

ORIGINAL ARTICLE

Hoofed animal detection in UAV thermal images using Balanced Random Forest and CNN features

Dorota Włodarczyk ^{1*} and Grzegorz Józków ¹¹Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences, Grunwaldzka 53, 50–357, Wrocław, Poland

*120774@student.upwr.edu.pl

Abstract

Wildlife monitoring is vital to conservation efforts and the prevention of animal-related negative impacts on human activities and ecosystems. The use of Unmanned Aerial Vehicles (UAVs) enables data collection with no harm to wildlife and in difficult field conditions. This study proposes a method of detecting hoofed animals in UAV-acquired thermal images, addressing the challenges of low-resolution thermal imaging and the presence of other heated objects hindering simple temperature analysis and image segmentation. The proposed method uses machine learning algorithms and is designed to work with a limited size of training dataset. The method consists of an initial segmentation step that detects potential animals based on thermal and geometrical signatures, followed by classification using a Balanced Random Forest (BRF) algorithm. One of the key aspects of the proposed method is the use of geometric and thermal features along with multi-scale Convolutional Neural Network (CNN) extracted feature representations in BRF. The benefit of the BRF is its speed, little requirement regarding the amount of training data, and its capacity to work with an imbalanced number of objects in different classes. The dataset was collected during two UAV flights over a fenced enclosure with wild hoofed animals. The proposed approach showed high efficiency, achieving an overall accuracy of 90%. These results confirm the feasibility of UAV-based animal detection based solely on thermal images collected during the day and showing many other heated objects. The method provides a solution for wildlife monitoring, with potential adaptability to different species and further applications.

Key words: classification, object detection, machine learning, unmanned aerial vehicle, remote sensing, wildlife monitoring

1 Introduction

Wildlife, an integral part of ecosystems, has long struggled to cope with the difficulties related to human impact. Expanding human activity into natural areas intensifies human-wildlife interactions (Nyhus, 2016), leading to potential conflicts and challenges in environmental and safety management (Ridwan et al., 2023). The migration and presence of animals in agricultural areas cause significant economic damage and increase the risk of field accidents (Ramadhan, 2024). In addition, the problem of spreading diseases, such as African Swine Fever (ASF), underscores the need to develop reliable methods of monitoring wildlife populations (Woźniakowski et al., 2021). Current methods of fighting this disease, such as preventive shooting,

are controversial from an ecological and ethical point of view, forcing the search for alternatives (Ridwan et al., 2023; Woźniakowski et al., 2021). Before undertaking certain actions that interfere with the natural environment of animals, it is crucial to implement animal monitoring to better understand their behavior, which in turn can help to plan and predict the effects of human activities.

Traditional methods of wildlife monitoring, such as field observations and photo traps, are widely used but have significant limitations (Pollock et al., 2002). Field observations, which involve counting animals or their tracks, require the participation of experienced field researchers and a great deal of time and effort (Witmer, 2005). Additionally, it can involve the risk of encountering predators (Pollock et al., 2002). Photo traps allow

for remote animal observation, but their effectiveness depends on the strategic placement of the devices, which remain stationary (Witmer, 2005). These methods have a limited range and do not guarantee full accuracy in estimating population numbers (Pollock et al., 2002; Burton et al., 2015). Alternatively, GPS tags allow precise tracking of individuals, but their high cost and the need to locate the animal in advance and tag application are major barriers (Foley and Sillero-Zubiri, 2020; Ford et al., 2009).

Increasingly, drones (Unmanned Aerial Vehicles – UAVs) are being used for this purpose, offering cutting-edge solutions for monitoring wildlife (Christiansen et al., 2014; Hodgson et al., 2016; Kalinowski et al., 2023; Lee et al., 2021; Lyu et al., 2024; Rančić et al., 2023; Tuia et al., 2022; Witczuk et al., 2017). Compared to traditional methods, drones allow for quick and efficient mapping of larger areas and acquisition of data in different lighting conditions through the use of RGB and thermal cameras (Hodgson et al., 2016; Lee et al., 2021; Witczuk et al., 2017). Thanks to advanced image processing and data analysis technologies, UAVs make it possible to identify animals in their natural environment with minimal disruption to their behavior (Witczuk et al., 2017). Integrating these technologies with machine learning methods allows automatic analysis of the collected data, which enhances the reliability of monitoring (Tuia et al., 2022).

Despite its many advantages, the use of drones also comes with some challenges. These include the high costs of the equipment (Witczuk et al., 2017), the need for adequate operator training (Jewell, 2013), and limitations due to weather (Verma et al., 2016), technical, and law factors. Nevertheless, UAVs are promising tools that can effectively support the management of wildlife and reduce the negative impact of humans on their natural environment.

This study focuses on image analysis and presents an automatic method of detecting animals solely in thermal images collected using UAV, which can also be further adapted for counting individuals, monitoring their presence in specified areas, or searching for missing persons. The proposed approach leverages machine learning, combining image features used in Convolutional Neural Networks (CNNs) with a Balanced Random Forest (BRF) classifier, offering a computationally efficient solution requiring minimal training data. The following sections provide a structured presentation of the research: a review of related work on detection and classification techniques (Section 2), a description of the dataset and data acquisition process (Section 3), and a detailed explanation of the proposed methodology, including preprocessing, feature extraction, and classification (Section 4). The results of the conducted experiments are analyzed in Section 5, followed by a discussion (Section 6) assessing model performance, potential improvements, and limitations. The final Section 7 presents concluding remarks and outlines directions for future research.

2 Related work

Research utilizing UAVs to identify animals in their natural environments uses a variety of different imaging techniques, including both thermal and RGB cameras (Christiansen et al., 2014; Kalinowski et al., 2023; Lee et al., 2021; Rančić et al., 2023), often used simultaneously to leverage the strengths of each type under different lighting conditions. RGB images are valued for their high resolution and efficiency during the day and in open areas. However, their usability decreases at night and in dense forests, where access to sunlight is limited. In such conditions, thermal images, capturing temperature values, tend to be more useful (Kalinowski et al., 2023).

Despite a considerable track record in automating the de-

tection of animals in UAV images, research faces numerous technological and environmental difficulties. The visibility limited by dense vegetation (Kalinowski et al., 2023; Rančić et al., 2023) and low resolution of thermal images (Christiansen et al., 2014; Kalinowski et al., 2023; Rančić et al., 2023) pose significant challenges. Additionally, during daylight, the presence of heated objects such as soil, branches, or buildings can increase the rate of false detections (Christiansen et al., 2014; Kalinowski et al., 2023). Variations in lighting conditions and background changes, including moving leaves or shifting shadows, further complicate the process, making it challenging for computer vision models like CNNs, which are sensitive to such anomalies (Rančić et al., 2023). Furthermore, ensuring a sufficient quantity and quality of training data (Christiansen et al., 2014; Lee et al., 2021) poses an additional challenge, particularly for rare species or peculiar environments.

A wide range of automatic image processing techniques are being used to address those limitations. The most advanced include CNN algorithms such as You Only Look Once (YOLO) which offers a high performance rate (Kalinowski et al., 2023; Rančić et al., 2023). For example, YOLOv4, used in deer identification studies on RGB images, achieved 86% precision and 75% recall (Rančić et al., 2023). However, shallow machine learning algorithms are used as well. An example is the use of the k-Nearest Neighbors (kNN) algorithm to classify small animals in agricultural fields, such as rabbits and chickens, in RGB and thermal images. In this approach, features extracted using morphological operations and Discrete Cosine Transform (DCT) achieved a balanced classification accuracy of 84.7% (Christiansen et al., 2014). Other studies propose combining simple algorithms of image analysis with deep learning. An example is the use of Canny's edge detection algorithm to identify shapes in thermal images, followed by classification using the LeNet-5 convolutional neural network. This approach has achieved animal detection accuracy of 98.6% and 96.5% in classification into wild boar and other animal classes, respectively (Kalinowski et al., 2023). In another study (Lee et al., 2021), a real-time detection method using both thermal and RGB images was developed to overcome the limitations of the low resolution of thermal images and the small amount of training data. The technique was based on Sobel's edge detector algorithm and analysis of thermal signatures processed using the DCT. By automatically removing features such as buildings and roads, the method reduced false detections, reporting a best precision of about 80.4% and recall of 69.9% in tests conducted from altitudes below 100 meters (Lee et al., 2021).

