Bioimage Analyses Using Artificial Intelligence and Future Ecological Research and Education Prospects: A Case Study of the Cichlid Fishes from Lake Malawi Using Deep Learning

Deokjin Joo[1]*[iD], Jungmin You[2][iD], Yong-Jin Won[3]*[iD]

[1]Hashed, Seoul, Korea
[2]Research Institute of Ecoscience, Ewha Womans University, Seoul, Korea
[3]Division of EcoScience, Ewha Womans University, Seoul, Korea

ABSTRACT

Ecological research relies on the interpretation of large amounts of visual data obtained from extensive wildlife surveys, but such large-scale image interpretation is costly and time-consuming. Using an artificial intelligence (AI) machine learning model, especially convolution neural networks (CNN), it is possible to streamline these manual tasks on image information and to protect wildlife and record and predict behavior. Ecological research using deep-learning-based object recognition technology includes various research purposes such as identifying, detecting, and identifying species of wild animals, and identification of the location of poachers in real-time. These advances in the application of AI technology can enable efficient management of endangered wildlife, animal detection in various environments, and real-time analysis of image information collected by unmanned aerial vehicles. Furthermore, the need for school education and social use on biodiversity and environmental issues using AI is raised. School education and citizen science related to ecological activities using AI technology can enhance environmental awareness, and strengthen more knowledge and problem-solving skills in science and research processes. Under these prospects, in this paper, we compare the results of our early 2013 study, which automatically identified African cichlid fish species using photographic data of them, with the results of reanalysis by CNN deep learning method. By using PyTorch and PyTorch Lightning frameworks, we achieve an accuracy of 82.54% and an F1-score of 0.77 with minimal programming and data preprocessing effort. This is a significant improvement over the previous our machine learning methods, which required heavy feature engineering costs and had 78% accuracy.

Keywords: Animal detection, Artificial intelligence, Citizenscience, Deep learning, Education, Machine learning

Introduction

Ecological research has recently depended on a tremendous increase in the usage of visual data. Motion-sensor cameras called "camera traps" produce hundreds of millions of images worldwide (Swanson *et al*., 2015). Uncrewed aerial vehicles (UAVs) and drones cover large areas with sub-decimeter resolution produce visual data in terabytes for a single project Kellenberger *et al*. (2020). While the camera traps can take millions of images, extracting information from these is traditionally done by humans (i.e., experts or community of volunteers), making the task time-consuming and costly that much of the valuable knowledge in these data repertoires remains untapped (Norouzzadeh *et al*., 2018). Moreover, humans suffer from the consistency problem and the trade-off issue between accuracy and computational efficiency. Deep learning is a revolutionary paradigm not only in the machine-learning field but also in ecological research. Recently, as deep-learning-based methods—particularly convolutional neural networks (CNNs)—have outperformed conventional methods in object detection, they are increasingly used in ecological research. In the future, attempts to utilize deep learning-based object recognition technology are expected to become more and more widespread in ecological research.

Ecological Research Using Artificial Intelligence (AI)

This section briefly summarizes several recent ecological research using image data processing technologies that leverage AI, such as CNNs. First, researchers can monitor various wildlife species and their numbers using a deep learning algorithm on the vast amount of data collected through camera traps. As digital camera technology advances in ecological research that requires the interpretation of large amounts of visual data obtained from extensive wildlife surveys, the number of studies using camera traps has increased to at least 125 in 2011 (Fegraus *et al*., 2011). However, when a motion sensor "camera traps" is used to collect wildlife pictures, it is time-consuming and expensive to extract images from video data manually. Norouzzadeh *et al*. (2018) collected wildlife pictures from the camera traps and then trained the images extracted from the image data through a deep neural network (DNN). The Snapshot Serengeti project is a large-scale camera trap project that has been shooting wildlife with 225 motion-sensing camera traps in Serengeti National Park, Tanzania, since 2011. It contains 1.2 million photos of 48 species of wildlife. The trained model classified 48 species, counted wildlife, and described animal behavior and companionship with young individuals with over 93.8% accuracy on the 3.2 million Snapshot Serengeti

Dataset. If only species classification is done, species identification can be automated with an accuracy of 99.3%, and a crowdsourcing team made up of volunteers showed an accuracy of 96.6% in species classification of wildlife data.

