International Journal of Engineering and Technology Innovation, vol. 16, no. 2, 2026, pp. 199-210

Intelligent TNVR Ear-Tag Recognition and Monitoring System for Stray Animal Management

Zhi-Yu Wang, Yung-Hoh Sheu*, Bo-Kai Yang, Chi-Jen Chen, Chi-Wen Chen

Department of Computer Science and Information Engineering, National Formosa University, Yunlin, Taiwan, ROC

Received 30 August 2025; received in revised form 22 January 2026; accepted 26 January 2026

DOI: https://doi.org/10.46604/ijeti.2026.15640

Abstract

This study aims to enhance stray animal management by improving the efficiency and sustainability of the trap-neuter-vaccinate-return (TNVR) process. An intelligent monitoring system integrating image recognition and radar sensing is proposed for real-time detection and identification. The system utilizes a domain-specific ear-tag recognition model based on OpenCV preprocessing and YOLOv8, achieving an accuracy of 91% under various environmental conditions. Captured data are automatically uploaded via 4G to a centralized server, supporting continuous monitoring and instant alerts. Designed for high-density urban settings, the system mitigates manual workload and enhances decision-making efficiency, contributing to sustainable and humane stray animal control. Although the proposed system demonstrates high detection accuracy and robust performance under real-world conditions, the current evaluation is conducted on a moderate-scale dataset; future work will focus on large-scale deployment and cross-context validation to further examine system generalizability.

Keywords: TNVR, stray animal, YOLOv8, image recognition

1. Introduction

With the rapid progression of urbanization, the population of stray animals has been increasing annually, posing pressing social and public health challenges for cities worldwide. Stray animal management has consequently become a focal issue at the intersection of animal welfare and public health, garnering increasing attention from both governments and nonprofit organizations prioritizing humane and effective control measures. Among various approaches, the trap-neuter-vaccinate-return (TNVR) program has emerged as one of the most widely adopted international practices. Through capturing, neutering, vaccinating, and returning stray cats and dogs to their original locations, TNVR simultaneously reduces reproductive potential and stabilizes local populations [1]. A pilot program in Auckland, New Zealand, demonstrated that TNVR significantly decreased the number of animals recaptured after abandonment while enhancing public support and participation [2].

Long-term TNVR initiatives in Europe, the United States, and New Zealand have yielded promising results, including documented reductions in euthanasia rates at shelters and declines in public reports in areas where TNVR has been systematically implemented [2-3]. Nevertheless, operational challenges remain, particularly in the domains of individual identification and data tracking. Traditional methods, such as ear clipping or metal tagging, offer limited precision and are difficult to integrate into digital management systems. At the community level, combining TNVR promotion with systematic tracking, data transparency, and individual identification technologies [3] has been shown to strengthen public trust and enhance governance efficiency. International experiences consistently underscore that embedding intelligent management within TNVR frameworks not only stabilizes stray populations but also provides a robust, data-driven foundation for policy decision-making.

* Corresponding author. E-mail address: yhsheu@nfu.edu.tw

Conventional identification methods, including ear clipping, metal or plastic tags, and microchip implantation, typically require close-range handling and manual recording, making them unsuitable for large-scale or dynamic monitoring. Research indicates that metal tags are prone to detachment, corrosion, or physical damage [4]. Furthermore, microchips, although capable of providing stable information, require specialized scanning equipment and human intervention, thereby lacking the capacity for automated remote identification.

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have significantly transformed the field of image recognition, establishing deep learning as the dominant framework for handling complex visual tasks. These methods have demonstrated strong performance in extracting hierarchical visual features and achieving high accuracy across diverse recognition scenarios. In animal identification, the availability of large, well-annotated datasets has enabled a wide range of applications, including species classification, breed recognition, and individual-level identification. Initial breakthroughs based on the ImageNet dataset established the framework for deep learning–based image classification, which were later extended to animal recognition tasks [5]. Subsequent datasets, such as the Oxford Pet Dataset, which comprises images of 37 cat and dog breeds, further demonstrated the ability of deep learning models to learn fine-grained, animal-specific features and distinguish subtle inter-breed differences [6]. These studies confirmed the effectiveness of CNNs in recognizing complex biological visual patterns.

