Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6813-6820 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517203

Design of Animal Detection and Control System Based on Image Processing and Ultrasonic Sound Device

Muhammed Atmaca
Department of Software Engineering, Firat University, Elazig, Turkey
Email: is.muhammedatmaca@gmail.com
Mine Arslan
Department of Software Engineering, Firat University, Elazig, Turkey
Email: minearslan541@gmail.com
Yunus Emre Ertürk
Department of Software Engineering, Firat University, Elazig, Turkey
Email: ynusemreerturk@gmail.com
Bilal Alatas
Department of Software Engineering, Firat University, Elazig, Turkey
Email: balatas@firat.edu.tr

---***ABSTRACT***---
In this study, an integrated system combining image processing and ultrasonic sound technologies is proposed to remove unwanted animals from certain areas without causing harm. Using a deep learning-based object recognition algorithm (YOLOv11), animal species such as pigs, dogs and pigeons are detected in real-time with high accuracy. After detection, ultrasonic frequencies specific to each species - 13 kHz for pigeons, 22 kHz for dogs, 29 kHz for pigs - are transmitted to the environment to remove the animals without harming them. The system was trained with open-source datasets and images obtained from the field, and prototyped to run on Raspberry Pi 4B considering portability and energy efficiency. In field tests conducted under real-life conditions, the detection accuracy of the system was measured as 95.5% with mAP@0.5, while the removal success was above 85% on average. With its real-time response time (40 ms) and species-based frequency adjustment, the system offers an ethical, low-cost and environmentally friendly solution that can be used in agriculture, urban security and industrial areas.

Keywords **- animal detection, computer vision, deep learning, ultrasonic sound, YOLOv11.**

Date of Submission: June 19, 2025 Date of Acceptance: August 04, 2025
----

1. INTRODUCTION

1.1. BACKGROUND

The expansion of human settlements into natural habitats leads to an increase in human-animal interactions and problems such as damage to agricultural production, ecosystem imbalance and zoonotic disease risk. Especially in rural areas, controlling species such as pigs, dogs and pigeons has become a critical issue for sustainable agriculture and urban safety [1].

Traditional animal removal methods include tools such as electric fences, gas detonators and audible deterrents, but these solutions are both costly and contradict animal welfare principles [1].

In recent years, the development of artificial intelligence and image processing technologies has made it possible to develop autonomous and animal-friendly systems. Among deep learning-based object recognition algorithms, the YOLO (You Only Look Once) architecture stands out especially in terms of real-time object detection. Versions such as YOLOv3 have become applicable on mobile devices by offering higher accuracy and lower latency with less computational power [2].

The fact that the hearing systems of animal species differ from those of humans allows for the use of species-based ultrasound technologies in animal control. Research shows that dogs respond to sound frequencies up to approximately 60 kHz and pigs up to 40 kHz [3], [4].

In the studies developed in line with this information, it is aimed to deter animals with ultrasonic sound without harming them. In this study, as a result of the field tests, the optimal frequencies for each species were determined and the system design was realized accordingly. Accordingly, 13 kHz for pigeons, 22 kHz for dogs and 29 kHz for pigs were emitted at frequencies specifically defined for each species.

Image processing-based animal recognition and accompanying species-based ultrasonic sound emission work together to provide an environmentally friendly, ethical and effective animal control mechanism. This approach enables the development of real-time and scalable systems that include not only detection but also intelligent intervention.

1.2. PROBLEM STATEMENT AND SIGNIFICANCE

The presence of unwanted animals in agricultural areas, parks, campuses and urban peripheries causes serious problems such as damaging crop production, destroying infrastructure equipment and in some cases even threatening human safety. Species such as pigs, dogs and pigeons can enter these areas, especially with the motivation to access food resources, which has economic and social impacts [5].

Traditional removal methods include techniques such as electric fences, gas detonators and sound repellents, but these solutions are often species-indiscriminating, passive and can negatively impact both non-target animals and the environment [6]. At the same time, these systems have high energy consumption, high maintenance costs and, in many cases, low site-specific efficiency. Furthermore, the ethical controversy of these methods is a major limitation.

At this point, the main problem is that current animal removal systems are neither able to correctly distinguish between animal species, nor to perform species-specific interventions. This deficiency results in both ineffective deterrence and practices contrary to animal welfare. Therefore, what is needed is a system that works in real time, recognizes the animal species through image processing, and deters it harmlessly by producing ultrasonic sound at a frequency specific to the detected species.

The importance of this study lies in the fact that it offers an ethical, environmentally friendly and low-cost solution that enables species detection with high accuracy using image processing and deep learning technologies, followed by the removal of animals with ultrasonic frequencies specific to their hearing range without harming them. In this way, it will be possible both to prevent agricultural losses and to manage human-animal interactions more safely.

