Ecological Informatics 92 (2025) 103467

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Bridging AI and wildlife conservation: Classifying wildcats from genetically validated images

Georgia Fargetta [a,*], Stefano Anile [b,c,d], Luigi Priano [a], Massimo Orazio Spata [a], Alessandro Ortis [a], Sebastiano Battiato [a]

[a] Department of Mathematics and Computer Science, University of Catania, Catania, Italy
[b] Department of Biological, Chemical and Pharmaceutical Sciences and Technologies (STEBICEF) - Section of Botany, Anthropology, Zoology, University of Palermo, Palermo, Italy
[c] National Biodiversity Future Center (NBFC), Palermo, Italy
[d] Wildlife Initiative Italy, Department of Research and Conservation, Verona, Italy

ARTICLE INFO

Dataset link: https://www.dmi.unict.it/ortis/wildcat/

Keywords:
Artificial intelligence
Computer vision
Wildcat classification
Mobile application
Usability study
Genetic validation

ABSTRACT

This study presents a multidisciplinary approach for classifying three cat categories, European wildcats (*Felis silvestris*, WC), hybrids (*F. silvestris* × *catus*, HY), and domestic cats (*F. catus*, DC), using convolutional neural networks (CNNs) on real-world images collected from camera traps, roadkills, and research teams. The initial dataset included 3,392 images (WC = 1,738; HY = 461; DC = 1,193) and was filtered using YOLOv12 to retain clearly visible cats, resulting in 2,518 images (WC = 1,201; HY = 461; DC = 856). We achieved the highest CNN's accuracy 86% on the balanced dataset and 71% on the genetically validated dataset. When trained on the larger, unbalanced dataset, accuracy on the genetic test increased to 81%, indicating that enlarging datasets may effectively compensate for class imbalance. An offline-capable mobile app was developed for field use, supporting local photo capturing and usability feedback via an included and structured questionnaire for future analysis. This work is the result of a synergistic collaboration among computer scientists, wildcat researchers, and citizen scientists. The proposed model offers a consistent and reproducible field-ready tool for European wildcat conservation.

1. Introduction and motivation

The European wildcat (*Felis silvestris*, hereafter wildcat) faces one of the most insidious threats to its survival: hybridization with the domestic cat (*Felis catus*). While the current biodiversity crisis calls for urgent action to mitigate species loss (Jaureguiberry et al., 2022; Ceballos and Ehrlich, 2023), the case of the wildcat exemplifies how genetic introgression can accelerate this decline. Hybridization, often underestimated compared to more visible threats such as habitat loss or poaching (Simberloff, 1996; Randi, 2008), has already driven the Scottish population of wildcats to functional extinction, with individuals now largely represented by hybrids (Senn et al., 2019; Nieto-Blázquez et al., 2022). This situation highlights the urgency of addressing hybridization as a conservation priority for wildcats across Europe. Hybridization between wildcats and domestic cats is not confined to Scotland but is widespread across Europe, with hybrids detected in Portugal, Spain, France, Germany (Matias et al., 2022), Switzerland Nussberger et al. (2018), and Italy (Gaudiano et al., 2022; Savazza et al., 2023; Anile et al., 2025). The wildcat is a small feline weighing 3.7–4.9 kg (Macdonald and Loveridge, 2010) present in several European countries and is currently classified as "least concern" by the IUCN (Gerngross et al., 2022), although hybridization represents a major conservation challenge (Nussberger et al., 2023). Genetic analyses of cat genomes are invaluable for detecting hybridization events, even those that occurred many generations ago (Mattucci et al., 2019). Yet obtaining wildcat samples (i.e. biological material from which extracting DNA) is not a trivial task because wildcat population density are typically low (Anile et al., 2014), hence non-invasive sampling methods like hair-trapping and scat collection can results in samples from few individuals (Velli et al., 2015; Viviani et al., 2024) when not applied at large scale (Anile et al., 2019a). Before the advent of genetic tools, researchers relied on anatomical measurements like the cranial indices (Lesch et al., 2022) and the intestinal length (Devillard et al., 2013) to distinguish between wildcats, hybrids, and domestic cats until (Ragni and Possenti, 1996) first described the coat markings of wildcats by identifying four diagnostic body regions (i.e. *Gularis*, *Occipitalis-Cervicalis*, *Dorsalis*, *Caudalis*) which were effective for distinguishing between wildcats and domestic cats. Later, Kitchener and colleagues expanded this model by

* Corresponding author.
E-mail address: georgia.fargetta@unict.it (G. Fargetta).

https://doi.org/10.1016/j.ecoinf.2025.103467
Received 22 July 2025; Received in revised form 5 September 2025; Accepted 6 October 2025
Available online 17 October 2025
1574-9541/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

including more discriminants morphological features (e.g. tail shape, band distinctiveness), spots over the flank and hindquarter (Kitchener et al., 2005) which were also effective for classifying hybrids. Interestingly, studies combining genetic and morphological analyses have shown encouraging concordance (Kilshaw et al., 2010; Ballesteros-Duperón et al., 2015; Senn et al., 2019) although the reliability of any morphological classification heavily depends on both the observer's expertise and the quality of the images (Eichholzer, 2010), making reproducibility challenging.

Artificial Intelligence is increasingly applied in wildlife research, offering quick, robust, and scalable solutions to several time-consuming processes, like labeling and understanding images or videos from camera-trapping surveys at both the species (Austen et al., 2016; Akbar et al., 2025; Wang et al., 2025) and individual level (Clapham et al., 2020), while also counting the number of individuals for group-living animals (Norouzzadeh et al., 2018) and describing their behaviors (Schindler and Steinhage, 2021). Similar classification-based approaches have also been applied to other species, highlighting the versatility of AI methods in wildlife monitoring. For instance, convolutional neural networks and vision transformers have been used for plankton classification (Nanni et al., 2025), automated pollen identification for monofloral honey classification (González et al., 2025), and consistent deep learning models for plant leaf disease identification from smartphone images (Pandey and Jain, 2022). Convolutional neural networks are capable of learning not only general evident anomalies in the coat pattern of the wildcat (i.e. large patch of white area over the body), but also subtle morphometric features (i.e. cranial measurements) which can contribute to performing this classification task robustly (Nussberger et al., 2023). Here, we assembled a dataset of wildcats, hybrids, and domestic cats for training CNNs for classifying images of cats from real-world scenarios (Fargetta et al., 2024). Furthermore, we evaluated our *a priori* morphological classification, as provided by expert researchers, against a dataset of cat images complemented with the genetic analysis of the depicted cats. Our innovative classification model, further optimized for mobile platforms and remote areas (i.e. Internet connection is not necessary to perform a classification once the app is installed on a mobile device), reached an average accuracy of 81% across the three classes (WC = 83%, HY = 72%, DC = 87%), making it a valuable and novel tool for detecting hybridization in cat images. Lastly, the current limitations and the future improvements which we envisioned of our model are discussed in order to obtain better classification performances in real-world scenarios (i.e. camera-traps, road-kills or images taken by photographers).

2. Related works on classification task

Classification is a fundamental task in computer vision and has been extensively explored across a wide range of datasets and domains, often serving as a benchmark for evaluating model performance (Ragusa et al., 2016; Battiato et al., 2020; Zhao et al., 2021; Wang et al., 2025). In recent years, research has increasingly focused on developing efficient and lightweight models suitable for deployment in resource-constrained environments.

