, it is us to design and execute AI, and therefore we should not fear of it. Understanding the nature of AI technology and learning how it can help us are very important for that reason.

Deep learning and big data are two significant areas in the field of AI that have seen substantial growth and development in recent years. Here's a conceptual background of how they interact.

Deep learning

Deep learning is a subset of machine learning, which itself is a subset of AI. It involves the use of artificial neural networks with several layers - these are the 'deep' structures that gave deep learning its name. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—in order to 'learn' from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help optimize the accuracy.

Received October 4, 2023; **Revised** October 10, 2023; **Accepted** October 11, 2023

***Corresponding author:** Ohseok Kwon
e-mail ecoento@knu.ac.kr
https://orcid.org/0000-0001-9075-3994

Big data

Big data refers to extremely large datasets that can be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions. The key aspects of big data, often referred to as the 3Vs, are volume (large amounts of data), velocity (speed at which data is generated and processed), and variety (different types of data, from structured to unstructured).

Interaction of deep learning and big data in AI

Deep learning algorithms, due to their capacity for high levels of abstraction, are particularly well-suited to make use of big data. They can use this wealth of information to train themselves and improve their accuracy. The more data fed to these algorithms, the better they usually perform. This is where big data comes into play. With the vast amounts of data available in today's digital world, deep learning algorithms can be used to make increasingly accurate predictions.

In essence, deep learning provides the brain that makes sense of the data, and big data provides the nutritious food that feeds this brain. Together, they are a powerful combination in the field of AI, driving many of the most exciting and promising applications of AI we see in the world today.

Big Data in Ecological Research

Big data in ecology refers to the vast amounts of data generated from various sources such as remote sensing, genomics, citizen science projects, and long-term ecological studies. These data sets provide a holistic view of ecosystems, capturing the intricate interactions between species and their environment. They allow researchers to track changes in biodiversity, monitor species distributions, and predict the impacts of climate change on ecosystems. One of the early studies implementing big data in ecological research is by Varotsos and Krapivin (2017).

There are many ways to get the big data for ecological research with help from modern technology.

Remote sensing

Satellite imagery: High-resolution images from satellites like Landsat and Sentinel can provide valuable ecological information on land cover, vegetation health, and changes in ecosystem structure over time.

Drones: Unmanned airial vehicles equipped with cameras or other sensors can capture detailed data over smaller areas, which can be especially useful for difficult-to-reach locations.

In-situ monitoring

Sensor networks: Automated sensor networks can collect large datasets on climate variables, soil conditions, and water quality.

Acoustic monitoring: Automated audio devices can collect big data on animal calls and other sounds, which can be used to monitor biodiversity.

Citizen science

Crowdsourced observations: Mobile apps like iNaturalist allow citizen scientists to upload observations of flora and fauna, contributing to large datasets.

Online databases: Websites like eBird and BugGuide allow users to contribute ornithological and entomological data, respectively.

Data repositories

Global biodiversity information facility: This open-access platform provides data on a wide range of biodiversity aspects.

National centers for environmental information: Holds datasets related to climate and environment.

Traditional surveys

Long-term ecological research: Traditional field studies, when conducted over long periods, can provide large volumes of data. However, the data might be specific to certain parameters or regions.

Literature mining: Existing academic literature often contains data that can be aggregated for meta-analysis.

Collaboration and data sharing

Interdisciplinary collaboration: Working with experts from fields like computer science or geoinformatics can help ecologists manage and analyze big data.

Data consortiums: Institutions often share data for mutual benefit.

Computational methods

Web scraping: Collecting publicly available data from websites and online repositories.

Social media monitoring: Platforms like Twitter can provide real-time data about human-wildlife interactions, etc.

It's crucial to consider the ethics of data collection, especially when it involves sensitive or endangered species, or when utilizing citizen science data. Proper permissions and protocols should be followed. By leveraging a combination of these methods, ecologists can acquire the large datasets necessary for comprehensive analysis and impactful research.

