Species Recognition Technology Based on Machine Learning

Qiming Wu[1]

[1] International School, Beijing University of Posts and Telecommunications, Beijing, China
wuqiming@bupt.edu.cn

Abstract. Animal Recognition Technology has an important role to play in identifying and conserving species. As society continues to progress and people's standard of living improves, it intensifies environmental pollution and ecological damage, which inevitably increases the risk of species extinction, thus increasing the urgent need for species protection. In the face of this challenge, researchers have continued to propose, improve, and refine animal identification techniques to achieve more accurate, faster, and simpler species identification techniques. The purpose of this paper is to explore and summarize existing species identification techniques and provide reference materials for future research. This paper will focus on the following areas: First is the selection of datasets, followed by algorithms for animal detection and recognition techniques, including traditional image processing methods and the latest deep learning techniques. The accuracy and performance of these models will then be evaluated to see how they perform in real-world applications. Finally, the model selection strategy will be explored. This paper aims to provide a detailed reference for researchers in the field of animal identification technology, to help subsequent researchers better understand the strengths and weaknesses of existing techniques, and to provide reference and inspiration for future research. At the same time, the author will present ideas and suggestions to contribute to the technological development of the field to promote the conservation of species and the realization of ecological balance.

Keywords: Animal Identification, Machine Learning, Deep Learning, Convolutional Neural Network.

1 Introduction

Accurate identification is fundamental to all aspects of taxonomic research and is an important part of the workflow of research, including medical, ecological, and evolutionary studies [1]. The continuous advancement of technology has inevitably caused some degree of damage to the environment. Out of the over 120,000 species monitored on the IUCN Red List of Threatened Species, approximately 17,000 are classified as "data deficient" [2]. The environment has become hostile resulting in declining animal populations. Higher requirements have been placed on identification techniques. Accurate identification is important for the discovery and conservation of new species.

Y. Wang (ed.), *Proceedings of the 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, Advances in Intelligent Systems Research 185,
https://doi.org/10.2991/978-94-6463-512-6_50

Technologies such as machine learning have an important practical role to play in species conservation. For example, Ana Cristina Mosebo Fernandes et al. used four machine-learning algorithms to locate current locations of wildlife habitat and predict suitable future locations to which wildlife may relocate based on the impacts of climate change and based on time frames of scientifically supported estimates of temperature increases [3]. J. Jenrette et al. Using machine learning and automation to identify and categorize sharks from images and video to fill information gaps in shark stock assessments [4]. Tobias Jensen et al. used machine learning and remote sensing data to detect invasive plants, particularly kudzu, in response to the economic and environmental problems posed by invasive plants worldwide. The superiority of random forests, neural networks, and support vector machines was determined by testing five different algorithms. [5]. Endangered animals can only be protected by choosing the right convolutional neural network structure.

With the rapid advancement of technology, the application of machine learning in various fields is gradually revealing its immense potential. In the fields of ecology and conservation science, traditional methods, while effective, often have limitations in data processing and model optimization. Integrating machine learning into ecological workflows can optimize inputs to ecological models and promote the development of integrated hybrid modeling tools [6]. The approach will necessitate close interdisciplinary collaboration to ensure qualities of new methods and to train a new generation of data scientists specializing in ecology and conservation [6].

The author has analyzed the achievements of species identification techniques based on different structures. This paper will provide important research material for future researchers of species identification techniques and contribute to the conservation of species. An in-depth analysis of species identification techniques with different structures allows for a better understanding of their strengths and weaknesses and provides guidance for the improvement and application of un techniques. This paper can contribute to the progress of species conservation and improve the efficiency and accuracy of species identification techniques in conserving biodiversity.

2 Dataset

Driven by contemporary technology, animal feature recognition technology plays an increasingly important role in the fields of biology, ecology, and environmental protection. By analyzing data on animals' external features, sounds, and behaviors, researchers can better understand their ecological habits, population dynamics, and ecosystem health, and then take effective conservation measures. The dataset used for training is indispensable to better train the model and apply it in practice. Technological advances have led to a variety of ways to collect data, as well as more in-depth research on data sets. The researchers constructed a large dataset of rare animals. As shown in Table 1.