The MobileNet family. The MobileNet series has been central to this progress, with each version introducing architectural refinements to balance accuracy and computational efficiency. MobileNetV2 (Sandler et al., 2018) has proven particularly effective, achieving state-of-the-art results in benchmark classification tasks and outperforming its predecessors thanks to depthwise separable convolutions and optimized residual blocks (Lin, 2024). MobileNetV3 (Howard et al., 2019), designed through neural architecture search and incorporating squeeze-and-excitation modules, further improves efficiency. However, its larger variant has shown overfitting issues on small datasets, indicating that increased complexity is not always beneficial (Lin, 2024).

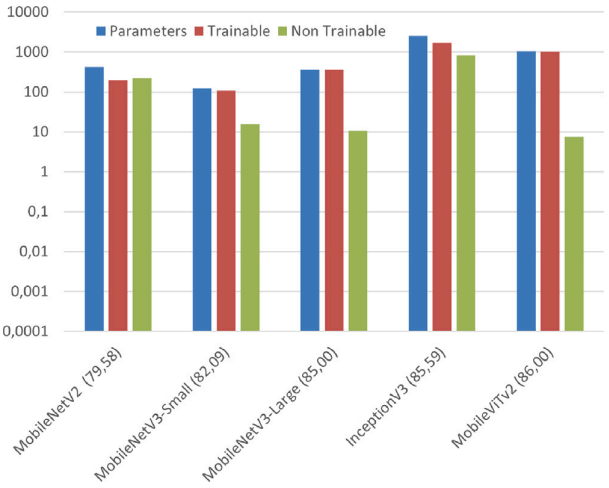


Fig. 1. Comparison of total, trainable, and non-trainable parameters across selected CNN architectures. The y-axis represents the number of parameters in thousands, plotted on a logarithmic scale to highlight the large differences in model complexity, with MobileNet variants requiring fewer parameters compared to InceptionV3 and MobileViTv2.

Alternative CNN architectures. InceptionV3 (Szegedy et al., 2016) offers a different approach, using multiple convolution filters of varying sizes to capture multi-scale features while reducing computational costs. In cat breed classification tasks, InceptionV3 reached competitive accuracy (79%–80%), but MobileNetV2 performed slightly better (82%). Ramadhan and Setiawan (2023), highlighting the advantage of models tailored to smaller datasets and mobile scenarios.

CNN–transformer hybrids. Building upon CNN efficiency, MobileViTv2 (Mehta and Rastegari, 2022) integrates convolutional layers with vision transformers to capture both local and global context. By combining separable convolutions with self-attention, it achieves long-range dependency modeling with reduced overhead. In plant disease classification, it outperformed EfficientNet-B7 and ensemble baselines, demonstrating high accuracy and strong generalization across conditions (Kheir et al., 2025). These results suggest that MobileViTv2 not only inherits the deployment advantages of MobileNetV2 but also introduces transformer-based enhancements critical for modern recognition tasks.

Applications in animal classification. The classification of animal species and breeds has long been a subject of scientific interest, with CNNs consistently delivering strong performance. For dogs, traditional approaches based on Local Binary Patterns and Histogram of Oriented Gradients have been surpassed by CNN-based transfer learning. ResNet50 has shown high accuracy across multiple datasets (Varghese and Remya, 2021), while VGGNet obtained error rates as low as 7.3% in animal classification (Simonyan and Zisserman, 2015). In cats, early studies used VGG models but were limited to only five breed classes (Zhang et al., 2020), and alternative approaches explored vocal-based classification with k-NN and MLPs (Raccagni and Ntalampiras, 2021). More recently, EfficientNet-B0 (Karlita et al., 2022) and Xception (Qatrunnada et al., 2022) have been employed for cat breed classification, extending the range of architectures applied to this task.

Summary and challenges. Overall, CNN-based methods have demonstrated strong viability for classifying animals at both species and breed level (Sandler et al., 2018; Szegedy et al., 2016). However, challenges remain, particularly regarding dataset size, intra-class similarity, and the need for effective fine-grained classification methods in real-world scenarios such as wildlife monitoring (Divya Meena and Agilandeeswari, 2019; Leorna et al., 2022). The growing integration of CNN

Fig. 2. Examples of the three cat types: the European Wildcats (WC), the Hybrids (HY), and the Domestic Cats (DC).

and transformer architectures offers promising directions for addressing these limitations while maintaining efficiency in deployment-critical contexts.

3. Materials and methods

3.1. Architectures description

Five Convolutional Neural Network (CNN) architectures have been tested and compared to classify cat images, each architecture showing specific characteristics in terms of efficiency and accuracy, see Table 2 and Fig. 1 for a comparison of classification accuracy and model complexity. These five architectures were selected and tested due to their light weight and computational efficiency, which are key features for the implementation of a classification model on mobile devices. The details of each architecture are given below, highlighting their strengths and weaknesses.

- MobileNetV2 (Sandler et al., 2018): specifically designed for mobile devices, this architecture introduces two key innovations: linear bottlenecks and inverted residuals. These mechanisms preserve computational efficiency by reducing information loss in the tightest layers. Depthwise separable convolutions reduce the number of operations without compromising performance, making the network ideal for resource-constrained environments.
- MobileNetV3 (Howard et al., 2019): this architecture is a direct evolution of MobileNetV2, optimized through automatic techniques (Platform-Aware NAS and NetAdapt) and manual modifications, such as the use of the h-swish activation function instead of the ReLU and the optimization of the Squeeze-and-Excite modules. Two main variants are available:
  - MobileNetV3-Large, which has higher accuracy but slightly higher weight, designed for devices with high computing power;
  - MobileNetV3-Small, designed for low-power environments, presents a simplified structure, making it ideal for real-time applications.
- InceptionV3 (Szegedy et al., 2016): this architecture is based on the use of Inception modules that combine convolutions of different sizes in parallel, enabling feature extraction at different scales. The use of techniques such as batch normalization and label smoothing reaches high performance while containing the number of parameters. InceptionV3 is heavier than MobileNet, but in general it offers superior representational capability, proving useful for fine classification of complex images.
- MobileViTv2 (Mehta and Rastegari, 2022): this hybrid architecture integrates local convolutions and transformer modules, overcoming the computational limitations of traditional transformers through the use of separable self-attention, which reduces complexity from quadratic to linear. This approach proves particularly suitable for scenarios where a balance between accuracy and speed of execution is needed. The algorithm offers excellent modeling of global relationships, proving particularly effective even on high-resolution images.

3.2. Datasets description

In this work, CNN model are used for classifying wild-living cats WC = *Felis silvestris*, HY = *Felis silvestris x catus* and DC = *Felis catus* from real-world images (e.g. from camera traps, road-kills or opportunistic images). Initially, we assembled two distinct and independent datasets. The first, referred to as the initial dataset, consists of cat images for the three cat classes, with a subset of these images associated with a genetic analysis (see Fig. 2). In contrast, the second dataset, termed the genetic dataset, comprises only cat images associated with a genetic analysis. Images for the initial dataset were collected through different sources: (i) a call across the network of researchers involved in the Eurowildcat group[1]; (ii) a citizen-science project (CSP) currently running in Italy (Savazza et al., 2023); (iii) the personal dataset of one of the author consisting of wildcat images opportunistically taken in Italy (Anile et al., 2019); (iv) for the domestic cat class, a selection of images was chosen from the Dogs vs. Cats Kaggle Dataset,[2] by selecting images of domestic cats in which human body parts were not visible and cats were located in an outdoor environment. In contrast, the genetic dataset was solely assembled due to the collaboration of the wildcat researchers working together under the Eurowildcat network, who have kindly shared their cat images and the results of the genetic analysis, with us. For example, Fig. 3 analyses the distribution of the number of images by country or data source in the dataset. Furthermore, the initial dataset assembled is composed by cat images (WC = 1738; HY = 461; DC = 1193) which was then filtered using the YOLOv12 model, resulting in an unbalanced dataset (WC = 1201; HY = 461, DC = 856); this unbalanced dataset was further manually filtered with the concomitant purposes of retaining high quality images and balancing the number of images for each cat class, given that images for the hybrid class were less abundant than for the other two cat classes (WC = 500; HY = 461; DC = 500). To summarize the pipeline of the work we trained five different CNN architectures (MobileNetV2, MobileNetV3-Small, MobileNetV3-Large, InceptionV3 and MobileVitV2) on the balanced dataset, then the model which obtained the highest accuracy (i.e., MobileVitV2 = 86%), was further tested (i.e. label predictions are consistent with the genetic status) against a genetic dataset (i.e. images of cats with a taxonomy status ascertained via genetic analysis; WC = 50; HY = 50; DC = 50) and achieved an accuracy of 71%. However, when this architecture was trained on the unbalanced, and larger, dataset the accuracy for the genetic test increased to 81%.