The concept of combining different methods for object detection in images extends beyond wildlife detection. Such combinations have been successfully applied in other applications. A particularly notable example is the integration of deep learning techniques, such as CNNs, with shallow learning algorithms. For example, in the field of agricultural remote sensing, Random Forest (RF) was used to classify crop types based on CNN-created features, achieving an accuracy of 94.3%, significantly outperforming results achieved using CNNs alone (Yang et al., 2020). The same concept has been applied to sinkhole and landslide detection studies with an equally high accuracy of 92% (Lee et al., 2016). A similar approach has been used in urban monitoring (De Oliveira and Wehrmeister, 2018). In studies on improving public safety and optimizing traffic management, pedestrians were successfully identified even in low-light conditions thanks to data obtained from thermal and RGB cameras mounted on drones (De Oliveira and Wehrmeister, 2018). Such applications confirm that combined approaches are not only versatile but also efficient.

In response to the difficulties of monitoring wildlife in their natural habitat, this study focuses on developing another approach for animal detection using only thermal images ac-

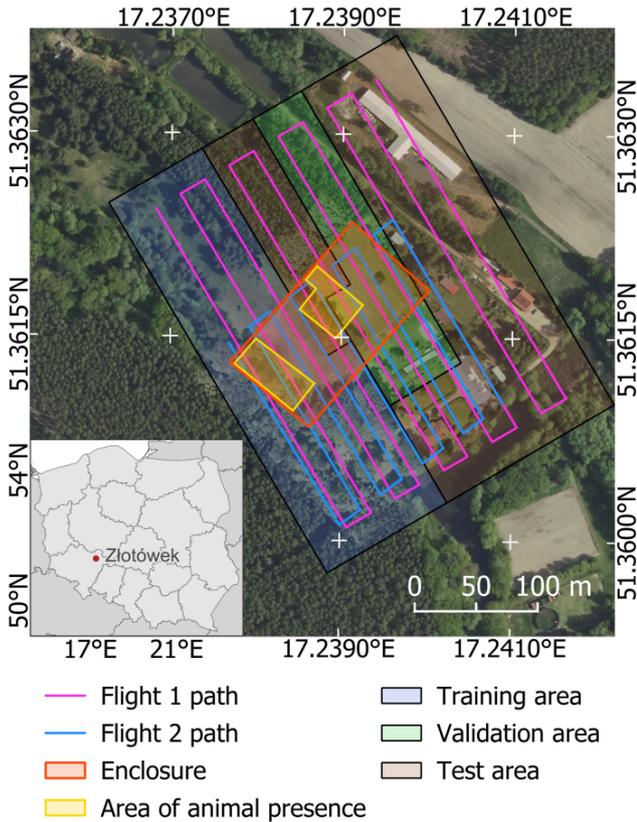


Figure 1. Localization of study area, flight paths and reference data partitioning

quired from UAVs. The goal was to develop a method that combines efficient image analysis and machine learning methods to identify animals with minimal computing power requirements and a limited amount of training data. The method is designed to handle daytime thermal imagery, where solar-heated objects often create additional challenges. Such images are easier to collect than nighttime images and are more suitable for animals showing their activity during the day. The combined approach incorporates the advantages of the convolutional features from deep learning (CNN) with classification using a modification of the RF algorithm – Balanced Random Forest. The work focuses on the development of the method from the perspective of thermal image analysis and does not deal with aspects of data acquisition or subsequent applications, such as animal localization, tracing, or counting.

3 Data

The test data were collected in late winter at a forest experimental facility in Złotówek (Poland) at the enclosure for wild animals (Figure 1). Recordings were acquired with an uncooled thermal imaging camera Optris PI 450 lightweight mounted on a DJI Matrice 600 Pro drone during two UAV flights. The flights were planned at 53 m above the ground level and resulted in a Ground Sampling Distance (GSD) of thermal images equal to about 17 cm.

Each thermal image had a resolution of 382×288 pixels, covering an area of approximately 65×49 m. The flight was planned according to typical UAV mapping requirements along parallel lines (Figure 1), with a front overlap of 90% between consecutive images and a side overlap of 65% between flight lines. The large front overlap ensured that even a moving animal

was visible in several consecutive images. The flights lasted about 14 and 7 minutes and resulted in 870 and 420 thermal images, respectively.

The fenced enclosure covers 1.263 ha. At the time, it housed 11 hoofed animals, including red deer (*Cervus elaphus*), roe deer (*Capreolus capreolus*), mouflon (*Ovis ammon musimon*), and European fallow deer (*Dama dama*). In the surroundings of the enclosure, there were: farm buildings, a forest, a pond, and a stream flowing from the pond into the forest (southwest of the enclosure) (Figure 1). During the flights, a small number of other animate objects (swans, peacocks, dogs, humans) were also captured in the footage; these were incidental and not part of the main target group of hoofed animals, but were useful for validating the robustness of the proposed method.

Additionally, in order to preliminarily assess the model performance in uncontrolled, more natural conditions, an auxiliary dataset was collected in late spring over a semi-open area located adjacent to the main study site. The surveyed area (approximately 20 ha) was characterized by fully developed vegetation, typical for this season. During field reconnaissance, traces indicating the recent presence of wild ungulates (likely roe deer) were identified, although the number and species could not be determined with certainty. A UAV equipped with the same thermal imaging camera (Optris PI 450) was used to acquire images under similar flight parameters (altitude, overlap, time of day). In addition, a wide-angle RGB camera (smartphone) was mounted on board to enable potential visual verification of detected objects. This supplementary dataset was not used for training or labeling, but served as an independent material to verify model generalization in a less controlled environment.

It was decided to collect thermal images in the morning during daylight. This decision was justified mostly due to the behavior of target animal species. The ungulate species observed in this study are known to exhibit significant activity during daylight and twilight hours, particularly in spring and summer. They often forage in open habitats such as meadows and forest edges (Baskin and Danell, 2003), where they are more visible to aerial sensors. At night, in contrast, these animals may tend to stay hidden or rest in dense vegetation, which can mask their thermal signatures and reduce detectability. Additionally, at night they may often group together, making it difficult to distinguish individual animals, especially in thermal imagery. For these reasons, the use of daytime thermal imagery can not only improve detection capability but also increase the applicability of the method in open and semi-open environments where ungulates are most active. In addition, UAV flights at daytime are safer and can be performed without additional lights required by law for night flights.

The thermal images were not subjected to additional radiometric calibration to adjust the observed temperatures to the real temperatures of objects. The manufacturer provides individual calibration files for each camera (based on serial number) that allow to achieve temperature measurement accuracy of $\pm 2^\circ\text{C}$ (or $\pm 2\%$ whichever is greater) at an ambient temperature of $23 \pm 5^\circ\text{C}$, however, the calibration conditions are unknown. In addition, the calibration model and raw image intensities are not provided, thus the user cannot perform the calibration. The parameters influencing observed temperature that can be specified by the user are the ambient temperature (in this study measured by the camera sensor) and the radiation emissivity factor (we used the camera default value). The tests showed that the observed temperature is lower than the real temperature and slightly decreases with increasing distance from the object (Figure 2). The camera offers an option for the in-situ calibration if the reference object with known temperature is observed, however, it applies only for static images and cannot be used for UAV images. This makes precise cali-

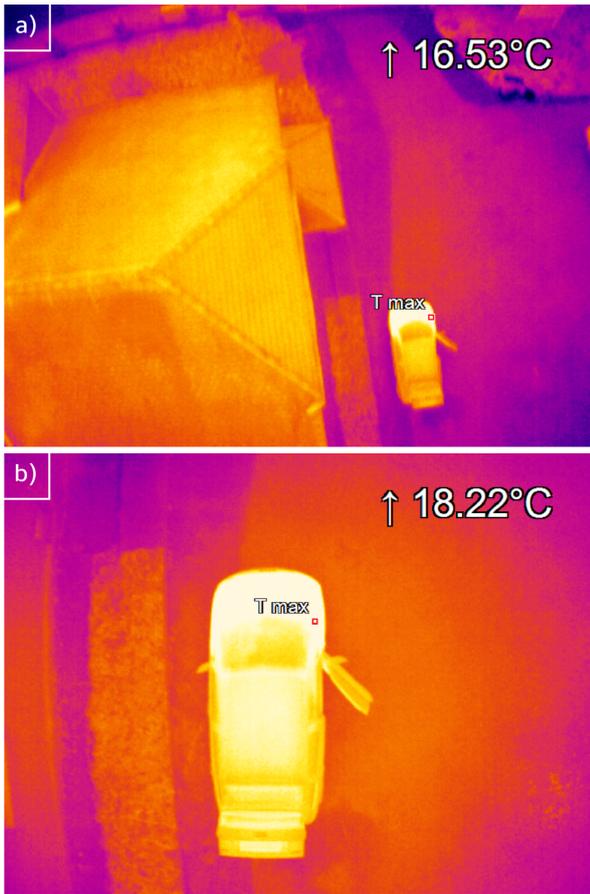


Figure 2. Temperature measurement variability with the distance to the object. Images of the car were collected during the landing procedure from a height of about (a) 30 m and (b) 10 m. The same (hottest) point of a car is marked with a red rectangle, and its observed temperature is displayed in the top right corner of each frame.

bration of collected temperature images extremely challenging and reinforces our decision to develop a detection method that is based on relative thermal contrast, and not in the absolute temperatures of the objects.