1.3. Objective of the Study

The main objective of this study is to develop a real-time system that uses integrated image processing and ultrasound technologies to remove unwanted animals (such as pigs, dogs, pigeons) from certain areas without causing harm. With the help of deep learning-based object recognition algorithms, target species are detected with high accuracy and ultrasonic sounds are transmitted to the environment at frequency ranges determined in accordance with the auditory sensitivity of each species.

For this purpose, a solution that does not harm animals, is environmentally friendly, energy efficient and can work on mobile hardware has been designed. The developed system is aimed to provide a more ethical and sustainable alternative to existing conventional methods with its field applicability, reaction time, removal success and user-friendly structure.

1.4. Contributions

This study presents an innovative system that integrates image processing and ultrasound technologies to remove animals without harming the environment and themselves. Thanks to the deep learning-based YOLOv11 architecture, different species such as pigs, dogs and pigeons are accurately detected in real time, and a target-oriented and selective removal mechanism is developed by emitting ultrasonic sound at frequencies specific to each species. In this respect, unlike existing systems, an ethical approach has been adopted that does not intervene in all living things, but only in the identified species.

The developed system offers a portable, low-cost and energy efficient solution not only in terms of software performance but also at the hardware level. It can run on a Raspberry Pi 4B, which provides a significant advantage in terms of practical field adaptability by enabling installation in large areas. In addition to the animal hearing threshold data reported in the literature, effective deterrence was achieved by optimizing the frequencies as a result of field tests specific to each species.

The system contributes to the literature as an environmentally friendly, scalable and sustainable solution alternative that can be used effectively in all living areas that should be kept closed to animal entry, such as agricultural lands, campus areas, parks and industrial zones.

2. Literature Review

2.1. Image Processing And Deep Learning For Animal Detection

Recent advancements in image processing technologies have significantly improved the detection of animals in various environments. Particularly, deep learning-based object recognition algorithms have enabled highly accurate real-time animal detection. The development of these algorithms has facilitated novel applications in animal control and management. Among the object detection architectures, the YOLO model stands out due to its single-stage detection approach, offering both high speed and precision. YOLOv3 and its successors have become deployable even on mobile devices with limited computational resources [2]. This feature is especially important for the widespread adoption of animal detection systems in rural areas and regions with limited infrastructure.

In training deep learning models, dataset diversity and class balance are critical factors that enhance the model's generalization capacity. In particular, for animal species classification tasks, data augmentation techniques and class balancing strategies improve accuracy and help prevent overfitting [7]. These techniques have been shown to positively impact model performance, especially when working with limited datasets.

It has been observed that the performance of animal detection systems is affected by camera positioning and environmental conditions. Detection accuracy tends to decrease under low-light conditions and in complex backgrounds, although data augmentation methods can partially mitigate these challenges. Studies conducted in agricultural areas have emphasized that traditional deterrent methods (e.g., electric fences, acoustic deterrents) are not only costly but also conflict with animal welfare principles [1].

In a comprehensive review addressing the development and challenges of deep learning-based object detection, the trade-off between speed and accuracy in real-time applications was highlighted, and various architectural strategies were compared to optimize this balance [8]. This research demonstrated that general-purpose object detection models, when fine-tuned for specialized applications such as animal detection, can achieve high performance even with limited datasets.

2.2. Ultrasonic Sound Technologies For Animal Deterrence

Ultrasonic sound technologies are widely used to repel animals in a non-harmful manner. Operating at frequencies above the human hearing threshold, these systems produce more effective results when customized according to the auditory sensitivities of different animal species. Studies examining the behavioral impact of ultrasonic waves on animals have shown that systems using fixed frequencies tend to lose effectiveness over time, whereas those incorporating dynamic frequency shifts maintain their deterrent effects for longer periods [9].

This finding indicates that animals tend to habituate to repetitive stimuli, and adaptive approaches are necessary to sustain the effectiveness of deterrent systems. Scientific studies have confirmed that ultrasonic sound waves do not cause permanent harm to animals, yet they are capable of inducing effective avoidance behavior. This technology aims to repel animals through behavioral response rather than physical barriers, particularly in agricultural and urban environments.

2.3. Integrated Systems For Animal Detection And Control

Systems integrating image processing and ultrasonic sound technologies offer more intelligent and selective solutions for animal control. This integrated approach enhances effectiveness while minimizing the impact on non-target species by applying species-specific intervention strategies.

In a study focusing on IoT-based animal monitoring, a wireless sensor network was employed to detect and report animal intrusions in agricultural areas [10]. The system proposed an energy-efficient architecture suitable for long-term operation, contributing significantly to the literature. Another study on a fog computing-based smart agriculture system developed a mechanism that provided real-time alerts to farmers and enabled automatic interventions by detecting animal intrusions [6]. In this approach, sensor networks were integrated with image processing technologies to ensure prompt detection and processing of animal presence in agricultural zones.