Artificial Intelligence in Ecological Research

AI, particularly machine learning, has emerged as a powerful tool for analyzing big ecological data. Machine learning algorithms can identify patterns and make predictions based on large, complex datasets, a task that would be nearly impossible for humans. For instance, AI can be used to predict species distribution based on environmental variables, identify species from images or sounds, and model ecosystem dynamics. Some of the possible applications of AI in ecological research are presented below.

Data analysis and pattern recognition

Species identification: Machine learning algorithms can identify species from images, sounds, or other types of data.

Climate modeling: AI can be used to develop more accurate models for understanding climate change effects on ecosystems.

Anomaly detection: Identifying unusual patterns in environmental data, which could signify a problem such as an invasive species or disease outbreak.

Data collection and automation

Remote sensing: AI can analyze satellite and drone imagery to monitor changes in land use, deforestation, or coral reef health.

Automated counting: Machine learning models can automatically count animal populations from camera trap images or aerial photographs.

Sensor data interpretation: AI can process large volumes of data from sensor networks, making real-time monitoring of ecosystems more efficient.

Conservation and management

Habitat assessment: AI can help identify critical habitats that need protection or restoration.

Resource allocation: Optimization algorithms can aid in efficient allocation of resources for conservation efforts.

Poaching prevention: AI algorithms can predict poaching activities and help in developing effective anti-poaching strategies.

Population dynamics

Predictive modeling: Machine learning can help forecast population sizes and distributions based on various factors like food availability, climate change, and human activity.

Animal behavior analysis: AI can help understand intricate animal behaviors through the analysis of movement data, vocalizations, and social interactions.

Citizen science and engagement

Data validation: AI can help in the quick validation of data gathered through citizen science initiatives.

Public awareness: Natural language processing (NLP) can analyze public sentiment and awareness about ecological issues.

Human-wildlife interactions

Conflict mitigation: AI can predict potential human-wildlife conflicts by analyzing data on animal movements and human activities.

Disease monitoring: AI can help in tracking the spread of diseases that could transfer between animals and humans.

Ecosystem services and economic valuation

Assessment of ecosystem services: AI can evaluate the economic and social value of ecosystem services, aiding in conservation planning.

Supply chain analysis: AI can track the sustainability of products sourced from ecosystems, like timber and fish.

Meta-analysis and literature mining

Systematic review: NLP can help review a large number of scientific articles to extract meaningful conclusions.

Bioacoustics

Sound classification: AI can analyze ecological audio recordings to classify sounds, identify species, and assess biodiversity.

Simulation and virtual reality

Ecological simulations: AI can be used to simulate various ecological scenarios, which can help in understanding complex dynamics and interactions.

These are just examples, and as technology continues to evolve, the range of applications is likely to expand, offering even more ways to leverage AI in ecological research.

Linking Species with the Ecosystem

The combination of big data and AI has significantly enhanced our ability to understand the linkages between species and ecosystems. For example, researchers can now model how changes in an individual species might affect the entire ecosystem, or how changes in the ecosystem might affect individual species. This has important implications for conservation, as it allows us to identify species that are particularly important for ecosystem function, and to predict how ecosystems might respond to the loss of these species.

Norouzzadeh *et al*. (2018) showed the application of AI in acquiring the ecological data from camera-trap images with deep learning. Buckland *et al*. (2023) showed the possibility in using modern technology to estimate population dynamics. There are rapidly increasing number of articles emerging from ecology sector that utilize these new scientific advances.

Some ecology journals are oriented in this direction. The most well-known ecology journals of this subject are given below.

Ecological modelling

Publisher: Elsevier.

Scope: This journal is focused on the use of computational methods and models to understand ecological systems. It covers a wide range of ecological topics, from population dynamics to ecosystem studies. The journal aims to promote the development and use of mathematical models and computational techniques in ecological research.