Additionally, it would take 8.4 years for 30,000 volunteers to process 3.2 million images over a 40-hour week, but an automated model saves 99.3% of the time. This improvement in efficiency underscores the importance of using AI to automate data extraction and demonstrates the potential to transform ecology, wildlife biology, zoology, conservation biology, and wildlife behavior through data science. However, this study did not deal with images containing more than one species, and the performance was low for non-classification results such as the number of wildlife in the images. In addition, learning was inevitably performed on a background including wildlife, and when applied to object image data in a new place, the accuracy was lowered. For example, an image data model trained in the United States was less accurate in identifying the same species in Canada (Tabak *et al*., 2019). Norouzzadeh *et al*. (2021) conducted research using an *object detection* model to solve the limitations of existing researches. Object detection algorithms present a bounding box on each object and the model's certainty of the object class. The algorithm can handle multiple species in one image and can be efficiently applied to new places by learning and focusing only on wildlife. Norouzzadeh *et al*. (2021) analyzed eMammal Machine Learning data, in which researchers and citizen scientists classified 450,000 camera trap images from 270 species worldwide, and Caltech Camera Traps data, in which 245,000 camera trap images from 22 species were classified in the southwestern United States. It was used as pre-training data for image detection. The trained model was tested on the Snapshot Serengeti and North America Camera Trap Images (NACTI) data of 3.7 million camera trap images from 28 species in five regions of the northern United States. This model showed 91.92% accuracy in wildlife detection and 97.7% accuracy in species classification, and 93.2% accuracy in species classification in NACTI data.

Meanwhile, AI technology is being applied to individual identification studies of wildlife. Recently in Korea, a study of automatic individual identification of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) using AI is being conducted (personal communication). In dolphin research, various external features of dolphins such as body, pectoral fin, caudal fin, dorsal fin, head shape, and wounds are used to identify individual dolphins. Since the dorsal fin is constantly exposed when a dolphin rises above the water to breathe, the shape of the wounded dorsal fin is generally used for identification. Identifying dolphins from dorsal fin data is an easy but time-consuming task. In the past, researchers directly cropped and

compared the dorsal fins, but now, using deep learning, the time taken in the recognition-identification process is significantly reduced, thereby increasing the efficiency of analysis and accuracy.

Second, AI technology is being introduced to wildlife conservation. With the number of elephants and rhinos rapidly declining in Africa, various strategies exist to combat poaching using UAVs or drones. In particular, UAVs equipped with thermal infrared cameras can be used for night surveillance, combating night-time poaching activities (Bondi *et al.*, 2018). However, it is complicated for park managers to monitor these UAV video images continuously. In addition, the more the drones are added, the burden of additional video monitoring increases. Bondi *et al.* (2018) developed an integrated cloud-based framework SPOT (Systematic POacher deTector) that detects the location of poachers in real-time using Faster R-CNN (Ren *et al.*, 2015) using thermal infrared camera images collected by UAVs or drones. The program trains the AI offline using the image frames as the training data set with labeled images. The trained model automatically detects poachers online and displays the wildlife in a live video stream. Several experiments were performed to ensure that the online detection could be completed in the cloud and a balance between local computer and remote could be maintained. The developed SPOT system will be deployed in several national parks in Africa, including Botswana.

Finally, ecological surveys collect terabytes of images in a single survey, and such large-scale surveys require a large amount of photographic interpretation. Conventional models require programming and labeled image data sets, but most data collected from ecological studies are not annotated, making CNN training difficult. Kellenberger *et al.* (2020) developed Annotation Interface for Data-driven Ecology (AIDE) (accelerating image-based ecological surveys with interactive machine learning) to solve this problem, an open-source web framework designed for image annotation for ecological surveys based on deep learning. AIDE is a web-based, open-source collaboration platform that integrates machine learning models and multi-purpose labeling tools for image annotation without writing any code. AIDE utilizes active learning (AL) to do this through a feedback loop. AL in AIDE enables the machine learning model to be iteratively trained on the latest annotations provided by the user, and when training is complete, the model is used to obtain predictions for yet unlabeled images. AL allows machine learning models to be trained with large-scale image annotation processing and potentially small training data. In addition, AIDE provides an easy-to-use, customization labeling interface and supports multiple users, database storage, and capability to the cloud or multiple systems. Through this web-based open source collaboration plat-

form, citizens can easily participate in ecological research, provide more observers and ideas, increase ecological literacy, save research costs and time, and learn about science and scientific research processes. As a result, science literacy will be strengthened, and social activities for ecosystem conservation will begin.