A study on a MobileNet SSD-based scarecrow monitoring system proposed a low-cost solution that detects animal presence and activates ultrasonic deterrents in agricultural fields [5]. This system can produce distinct ultrasonic frequencies based on the detected species, thereby enabling species-specific deterrence strategies.

Another important advantage of integrated systems is their energy efficiency and long operational lifespan. While traditional deterrent systems operate continuously, those triggered by image processing become active only when animal detection occurs, significantly reducing power consumption.

In line with these efforts, a recent study demonstrated the use of lightweight CNN models combined with IoT and data analytics for accurate real-time detection of agricultural anomalies. The system achieved detection rates exceeding 98%, underscoring the viability of integrated smart agriculture solutions in low-resource settings [11].

3. Methodology

3.1. Data Set

In this study, a specially prepared image dataset compiled from different sources was used to detect unwanted animals (pigs, dogs, pigeons) in agricultural and urban environments through image processing. The purpose of the dataset is to create a learning environment based on various scenarios to enable the developed deep learning model to operate with high accuracy in different environmental conditions.

3.1.1. Data Sources

The image dataset used in this study is structured in such a way that pig, dog and pigeon species in different environmental conditions can be detected with high accuracy. The dataset was compiled from both academic open databases and real images obtained from the field.

Open access image datasets were used as the first source. In particular, images from widely accepted databases such as Google Open Images Dataset and Animals with Attributes 2 were integrated into the model with class labels and coordinate information corresponding to the target species. The data from these sources were converted into YOLO format and made directly usable for training.

Second, actual site-specific images were included in the dataset. These images were captured through fixed and moving cameras in natural habitats in agricultural areas, rural areas and urban environments. The real-world images were used to increase the robustness of the model to environmental variations with varying lighting conditions, different backgrounds and various camera angles. Furthermore, this data is critical to test the practical applicability of the system.

Images from these two main sources were combined to optimize visual diversity, class balance and realism, and were subjected to pre-processing steps such as quality control, resizing and normalization before being used in training.

3.1.2. Labeling Process

The image and audio data used in this study have a multimodal structure for animal detection. The image data were labeled using the Bounding Box method in accordance with the YOLOv11 algorithm, and each animal class (e.g. "dog", "boar", "pigeon") was associated with the label it belongs to. The audio data was categorized according to the characteristic animal sounds. Labeling was performed with the help of open-source tools such as Roboflow Annotate and LabelImg.

In order to increase accuracy and minimize observer errors in the labeling process, a double observer method was applied and the labels were verified by cross-checking. This approach improves the labeling quality and provides highly reliable data for model training. A recent study in the literature has shown that the double observer-based labeling strategy significantly reduces false positive rates and improves accuracy [12].

3.1.3. Visual Diversity and Balance

In order to increase the generalizability of the model, the dataset contains a balanced set of images with a variety of environmental and spatial conditions. These include:

- Different camera angles (high, side, fixed, moving),
- Changing light conditions (day, night, low light),
- Different background environments (field, woodland, city park),
- There are various distance and size scales (distant/non-distant object locations).

This diversity not only increases the richness of the training data, but also ensures that the model produces consistent results in different scenarios. In addition, balanced sampling across classes prevents learning bias due to over-representation of certain classes. A study in the literature has shown that data diversity and class balance strategies prevent overfitting and improve model accuracy, especially in animal species classification applications [13]. The number of images for each animal class in the dataset is presented in Table 1.

TABLE 1 DISTRIBUTION OF IMAGES PER ANIMAL CLASS IN THE DATASET

Class Name	*Number of Images*
Pig	3996
Dog	3452
Pigeon	3332

3.1.4. TRAINING/TEST SPLIT

In order to train the model effectively and evaluate its performance accurately, the dataset was randomly split into 80% training and 20% testing. This ratio is in line with the training/test splitting strategies proposed in the literature and is reported to be effective in increasing the accuracy and generalizability of the model [14], [15]. In addition, 10% of the training data were allocated as a validation subset in order to optimize the hyperparameter settings of the model and prevent overfitting.

As a result, 5512 out of 9646 images were assigned to training, 2756 to validation and 1378 to test subsets. This configuration reduces the problem of unbalanced sampling during the model learning process and contributes to more reliable results in the evaluation process.

3.2. DEEP LEARNING MODEL: YOLO ALGORITHM

In the field of image processing and object detection, models that can provide both accuracy and speed performance in real-time applications are of great importance. The YOLO algorithm is one of the most common examples of single-stage object detection architectures developed to address this need. Especially the latest version, YOLOv11, stands out with its architectural innovations and resource efficiency compared to previous versions [16].

The YOLOv11 architecture offers a more flexible location estimation system by eliminating traditional fixed anchor boxes with its anchor-free structure. In addition, the decoupled head approach, which performs classification and regression operations in different network branches, provides advantages in terms of accuracy and overall learning stability [16]. Thanks to these architectural updates, the model exhibits significant performance improvements in both small object detection and multi-class decomposition.