Audience: Ecologists, environmental scientists, mathematical modelers, and computational scientists.

Noteworthy: Given its broad scope, "Ecological Modelling" has been a leading journal in the field of ecological modeling for several decades.

Methods in ecology and evolution

Publisher: British Ecological Society.

Scope: This journal aims to promote the development of new methods in ecology and evolution and facilitate their dissemination across the scientific community. It is focused on methodological papers that introduce new ways to collect, model, and analyze ecological and evolutionary data.

Audience: Researchers in ecology, evolution, and related disciplines interested in methodological advancements.

Noteworthy: It often features articles that apply advanced computational techniques, such as machine learning and AI, to ecological and evolutionary questions.

Computational ecology and software

Publisher: International Academy of Ecology and Environmental Sciences.

Scope: This open-access journal focuses on software, models, and computational methods specifically designed for ecological research. It serves as a platform for ecologists to share software tools and computational methods to advance the field.

Audience: Ecologists, computational biologists, and software developers interested in ecological applications.

Noteworthy: Being open-access, this journal makes its articles freely available, allowing for wider dissemination of computational methods and tools in ecology.

These journals would be beneficial resources for someone interested in the intersection of computational methods and ecology. Given your background in ecology, you may find articles in these journals that are closely aligned with your field of expertise.

Challenges and Future Directions

Despite the significant advances brought about by big data and AI, there are still many challenges to be addressed. These include issues related to data quality and accessibility, the interpretability of AI models, and the need for interdisciplinary collaboration.

Applying AI in ecological research presents numerous challenges, from data collection and quality issues to ethical and interpretability concerns. However, solutions do exist for overcoming these obstacles. Here are some challenges and ways to address them.

Data availability and quality

Issue: Lack of high-quality, labeled data for training machine learning models.

Solution: Leverage citizen science, collaborations, and existing datasets to enhance the availability of high-quality data. Data augmentation techniques can also be used.

Computational resources

Issue: AI, particularly deep learning models, often require significant computational power.

Solution: Utilize cloud computing services, or collaborate with departments or organizations that have the necessary computational resources.

Interpretability and transparency

Issue: Many advanced AI algorithms, like neural networks, are often seen as "black boxes," making it difficult to interpret their decisions.

Solution: Use explainable AI techniques or simpler models where interpretability is crucial for scientific validity.

Ethical considerations

Issue: Ethical dilemmas around data privacy, especially in citizen science initiatives, and the use of AI in conservation policing.

Solution: Develop ethical guidelines and obtain informed consent when collecting data. Ensure that AI applications comply with legal regulations.

Domain expertise

Issue: The interdisciplinary nature of applying AI in ecology means that expertise is needed in both domains.

Solution: Foster interdisciplinary teams comprising ecologists, data scientists, and AI experts to better formulate research questions and solutions.

Scalability

Issue: Models trained on specific datasets may not generalize well to other ecosystems or species.

Solution: Use transfer learning or domain adaptation techniques to make models more adaptable to different but related problems.

Bias and fairness

Issue: Data and model biases can lead to unfair or inaccurate ecological predictions or assessments.

Solution: Use techniques to identify and mitigate biases in both the training data and the model itself.

Cost and time

Issue: Data collection, labeling, and analysis using AI can be expensive and time-consuming.

Solution: Apply for grants and partnerships that can provide both financial and logistical support.

Data integration

Issue: Ecological research often involves diverse data types (e.g., audio, image, numerical), making it challenging to integrate them into a unified AI model.

Solution: Use multimodal machine learning techniques designed to handle multiple types of input data.

Real-world applicability

Issue: There might be a gap between what a model predicts in a controlled setting and its efficacy in the field.

Solution: Validate models using independent datasets and in real-world settings to ensure their robustness and applicability.

Overcoming challenges

Collaboration: Work with interdisciplinary teams to bring expertise from various domains.