Ecological Education Prospects

Just as AI has begun to be used as a powerful tool for ecological research in the 21st century, it is expected that AI will become increasingly relevant to ecological education. Today, the global ecosystem is facing an ecological crisis. Air and water pollution, biodiversity reduction, resource depletion, climate change, etc., are important issues that humankind, including ecosystems, faces. Ecological education is the surest and essential solution to overcome the environmental and ecological crises and solve the conservation of the ecosystem and the existence of humankind (Capra, 1996). Ecological education is an ecosystem-based education and started when ecology began as a discipline of science and contains the principles of maintaining the Earth's ecosystem. It involves shifting the learner's cognitive base to an ecological paradigm. The scope of eco-education extends education related to ecology in the humanities or social sciences originating from ecology and ecology education. In addition, eco-education aims to promote ecological literacy, the ability to understand the natural systems that enable living things on the planet, as the basic knowledge to solve the crisis of the ecosystem (Kim, 2015).

Since the 2009 revision of the education curriculum in Korea, climate change and biodiversity have been accepted as essential topics, and related education has been expanded, increasing the need for education. Biodiversity conservation is a core value of sustainable development, leading human life to an ecological paradigm. Therefore, it is required to recognize the seriousness of the ecological crisis caused by the loss of biodiversity that humanity is facing and make efforts to conserve biodiversity (Lim & Lee, 2018). As a solution for the conservation of biodiversity, the voluntary actionability of members of society is required. Therefore, it is essential to provide opportunities for students to conclude on their own through participation in self-directed exploration and exploration of information through education (Noh, 2003).

In 2019, the Korean government announced the 'Artificial Intelligence National Strategy' (Ministry of Science and ICT, 2019). Furthermore, the 'Artificial Intelligence Talent Nurturing' strategy includes an education plan to strengthen AI education competency. In addition, 'Artificial Intelligence Education' will be introduced through the 2022 revised curriculum, which will be effective from 2025 (Ministry of Education, 2020). AI education consists

of 'programming' education that develops creativity by recognizing and solving problems and demonstrating it on a computer, 'AI principles' and 'AI utilization' education as basic communication for the future, and 'AI ethics' as critical thinking skills. However, despite the policy support related to AI education, the relevant curriculum structure and environment creation are still insufficient (Lee, 2020).

Ecological education integrated with AI can allow students to creatively solve the social problem of ecosystem destruction based on the principles of AI through educational programs that can develop AI literacy. Data recognition technology, which has entered the most stable stage among various AI models, can be used for future ecological education. For example, image classification/recognition technology can identify living things that appear in photos or videos and determine the species. Through an educational program that learns and classifies images of endangered wildlife and ecosystem-disrupting species through machine learning, it is possible to know the principles of machine learning and creatively use them in activities to conserve the ecosystem. In addition, by introducing ecological research using AI, it is possible to guide students to various careers and occupations related to ecology. If the ecological education program integrated with AI is applied, students' ecological literacy will increase, contributing to ecosystem conservation and environmentally sustainable society. In addition, as explained above with the case of AIDE, citizen science, which connects scientists and citizens by building an ecological research platform using AI, is expected as a powerful method for the education of AI and biodiversity and ecological research. Through this, we can accumulate more observation data and reduce the cost of scientific research.

Identification of Cichlid Fishes from Lake Malawi Using Deep Learning

The information that can be extracted from the bio-image data of wild animals and the purpose of use are very diverse. The representative case is species identification. Ecological research works which use AI to obtain useful information from bio-images are now moving into the application stage. At this point, we would like to look back on our 2013 past study. At that time, we developed a pipeline that extracts features from photographic data of African cichlid fishes through a computer vision method which automatically classifies them through a Support Vector Machine and Random Forests. The program we developed was able to classify photographic images of 594 cichlids belonging to 12 different classes (species and sex) with an average accuracy of 78%.

In this paper, we intended to compare the results of the early 2013 study to that of more modern deep learning methods. We reanalyze the 594 samples with the CNN deep learning method. For this purpose, we had to develop new programs and tutorials that incorporate newly developed deep learning methods since 2013. We hope that our experience and results will be helpful to researchers who want to analyze large amounts of ecological image data using AI.

For reproducibility, we made both the code and the 594 samples publicly available on the Github repository (https://github.com/forcecore/ghoti-2021). The full raw data will be also made available on Zenodo in the near future, after a thorough review. The URL will be published on the Github repository.