3.1 Reference data

The reference data used for BRF model training and evaluation were derived from the first UAV flight. The assignment of images to training, validation, and test datasets was conducted according to flight lines or their parts. Because animals formed two separate groups in the enclosure (Figure 1), such division of the data ensured that each animal was present in one dataset type only. Given that the BRF algorithm operates at the object level, data partitioning was performed after image segmentation (see Section 4.1) rather than at the image pixel level. Each detected image segment showed a single animal or other heated object. These segments were assigned manually to one of two classes: hoofed animals (class 1) or other objects (class 0). Note that each animal visible in many images resulted in many segments. Such segments were treated as separate objects in BRF regardless of whether they presented the same animal or different ones. This approach was used to develop a detection method that is independent of flight plan parameters, especially image overlaps or animal movement. The created model should be able to detect an animal visible in UAV thermal images regardless of the looking angle. The problem

of identifying the same (possibly moving) animal in different images is beyond the scope of this study. However, the large image overlap was helpful during the manual assignment of the object to one of two classes. It was easy to assign a segment to class 1 (hoofed animals) if it was in the central part of the image. Segments in the image edges or corners were more difficult to assign, but checking the segment showing the same animal or other heated object but in a different image helped to prepare reference data. The large image overlap also helped to check if all image segments containing animals were created.

Another drawback of the low resolution of thermal images and lack of RGB images was the inability to distinguish individual species in the collected data. For that reason, the aim of the method was to detect animals belonging to the same biological order that looks similar in thermal images.

The BRF model learned on the training dataset prepared from images collected during the first UAV flight was tested also using the data collected during the second UAV flight to validate the proposed method. The reference data for the second flight was prepared in the same way as for the first flight, but all images collected during the second UAV flight were treated as the test dataset. In addition, in class 0 (other objects – non-hoofed animals) there were marked segments showing other animals, including humans. This allowed us to check for possible false positive assignments of other animals. The summary of the segmentation process and the number of reference objects belonging to each class for both flights are given in Section 5.

4 Methodology

The method used to detect animals in this study consists of two main steps. First, an analysis of pixel temperatures is performed to detect image segments that possibly show animals and to extract segment features. In the second step, these segments are classified using the BRF algorithm to eliminate image segments that do not show animals. A key component of this classification is the set of features used, which includes temperature and geometrical features calculated in the first step, as well as additionally computed CNN features (Figure 3).

The details presented in the sections below describe the key steps in the segment extraction and classification process, with an emphasis on minimizing false detections of hoofed animals.

4.1 Detection of image segments potentially showing animals

Parameters of the equipment used (thermal camera) and the flight plan (flying height equal to about 53 m above ground level) caused the detected animals to appear as small, bright groups of pixels (ranging from several to dozens of pixels), clearly distinguishable from the background. The extraction of these groups of pixels as image segments was executed in three steps:

- i. image preprocessing,
- ii. image thresholding by analyzing neighboring pixel intensities,
- iii. segment extraction and filtering.

The image preprocessing step included converting each image to 8 bits and smoothing. The normalization of temperature images to 8 bits intensity images was executed to accelerate image segmentation process because original images were recorded as 16 bits temperature images (from about -328°C to about $+328^{\circ}\text{C}$ with a step of 0.01°C) which resulted in a low dynamic range. Note that no temperature calibration was ap-

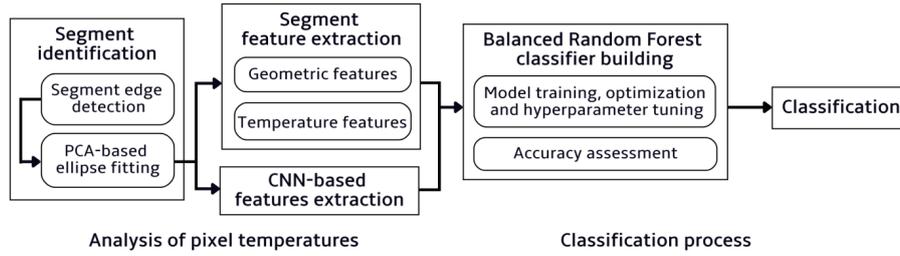


Figure 3. Key steps in the workflow for hoofed animal detection

plied to thermal images, thus exact temperature analysis was impossible. The normalization was applied individually to each image based on the minimum and maximum temperatures recorded in this image. The image smoothing is a standard procedure for noise removal. Similarly to many other studies, a Gaussian filter was used for this purpose as it effectively smooths noise while introducing minimal distortion to the image (Afshari et al., 2017).

The next step was adaptive thresholding (Deng and Cahill, 1993). Since a global threshold often leads to incorrect segmentation under non-uniform lighting conditions, for the adaptive thresholding we used a local Gaussian kernel (size $k = 13$, corresponding to ~ 2 m on the ground) to compute a new image T . The binary image was created based on the intensity differences between the smoothed image I and image T - Equation (1). The threshold value C was selected empirically by visually verifying whether the binary image retained animal pixels in the training data. The threshold was set to -8 to ensure segments containing animals were detected. The parameters used (the size of the Gaussian kernel and threshold value) can be adjusted depending on dataset characteristics such as GSD, animal size, and temperature differences.

$$B(i, j) = \begin{cases} 1, & I(i, j) - T(i, j) > C \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where:

I : smoothed image,

$T(i, j) = \sum_{m=-k}^k \sum_{n=-k}^k I(i + m, j + n) \cdot G(m, n)$: pixel value after applying Gaussian kernel,

C : threshold value.

The contours were generated from binary images by creating closed polygons around groups of pixels (segments) that met the threshold criteria (white pixels in Figure 4). The extracted contours were further analyzed and filtered to identify segments potentially corresponding to hoofed animals (Figure 5a).

The first filtering step focused on the size of the segment area. Only segments whose area fell within a specified range were retained. Specifically, segments with an area between $A_{min} = 5$ pixels and $A_{max} = 60$ pixels were considered. This range corresponds to ground areas of approximately 0.14 m^2 to 1.39 m^2 . The lower threshold allowed for the identification of small segments likely representing partially visible animals, while the upper threshold was based on a manual analysis of the maximum segment size observed for animals in the training dataset. Segments exceeding A_{max} could correspond to non-animal objects such as parts of buildings, tree stems, or exposed soil, and were excluded from further analysis (Figure 5b).

The second filtering step used temperatures from the original images and considered temperature variability within the

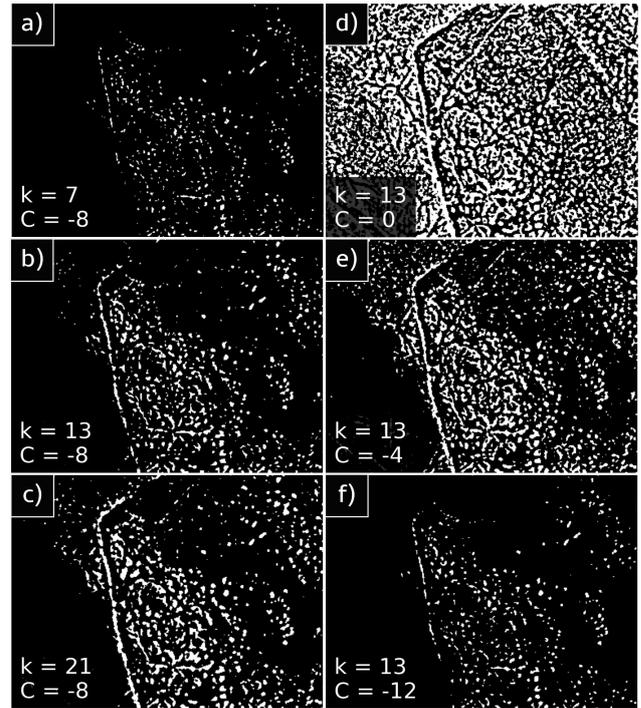


Figure 4. Binary images computed using adaptive thresholding with varying parameters. (a-c) Varying size of Gaussian kernel (k) with fixed threshold ($C = -8$); (d-f) Varying threshold (C) with fixed size of Gaussian kernel ($k = 13$). Panel (b) shows the selected parameters ($k = 13, C = -8$).

segment. Segments corresponding to sun-heated objects (e.g. tree branches, soil) usually do not exhibit significant temperature changes, while animal bodies typically show noticeable temperature variations due to differences in blood circulation, skin thickness, or fur density (e.g., the head being warmer than other, less vascularized, or more insulated body parts). Although this variability is difficult to observe in thermal image visualizations, it can be easily calculated and is practically unaffected by the temperature calibration. In this study, the standard deviation (relative to its median) was used as a measure of temperature variability within the segment. The filtering threshold of temperature variability was empirically set to 0.5°C , which represents the maximum standard deviation that did not filter out segments showing animals in the training dataset.