YOLOv11 is available in different sizes (nano, small, medium, large), making it applicable in both high-performance server environments and embedded systems. For example, YOLOv11m achieved 51.5% mAP@0.5-0.95 performance on the COCO dataset, while achieving a 22% reduction in the number of parameters and a 19% increase in inference rate [17].

Some important hyperparameters used in the training process of the model include 640×640 pixel input size, 25 epochs, 64 batch size and 0.001 learning rate. In the model optimized with CIoU + BCE loss functions, data augmentation techniques such as mixup, mosaic, horizontal projection are also applied to reduce the risk of overfitting [16].

Comparative studies in the literature show that YOLOv11 outperforms predecessor models (YOLOv8, YOLOv10) and alternative architectures (Faster R-CNN, RetinaNet) in both object detection accuracy and inference time. In an experimental study conducted by Sapkota et al., it was reported that YOLOv11 achieved higher F1 scores and overall success than previous versions in small object detection in complex environments, especially in fruit classes [18].

In conclusion, these advanced structures and performance increases offered by the YOLOv11 architecture provide a powerful solution for real-time and low resource consumption systems, such as Raspberry Pi based animal detection and removal projects.

3.3. MODEL TRAINING

In this study, a deep learning model built on the YOLOv11 architecture is trained to detect and then remove animals such as pigs, dogs and pigeons using species-specific ultrasonic signals without harming the environment. The training process includes data preprocessing, hyperparameter optimization, model training and performance evaluation.

3.3.1. DATA PREPROCESSING

Before training the model, all labeled images were resized and normalized to 640×640 pixel resolution in accordance with the YOLO format. In order for the model to work reliably under different environmental conditions, various data augmentation techniques were applied. These include random rotation, horizontal reflection, brightness and contrast adjustment, color distortion and Gaussian noise addition. This diversification strategy enabled the model not only to memorize the training data, but also to learn structures that can be generalized to a wider range of scenarios.

Recent studies show that data augmentation methods significantly improve model performance and reduce the risk of overfitting, especially in image classification problems. In addition, these techniques are also widely used to increase the success rate in classes with data imbalance [7].

3.3.2. HYPERPARAMETER SETTINGS

Correctly configuring the hyperparameters during the training process of the model is one of the key factors that directly affect the success of deep learning models. The hyperparameter settings for the YOLOv11 architecture used in this study were determined based on experimental accuracy observations and recommended practices in the literature [7]. These settings, which are presented in detail in Table 2, ensured that the model was balanced in terms of both accuracy and overall performance.

TABLE 2 Hyperparameter settings used during model training

Parameter	***Value***
Batch size	64
Number of epochs	25
Learning rate	0.001429
Optimization	AdamW
Image dimension	640×640
Loss function	CIoU + BCE + DFL
Augmentation	Active (RandAugment)

In particular, the optimization algorithm used, AdamW, increases the generalization capability of the model and stabilizes the learning process by applying weight decay in a more controlled manner [19]. This algorithm is often preferred for complex tasks such as object detection.

During model training, the transfer learning method was applied, using YOLOv11 weights that were previously trained on the COCO dataset. This approach allowed the model to learn low-level visual features (edges, corners, textures) in advance, significantly reducing the training time [8]. During transfer learning, only the last layers of the model were retrained, while the weights in the previous layers were fixed to preserve information transfer.

One of the most important benefits of this strategy is that it is a very effective method to reduce the risk of overfitting in projects working with limited data, to shorten the training time and to increase the generalization capacity of the model [20]. In this study, only the last layers are updated and the other layers are frozen to preserve the information transfer. This approach makes the training process more efficient and preserves the quality of the generalizations learned.

Furthermore, an "early stopping" strategy was applied to reduce the risk of overfitting during the training process. When the performance of the model on the validation set did not improve for a certain number of epochs, the training process was automatically terminated. This method not only made the training of the model time-efficient, but also greatly reduced generalization errors due to overlearning [21].

Thanks to this integrated approach, the YOLOv11 model obtained within the scope of the project not only achieved high object detection accuracy but also completed training in a short time. The augmentation techniques applied during the training process (especially RandAugment) have also made the model more flexible by increasing data diversity [7].

3.3.3. TRAIN PROCESS

The model was trained on a custom dataset divided into 70% training, 20% validation and 10% testing. The training process was carried out with PyTorch 2.0 infrastructure using NVIDIA L4 GPU accelerator in Google Colab environment. This environment provided a sufficient infrastructure for training large-scale deep learning models in terms of both processing power and memory capacity.

At the end of each epoch during training, metrics such as training loss, validation loss, mean accuracy (mAP) and class-based accuracy were recorded to evaluate the learning performance of the model. The overall performance of the model was monitored based on these metrics, and guidance was given based on the validation performance throughout the training process.