In addition, an exploratory analysis of the visual characteristics of the dataset was conducted to guide the next steps. We focused on three basic parameters: brightness, contrast and saturation. Brightness shows a normal distribution in all classes, with a central peak between 100 and 130. Images belonging to the WC class show higher mean values than those in the DC and HY classes. This observation suggested a possible difference in exposure or illumination conditions during the image acquisition phase, which could manifest itself in a more pronounced visual perception of the scene. It is also observed that contrast

[1] Eurowildcat group: https://euromammals.org/eurowildcat/.
[2] Dogs vs. Cats: https://www.kaggle.com/competitions/dogs-vs-cats/data.

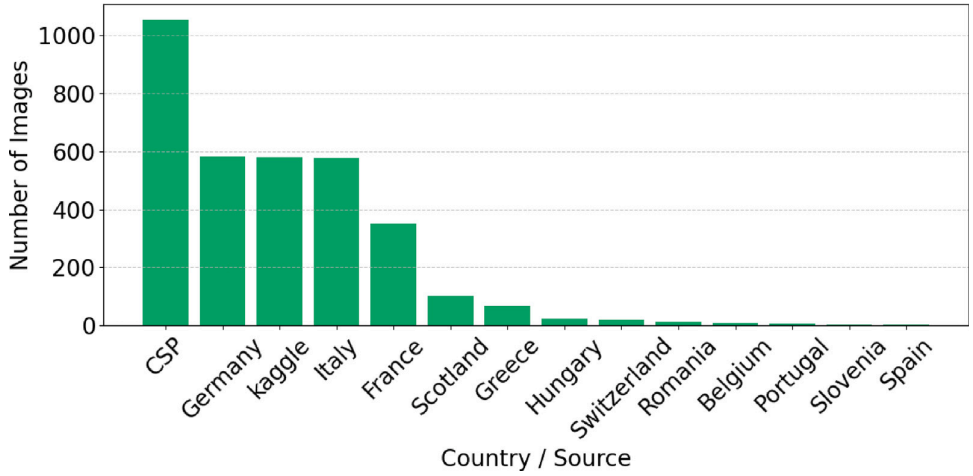


Fig. 3. Distribution of the number of images by country or data source in the dataset.

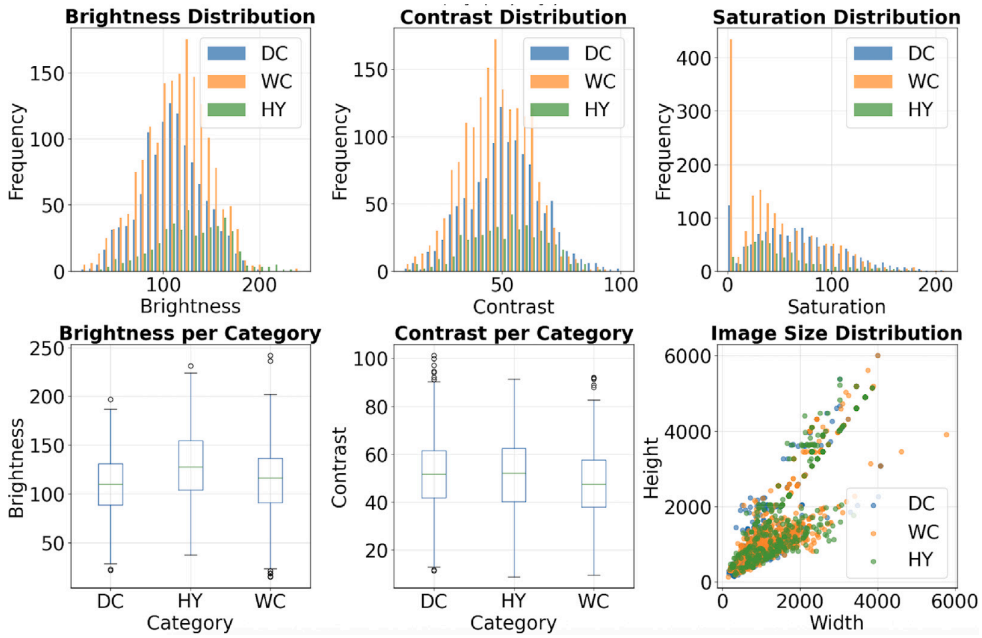


Fig. 4. Statistical analysis of the dataset according to different image features. Top row: distribution of brightness, contrast, and saturation grouped by class (WC, HY, DC). Bottom row: corresponding boxplots for brightness and contrast per class, and a scatter plot showing the distribution of image sizes (height vs. width) for each class.

follows a symmetrical distribution, with values predominantly between 30 and 70. The three categories show similar behavior, but with slight variations: images in the HY class tend to show slightly higher mean values, while those in the DC class include more instances with high contrast, which may indicate well-defined subjects compared to the background. Regarding saturation, a remarkably skewed distribution is observed, with a predominance of low values (below 50). Images belonging to the WC class are distinguished by higher color intensity than those labeled as DC and HY, which generally exhibit lower color saturation. These discrepancies may be attributable to variations in environmental conditions or semantic characteristics of the images. The analysis revealed significant differences between the cat classes in terms of brightness and saturation, with relatively uniform behavior in contrast. These differences, although minor, are significant for data interpretation and should be considered during the preprocessing and machine learning model development phases (Fig. 4).

3.3. Workflow description

First, we screened our initial dataset using the YOLOv12 model to detect the presence of a cat; this preliminary step was implemented to discard the images in which the cat shape is not recognizable (e.g poor quality images, where the cat body was obscured by vegetation or covered by blood leaks in case of road-kills), thus optimizing performances and reducing the computational load. After this filtering, the resulting unbalanced dataset (i.e. number of images per cat class was highly variable) was further manually filtered to retain only good quality images (i.e full body of the cat was visible and the coat was in optimal condition), leading to a balanced dataset (i.e. number of images per cat class was almost uniform, with slightly fewer HY images, due to the limited availability of this class); this step was conducted in order to reduce the bias of an unbalanced dataset when comparing the different set of candidates architectures ($n = 5$) for performing the classification