4.2 Geometric and thermal feature extraction

To describe the resulting segments, an analysis of their geometric and thermal characteristics was carried out. Simple features, such as segment size, shape, and thermal properties,

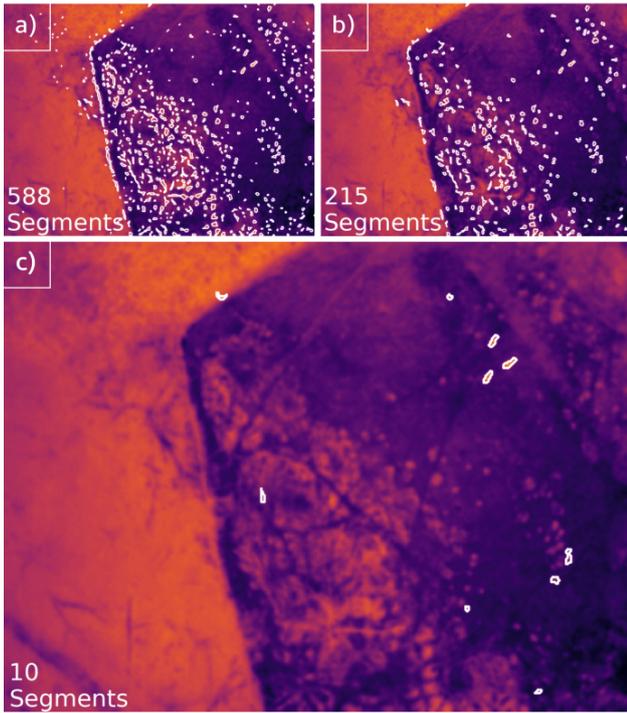


Figure 5. Segment filtering process: (a) all detected segments; (b) segments retained after area-based filtering (5–60 pixels); (c) segments retained after additional filtering based on the standard deviation of temperature from the median within each segment.

provide important information for Random Forest classification. It should be emphasized that these features were determined from the original thermal images, as preprocessing may have altered certain object characteristics, such as shape and temperature distribution.

To reduce the impact of segment irregularities and simplify their shape, each segment was approximated by an ellipse (Figure 6a). This approach was chosen due to the characteristic elongated shape of hoofed animals seen in UAV images. The ellipse seems to be a good generalization of their shape in low-resolution thermal images. The process of estimating ellipse parameters from irregular segment shapes was carried out using Principal Component Analysis (PCA).

Key ellipse parameters, such as semi-axis lengths, area, perimeter, and eccentricity, were determined and used as geometric features (Table 1). Furthermore, based on PCA, additional features categorized as geometric were obtained by calculating mean distance variances of pixels belonging to the segment along the two principal components, as defined in Equation (2). The standard deviations σ_1 and σ_2 were included as geometric features.

$$\sigma_1^2 = \frac{1}{N_c} \sum_{p=1}^{N_c} (z_{p1} - \bar{z}_{p1})^2, \quad (2)$$

$$\sigma_2^2 = \frac{1}{N_c} \sum_{p=1}^{N_c} (z_{p2} - \bar{z}_{p2})^2$$

where:

N_c : number of pixels in the segment,

z_{p1}, z_{p2} : coordinates of pixels on the principal axes p_1, p_2 determined by PCA,

$\bar{z}_{p1}, \bar{z}_{p2}$: average values of the coordinates of all segment pixels along these axes.

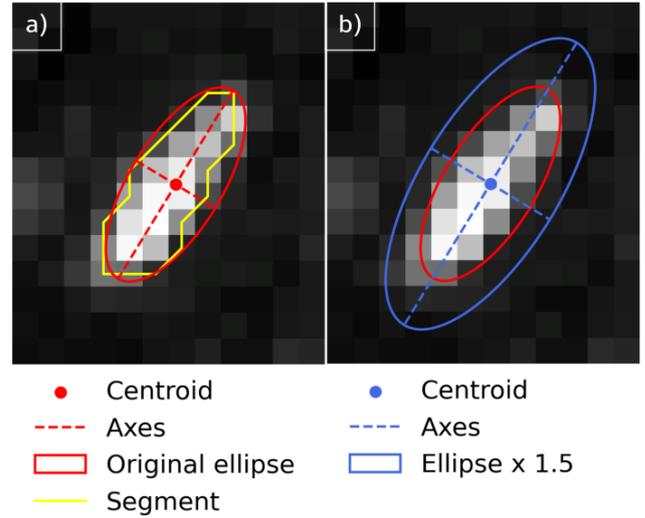


Figure 6. Visualization of a segmented region: (a) original segment and PCA-fitted ellipse; (b) PCA-fitted ellipse and scaled ellipse (Ellipse \times 1.5) defining surrounding area.

Table 1. Summary of geometric features with their equations

Geometric features name	Symbol
Length of major axis	a
Length of minor axis	b
Area	A
Circumference	L
Eccentricity	E
Standard deviation for the first principal component	σ_1
Standard deviation for the second principal component	σ_2

The thermal analysis included only differential temperature features (Table 2) to prevent overtraining the model by limiting the influence of external factors on temperature values. The values were analyzed both within the ellipse and in its immediate vicinity, defined by enlarging the ellipse by 50% of the semi-axis lengths (Figure 6b).

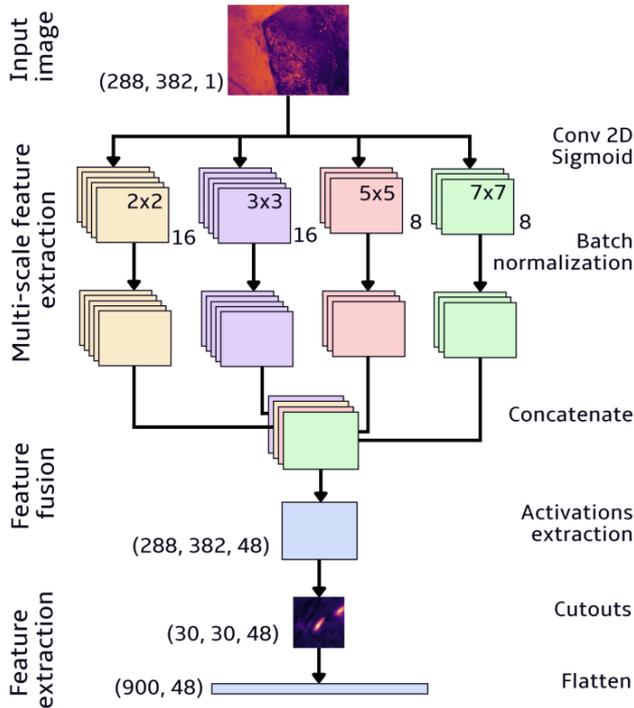
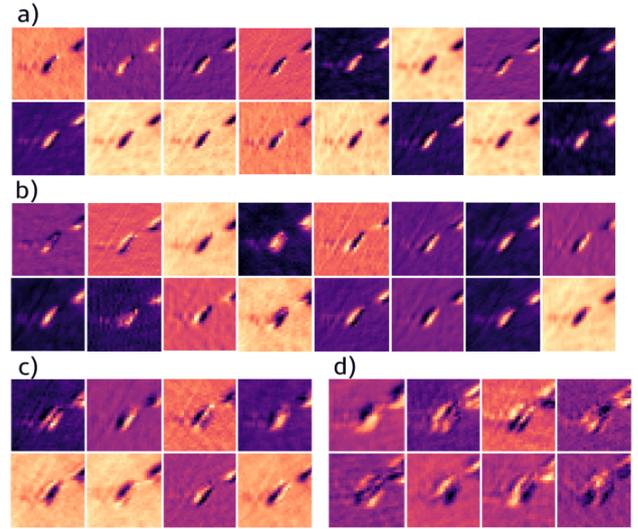
4.3 Convolutional feature extraction

Since the dataset was based on a single flight with a limited number of training samples, the available data could be insufficient for a CNN-only detection approach. Therefore, the CNN was used exclusively to generate additional segment features for the BRFC classification. We used CNNs to capture patterns and textures in images that do not have a direct physical representation, due to their ability to automatically learn hierarchical image features. The network architecture was designed specifically for the analysis of small objects, such as animals.