Furthermore, an early stopping strategy was applied in the training. When the performance of the model on the validation set did not improve over consecutive epochs, the training process was terminated. This practice was preferred to prevent overfitting and to optimize the training time. As a matter of fact, studies show that stopping early prevents the model from memorizing the training data and contributes to better generalization [22]. The training was structured over 25 epochs and the weights with the best validation performance were automatically saved.

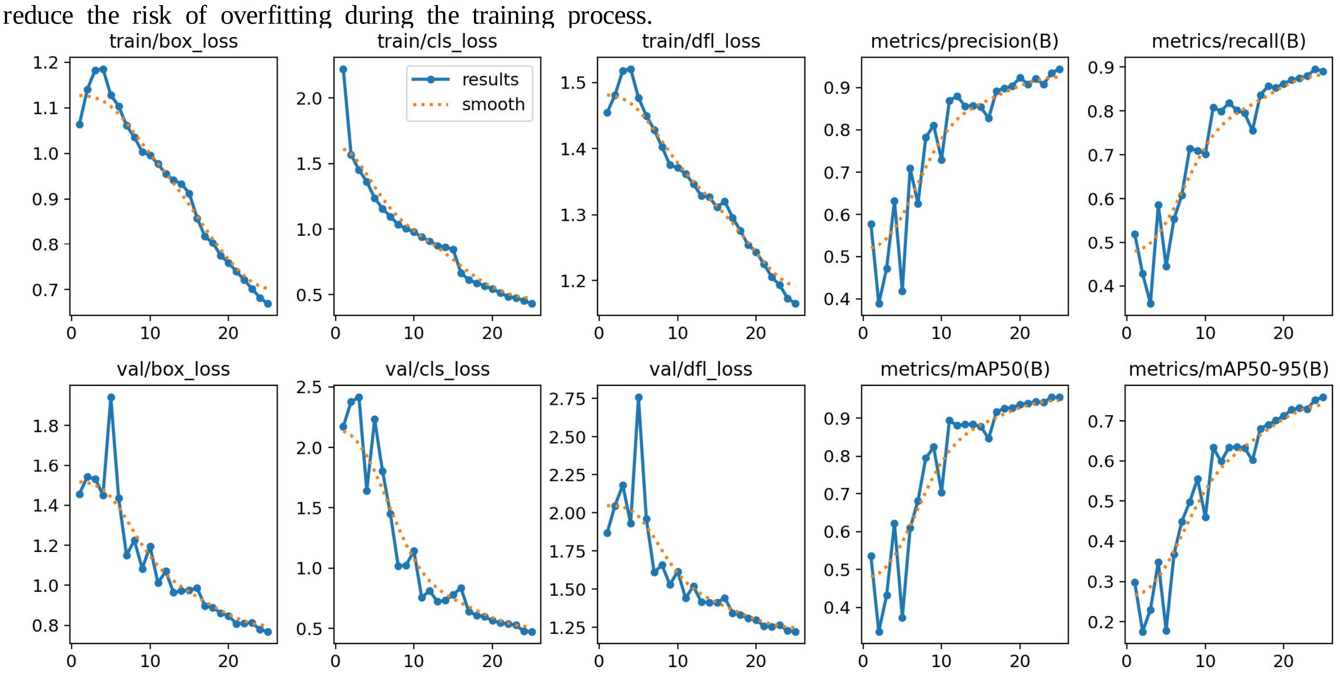

Figure 1. Epoch-wise training and validation losses (box, cls, dfl) and performance metrics (precision, recall, mAP@0.5 and mAP@0.5-0.95) observed during the model training process.

Int. J. Advanced Networking and Applications
Volume: 17, Issue: 02, Pages : 6813-6820 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.202517203

3.3.4. PERFORMANCE EVALUATION

The accuracy and performance of the model were evaluated using commonly used metrics such as mAP@0.5, precision, recall and F1 score. mAP@0.5, defined as the average precision value, measures whether the object is correctly detected with a 50% overlap rate. Precision is the ratio of correct positive predictions to all positive predictions, while recall is the ratio of correct positive predictions to total true positive samples. The F1 score, calculated as the harmonic mean of these two values, is used as a summary indicator of overall model success.

To provide a more comprehensive understanding of the model's learning dynamics, epoch-wise variations of training and validation losses (box, cls, and dfl) as well as key performance metrics (precision, recall, mAP@0.5, and mAP@0.5:0.95) were monitored and visualized throughout the training process. As shown in Figure 1, the model exhibited a consistent decrease in all loss components while steadily improving in evaluation metrics across epochs. Notably, the convergence trend observed after epoch 15 supports the effectiveness of the applied early stopping strategy and highlights the model’s stability in learning.

TABLE 3 Performance metrics of the trained YOLOv11 model on the test dataset.