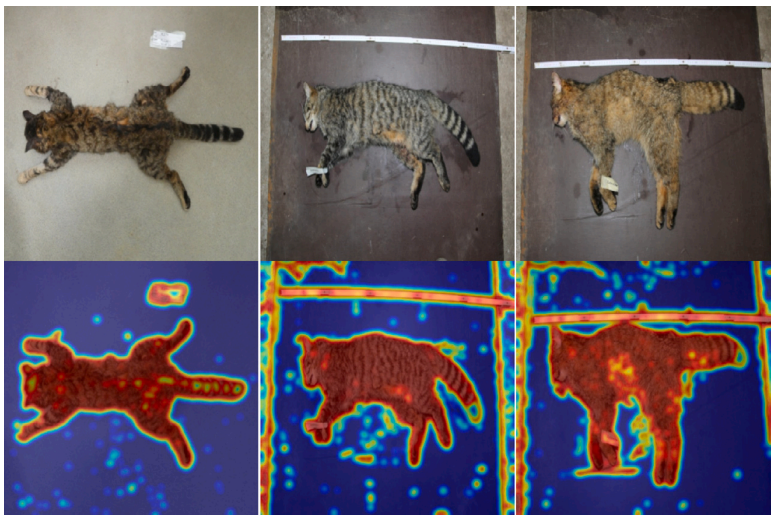


Fig. 5. Comparison of Grad-CAM outputs for DC (left), HY (center), and WC (right).

tasks. Eventually, the most accurate architecture, according to the results of the previous steps, is trained again on the unbalanced dataset, and ultimately tested on the genetic dataset. This approach ensured that the most promising architecture was validated on a set of images not used during the training, with the ultimate goal of evaluating its true generalization ability to classify unknown images. Each dataset was split into 85% for training and 15% for testing, and the candidate architectures were evaluated using configurations optimized for mobile devices, maintaining a balance between classification performance and computational efficiency, with the exception of the InceptionV3 model. In this case, a K-Fold cross-validation approach was applied to 90% of the dataset, divided into 4 folds, while the remaining 10% was reserved as a final independent test set to evaluate the model's generalization performance. The standard deviation across the folds was approximately 0.01 (0.00994).

3.4. Fine-tuning of the most accurate architecture

The most accurate model was trained using images scaled to 384 × 384 pixels. In order to maximize learning, the following features were implemented: (1) a class weighting strategy to compensate for the slight imbalance present in the data; (2) the use of ReLU and Dropout features to improve the generalization of the model; (3) training was conducted through the use of (i) the CrossEntropy loss function with weights; (ii) the AdamW optimizer for more stable learning; (iii) the OneCycleLR scheduler for dynamic learning rate management. Overall, the synergistic use of these features attained a stable and effective system during the training sessions (see Fig. 8(b)), while also allowing for a consistent classification of wild-living cats. Lastly, Grad-CAM heat maps were generated whenever an image is tested; Indeed, the Grad-CAM technique uses the gradients of any target concept, flowing into the final convolutional layer to produce a coarse localization map highlighting important regions in the image for predicting the concept (Selvaraju et al., 2017), providing visual explanations that make deep learning models more interpretable and transparent. This implementation allows users to inspect the output of the model and verifies that the model's attention is focused on relevant features (e.g., body region particularly useful for the morphological classification, like the *Dorsalis* and *Caudalis*, as well as more subtle body regions, like the *Gularis* and *Rhinarium*).

One of the primary objectives of this research was the development of a user-friendly mobile application that can be employed by researchers in real-world scenarios. To obtain this, the best classification model developed during the study was further improved and adjusted for use on mobile platforms, including smartphones and tablets. This adjustment involved breaking the classification process into two separate stages.

- *Detection phase*: in the initial phase, the system is configured to detect the presence of a cat within the camera's field of view by employing the YOLOv8n object detection model. The selection of YOLOv8n was chosen for its lightweight architecture, making it particularly suitable for mobile devices. This first filtering step is important for excluding images where a cat is not visually noticeable. This includes poor-quality images or those where the cat's body is blocked by surrounding elements or partly covered (e.g., in cases of roadkill). This method avoids extra calculations and improves the system's efficiency by starting classification only when a valid cat detection occurs.
- *Classification phase*: once the system successfully detects a cat, it classifies it into one of three defined categories: wild cat (WC), hybrid (HY), or domestic cat (DC). It does this using the image classification model trained during the project. This model can clearly tell the differences between these cat classes with high accuracy. It provides a probability distribution across the three classes, offering a clear decision-making result (see Table 1).

3.5. System requirements

Android is a heterogeneous platform, with devices having highly variable hardware performance. In order to ensure a smooth and consistent user experience, it is necessary to define minimum system requirements. The application was tested on several devices, including a Samsung Galaxy A35 equipped with 6 GB RAM, Exynos 1380 processor and Mali-G68 MP5 GPU. Despite compatibility, slowdowns and a less responsive rating compared to higher-performance devices emerged on this model. Therefore, it is recommended for use on mid- to high-end smartphones, such as the Google Pixel 8 or Samsung Galaxy S24 and later. On devices with limited computational capabilities, the application retains its functionality, albeit with longer response times and reduced performance. By using appropriate hardware, the full functionality offered can be exploited, significantly enhancing the user experience (see Table 1).

Table 1
Tools used in the project to design the mobile application for cat recognition.

Tool	Author/Model	Version
Smartphone	Samsung Galaxy S24 Ultra SM-S928B/DS	Android 15 (S928BXXU4BYD9)
Unity	–	2022.3.60f1
AR Foundation	Unity Technologies Inc.	5.1.6
Google ARCore XR Plugin	Unity Technologies Inc.	5.1.6
Apple ARKit XR Plugin	Unity Technologies Inc.	5.1.6
Sentis	Unity Technologies Inc.	1.2.0-exp.2
Native Gallery for Android & iOS	yasirkula	1.8.0

4. Results

We evaluated five CNN-based architectures: MobileNetV2, MobileNetV3-Small, MobileNetV3-Large, InceptionV3, and MobileViTv2-150 on the balanced dataset of cat images. As shown in Table 2, MobileViTv2-150 reached the highest classification accuracy, achieving 86.00%, followed closely by InceptionV3 at 85.59%. As shown in Table 2, the InceptionV3 model has a significantly higher number of parameters than the other neural architectures examined in the study. The high computational complexity of this network led to its exclusion from the final implementation, as its size and inference times are incompatible with the resource and performance constraints required by mobile devices. Despite its superior performance, the confusion matrix analysis (see Fig. 6) revealed that this model struggled to distinguish between the HY and WC classes, probably due to overlapping morphological traits and similar coat patterns. We, therefore, further evaluate the model's ability to generalize by testing it on the genetic dataset and the accuracy for this model lowered to 71.00%. However, when MobileViTv2-150 was trained on the unbalanced dataset and tested again on the genetic dataset, the accuracy increased to 81.00%, representing an absolute increase of 10 percentage points in accuracy (from 71% to 81%) between these two versions of the model (see Fig. 7(b)). In addition, by comparing the confusion matrices of MobileViTv2-150 trained on the balanced and unbalanced dataset, a reduction in misclassifications between the HY and DC classes was evident. Furthermore, although WC class recall slightly decreased, the overall precision increased, indicating that the unbalanced dataset better captured inter-class variability relevant to real-world scenarios. Table 3 reports the class-wise performance of the MobileViTv2-150 model on the genetic dataset. The model obtained an overall accuracy of 81%. Among the three classes, the DC class exhibited the highest precision (0.872) and F1-score (0.845), indicating reliable classification. The WC class also showed strong and balanced performance with both precision and recall above 0.82. In contrast, the HY class had the lowest precision (0.722), suggesting a higher rate of false positives, though its recall (0.780) indicates moderate sensitivity. These results highlight the model's ability to generalize well across classes, although showing some confusion in identifying hybrids. Fig. 8 illustrates the training dynamics of the MobileViTv2-150 model on the unbalanced dataset. In subplot (a), the accuracy curves show steady improvement for both training and test sets, with the test accuracy plateauing around 81% and maintaining a consistent gap from the training curve, indicating good generalization without signs of overfitting. Subplot (b) presents a continuous decrease in both training and test loss, with convergence occurring around epoch 60. The smooth and parallel trend of the loss curves further confirms the model's stable learning behavior and effective regularization.