Trial and Error with CameraTraps Project

Microsoft is hosting AI for Earth (Microsoft AI, 2021) project, and CameraTraps (Microsoft, 2021) is one of the subprojects. The repository contains many valuable tools for detecting images collected from motion-triggered camera traps. In this section, we record an attempt to adapt an open-source project to our needs. Such a task may or may not succeed, as most open-source AI projects are unstable and usually become un-runnable in a couple of years. Eventually, we failed to meet our needs. However, our failure may provide a useful debugging process that can be valuable to the researchers. We briefly outline what we did here.

Following the CameraTraps Tutorial

CameraTraps so far seems to depend on Tensorflow, which is one of the most popular deep learning frameworks and graphics processing unit (GPU) may be used for faster training of the neural networks.

The first step is to set up a new "virtual environment" for running Python programs. We do this as each Python project may require different package (library) requirements, which can induce "dependency hell" when not careful. Programmers bypass this problem by isolating projects with virtual environments, one environment for each project. After the creation of the virtual environment, the required packages may be installed. The requirements are usually stated in requirements.txt by the project maintainers.

To feed our dataset, we need to figure out the data processing pipeline the project has. We need both the project documentation on the data format and the example dataset. It is common for deep learning research projects to fail over time for two reasons. One is, the researchers move onto the next phases of their research, and the previous programs get out of sync with the later stages. The other reason is that, even when the project

stays consistent, the underlying deep learning framework is developed rapidly, and the project becomes incompatible. In our case, we had both of the problems. Moreover, we later found out that the document of the project was out-of-sync with the code. We reported this issue to the authors and confirmed that the repository is currently specific to the original authors' needs and the tutorials are outdated. In this case, we decided to write our own code from scratch, instead of trying to fix the problem.

Starting from a PyTorch Example

We highly recommend PyTorch over Tensorflow in your research phase. Most state-of-the-art deep learning algorithms are implemented in PyTorch as it is much easier to implement ideas than Tensorflow. PyTorch has an extensive tutorial so that researchers may cherry-pick them as a starting point. Since image classification is our task, we pick the image classification task tutorial and employ a pre-trained ResNet50 as the neural network architecture of choice, which is one of the most well-known for the image classification task. Then we apply a transfer-learning technique, which is to replace the final fully connected (FC) layer of the pre-trained neural network. This reduces the number of required training samples. To feed the neural network with data for training requires a matching dataset class, ImageFolder class, in PyTorch.

Experimental Results

To evaluate the performance of the deep-learning-based classifier, we ran ten randomized split/train/test runs. The training and test sets had 535 and 59 samples, respectively. Each training run halted in roughly 300 epochs. We got the results as shown in Table 1.

Table 1. Accuracies and F1-scores on the test set of the 10 randomized classification trials

Trial no.	Accuracy	F1-score
1	0.76271	0.7609
2	0.84746	0.8596
3	0.86441	0.6936
4	0.88136	0.77163
5	0.77966	0.76917
6	0.84746	0.80512
7	0.76271	0.6706
8	0.84746	0.82617
9	0.83051	0.73485
10	0.83051	0.81271
Average	0.8254±0.0423	0.770435±0.059

In the previous study by Joo *et al*. (2013), they report accuracies of 0.6649 and 0.7562 for 48-feature-classifier and 82-feature-classifier, respectively. Our results in Table 1 indicate the superiority of deep learning over the classifier based on hand-crafted features. Moreover, we expect further improvement if more deep learning techniques are applied and the hyperparameters are tuned carefully.

Discussion

The African Lake Malawi cichlid fish species, which have undergone radiative evolution in a very short evolutionary time, show a remarkable diversity in colors and various striped patterns in the appearance. From a biological point of view, biological discrimination against closely related species, such as cichlid fishes, which have diversified relatively recently and still have interspecific gene flow, cast a challenge to automatic classification methods. We adopted recently progressed AI methods to solve the difficulty and connected them for seamless automatic analysis of the cichlid fish photographic data that we used in 2013 (Joo *et al*., 2013). In this work, we employed a transfer-learning technique to replace and fine-tune the final FC layer of a pre-trained neural network. Note that the embeddings (the extracted features) of the pre-trained neural network are not real-world and may not work on some datasets. The reason is, "images found in current image datasets are not drawn from the same distribution as the set of all possible images naturally occurring in the wild" (Chang & Lipson, 2019), and we advise the readers to try multiple neural network architectures which are pre-trained on different datasets so that the reader may find better performing neural nets. Please refer to the supplementary material appended to this paper for more detailed information about our methods and image data of cichlid fishes. The supplementary material includes Linux commands and program codes to help the readers understand the deep learning research process.