2.1 Introduction to the dataset

The panda dataset is derived from 25 pandas from Chengdu Research Base of Giant Panda Breeding and CCRCGP Dujiangyan Base, totaling about 65,000 giant panda facial images, with about 4,300 images for each giant panda [7]. This dataset is divided into validation, training and calibration sets in the ratio of 3:6:1, aiming to provide sufficient data support for model training, validation, and tuning [7].

The Risso's dolphin dataset was obtained from boat surveys and camera photography and contains approximately 1,000 photographs [8]. The relatively small number of dolphins and the difficulty of collecting them makes the dolphin dataset relatively small, but the dolphin dataset covers a wide range of different dolphin species and provides feasibility for the study [8].

The Animal's set was captured by an unmanned aerial vehicle (UAV) and contains approximately 3,400 high-quality photographs [9]. These photos cover five unique categories, namely mammals, birds, reptiles, amphibians, and fish, providing a rich sample for a multi-category classification task [9].

The birdsong dataset uses the Cornell Bird Challenge (CBC) 2020 dataset and its extensions [10]. This dataset includes more than 260 bird species with approximately 10 to 1800 audio samples per species [10]. Ultimately, each audio was segmented into 7-8 second segments totaling approximately 15,000 samples [10].

Table 1. Description of data set sources

Data set name	Data set size	Number of species	Type of data set	Acquisition method
Panda [7]	65000	1	photos	Camera or cell phone's zoom function to get a picture of the panda's face
Dolphins [8]	1000	28	photos	Vessel surveys and use of cameras to take photographs
Animals [9]	3400	5	photos	Photography by drone
irdsong [10]	15000	264	audio frequency	"Cornell Bird Challenge" (CBC) 2020 dataset22 and its extensions

2.2 Analysis and comparison of data sets

The advantage of the panda dataset is that it contains many photos of the panda's face from different angles. The multi-angle shots of the panda dataset can help the model learn and extract the panda's features in a more comprehensive way. In addition, the size of the dataset is quite large with tens of thousands of images, providing ample samples for model training. However, the disadvantage of the panda dataset is its restricted data source. The data only came from 25 individual pandas from Chengdu Research Base of Giant Panda Breeding and CCRCGP Dujiangyan Base. This limitation may lead to a decrease in the model's performance when processing images of pandas from other regions or other populations, as there is a degree of variation between individual pandas.

In contrast, the Risso's Dolphins dataset contains images of multiple species of dolphins, which provides an opportunity to study diversity. However, the relatively small size of this dataset may have limited model performance. In addition, the number of samples is uneven across dolphin species, which may result in poorer model performance on a few categories.

The Animal's set dataset is of a moderate size, covering five main animal classes: mammals, birds, reptiles, amphibians, and fish. Although the Animal's dataset is broadly categorized, it lacks more specific categorical subdivisions, potentially impacting the model's capability to distinguish between various species of animals, particularly during feature extraction. Accuracy would have been high if the dataset had been selected similarly for the five major species of animals, but it was off topic. The author suggests expanding the number of datasets and again reviewing and improving the model.

The Birdsongs dataset is derived from the official website and has a large amount of audio data and a rich variety of species. These audio data are of high quality and are well suited for conducting research on the classification and identification of birds based on their calls. The professionalism and richness of the bird song dataset sources provide a reliable basis for research.

3 Animal recognition technology architecture

3.1 Introduction to the model

VGG for panda set

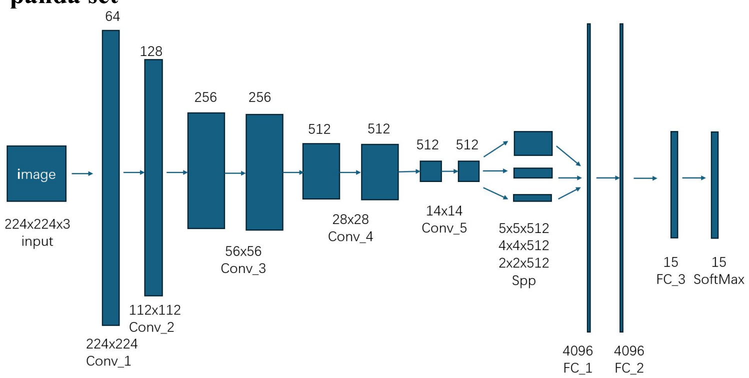


Fig. 1. VGG neural network (Picture credit : Original)