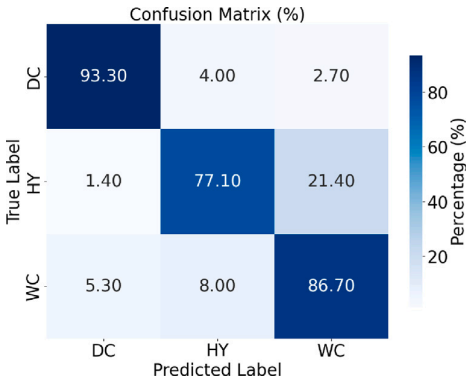


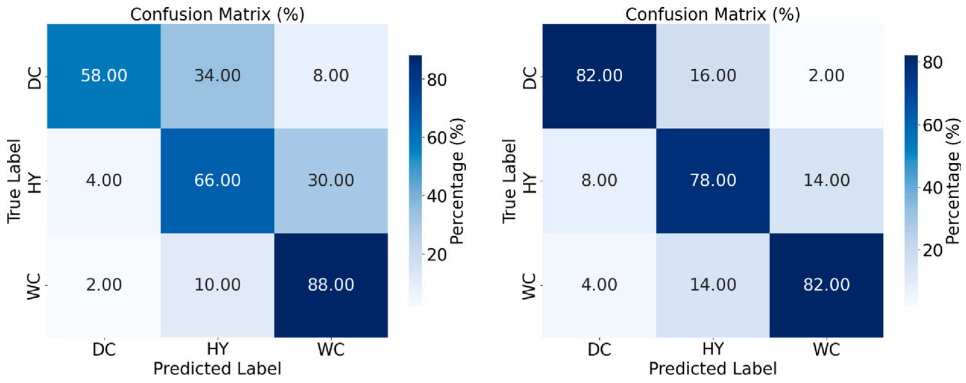
Fig. 6. MobileViTv2-150 trained on the balanced dataset. Confusion observed primarily between HY and WC classes.

5. Discussion

The collaboration of researchers from different scientific fields (zoology, informatics, and social science), further supported by two independent networks, one of wildcat researchers and the other of citizen scientists, has led to the development of an innovative and reliable model for classifying images of wild-living cats. Our classification model was further assembled for distribution as a mobile app, capable of working off-line, for the two most common mobile operating systems. The MobileViTv2-150 architecture outperformed all other tested architectures in terms of accuracy, while also maintaining a favorable trade-off between model complexity and performance, making it well-suited for mobile applications. Although MobileViTv2-150 reached the highest accuracy, the differences between the architectures were not subjected to statistical validation due to limited test samples. Future work should consider cross-validation, confidence intervals, and F1-scores to ensure statistical significance when selecting the most accurate architecture for this classification problem.

The use of convolutional neural networks (CNNs) and vision transformers in ecological informatics has increased markedly in recent years, with applications ranging from plankton classification (Nanni et al., 2025), pollen identification for honey authentication (González et al., 2025), to plant disease detection from smartphone images (Pandey and Jain, 2022). Within the domain of wildlife research, deep learning has been successfully applied for the identification of feral cats from camera trap images by exploiting body-part cues (Akbar et al., 2025), as well as for more general wildlife species classification using context-aware architectures (Wang et al., 2025). Compared to these studies, our pioneer work contributes a portable and lightweight recognition tool specifically tailored to the ecological challenge of distinguishing wild-living cats from hybrids and domestic individuals. This focus not only addresses a biologically relevant and taxonomically challenging classification problem but also highlights the potential of combining citizen science and conservation research in practical field applications.

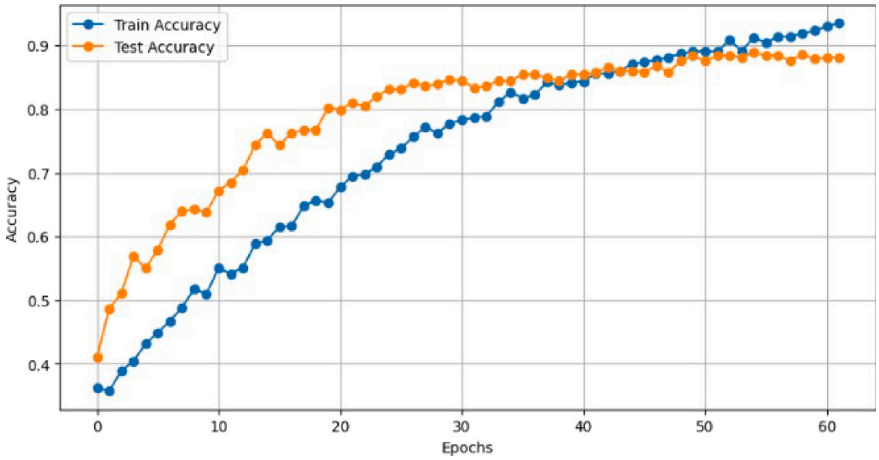
The broader ecological significance of our findings lies in enabling non-specialists, such as citizen scientists or wildlife amateur, to contribute reliable data for wildlife monitoring. The app provides an accessible tool for large-scale, distributed data collection, which is particularly relevant for species like the wildcat, whose conservation status is threaten by hybridization. By embedding artificial intelligence into



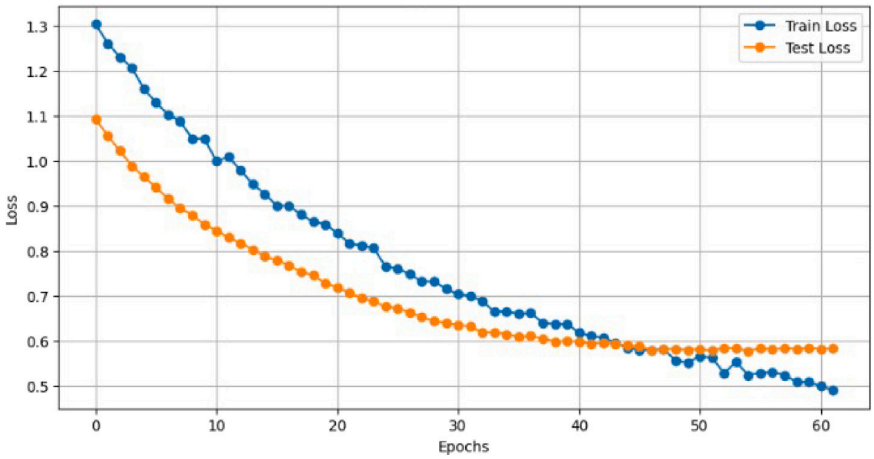
(a) MobileViTv2-150 trained on the balanced dataset and tested on the genetic dataset. Most misclassifications occur between DC-HY and HY-WC classes.

(b) MobileViTv2-150 trained on the unbalanced dataset and tested on the genetic dataset. Reduced misclassifications between DC-HY and HY-WC classes.

Fig. 7. Comparison of confusion matrices for MobileViTv2-150 under different training conditions with the same test set.



(a) MobileViTv2-150: training and test accuracy across epochs.



(b) MobileViTv2-150: training and test loss across epochs.