Education: Train ecologists in basic AI and data science skills, and vice versa.

Community involvement: Engage the community for citizen science efforts to aid in data collection and validation.

Regular updates: Keep abreast of the latest technologies and methodologies in AI and ecology.

By proactively addressing these challenges and continually refining methodologies and models, researchers can more effectively harness the power of AI in ecological research.

The application of AI in ecological research is still an evolving field, and several future directions can be foreseen, including but not limited to.

Advanced data analytics

Real-time analysis: As sensor and internet of things technologies improve, real-time analysis of ecosystems using AI will become more feasible.

Multimodal data integration: Future research is likely to focus on developing AI algorithms capable of integrating multiple types of data (e.g., audio, visual, genetic) to offer a more comprehensive analysis.

Improved models and algorithms

Explainable AI: Given the need for interpretability in scientific research, developing AI models that are both powerful and explainable will be crucial.

Uncertainty quantification: AI models that can estimate the uncertainty of their predictions will be increasingly important, especially for decision-making in conservation and management.

Enhanced monitoring and conservation

Automated monitoring: AI-powered drones and automated traps and sensors could revolutionize how species and habitats are monitored.

Predictive policing: AI could help in predicting illegal activities like poaching, thereby optimizing patrol routes for rangers.

Citizen science and public engagement

Crowdsourcing and validation: AI can assist in automatically validating the large amounts of data generated through citizen science projects.

Public awareness: AI-powered chatbots or platforms could educate the public and engage them in ecological issues and data collection.

Interdisciplinary collaboration

Integrated platforms: Development of comprehensive platforms that can analyze, interpret, and visualize ecological data, thereby enabling better collaboration between ecologists, data scientists, and policymakers.

Curriculum development: Incorporating AI into ecological research training to generate a workforce proficient in both fields.

Ethical and policy considerations

Ethical AI: Given that conservation decisions can have significant social and ethical implications, the ethical dimensions of using AI in ecology will likely receive more attention.

Policy and regulation: As AI plays a more significant role in ecological research, policies may be developed to guide its ethical and effective use.

Climate change and sustainability

Climate modeling: AI can help create more accurate and dynamic models to predict how climate change will impact biodiversity and ecosystems.

Sustainable practices: AI could help in monitoring and enforcing sustainable fishing, logging, and other practices that impact ecosystems.

Personalized approaches

Local ecosystem modeling: AI algorithms tailored to specific local conditions could become more common, allowing for more accurate ecological assessments on a smaller scale.

Global collaborations

Global ecosystem monitoring network: AI could facilitate the development of a global network for ecosystem monitoring, bringing in data from around the world for real-time analysis and insights, such as the International Long-Term Ecological Research.

The field is ripe for innovations that can significantly accelerate the pace of research and the effectiveness of conservation efforts. However, the future will also require addressing various challenges, from data quality to ethical considerations, to fully leverage the potential of AI in ecological research.

Conclusion

The integration of big data and AI in ecological research has opened up new avenues for understanding the complex relationships between species and their ecosystems. As these technologies continue to evolve, they promise to provide even deeper insights into these relationships, ultimately enhancing our ability to conserve and manage our planet's biodiversity.

Conflict of Interest

The author declares that he has no competing interests.

References

Buckland, S.T., Borchers, D.L., Marques, T.A., and Fewster, R.M. (2023). Wildlife population assessment: changing priorities driven by technological advances. *Journal of Statistical Theory and Practice*, 17, 20.

Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., *et al*. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, 115, E5716-E5725. https://doi.org/10.1073/pnas.1719367115

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., *et al*. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529, 484-489. https://doi.org/10.1038/nature16961

Varotsos, C.A., and Krapivin, V.F. (2017). The effects of large climate fluctuations on forests as a Big Data approach. *ArXiv*, 1-14. https://doi.org/10.48550/arXiv.1709.08257