One of the most important differences between the previous approach and the new deep learning method is that the direction of the feature extraction. The deep learning approach found the features in bottom-up fashion from the training samples, as opposed to the previous approach where the human engineers have somewhat arbitrarily applied a small set of features which works, by trial and error in top-down manner. To make matters worse, these top-down features had limited expressivity, as they could only be defined mathematically through a programming language.

This in turn leads to another important implication. To make the bottom-up approach feasible, the researchers must first collect sufficient high-quality data. This may require adjusting the experiment design and/or process. We suggest that this data-driven approach with AI tools

will lead to deeper understanding of the nature. Sutton (2019) argues that with current path of technology development, human feature engineering efforts are of little help and more computing resources and data are the keys of the improvements. In this spirit, we propose to employ more AI/data-driven approaches to overcome our potential narrow understanding of the nature.

In future research, the Cichlid Fishes image identification model could improve that image data set extended by reconstruction, convolution of feature extraction by neural network (ResNet50 and AlexNet, VGG19, GoogLeNet, etc.), deep learning algorithm comparison (CNNs, R-CNN, and YOLO), activation function comparison, etc.

Conflict of Interest

Acknowledgments

References

Bondi, E., Fang, F., Hamilton, M., Kar, D., Dmello, D., Choi, J., et al. (2018). SPOT poachers in action: augmenting conservation drones with automatic detection in near real time. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32, 7741-7746.

Capra, F. (1996). *The Web of Life: A New Synthesis of Mind and Matter*. London: Harper Collins.

Chang, O., and Lipson, H. (2019). *Seven Myths in Machine Learning Research*. Retrieved December 9, 2021 from http://arxiv.org/abs/1902.06789.

Fegraus, E.H., Lin, K., Ahumada, J.A., Baru, C., Chandra, S., and Youn, C. (2011). Data acquisition and management software for camera trap data: a case study from the TEAM Network. *Ecological Informatics*, 6, 345-353.

Joo, D., Kwan, Y.S., Song, J., Pinho, C., Hey, J., and Won, Y.J. (2013). Identification of cichlid fishes from Lake Malawi using computer vision. *PloS One*, 8, e77686.

Kellenberger, B., Tuia, D., and Morris, D. (2020). AIDE: accelerating image-based ecological surveys with interactive machine learning. *Methods in Ecology and Evolution*, 11, 1716-1727.

Kim, K.D. (2015). Contents and prospects of ecological education. *Journal of Holistic Convergence Education*, 19, 1-19.

Lee, E.K. (2020). A comparative analysis of contents related to artificial intelligence in national and international K-12 curriculum. *The Journal of Korean Association of Computer Education*, 23, 37-44.

Lim, H.M., and Lee, S.W. (2018). A study on pre-service elementary school teachers' knowledge, awareness and attitude of the biodiversity conservation. *Journal of Korean Practical Arts Education*, 31, 19-44.

Microsoft. (2021). *CameraTraps*. Retrieved July 9, 2021 from https://github.com/microsoft/CameraTraps.

Microsoft AI. (2021). *AI for Earth*. Retrieved July 9, 2021 from https://www.microsoft.com/en-us/ai/ai-for-earth.

Ministry of Education. (2020). *Comprehensive Plan for Convergence Education that Changes the Learning Paradigm (20-24)*. Sejong: Ministry of Education.

Ministry of Science and ICT. (2019). *National Strategy for Artificial Intelligence*. Sejong: Ministry of Science and ICT.

Noh, H.J. (2003). Intrinsic value in biodiversity and moral education. *Journal of Korean Philosophical Society*, 86, 115-137.

Norouzzadeh, M.S., Morris, D., Beery, S., Joshi, N., Jojic, N., and Clune, J. (2021). A deep active learning system for species identification and counting in camera trap images. *Methods in Ecology and Evolution*, 12, 150-161.

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.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster R-CNN: towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91-99.

Sutton, R. (2019). *The Bitter Lesson*. Retrieved December 9, 2021 from http://www.incompleteideas.net/IncIdeas/Bitter-Lesson.html.

Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., and Packer, C. (2015). Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data*, 2, 150026.

Tabak, M.A., Norouzzadeh, M.S., Wolfson, D.W., Sweeney, S.J., Vercauteren, K.C., Snow, N.P., et al. (2019). Machine learning to classify animal species in camera trap images: applications in ecology. *Methods in Ecology and Evolution*, 10, 585-590.