Fig. 8. Performance of the MobileViTv2-150 model trained on the unbalanced dataset. (a) Accuracy curve showing progressive improvements in both training and test sets, reaching convergence around epoch 60. (b) Loss curve illustrating a consistent reduction in both training and test loss, indicating stable learning and limited overfitting.

Table 2
Classification accuracy and complexity comparison of selected CNN architectures on the balanced dataset.

Metric	MobileNetV2	MobileNetV3-Small	MobileNetV3-Large	InceptionV3	MobileViTv2
Accuracy (%)	79.58	82.09	85.00	85.59	**86.00**
Total parameters	4,196,099	1,224,995	3,597,619	25,397,222	10,353,116
Trainable parameters	1,972,227	1,070,131	3,586,939	17,096,262	10,278,476
Non-trainable parameters	2,223,872	154,864	10,6802	8,300,960	74,640
Total FLOPs	960.622M	114.883M	680.459M	5.748G	9.095G

Table 3
Precision, recall, and F1-score for each cat class using the MobileViTv2-150 architecture. The model achieved an overall classification accuracy of 81% on the genetically verified dataset.

Class	Precision	Recall	F1-Score
Domestic (DC)	0.872	0.820	0.845
Hybrid (HY)	0.722	0.780	0.750
Wildcat (WC)	0.837	0.820	0.828

mobile technology, our approach lowers the barrier to entry for ecological data acquisition and supports evidence-based conservation planning. Moreover, the adoption of standardized classification tools may facilitate data reproducibility and hence comparability across projects and countries, contributing to transnational conservation strategies.

When comparing the confusion matrices of the model trained on the two datasets (balanced vs. unbalanced), the better performances of the model trained on the unbalanced dataset suggested that the number of images for each cat class play a crucial role in enhancing generalization, especially for the less represented classes, like HY in our case. To further understand the model's decision-making process, Grad-CAM visualizations generated for the genetic dataset (see Fig. 5) were inspected, confirming that the model captures discriminative patterns (e.g. ear, shape, coat color, tail) that align with human-interpretable regions.

The confusion matrix revealed a higher rate of misclassifications for HY (72%) with respect to the other two classes (WC = 83% and DC = 87%): this lower accuracy was expected for two reasons. First, the HY class has the lowest number of images, with the majority of HY images taken with the cats being *norma lateralis* or *ventralis* during veterinary necropsies. Hence, the limited range of body positions present in the HY images might be one of the reason underpinning this relatively low accuracy for this class when generalizing to real-world images. The second reason lies on the visual similarities in the coat patterns between tabby domestic cats (i.e. a *Felis catus* with the tabby pattern), HY and WC, which is an inherent characteristic of the classification problem that we aimed to tackle with our model; despite the above mentioned limitations, the model demonstrated satisfactory learning capabilities, although it is essential to acquire more HY images for achieving a higher degree of accuracy for this class.

Another critical issue for the ecological informatics community is reproducibility. Following the recommendations of Huettmann and Arhonditsis (2023), our study emphasizes the need to make data and code accessible for validation and reuse. While privacy and ethical constraints prevent the full release of some training images, the trained model weights and implementation pipeline will be made available to ensure that the study can be independently reproduced and extended by other researchers. This aligns with current best practices in open ecological data science and strengthens the credibility and utility of our results.

Overall, our model is a valuable tool which can be used as standard bench-mark for camera-trapping studies when there is the need of classifying wild-living cats (Anile et al., 2025). Our work's novelty derives from the seamless integration of a portable recognition tool, with a comprehensive statistical analysis, for usage in conservation zoology. Despite privacy constraints preventing full public release of the dataset used for training the CNN architecture, the support gained from the network of European wildcat researchers was crucial and underscored the urge for a deep learning model capable of handling this classification problem. Future work will focus on expanding the dataset, enhancing hybrid detection algorithms, and conducting in-depth analysis of user feedback to refine both usability and classification performance.

6. Conclusion

This paper presents a mobile application for the taxonomy classification of wild-living cats and offers valuable insights for wildcat research and conservation (Nussberger et al., 2023). The application exploits a customized, modern artificial neural network for image classification, with careful consideration of the trade-off between accuracy and computational requirements. The model runs entirely on the device, enabling fully offline operation. The system has been shown to classify input images under a variety of real-world conditions. These results were obtained using a novel dataset created for this study, and further validated on a genetically verified dataset. Together, they demonstrate the feasibility of deploying advanced neural models on resource-constrained mobile hardware. Future work will focus on expanding the genetic dataset, enhancing hybrid detection algorithms, and conducting in-depth analysis of user feedback to refine both usability and classification performance. A slight confusion between HY and WC classes was observed, suggesting an inherent difficulty in distinguishing their coat patterns. Despite the limited dataset, the model demonstrated satisfactory learning capabilities. The lower accuracy of the DC class can be attributed mainly to two factors. First, differences in the position of the subjects are evident: while in the training dataset domestic cats are portrayed in natural poses, in the genetically tested images these appear partly in standardized positions (*norma lateralis* or *ventralis*), which are rarely represented in the training set. Second, the visual similarity in coat patterns generated ambiguity, as parts of the tested images depict subjects with characteristics similar to those of hybrids. Slight confusion between HY and WC classes is also observed, which could indicate a tendency for the model to confuse images of these two categories. Despite the limited dataset, the model demonstrated satisfactory learning capabilities. However, the paucity of data limited the generalization ability. To optimize performance, it is essential to expand the dataset, especially for the DC class, by including images with the poses and visual features typical of genetic testing.

CRediT authorship contribution statement

Georgia Fargetta: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Data curation, Conceptualization. **Stefano Anile:** Writing – review & editing, Validation, Data curation. **Luigi Priano:** Writing – original draft, Visualization, Validation, Software, Methodology, Data curation. **Massimo Orazio Spata:** Writing – review & editing, Supervision, Software, Methodology, Data curation, Conceptualization. **Alessandro Ortis:** Writing – review & editing, Supervision, Conceptualization. **Sebastiano Battiato:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Georgia Fargetta reports financial support was provided by University of Catania. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

G.F. is supported by PNRR MUR project PE0000013-FAIR.

S.A is funded under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.4- Call for tender No. 3138 of 16 December 2021, rectified by Decree n. 3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union - NextGenerationEU Project Code CN_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP B73C22000790001, Project Title National Biodiversity Future Center-NBFC.

M.O.S. is supported by HICONNECTS, CUP B63C23000380009.

A.O. is supported by research PIAno di inCEntivi per la Ricerca di Ateneo 2024/2026 — Linea di Intervento i "Progetti di ricerca collaborativa" - SIAM Project - University of Catania, Italy.

S.B. is supported by Spoke 1 Future HPC & Big Data of the Italian Research Center on High-Performance Computing. We sincerely thanks *Simone Calandri* and all the others Italian citizen scientists whose have contributed to this work with their images of cats taken with camera-traps. We are also extremely grateful to the following Eurowildcat researchers (*Ana Barros, Matteo Bastianelli, Roo Campbell, Sebastien Devillard, Luisa Fischer, Peter Gerngross, Malte Goltz, Zsolt Hegyeli, Mathias Herrmann, Marco Heurich, Rene Jansen, Kerry Kilshaw, Andrew Kitchener, Miha Krofel, Johannes Lang, Jozsef Lanszki, Despina Migli, Pedro Monterroso, Beatrice Nussberger, Ferran Sayol, Sabrina Streif*) for sharing their valuable data with us; without their kind support the development of this model would have not been conceivable and possible.

Data availability

The pre-trained models and the majority of the dataset used in this study are publicly available at the following URL: https://www.dmi.unict.it/ortis/wildcat/.