The CNN used consisted of four convolutional layers, each using filters of different sizes (2×2 , 3×3 , 5×5 and 7×7) (Figure 7). Each convolutional layer contained a specific number of filters: 16 filters for 2×2 , 16 filters for 3×3 , 8 filters for 5×5 , and 8 filters for 7×7 kernel size. This multi-scale approach enabled the network to capture both fine details and more general spatial patterns. The function of the convolutional filter is analogous to the Gaussian filter described earlier, except that, in the case of the convolutional filter, the kernel weights $G(m, n)$ are not predetermined but are optimized during the neural network training process (O'Shea and Nash, 2015).

Table 2. Summary of thermal features with their equations, where: e – ellipse, s – surrounding

Thermal feature name	Equation
Difference in maximum temperatures between the ellipse and the surrounding area	$\Delta T_{max} = T_{max,e} - T_{max,s}$
Difference in minimum temperatures between the ellipse and the surrounding area	$\Delta T_{min} = T_{min,e} - T_{min,s}$
Difference in median temperatures between the ellipse and the surrounding area	$\Delta T_{median} = T_{median,e} - T_{median,s}$
Range of temperatures in the ellipse	$R_e = T_{max,e} - T_{min,e}$
Range of temperatures in the surrounding area	$R_s = T_{max,s} - T_{min,s}$


Figure 7. Convolutional network architecture for multi-scale feature extraction

Figure 8. Examples of activation map cutouts for a single detected segment in different convolution filter sizes: a) 2×2 , b) 3×3 , c) 5×5 , d) 7×7 .

This network was applied to all images in which at least one segment had previously been detected, allowing for the generation of global activation maps. Rectangular areas of 30×30 pixels were then cut out from these maps (Figure 8), based on the coordinates of the centers of the fitted ellipses, allowing the analysis of local patterns to be consistent. These cutouts were subsequently flattened into feature vectors, providing a broader representation of the object. As a final result, each segment corresponded to 48 feature vectors, each consisting of 900 values (Figure 7).

4.4. Balanced Random Forest model building

The segment detection process described earlier results in a large number of segments that mostly correspond to non-animal objects. The number of animals, and consequently the number of segments representing them is much lower. This creates a problem of classifying objects with an unbalanced number of samples in each class. Traditional machine learning algorithms, such as the classic RF, very often fail under such conditions, as they tend to favor the dominant class, leading to incorrect predictions (More and Rana, 2017).

To solve this problem, the BRF algorithm, a modification of classical RF, was used. BRF employs data balancing techniques during sample bootstrapping from the training dataset, automatically adjusting the number of samples in minority classes to the level of the dominant class by oversampling the minority class or undersampling the majority class. This keeps the number of samples of each class considered for building each tree close to each other, helping to build balanced trees without actually changing the sample size. As a result, the algorithm is more effective at predicting the minority class (More and Rana, 2017). The entire process of BRF model optimization and evaluation, including the use of training, validation, and test datasets, is illustrated in Figure 9.

To optimize the model, a subspace of hyperparameters was explored using the Tree-structured Parzen Estimator (TPE) sampler (Watanabe, 2023). The tuned hyperparameters included the number of estimators, the maximum tree depth, the minimum number of samples required to split a node, the minimum number of samples per leaf, and the proportion of samples used to construct each tree. The search was conducted us-

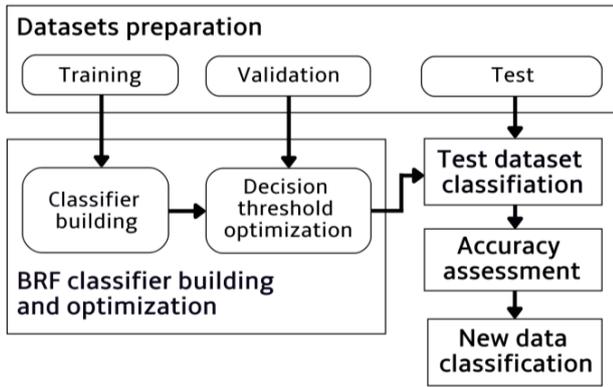


Figure 9. Workflow of the Balanced Random Forest classification

ing an adaptive Bayesian optimization approach to efficiently navigate the hyperparameter space and identify the most effective configuration. The hyperparameter optimization procedure was conducted on the training dataset using cross-validation (Browne, 2000) to determine the best-performing configuration for the BRF model.

The decision threshold was optimized using the validation dataset to account for class imbalance. Instead of relying on the default majority voting in BRF, which may not be optimal for imbalanced classes, a specific threshold was chosen to improve classification reliability. Objects were classified as class 1 if the proportion of votes exceeded this threshold.

4.5 Accuracy assessment

To estimate the accuracy of the model, a classification evaluation metric based on a confusion matrix was used. According to its general concept, it allows counting samples classified correctly (true positive – TP, true negative – TN) and incorrectly (false positive – FP, false negative – FN) (Vujovic, 2021). From these values, key performance measures such as precision, recall, F1-score, overall accuracy (OA) (Vujovic, 2021) and Cohen’s kappa coefficient (Chicco et al., 2021) were calculated.

Overall accuracy is one of the most commonly used metrics, representing the percentage of all correctly classified samples across both classes. Given the imbalanced nature of the dataset, the F1-score was used to balance precision and recall, ensuring a more suitable assessment of model performance in detecting rare cases (Jeni et al., 2013). Additionally, Cohen’s kappa was employed to account for agreement beyond random chance, making it particularly useful for evaluating models where the minority class is crucial (Crow and Watts, 2024).

To further assess the classification performance of the model, we used the Area Under the Curve (AUC) parameters for the Precision-Recall curve (PR AUC) and the Receiver Operating Characteristic curve (ROC AUC). PR AUC is particularly useful for unbalanced datasets, as it highlights the model’s ability to correctly classify the minority class, focusing on the trade-off between precision and recall. On the other hand, the ROC AUC provides a broader measure of distinction between positive and negative classes, illustrating the model’s ability to distinguish between them at various decision thresholds. These metrics complement the traditional evaluation results, offering a more detailed perspective on classification performance under imbalanced conditions (Vujovic, 2021).

Table 3. Summary of detected segments and their class for each dataset

Dataset	Class 0: Other objects	Class 1: Hoofed animals	Total
Training	1515	830	2345
Test	736	110	846
Validation	191	109	300
Total segments	1896	1049	3491

Table 4. Hyperparameter tuning: ranges, steps, and optimal values

Tuned hyperparameters	Range:			Optimal value
	From	To	Step	
Number of estimators	50	1500	50	800
Maximum tree depth	10	50	5	35
Minimum number of samples required to split a node	2	50	1	2
Minimum number of samples per leaf	1	20	1	1
Proportion of samples used to construct each tree	0.5	1.0	0.1	0.6

5 Results

Segmentation of the thermal images from the first UAV flight resulted in 3491 detected segments, which were classified into two categories: hoofed animals (class 1) and other objects (class 0). The high number of class 0 segments is primarily due to the daytime acquisition conditions, which led to many sun-heated objects being detected beside actual animals. The spatial distribution of training, test, and validation datasets is illustrated in Figure 1, while the sample count per dataset is summarized in Table 3. For each detected segment, a total of 60 features were extracted: 7 geometric, 5 thermal, and 48 CNN-derived features.

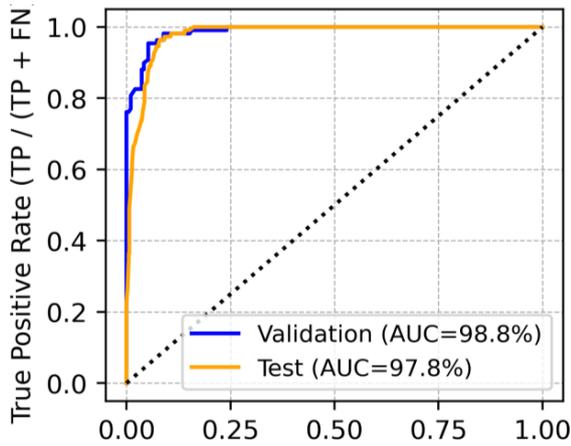
In BRF, during sample bootstrapping, an undersampling strategy was applied to the dominant class to address class imbalance. Samples were drawn without replacement, ensuring that unique data points were used in each iteration. This approach mitigated the risk of overfitting and allowed the model to learn a more balanced representation of both classes. The best-performing BRF classifier was obtained for the hyperparameters given in Table 4 (optimal value). The number of trees and the maximum tree depth were around the median of tested values, while the minimum number of samples required to split a node and the minimum number of samples per leaf were set to the lowest tested values. This indicates that the trees were allowed to grow as deep as necessary, capturing fine-grained decision boundaries in the data. However, since only 60% of the training data was used to construct each tree, individual trees were trained on different subsets, reducing the risk of overfitting while still preserving detailed feature representations.

The decision threshold was optimized on the basis of Cohen’s kappa metric, computed for various threshold values. The model maintained high classification performance across multiple thresholds, with K -values exceeding 80% for decision thresholds between 0.2 and 0.6. To enhance recall while preserving model precision, a threshold of 0.4 was selected, where $K = 88.5\%$.