Metric	*Value (%)*
mAP@0.5	95.5
Precision	96.0
Recall	92.7
F1 Score	92.0

The main results obtained at the end of the training process are presented in Table 3. Accordingly, the model's mAP@0.5 value is 95.5%, precision is 96.0%, recall is 92.7% and F1 score is 92.0%. These results show that the trained YOLOv11-based model can detect animal species such as pigs, dogs and pigeons with high accuracy and offers a successful performance in real-time applications. Especially when various scenarios such as different lighting conditions, camera angles, and whether the animal is moving or stationary are taken into account, the model provides consistent results.

The results show that the trained YOLOv11-based model can detect animal species such as pigs, dogs and pigeons with high accuracy and provides reliable performance in real-time applications. In particular, the model provides consistent results in various scenarios, including different lighting conditions, camera angles, and whether the animal is moving or stationary.

3.3.5. Hardware and System Configuration

The real-time animal detection and removal system developed in this study is configured using low-cost, portable and energy efficient hardware. All hardware components of the system are selected according to their tasks and integrated in a way that they can operate reliably in the field. The developed system has been configured by considering criteria such as portability, energy efficiency and field operability. All hardware components used were selected in accordance with their tasks and are presented in detail in Table 4.

TABLE 4 Hardware components used in the real-time animal detection and deterrence system.

Hardware	*Value (%)*
Processor Card	Raspberry Pi 4 Model B
Camera	Raspberry Pi Camera V3
Speaker	Piezoelectric Ultrasonic Speaker

Raspberry Pi 4 Model B was preferred as the main unit carrying the processing load of the system. This board is an ideal platform for the system as it has enough processing power to run image processing algorithms and can communicate directly with peripherals such as voice warning systems via GPIO pins. In addition, its low power consumption (~5W) and compatibility with Linux-based operating systems offer significant advantages in portable applications. The optimized version of YOLOv11 running on the system is tuned to work effectively on this platform despite limited resources.

The Raspberry Pi Camera V3 used as the image source has a resolution of 12 megapixels and autofocus support. Thanks to its high refresh rate (1080p @ 50FPS), it provides sufficient frame rate for capturing moving targets and performs well in low light conditions. This camera connects directly to the Raspberry Pi via the MIPI CSI interface, allowing low latency image transmission.

A piezoelectric ultrasonic speaker is used in the system to ensure the removal of animals after detection. This speaker operates in the range of 10-42 kHz and offers a warning mechanism suitable for the hearing threshold of species such as pigs, dogs and pigeons thanks to its center frequency of 40 ± 5 kHz. Operating with a nominal voltage of 10VP-P, the device is triggered by a software-controlled PWM signal via GPIO. Thus, a frequency range specific to each animal species can be produced and the system can dynamically intervene in different species. Thanks to its ABS body, it is both lightweight and resistant to field conditions.

All components were selected based on the criteria of energy efficiency, hardware compatibility, cost-effectiveness and ease of implementation in the field. The entire system is configured to work with portable power supplies (e.g. power banks), making it suitable for both stationary and mobile use scenarios.

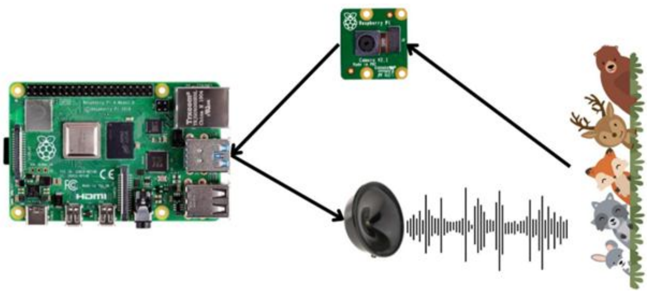

Figure 2. General architecture of the real-time animal detection and deterrence system.

The general architecture of the developed system is shown in Figure 2. The image data received from the environment through the camera module is analyzed with the YOLOv11 algorithm running on the Raspberry Pi; according to the detected animal species, an ultrasonic signal is generated at a defined frequency and transmitted to the environment through a piezo speaker. The hardware infrastructure of the system is configured to enable low-power and portable embedded systems to function in the field. In this context, image acquisition, processing and decision-making operations are centrally carried out on the Raspberry Pi 4 Model B. Image streaming is provided by the Raspberry Pi Camera V3, which is directly connected via the CSI (Camera Serial Interface) port, and visual data from the environment is continuously processed through this camera. With 12 MP resolution and effective performance even in low light conditions, the camera enables clear analysis of moving targets.

The acquired images are analyzed by the YOLOv11 algorithm running on the Raspberry Pi, which is able to classify the targeted animal species (pig, dog and pigeon) with high accuracy thanks to pre-trained weights. Each detected class is associated with a specific output action defined in the system. Based on the classification result, the piezoelectric ultrasonic speaker on the hardware is activated.