Fig. 1 shows the VGG convolutional neural network. VGGNet was enhanced by a convolutional neural network. It was designed by Jin Hou et al. and comprises three fully connected layers, a SoftMax layer, and five convolutional modules [7]. Convolution module comprises a convolutional layer and a pooling layer. These layers work together to extract image features and compress the input image, thereby simplifying network computations [7]. The training dataset includes multiple photos of panda faces taken from various angles to enhance the model's robustness and the ability to handle diverse variations. The model employs a Spatial Pyramid Pooling layer in place of the last pooling layer. To mitigate the Exploding Gradient Problem, the network layers are reduced to 11, and the Dropout layer is substituted with a Batch Normalization layer [7].

Conv-ReLu-MaxPool for Risso's Dolphins set

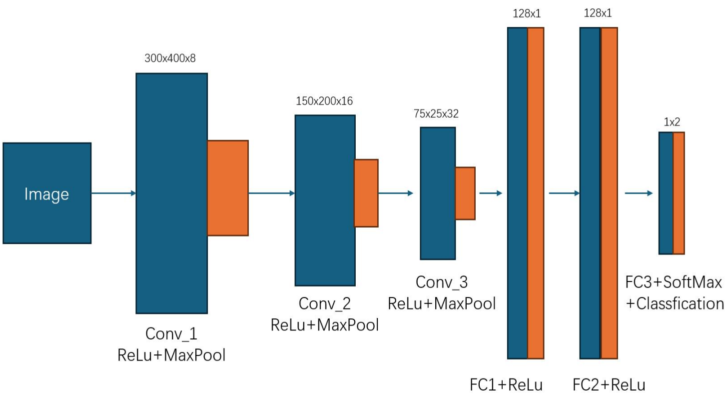


Fig. 2. Conv-ReLu-MaxPool neural network (Picture credit : Original)

Fig. 2 shows the Conv-ReLu-MaxPool neural network. Maglietta, R. et al. designed a CNN architecture for recognizing fin images. The CNN structure uses a combination of three different types of layers: the Conv-ReLu-MaxPool. In each combination, the number of filters learned on the convolutional layer increases from 8 to 16 to 32 [8]. After learning the features, the architecture of the CNN performs the classification through two fully connected layers (FC + ReLu) and the last layer provides the classification output using SoftMax function [8]. Aimed at identifying known and unknown dolphins rather than explicitly identifying known and tagged dolphins by name, Maglietta, R. et al. proposed a new strategy called NNPool [8].

CNN-SVM Deep Learning Model

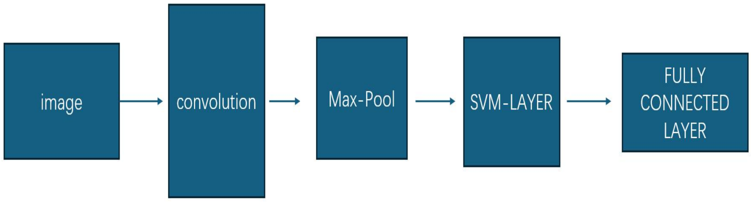


Fig. 3. CNN-SVM neural network (Picture credit : Original)

Fig. 3 shows the CNN-SVM neural network. Kukreja, V. et al. designed a CNN as a feature extractor for multi-class classification work [9]. Cable News Network became a comparative choice for detecting and recognizing animal species based on pictures due to its effectiveness in autonomously recognizing complex visual patterns [9]. To enhance the discriminative power of the model, Kukreja, V. et al. added an SVM classifier to the CNN framework. Kukreja, V. et al. established deep learning techniques to enhance parameter optimization. The core of the Kukreja, V. et al. approach is to divide the dataset into sections for thorough testing and validation [9].