Feature importance analysis revealed distinct contributions of thermal, geometric, and convolutional features to the classification process. Among thermal features, the most important (19.48%) was the difference in maximum temperatures between the ellipse and the surrounding area (ΔT_{max}), fol-

Table 5. Performance metrics of the model across test datasets from both flights in percentage values

Flight no.	1 (test data)		2	
	0	1	0	1
Precision	99.7	56.2	95.8	86.0
Recall	88.6	98.2	93.6	90.4
F1-score	93.8	71.5	94.7	89.2
OA	89.8		92.7	
K	65.9		82.9	
PR AUC	86.2		91.2	
ROC AUC	93.4		92.0	

**Figure 10.** Receiver Operating Characteristic (ROC) curve for validation and test datasets

lowed by the temperature range within an ellipse (R_e) having an importance of 9.02%. These results highlight the critical role of temperature contrast in detecting warm-bodied objects, as higher thermal variation increases the likelihood of distinguishing animals from other heated elements in the scene. In the geometric feature set, eccentricity (e) had the highest importance (7.02%), reinforcing the role of shape descriptors in classification. This result is consistent with the expectation that hoofed animals have a characteristic elongated or elliptical shape in thermal imagery.

Convolutional features, extracted at multiple spatial scales, had relatively low individual importance, with all values below 1.47%. The most relevant features were obtained using a 3×3 kernel, where the highest-ranked filter reached 1.47% importance, while those extracted with a 2×2 kernel consistently exhibited the lowest contributions, none exceeding 0.95%. Despite their lower individual importance, removing any subset of these features led to a significant drop in classification accuracy, confirming their essential role in capturing spatial patterns within thermal images. Moreover, retaining all convolutional features did not affect significantly computational time, as the relatively small size of thermal images ensured efficient processing.

A comprehensive evaluation of model performance for test datasets is presented in Table 5. For hoofed animals (class 1), the model achieved an F1-score of 71.5%, while the overall accuracy (OA) reached 89.8%. The ability of the model to differentiate between classes was further supported by a ROC AUC of 93.4% (Figure 10) and a PR AUC of 86.2%, while Cohen's kappa (K) reached 65.9%, indicating moderate agreement.

An additional independent dataset from flight 2 was used to further evaluate the model's generalization capability. This flight, conducted on the same day but at a different time, pro-

vided an opportunity to assess the model's performance on previously unseen data under similar, but not identical environmental conditions. Note that observed animals were active during the day and changed their location, though concentrating still in two regions of the enclosure (Figure 1). As in flight 1, the same segmentation parameters were applied in flight 2, and segments were manually assigned to reference classes for evaluation. This ensured consistency in the assessment of classification performance across both datasets.

Despite these variations in animal movement and environmental conditions, the model maintained high classification performance (Table 5), achieving 92.7% OA and an F1-score of 89.2% for class 1 (hoofed animals). The observed K -value of 82.9% suggests improved classification agreement compared to the test data from flight 1. The model also demonstrated a PR AUC of 91.2% and a ROC AUC of 92.0%, confirming its ability to distinguish between target and non-target objects across independent datasets. The improved results for flight 2 can be attributed to the fact that this dataset was collected only over the fenced enclosure, whereas flight 1 also included areas with buildings and other heated objects. The absence of man-made structures in flight 2 reduced the number of false positives, which explains the higher precision and overall accuracy observed in this dataset. It is expected that in the natural environment (without the enclosure) the object content in the scene will be similar to that visible in the data collected during flight 2. Wild animals will rather avoid approaching man-made objects.

Although the classification task was designed to distinguish hoofed animals (class 1) from other objects (class 0), the presence of additional animate objects in the test data influenced classification performance. In flight 1, three humans appeared in the images, along with swans, dogs, and peacocks (Figure 11), while flight 2 contained only a single human and a dog. These instances were too rare to form a dedicated class and lay outside the main scope of detecting hoofed animals, so they were collectively assigned to class 0. However, their presence allowed for an additional misclassification analysis to assess how well the model differentiates target animals from other warm-bodied objects.

The precision for class 1 in flight 1 was notably lower (54.8%) compared to flight 2 (86%), primarily due to the presence of non-hoofed animals that were misclassified as hoofed animals. Swans, recorded on a cooler water surface, were frequently misclassified as class 1 due to their strong thermal contrast with the background. Peacocks and dogs also contributed to classification variability, being inconsistently labeled as either class 0 or class 1. This increased the number of false positives, leading to lower precision for the positive class.

Despite these challenges, inanimate objects were consistently classified as class 0 with high accuracy, and the model correctly distinguished humans from hoofed animals in both datasets. The improved results for flight 2 were likely due to the absence of swans, which reduced confusion between similar thermal signatures. The classification results for these cases are summarized in Table 6.

To further explore the method's performance in more complex, natural conditions, we applied a developed detection model for the data collected in the real environment during different season. Despite more challenging environmental conditions, and without creating a new reference dataset, the model trained using the data collected for the enclosure successfully identified two individual animals (Figure 12). The visual analysis of thermal and RGB images did not show other animals. This result suggests that our approach can work outside controlled conditions and handle different environmental and lighting settings, particularly for diurnal species in semi-open or forest glade landscapes.

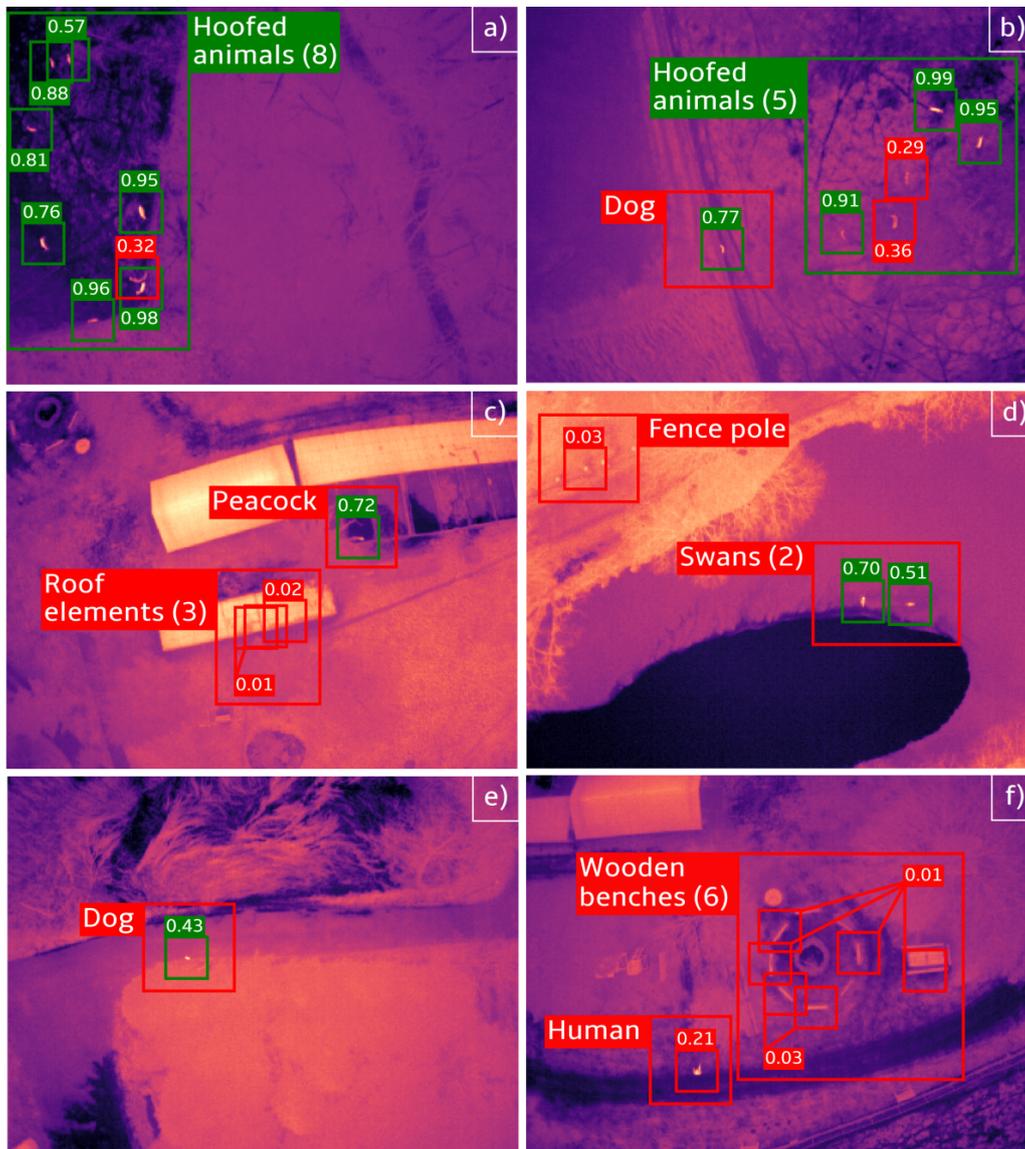


Figure 11. Examples of classification results. Green markers indicate class 1 (hoofed animals) while red markers indicate class 0 (other objects). Small squares highlight detected segments, and the value shows the BRF probability of belonging to class 1. Large rectangles with descriptions highlight real (reference) objects. A square with a color different from the surrounding rectangle indicates a misclassified object.