The speaker is controlled via GPIO 18, one of the Raspberry Pi's GPIO pins. It should be noted that this pin was chosen because it has the hardware capability to generate a PWM (Pulse Width Modulation) signal. The frequency of the signal sent to the speaker is dynamically determined by the software and adjusted to operate at a frequency specific to the type of animal detected. Thus, the system aims to remove the target animal from the environment by generating a different acoustic warning for each species.

The connection between the speaker and the Pi is not made directly, and the GPIO output is fed through a MOSFET driver circuit for safe driving. Although the device has a nominal voltage of 10V, its low-current operation and high frequency generation (40 ± 5 kHz) make it an efficient ultrasonic repeller that can be used in the field. The power and ground lines of the system were provided through the 5V and GND pins on the Raspberry Pi, thus making the entire system feedable from a single source.

Thanks to this structure, after receiving and analyzing the image, the system generates a warning at a frequency suitable for the target species with a short delay and the animal is targeted to move away from the environment. Since the entire processing chain runs in a closed circuit on the Raspberry Pi, the system maintains its portability and the ability to operate independently in the field. This integrated structure offers high reliability and modularity in applications that require both real-time data processing and physical interaction.

3.3.6. DATA SET REAL-TİME APPLİCATİON

After model training and performance analysis, real-time application scenarios were developed to test the functionality of the system under real-world conditions. By analyzing the live camera image, the system instantly detects the animals in the environment and reacts by generating an ultrasonic signal specific to the detected species. The main objective of the application is to demonstrate that the system works as an environmental response mechanism that can make fast and accurate decisions not only in the laboratory environment but also in the field.

The system architecture is structured to process the snapshot stream of the camera on the Raspberry Pi through a Python-based application with RTSP (Real Time Streaming Protocol) support. Each image frame acquired by the camera is analyzed by the optimized YOLOv11 object detection model; when target objects (pig, dog, pigeon) are identified, the system triggers an ultrasonic signal at the appropriate frequency via GPIO in line with the class label. This process is completed with a delay of seconds, preserving the system's real-time reflex capability.

The real-time implementation also provided important findings in terms of testing the efficiency of optimized models, especially in embedded systems with low hardware capacity. Scaling the camera images to 640x640 resolution speeds up the inference process, while the limited number of detected objects protects the model from unnecessary processing overhead. The metrics obtained during the implementation were consistent with the accuracy in the test scenarios, and sufficient detection performance was achieved, especially in low light conditions.

As a result of this application, it was observed that the developed system works reliably even under different camera positions, various lighting scenarios and moving targets. These findings show that the model can be adapted to areas such as agriculture, campus security and wildlife control where real-time environmental intervention is required. Moreover, the scalable implementation of the system on low-cost hardware provides a significant advantage in terms of practical field deployability.

4. CONCLUSION

In In this study, a real-time, ethical, and environmentally friendly system that combines image processing and ultrasonic sound technologies for animal detection and deterrence has been developed. Using the deep learning-based YOLOv11 architecture, different species such as pigs, dogs, and pigeons were detected with high accuracy. After detection, species-specific ultrasonic frequencies were emitted to deter animals without causing physical harm. The system was built with portable and low-cost hardware components and was functionally tested in field conditions.

Real-world tests demonstrated that the system achieved a detection accuracy of 95.5% and an average deterrence success rate of over 85%. The proposed structure stands out not only for its technical performance but also for its emphasis on ethical considerations and sustainability.

When compared to traditional animal deterrent systems, the proposed solution offers several distinct advantages. Electric fencing, while effective, often involves high energy consumption (ranging from 1.5 to 4.5 kWh/day) and poses a risk of physical harm to animals [1]. Additionally, systems such as gas detonators tend to lose their effectiveness within two weeks due to animal habituation, with deterrence rates falling below 70% [9]. These conventional systems are generally non-selective and can impact non-target species and the environment negatively.

In contrast, the system developed in this study operates based on real-time object recognition and triggers ultrasonic deterrents only when a specific target species is identified. This selective intervention mechanism enhances environmental sensitivity while maintaining high effectiveness. Furthermore, the system conserves energy by activating only when needed and offers long-term usability with minimal maintenance requirements.

However, the system's performance can be partially affected by factors such as low light conditions or complex backgrounds, indicating a need for further improvements at the hardware and algorithmic levels. In future studies, the system is planned to be extended to cover additional species, support remote control through IoT technologies, and integrate AI-assisted decision-making capabilities. In this context, incorporating more powerful microcontrollers (e.g. NVIDIA Jetson Nano or Coral Dev Board), integrating camera modules with improved low-light performance, and enhancing spatial awareness using depth sensing technologies are among the targeted improvements. Furthermore, adaptive ultrasonic frequency generation and cloud-based behavioral data analysis infrastructures are among the directions for future work.