These resources are sufficient to reproduce the results reported in the paper, as the pre-trained models encapsulate the knowledge learned from the full training data. A small portion of the training data, provided by an external collaborator who is not an author, cannot be publicly released. It was shared with us solely for the purposes of this study. Researchers may request access to this subset through the provided URL for legitimate research purposes.

References

Akbar, R.R.S., Rees, M.W., Fleming, P.A., Sohel, F., 2025. Body-part-based individual feral cat identification from camera trap images using deep learning. Ecol. Informatics 103258.

Anile, S., Devillard, S., Ragni, B., Rovero, F., Mattucci, F., Lo Valvo, M., 2019a. Habitat fragmentation and anthropogenic factors affect wildcat (*Felis silvestris silvestris*) occupancy and detectability on mt. etna. Wildl. Biology 1, 1–13. http://dx.doi.org/10.2981/wlb.00561.

Anile, S., Devillard, S., Ragni, B., Rovero, F., Mattucci, F., Valvo, M.L., 2019. Habitat fragmentation and anthropogenic factors affect wildcat felis silvestris silvestris occupancy and detectability on mt etna. Wildl. Biology 2019b (1), 1–13.

Anile, S., Greco, I., Marconi, M., Cappai, N., Salvatori, M., Rovero, F., 2025. Estimates of density, occupancy and activity patterns of wild-living cats in Foreste Casentinesi National Park (Italy) using data from a non-targeted camera array. Eur. J. Wildl. Res. 71 (3), 45. http://dx.doi.org/10.1007/s10344-025-01923-6.

Anile, S., Ragni, B., Randi, E., Mattucci, F., Rovero, F., 2014. Wildcat population density on the e tna volcano, i taly: a comparison of density estimation methods. J. Zool. 293 (4), 252–261.

Austen, G.E., Bindemann, M., Griffiths, R.A., Roberts, D.L., 2016. Species identification by experts and non-experts: comparing images from field guides. Sci. Rep. 6 (1), 33634. http://dx.doi.org/10.1038/srep33634.

Ballesteros-Duperón, E., Virgós, E., Moleón, M., Barea-Azcón, J.M., Gil-Sánchez, J.M., 2015. How accurate are coat traits for discriminating wild and hybrid forms of felis silvestris? Mammalia 79 (1), 101–110.

Battiato, S., Ortis, A., Trenta, F., Ascari, L., Politi, M., Siniscalco, C., 2020. Detection and classification of pollen grain microscope images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 980–981.

Ceballos, G., Ehrlich, P.R., 2023. Mutilation of the tree of life via mass extinction of animal genera. Proc. Natl. Acad. Sci. 120 (39), 2306987120. http://dx.doi.org/10.1073/pnas.2306987120, publisher: Proceedings of the National Academy of Sciences..

Clapham, M., Miller, E., Nguyen, M., Darimont, C.T., 2020. Automated facial recognition for wildlife that lack unique markings: A deep learning approach for brown bears. Ecol. Evol. 10 (23), 12883–12892. http://dx.doi.org/10.1002/ece3.6840.

Devillard, S., Jombart, T., Leger, F., Pontier, D., Say, L., Ruette, S., 2013. How reliable are morphological and anatomical characters to distinguish European wildcats, domestic cats and their hybrids in France? J. Zool. Sys. Evol. Res. 52 (2), 154–162.

Divya Meena, S., Agilandeeswari, L., 2019. An efficient framework for animal breeds classification using semi-supervised learning and multi-part convolutional neural network (mp-cnn). IEEE Access 7, 151783–151802. http://dx.doi.org/10.1109/ACCESS.2019.2947717.

Eichholzer, A., 2010. Testing the Applicability of Pictures Taken By Camera-Traps for Monitoring the European Wildcat *Felis* Silvestris Silvestris in the Jura Mountains of Switzerland. (Ph.D. thesis). University of Zürich, Switzerland.

Fargetta, G., Ortis, A., Anile, S., Battiato, S., 2024. Evaluation of Cnns for Wildcats Classification in Real World Scenario, in: Recent Trends and Advances in Artificial Intelligence. Springer Nature, Switzerland, pp. 15–25.

Gaudiano, L., Corriero, G., Villani, M., Anile, S., 2022. Gone Before It's Known? camera-Trapping Shows Alarming Levels of Putative Hybrids In The Wildcat (*Felis Silvestris*) Population of The Gargano National Park (Southern Italy). Mammalia, http://dx.doi.org/10.1515/mammalia-2021-0125.

Gerngross, P., Ambarli, H., Angelici, F.M., Anile, S., Campbell, R., Ferreras de Andres, P., Gil-Sánchez, J.M., Gotz, M., Jerosch, S., Mengüllüoglu, D., Monterroso, P., Zlatanova, D., silvestris, Felis, 2022. The IUCN Red List of Threatened Species 2022: e.T181049859A181050999. Tech. rep., IUCN, International Union for Conservation of Nature.

González, J.M.V., Osuna, J.J.M., Pinto, A.P., Roberto, I.E., 2025. Assisting monofloral honey classification by automated pollen identification based on convolutional neural networks. Ecol. Informatics 103340.

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q.V., Adam, H., 2019. Searching for mobilenetv3. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. ICCV.

Huettmann, F., Arhonditsis, G., 2023. Towards an ecological informatics scholarship that is reflective, repeatable, transparent, and sharable!. Ecol. Informatics 76, 102132.

Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D.E., Coscieme, L., Golden, A.S., Guerra, C.A., Jacob, U., Takahashi, Y., Settele, J., Díaz, S., Molnár, Z., Purvis, A., 2022. The direct drivers of recent global anthropogenic biodiversity loss. Sci. Adv. 8 (45), eabm9982. http://dx.doi.org/10.1126/sciadv.abm9982.

Karlita, T., Choirunisa, N.A., Asmara, R., Setyorini, F., 2022. Cat breeds classification using compound model scaling convolutional neural networks. In: Proceedings of the International Conference on Applied Science and Technology on Social Science 2021. iCAST-SS 2021, Atlantis Press, pp. 909–914.

Kheir, A., Koubaa, A., Kolluru, V., Mungara, S., Feike, T., 2025. Smart plant disease diagnosis using multiple deep learning and web application integration. J. Agric. Food Res. 21, 101948. http://dx.doi.org/10.1016/j.jafr.2025.101948.

Kilshaw, K., Drake, A., Macdonald, D.W., Kitchener, A.C., 2010. The Scottish Wildcat: A Comparison of Genetic and Pelage Characteristics. Tech. Rep., Vol. 356, Scottish Natural Heritage.

Kitchener, A.C., Yamaguchi, N., Ward, J.M., Macdonald, D.W., 2005. A diagnosis for the scottish wildcat (felis silvestris): a tool for conservation action for a critically-endangered felid. In: Animal Conservation Forum. Vol. 8, Cambridge University Press, pp. 223–237.

Leorna, S., Brinkman, T., vs, Human., 2022. Machine: detecting wildlife in camera trap images. Ecol. Informatics 72, 101876. http://dx.doi.org/10.1016/j.ecoinf.2022.101876.

Lesch, R., Kitchener, A.C., Hantke, G., Kotrschal, K., Fitch, W.T., 2022. Cranial volume and palate length of cats, Felis spp. under domestication, hybridization and in wild populations. R. Soc. Open Sci. 9 (1), 210477. http://dx.doi.org/10.1098/rsos.210477.