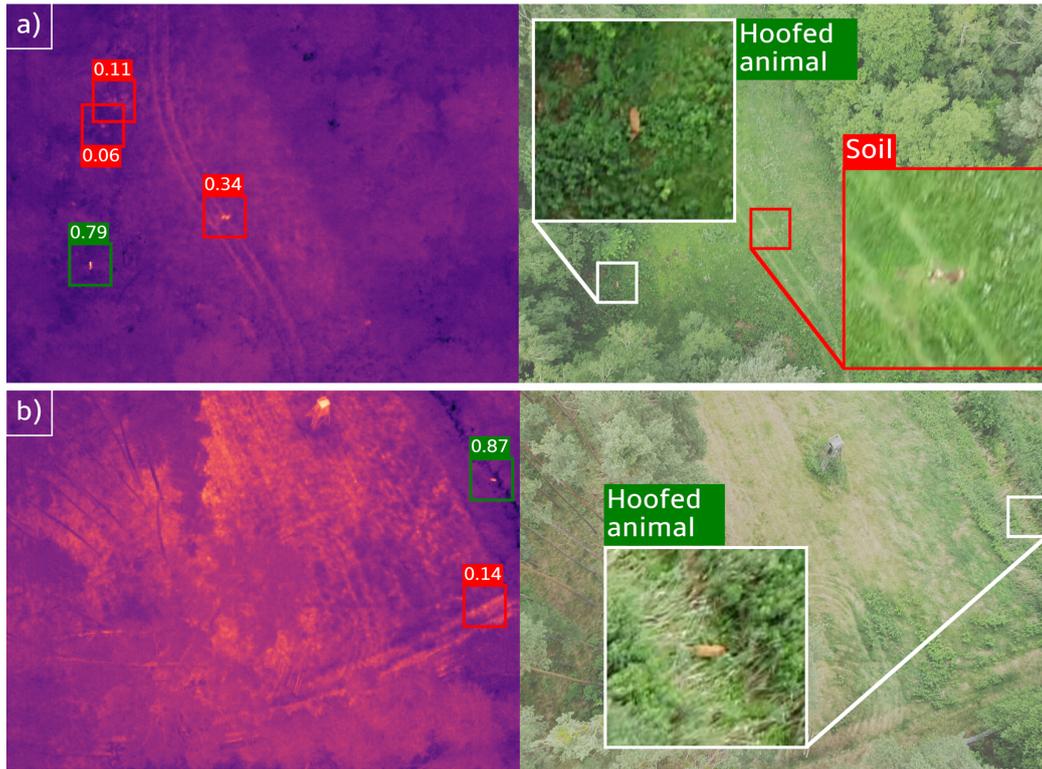


Figure 12. Sample frames from the late-spring forest test flight. Panels (a) and (b) show paired thermal (left) and RGB (right) images. As in Figure 11, green and red markers denote objects classified as hoofed animals (class 1) and non-target objects (class 0), respectively, along with their corresponding probability of belonging to class 1. On the RGB images, rectangles indicate the locations of high-probability detections: white for hoofed animals and red for non-target objects.

Table 6. Classification results for segments representing different objects. Values indicate the number of segments assigned by BRF to classes 0 and 1.

Flight no.	1 (test data)		2		
	0	1	0	1	
Predicted class					
Actual class 1: Hoofed animals	2	108	32	302	
Actual class 0	Inanimate object	592	22	659	12
	Human	39	0	26	9
	Dog	6	12	8	21
	Swan	5	40	-	-
	Peacock	10	10	-	-

6 Discussion

The results indicate that our approach performed well in detecting and classifying animals using only drone-acquired daytime thermal images. The overall accuracy ranged from 89.8% to 92.7%, which is consistent with, or surpasses, results from other machine learning-based wildlife detection methods, where reported accuracies typically lie between 80% and 98% (see Section 2). As discussed earlier, various studies have explored thermal or dual-sensor approaches (Christiansen et al., 2014; Lee et al., 2021; Kalinowski et al., 2023; Rančić et al., 2023). For instance, Lee et al. (2021) combined thermal and RGB images and achieved a best precision of 80.4% and a recall of 69.9% at low altitudes, while our single-sensor method reached a precision of 86% and a recall of 90.4% under comparable conditions. Meanwhile, Kalinowski et al. (2023) achieved 98.6% accuracy for wild boar detection using a thermal camera mounted at an oblique angle, which allowed them to capture side profiles—essential for distinguishing wild boar from other

species. In our case, the thermal camera was mounted vertically (nadir) at an altitude of 53 m, which, together with the image resolution, limits the ability to distinguish species reliably. Instead, the method detects hoofed animals as a broader group. Compared to Rančić et al. (2023), whose YOLOv4 approach on RGB data yielded 86% precision and 75% recall, our pipeline remains competitive and requires fewer labeled examples. Moreover, our method operates on thermal imagery alone and relies on simpler reference data, without the need for bounding box annotation. This reduces the time and effort required for dataset preparation and makes the pipeline more practical in large-scale or low-resource settings. Christiansen et al. (2014) reported a balanced classification accuracy of 84.7% at altitudes below 10 m for small-animal detection, which improved to 93.3% when temporal tracking was applied. Above 10 m, their performance dropped to approximately 75–78%. By contrast, our flights were conducted at a fixed altitude of around 53 m and focused on larger animals. Despite the higher altitude (mostly due to safety reasons to keep sufficient height over tree tops), our method maintained high accuracy. The pipeline combines segment-based detection with a BRF classifier enriched by CNN-derived features. This setup proved to be computationally efficient and well-suited for use with limited training data while maintaining detection performance comparable to more complex approaches. These comparisons suggest that the proposed method is robust and adaptable, even when data collection is limited or conditions vary between flights.

A crucial step in the process was the initial segmentation to detect potential animals in thermal images. Given the thermal characteristics of the scene, animal detection relied on temperature differences rather than structural or color-based features. The applied segmentation method effectively identified potential targets by leveraging temperature-based filtering and con-

tour analysis, ensuring that objects corresponding to animals were correctly extracted while minimizing false detections of static elements such as buildings or heated soil.

Lowering the decision threshold to 0.4 reduced false negatives but increased false positives, particularly for objects with similar thermal signatures. Swans, primarily detected on water, were often classified as class 1 due to the high temperature contrast between their bodies and the cooler surface. Other non-target animals, such as dogs and peacocks, were assigned inconsistently, likely influenced by their size, posture, and partial visibility. The minimum object size of five pixels ensured the detection of partially visible animals but also contributed to variability in classification. While the model effectively differentiates between animals and non-animal objects, its performance decreases for smaller non-target animals.

When applied to an independent dataset from flight 2, the model achieved even higher accuracy than in flight 1. This improvement was likely due to the lower presence of non-hoofed animals, which reduced the number of false positives and improved classification precision. The ability of the model to generalize effectively under similar environmental conditions further supports its applicability for wildlife monitoring using UAV thermal imagery. A supplementary test flight conducted in a forest area during late spring further confirmed that the method can perform under more natural and variable conditions.

Moreover, the use of daytime thermal images provides an additional advantage, as many hoofed animals were active during the day. In contrast, images captured at night could make it harder to detect them, as they might seek shelter or remain hidden in dense vegetation. Although the low resolution of thermal images is a drawback in comparison to high resolution RGB images, it becomes an advantage in terms of the processing speed.

7 Conclusions

The proposed method detects animals in UAV-acquired daytime thermal images using minimal computational resources and limited training data. It combines a straightforward image segmentation routine with a Balanced Random Forest classifier augmented by convolutional features. The approach achieved 89.8% overall accuracy, a Cohen's kappa of 65.9%, and an ROC AUC of 93.4% on the primary test dataset. An additional flight confirmed the method's robustness under different conditions, with performance metrics remaining consistent across datasets collected at different time of the day.

Although the use of low-resolution thermal images collected during daytime was initially perceived as a limitation, it ultimately proved advantageous: employing simple segmentation thresholds and a shallow machine learning model achieved robust performance, while the deep learning component (CNN) was confined to feature extraction and did not require a large dataset. This approach yielded results that could not have been attained using solely the geometric and thermal features derived from the segmented images.