At a more advanced level, solar-powered versions of the system can be developed to ensure sustainable operation in rural areas and regions with limited energy infrastructure. Plans also include increasing spatial awareness with depth sensors, using light-weight offline models for fast decision-making, and enabling advanced functionalities such as long-term behavioral tracking through cloud-based analytics and species identification using bioacoustic data.

In conclusion, the proposed system not only delivers a technically successful solution but also presents a holistic approach to managing human–nature interaction in a more balanced and ethical manner.

REFERENCES

[1] M. L. Reidy, T. A. Campbell, and D. G. Hewitt, Evaluation of electric fencing to inhibit feral pig movements, The Journal of Wildlife Management, 72(4), 2008, 1012-1018.

[2] J. Redmon and A. Farhadi, YOLOv3: An Incremental Improvement, arXiv preprint arXiv:1804.02767, 2018.

[3] L. Strain, Electrodiagnostic Evaluation of Auditory Function in the Dog, Veterinary Clinics of North America: Small Animal Practice, 42(6), 2012, 1241-1257.

[4] M. Talling, J. B. Waran, J. A. Wathes, and S. Lines, Hearing in domestic pig (Sus scrofa) and goat (Capra hircus), Applied Animal Behaviour Science, 62(1), 1999, 1-20.

[5] V. S. Balaji, A. R. Mahi, A. G. P. S. Anirudh, and M. Manju, Scarecrow monitoring system: employing mobilenet ssd for enhanced animal supervision, arXiv preprint arXiv:2407.01435, 2024.

[6] J. Miao, D. Rajasekhar, S. Mishra, J. S. Nayak, and R. Yadav, A Fog-based Smart Agriculture System to Detect Animal Intrusion, arXiv preprint arXiv:2308.06614, 2023.

[7] Z. Wu, Image data augmentation techniques based on deep learning: A survey, Mathematical Biosciences and Engineering, 21(6), 2024, 6190-6224.

[8] L. Liu et al., Deep Learning for Generic Object Detection: A Survey, arXiv preprint arXiv:1809.02165, 2018.

[9] R. J. Young, Effectiveness of varying acoustic stimuli as wildlife deterrents: A review of habituation and sensitization patterns, Applied Animal Behaviour Science, 181, 2016, 103–114.

[10] T. K. Narayan and D. W. MacDonald, Wildlife Monitoring Using an Arduino-Based Active Infrared System, Ecological Informatics, 41, 2017, 24–32.

[11] A. Abubakar and A. K. Jibril, Prospect of Smart Agriculture Using IoT and Data Analytics: A Perspective of Kebbi State, Northwestern Nigeria, International Journal of Advanced Networking and Applications (IJANA), 15(6), 2024, 6194–6203.

[12] M. Ponciano, A. Brambilla, S. Bassano, T. Smith, and A. von Hardenberg, A new double observer-based census framework to improve abundance estimations in mountain ungulates and other gregarious species with a reduced effort, Ecological Solutions and Evidence, 5(4), 2024, e12405.

[13] S. B. Islam, D. Valles, T. J. Hibbits, W. A. Ryberg, D. K. Walkup, and M. R. J. Forstner, Animal Species Recognition with Deep Convolutional Neural Networks from Ecological Camera Trap Images, Animals, 13(9), 2023, 1526.

[14] M. Gholamy, K. Kreinovich, and O. Kosheleva, Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation, Departmental Technical Reports (CS), University of Texas at El Paso, 2018.

[15] H. Bichri, A. Chergu, and M. Hain, Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets, International Journal of Advanced Computer Science and Applications, 15(2), 2024, 112-116.

[16] R. Khanam and M. Hussain, YOLOv11: An Overview of the Key Architectural Enhancements, arXiv preprint arXiv:2410.17725, 2024.

[17] L. He, Y. Zhou, L. Liu, and J. Ma, Research and Application of YOLOv11-Based Object Segmentation in Intelligent Recognition at Construction Sites, Buildings, 14(12), 2024, 3777.

[18] R. Sapkota et al., Comprehensive Performance Evaluation of YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments, arXiv preprint arXiv:2407.12040, 2024.

[19] B. Irawan, P. N. Andono, and R. S. Basuki, Optimization of YOLOv5 Hyperparameter Using Adam Optimizer in Vehicle Object Detection, Journal of Applied Intelligent System, 9(1), 2024, 40-50.

[20] F. Nweke, Y. W. Teh, M. A. Al-gard, and U. R. Alo, Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges, Expert Systems with Applications, 105, 2018, 233-261.

[21] C. Prechelt, Early Stopping - But When?, Neural Networks: Tricks of the Trade, Springer, 2012, 55-69.

[22] B. M. Hussein and S. M. Shareef, An Empirical Study on the Correlation between Early Stopping Patience and Epochs in Deep Learning, ITM Web of Conferences, 64, 2024, 01003.