Hybrid models for Birdsong

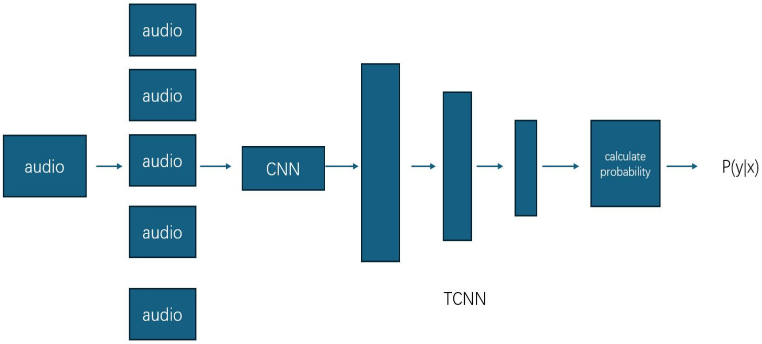
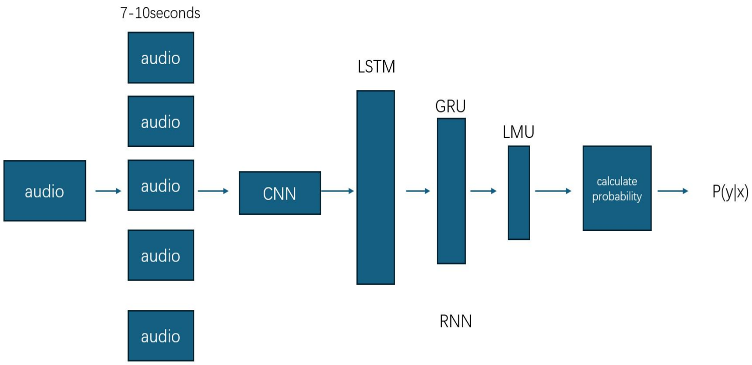


Fig. 4. TCNN model (Picture credit :



Original)

Fig. 5. RNN model (Picture credit : Original)

Fig. 4 and fig. 5 shows the Hybrid models for Birdsong. For categorizing birds based on their calls, Gupta, G. et al. first clipped the raw audio to 7 seconds long (32 kHz) [10]. A hybrid approach was used for the modeling, including both stand-alone and temporal models [10]. For the standalone model, an ImageNet-based model was used as a classifier with 3 channels (RGB) and 100 neurons in the output layer [10]. For the temporal model, a sliding window mechanism was used to process the input and convert it to a Mel spectral image [10]. The representation model employs three CNN architectures of different lengths (CNN1, CNN2 and CNN3) [10]. Each model ends with an adaptive averaging pool (AAvgPool) layer [10]. In the temporal module, three different networks (TCNN1, TCNN2, and TCNN3) are used, and each model ends with an adaptive averaging pool (AAvgPool) layer [10].

3.2 Analysis of results

Table 2. Experimental results

Data set name	Structure	Experimental ratio	Number of species	Accuracy	Sensitivity	Specifics
Panda [7]	VGGNet	6:3:1	1	0.95	—	—
Dolphins [8]	RUSPool	—	28	0.78	0.58	0.81
Dolphins [8]	NNPool	—	28	0.87	0.70	0.90
Animals [9]	CNN—SVM	—	5	0.96	—	—
Birdsong [10]	hybrid model	8:1:1	100	0.67	—	—

As shown in Table 2. In training the Panda dataset, the researchers used a VGG structure and an experimental ratio of 6:3:1, which is a very reasonable experimental ratio. The multiple camera angles and the large dataset make it possible to preserve the characteristics of individual pandas. The final model achieved 95% accuracy. The limitation is that the dataset only contains data from 25 pandas, and other panda datasets can be collected in the future to further improve the model.