This design eliminates the need for large labeled datasets and enables real-time or near-real-time processing. The image segmentation hyperparameters reflect the ground sampling distance and can be adapted for monitoring other animal species or human presence.

One of the future research plans is to use the developed detection method for an animal counting application. To avoid multiple counting of the same animal visible in different images, a photogrammetric processing may be executed to determine the image exterior and inner orientation. This will allow to determine image overlaps and avoid redundancy in animal

counting. Another future research direction is animal detection with distinguishing species, making it suitable for more detailed ecological monitoring. It should be possible by a simultaneous use of low-resolution thermal and high-resolution RGB images. More broadly, the framework can be adapted to other domains that rely on thermal imagery.

Funding

The APC/BPC is financed/co-financed by Wrocław University of Environmental and Life Sciences.

References

- Afshari, H., Gadsden, S., and Habibi, S. (2017). Gaussian filters for parameter and state estimation: A general review of theory and recent trends. *Signal Processing*, 135:218–238, doi:10.1016/j.sigpro.2017.01.001.
- Baskin, L. and Danell, K. (2003). *Ecology of ungulates: A handbook of species in Eastern Europe and Northern and Central Asia*. Springer Science and Business Media.
- Browne, M. W. (2000). Cross-Validation Methods. *Journal of Mathematical Psychology*, 44(1):108–132, doi:10.1006/jmps.1999.1279.
- Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., and Boutin, S. (2015). REVIEW: Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology*, 52(3):675–685, doi:10.1111/1365-2664.12432.
- Chicco, D., Warrens, M. J., and Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment. *IEEE Access*, 9:78368–78381, doi:10.1109/access.2021.3084050.
- Christiansen, P., Steen, K., Jørgensen, R., and Karstoft, H. (2014). Automated Detection and Recognition of Wildlife Using Thermal Cameras. *Sensors*, 14(8):13778–13793, doi:10.3390/s140813778.
- Crow, L. and Watts, S. J. (2024). Limits to classification performance by relating Kullback–Leibler divergence to Cohen's Kappa.
- De Oliveira, D. C. and Wehrmeister, M. A. (2018). Using Deep Learning and Low-Cost RGB and Thermal Cameras to Detect Pedestrians in Aerial Images Captured by Multirotor UAV. *Sensors*, 18(7):2244, doi:10.3390/s18072244.
- Deng, G. and Cahill, L. (1993). An adaptive Gaussian filter for noise reduction and edge detection. In *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference, NSSMIC-93*. IEEE, doi:10.1109/nssmic.1993.373563.
- Foley, C. J. and Sillero-Zubiri, C. (2020). Open-source, low-cost modular GPS collars for monitoring and tracking wildlife. *Methods in Ecology and Evolution*, 11(4):553–558, doi:10.1111/2041-210x.13369.
- Ford, A. T., Clevenger, A. P., and Bennett, A. (2009). Comparison of Methods of Monitoring Wildlife Crossing-Structures on Highways. *The Journal of Wildlife Management*, 73(7):1213–1222, doi:10.2193/2008-387.
- Hodgson, J. C., Baylis, S. M., Mott, R., Herrod, A., and Clarke, R. H. (2016). Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports*, 6(1), doi:10.1038/srep22574.
- Jeni, L. A., Cohn, J. F., and De La Torre, F. (2013). Facing Imbalanced Data—Recommendations for the Use of Performance Metrics. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, page 245–251. IEEE, doi:10.1109/acii.2013.47.

- Jewell, Z. (2013). Effect of Monitoring Technique on Quality of Conservation Science. *Conservation Biology*, 27(3):501–508, doi:10.1111/cobi.12066.
- Kalinowski, P., Szczepaniak, P., Ulanowicz, L., and Sibilski, K. (2023). Zastosowanie aerofotogrametrii w podczerwieni do śledzenia dzików w ich naturalnym środowisku oraz identyfikacji osobników zarażonych afrykańskim pomorem świń. cz. I – Architektura systemu i metodyka identyfikacji (Application of aerophotogrammetry in infrared to track wild boars in their natural environment and identify individuals infected with African swine fever. Part I – System architecture and identification methodology). *Mechanika w Lotnictwie*, page 165–183, doi:10.15632/ml2022/165–183.
- Lee, E. J., Shin, S. Y., Ko, B. C., and Chang, C. (2016). Early sink-hole detection using a drone-based thermal camera and image processing. *Infrared Physics and Technology*, 78:223–232, doi:10.1016/j.infrared.2016.08.009.
- Lee, S., Song, Y., and Kil, S.-H. (2021). Feasibility Analyses of Real-Time Detection of Wildlife Using UAV-Derived Thermal and RGB Images. *Remote Sensing*, 13(11):2169, doi:10.3390/rs13112169.
- Lyu, H., Qiu, F., An, L., Stow, D., Lewison, R., and Bohnett, E. (2024). Deer survey from drone thermal imagery using enhanced faster R-CNN based on ResNets and FPN. *Ecological Informatics*, 79:102383, doi:10.1016/j.ecoinf.2023.102383.
- More, A. S. and Rana, D. P. (2017). Review of random forest classification techniques to resolve data imbalance. In *2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)*, page 72–78. IEEE, doi:10.1109/icsim.2017.8122151.
- Nyhus, P. J. (2016). Human–Wildlife Conflict and Coexistence. *Annual Review of Environment and Resources*, 41(1):143–171, doi:10.1146/annurev-environ-110615-085634.
- O’Shea, K. and Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, doi:10.48550/arXiv.1511.08458.
- Pollock, K. H., Nichols, J. D., Simons, T. R., Farnsworth, G. L., Bailey, L. L., and Sauer, J. R. (2002). Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics*, 13(2):105–119, doi:10.1002/env.514.
- Ramadhan, A. L. (2024). Understanding Human–Wildlife Interactions in Urban Environments: Implications for Conflicts, Disease Transmission, and Conservation. *Law and Economics*, 18(2):99–109.
- Rančić, K., Blagojević, B., Bezdan, A., Ivošević, B., Tubić, B., Vranešević, M., Pejak, B., Crnojević, V., and Marko, O. (2023). Animal Detection and Counting from UAV Images Using Convolutional Neural Networks. *Drones*, 7(3):179, doi:10.3390/drones7030179.
- Ridwan, Q., Wani, Z. A., Anjum, N., Bhat, J. A., Hanief, M., and Pant, S. (2023). Human–wildlife conflict: A bibliometric analysis during 1991–2023. *Regional Sustainability*, 4(3):309–321, doi:10.1016/j.regsus.2023.08.008.
- Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., Mathis, A., Mathis, M. W., van Langevelde, F., Burghardt, T., Kays, R., Klinck, H., Wikelski, M., Couzin, I. D., van Horn, G., Crofoot, M. C., Stewart, C. V., and Berger-Wolf, T. (2022). Perspectives in machine learning for wildlife conservation. *Nature Communications*, 13(1), doi:10.1038/s41467-022-27980-y.
- Verma, A., van der Wal, R., and Fischer, A. (2016). Imagining wildlife: New technologies and animal censuses, maps and museums. *Geoforum*, 75:75–86, doi:10.1016/j.geoforum.2016.07.002.
- Vujovic, Z. D. (2021). Classification Model Evaluation Metrics. *International Journal of Advanced Computer Science and Applications*, 12(6), doi:10.14569/ijacsa.2021.0120670.
- Watanabe, S. (2023). Tree-structured parzen estimator: Understanding its algorithm components and their roles for better empirical performance. *arXiv preprint arXiv:2304.11127*, doi:10.48550/arXiv.2304.11127.
- Witczuk, J., Pagacz, S., Zmarz, A., and Cypel, M. (2017). Exploring the feasibility of unmanned aerial vehicles and thermal imaging for ungulate surveys in forests – preliminary results. *International Journal of Remote Sensing*, 39(15–16):5504–5521, doi:10.1080/01431161.2017.1390621.
- Witmer, G. W. (2005). Wildlife population monitoring: some practical considerations. *Wildlife Research*, 32(3):259, doi:10.1071/wr04003.
- Woźniakowski, G., Pejsak, Z., and Jabłoński, A. (2021). Emergence of African Swine Fever in Poland (2014–2021). Successes and Failures in Disease Eradication. *Agriculture*, 11(8):738, doi:10.3390/agriculture11080738.
- Yang, S., Gu, L., Li, X., Jiang, T., and Ren, R. (2020). Crop Classification Method Based on Optimal Feature Selection and Hybrid CNN–RF Networks for Multi-Temporal Remote Sensing Imagery. *Remote Sensing*, 12(19):3119, doi:10.3390/rs12193119.