Lin, M., 2024. A comparative analysis study of deep learning methods in cat and dog classification. Trans. Comput. Sci. Intell. Syst. Res. 5, 1595–1601. http://dx.doi.org/10.62051/yqccgn73.

Macdonald, D.W., Loveridge, A.J., 2010. The Biology and Conservation of Wild Felids. Vol. 2, Oxford University Press.

Matias, G., Rosalino, L.M., Alves, P.C., Tiesmeyer, A., Nowak, C., Ramos, L., Steyer, K., Astaras, C., Brix, M., Domokos, C., Janssen, R., Kitchener, A.C., Mestdagh, X., L'Hoste, L., Titeux, N., Migli, D., Youlatos, D., Pfenninger, M., Devillard, S., Ruette, S., Anile, S., Ferreras, P., Díaz-Ruiz, F., Monterroso, P., 2022. Genetic integrity of European wildcats: variation across biomes mandates geographically tailored conservation strategies. Biol. Cons. 268, 109518. http://dx.doi.org/10.1016/j.biocon.2022.109518.

Mattucci, F., Galaverni, M., Lyons, L.A., Alves, P.C., Randi, E., Velli, E., Pagani, L., Caniglia, R., 2019. Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. Sci. Rep. 9 (1), 11612. http://dx.doi.org/10.1038/s41598-019-48002-w.

Mehta, S., Rastegari, M., 2022. Separable self-attention for mobile vision transformers. arXiv preprint arXiv:2206.02680.

Nanni, L., Lumini, A., Barcellona, L., Ghidoni, S., 2025. Convolutional neural networks and vision transformers for plankton classification. Ecol. Informatics 103272.

Nieto-Blázquez, M.E., Schreiber, D., Mueller, S.A., Koch, K., Nowak, C., Pfenninger, M., 2022. Human impact on the recent population history of the elusive European wildcat inferred from whole genome data. BMC Genomics 23 (1), 709. http://dx.doi.org/10.1186/s12864-022-08930-w.

Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc. Natl. Acad. Sci. 115 (25), E5716–E5725.

Nussberger, B., Barbosa, S., Beaumont, M., Currat, M., Devillard, S., Heurich, M., Howard-McCombe, J., Mattucci, F., Nowak, C., Quilodrán, C.S., Senn, H., Alves, P.C., 2023. Eurowildcat consortium, a common statement on anthropogenic hybridization of the European wildcat (*Felis silvestris*). Front. Ecol. Evol. 11, 1156387. http://dx.doi.org/10.3389/fevo.2023.1156387.

Nussberger, B., Currat, M., Quilodran, C.S., Ponta, N., Keller, L.F., 2018. Range expansion as an explanation for introgression in european wildcats. Biol. Cons. 218, 49–56.

Pandey, A., Jain, K., 2022. A robust deep attention dense convolutional neural network for plant leaf disease identification and classification from smart phone captured real world images. Ecol. Informatics 70, 101725.

Qatrunnada, N., Fachrurrozi, M., Utami, A., 2022. Cat breeds classification using convolutional neural network for multi-object image. Sriwij. J. Informatics Appl. 3, http://dx.doi.org/10.36706/sjia.v3i2.46.

Raccagni, W., Ntalampiras, S., 2021. Acoustic classification of cat breed based on time and frequency domain features. In: 2021 30th Conference of Open Innovations Association FRUCT. pp. 184–189. http://dx.doi.org/10.23919/FRUCT53335.2021.9599975.

Ragni, B., Possenti, M., 1996. Variability of coat-colour and markings system in felis silvestris. Ital. J. Zoöl. 63 (3), 285–292.

Ragusa, F., Tomaselli, V., Furnari, A., Battiato, S., Farinella, G.M., 2016. Food vs non-food classification. In: Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management. pp. 77–81.

Ramadhan, A., Setiawan, A., 2023. Catbreedsnet: An android application for cat breed classification using convolutional neural networks. J. Online Inform. 8, 52–60. http://dx.doi.org/10.15575/join.v8i1.1007.

Randi, E., 2008. Detecting hybridization between wild species and their domesticated relatives. Mol. Ecol. 17, 285–293.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. CVPR.

Savazza, S., Rossi, P., Rebora, N., Lioy, F.G., Franculli, D., Roveyaz, A., Bruno, A., Sicuro, A., Dino, G., La Placa, E., Gaudiano, L., Cicognani, L., Monti, F., Serra, G., Gallipò, G., Draksler, V., Vecchietti, C., Filoni, L., Cairone, A., Krodaliu, E., Nardone, A., Dartora, F., Scotti, M., Rimassa, F., Vecchiolli, C., Calandri, S., Francescangeli, D., Parretta, G., Giosso, S., Barbieri, L., Ecotti, D., Piolanti, F., Piolanti, G., Comazzi, S., Filacorda, S., Nielsen, C.K., Anile, S., 2023. Cats and clouds: how a citizen camera-trapping project boosts wildcat (*Felis silvestris*) conservation. Eur. J. Wildl. Res. 69 (3), 49. http://dx.doi.org/10.1007/s10344-023-01670-6.

Schindler, F., Steinhage, V., 2021. Identification of animals and recognition of their actions in wildlife videos using deep learning techniques. Ecol. Informatics 61, 101215. http://dx.doi.org/10.1016/j.ecoinf.2021.101215.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision. ICCV, pp. 618–626. http://dx.doi.org/10.1109/ICCV.2017.74.

Senn, H.V., Ghazali, M., Kaden, J., Barclay, D., Harrower, B., Campbell, R.D., Macdonald, D.W., Kitchener, A.C., 2019. Distinguishing the victim from the threat: SNP-based methods reveal the extent of introgressive hybridization between wildcats and domestic cats in Scotland and inform future in situ and ex situ management options for species restoration. Evol. Appl. 12 (3), 399–414. http://dx.doi.org/10.1111/eva.12720.

Simberloff, D., 1996. Hybridization between native and introduced wildlife species: Importance for conservation. Wildl. Biol. 2 (3), 143–150. http://dx.doi.org/10.2981/wlb.1996.012.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. CVPR.

Varghese, S., Remya, S., 2021. Dog breed classification using cnn. In: Security Issues and Privacy Concerns in Industry 4.0 Applications. pp. 195–205. http://dx.doi.org/10.1002/9781119776529.ch10.

Velli, E., Bologna, M.A., Castelli, S., Ragni, B., Randi, E., 2015. Non-invasive monitoring of the European wildcat (*Felis silvestris silvestris* Schreber, 1777): comparative analysis of three different monitoring techniques and evaluation of their integration. Eur. J. Wildl. Res. 61 (5), 657–668.

Viviani, L., Nonnis Marzano, F., Mattucci, F., Sforzi, A., 2024. Non-invasive monitoring of the European wildcat in a recently colonized area in the Northern Apennines (Italy). Eur. J. Wildl. Res. 70 (2), 22. http://dx.doi.org/10.1007/s10344-024-01777-4.

Wang, L., Wang, S., Deng, C., Zhu, H., Tian, Y., Zhang, J., 2025. Deloco: Decoupled location context-guided framework for wildlife species classification using camera trap images. Ecol. Informatics 85, 102949.

Zhang, Y., Gao, J., Zhou, H., 2020. Breeds classification with deep convolutional neural network. In: Proceedings of the 2020 12th International Conference on Machine Learning and Computing. ICMLC '20, Association for Computing Machinery, New York, NY, USA, pp. 145–151. http://dx.doi.org/10.1145/3383972.3383975.

Zhao, W., Zhou, D., Qiu, X., Jiang, W., 2021. How to represent paintings: A painting classification using artistic comments. Sensors 21 (6), 1940.