For the discrimination of the five major classes of animals, Kaur, A. at al found that the model showed surprising accuracy [9]. Even though different types of animals have their own distinctive characteristics, the model is still able to effectively extract the characteristics of the five major species categories. Kaur, A. at al designed Conv-ReLu-MaxPool for Risso's Dolphins set provides a reference for identifying similar species together in the future. Given the diversity of species types, it is highly recommended that additional datasets be sought, and it is considered desirable to have dataset sizes of 10,000 or more. Models need to be tested with large data sets from a wide selection of animals, and models need to be refined based on the results of the experiments.

In the dolphin study, Maglietta, R. et al. designed a model that performed a summary of the data after each feature extraction and then passed it on to the next session and improved the output of the model [8]. The results of the model designed by Maglietta, R. et al. showed that the NNPool model was higher in accuracy than the RUSPool model. This research has made an important contribution to dolphin conservation efforts.

Gupta, G at al. analyzed existing convolutional neural networks by finally integrating and improving different networks [10]. For the classification of birds, although the accuracy of the model was only 0.67, this result is not easy considering that the classification task involved 100 bird species. Moreover, the researchers' model is based on audio for classification, adding to the difficulty of recognition. If the model can be refined in the future, it will provide a reliable basis for judging rare birds and protecting them.

4 Conclusion

The author has analyzed and summarized the research on species identification techniques based on recent years. For different research objects need to choose the appropriate processing which includes incoming data sets, extraction of features, and analysis of results. With the emergence of a variety of ways to collect data, there will be a larger, more diverse, and more detailed dataset established to provide a solid dataset for model training, thus improving the accuracy and generalization ability of species identification models. With the continuous evolution of deep learning and machine learning algorithms, the emergence of more efficient and flexible model structures and algorithms can be foreseen, which can better adapt to different kinds of species recognition tasks with different data types and improve the performance and applicability of models. Secondly, technological advances will promote the application of species identification technology in a wider range of fields, such as ecological research, environmental monitoring, and wildlife protection. In addition, the fusion application of species identification technology with other related technologies and data sources, such as combining remote sensing technology, bioacoustics, environmental sensors, etc., realizes the comprehensive analysis and cross-validation of multi-source data to improve the accuracy and reliability of species identification. Interdisciplinary cooperation and knowledge sharing will also promote the development of species identification technology and make greater contributions to biodiversity conservation and sustainable development. In conclusion, in the future, species identification technology will be further developed and improved in many aspects such as data, algorithms, applications, etc., which will provide more effective tools and methods for human beings to recognize and protect the earth's biodiversity.

References

1. Wäldchen, J. Mäder, P.: Machine learning for image based species identification. Methods EcolEvol.9: 2216–2225 (2018).
2. Committee, T. I. R. L.: The IUCN Red List of Threatened Species - Strategic Plan 2017-2020. IUCN (2017).
3. Mosebo Fernandes, A. C., et al.: Machine learning for conservation planning in a changing climate. Sustainability 7657 (2020).

4. Jenrette, J., Liu, Z. Y.-C., at al.: Shark detection and classification with machine learning, Ecological Informatics 69,101673 (2022).
5. Jensen, T., Seerup Hass, F., at al.: Employing Machine Learning for Detection of Invasive Species using Sentinel-2 and AVIRIS Data: The Case of Kudzu in the United States. MDPI 12, 3544 (2020).
6. Tui, D., Kellenberger, B., Beery, S.: Perspectives in machine learning for wildlife conservation. Nat Commune13, 792 (2022).
7. Jin, Hou., Yuxin, He., at al.: Identification of animal individuals using deep learning: A case study of giant panda, Biological Conservation, ScienceDirect Volume 242,108414 (2020),
8. Maglietta, R.: Convolutional Neural Networks for Risso's Dolphins Identification, IEEE vol. 8, pp. 80195-80206 (2020).
9. Kaur, A., Kukreja, V., Chattopadhyay, S., Verma, A., Sharma, R.: Empowering Wildlife Conservation with a Fused CNN-SVM Deep Learning Model for Multi-Classification Using Drone-Based Imagery2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). IEEE, 2: 1-4 (2024).
10. Gupta, G., Kshirsagar, M., Zhong, M. et al.: Comparing recurrent convolutional neural networks for large scale bird species classification. Sci Rep 11, 17085 (2021)