Beyond companion animals, deep learning–based image recognition has also been successfully applied to wildlife and livestock monitoring. For example, a CNN-based system enabled individual kangaroo identification in Australia [7]. In India, researchers combined ear-tag detection with image recognition techniques to monitor dairy cattle health, thereby improving farm management efficiency and disease prevention capabilities [8]. Overall, these studies highlight the growing potential of AI-driven image recognition and automated management systems in advancing animal identification, monitoring, and welfare, providing a strong foundation for further research and real-world deployment. Within this context, the You Only Look Once (YOLO) series has demonstrated exceptional performance in real-time object detection across diverse fields, including transportation, healthcare, security, and animal monitoring. Since the release of YOLOv1 in 2016 [8], subsequent iterations have introduced substantial architectural improvements aimed at balancing detection accuracy and computational efficiency.

YOLOv4 incorporated CSPDarknet53 as its backbone network, significantly enhancing both accuracy and processing speed [9]. YOLOv5, implemented using the PyTorch framework, further gained widespread adoption due to its modular design, ease of training, and flexibility in deployment across different hardware platforms. More recently, YOLOv8 has integrated Transformer-based components and anchor-free detection strategies, enabling improved recognition of small, dense, and occluded objects. Its application to wildlife monitoring in nature reserves has demonstrated successful real-time identification of multiple mammalian species, effectively reducing false detections and improving overall monitoring efficiency [10]. In addition to wildlife applications, YOLOv8 has been employed for automated identification and ear-tag detection of stray cats [11], where it has been integrated into urban surveillance systems to support TNVR operations.

However, existing studies primarily focus on stray cats, without extending to broader real-world field deployment or incorporating ear-clipping recognition for stray dogs, thus remaining limited. In parallel, Internet of Things (IoT) and Wireless Sensor Network (WSN) technologies have been widely implemented in environmental monitoring and analysis [12]. Despite these advances, no study to date has applied IoT and WSN frameworks to stray animal movement monitoring, nor has a comprehensive data analysis and management system been proposed for this context.

To address these limitations, this study introduces an intelligent TNVR ear-tag recognition and monitoring system for stray animal management that integrates AI-based visual recognition with IoT technologies. By employing camera devices to capture animal images and YOLOv8 models to automatically detect ear tags, the proposed system enhances identification efficiency and accuracy in field settings while reducing labor costs and error rates. The integration of data storage and management functionalities establishes a centralized database and tracking mechanism to support systematic stray animal population management. Moreover, the system enables real-time, scientific, and visual decision support for policymakers, thereby contributing to the efficiency and sustainability of TNVR implementation.

2. System Principle

To develop an ear-tag recognition system for stray cats and dogs based on the YOLOv8 model and improve both detection speed and accuracy, a comparative analysis was conducted across various deep learning–based object detection techniques, including YOLO, Single Shot MultiBox Detector (SSD), and Faster R-CNN. The results indicated that YOLO achieves the optimal balance between Mean Average Precision (mAP) and inference speed, making it the most suitable choice for real-time prediction tasks [13-15]. Therefore, YOLOv8 was selected as the CNN model for training in this study to achieve simultaneous recognition of ear-clipped cats and dogs within images. The research process was divided into three main stages:

(1) The first stage entailed building a dedicated dataset;

(2) The second stage focused on data augmentation and annotation;

(3) The third stage consisted of training the YOLOv8 model using the Ultralytics library.

As shown in Fig. 1, the overall workflow for public dataset acquisition and model training adopted in this study is illustrated, providing a clear overview of the complete process from data preparation to model development.

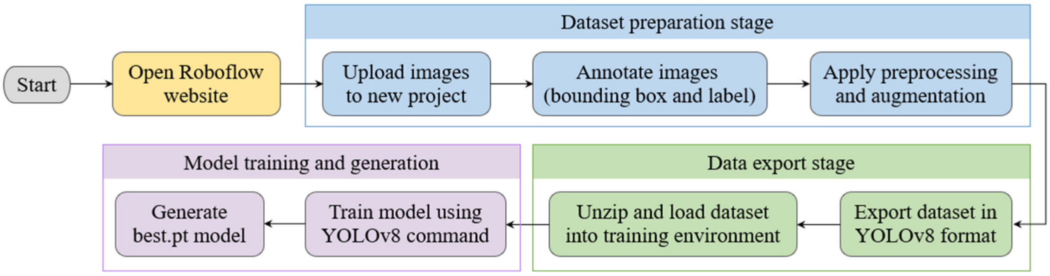


Fig. 1 The public dataset acquisition and model training process

2.1. Dataset collection

This study centers on the task of ear-clipping (ear-tag) recognition for stray cats and dogs, with the construction of a proprietary image dataset serving as the foundation for training the YOLOv8 model. In total, 1,973 images of cats and dogs were collected from multiple sources, including camera-based field photography and publicly available repositories. To enhance the model's generalization capability, the dataset incorporated diverse samples representing variations in breed, posture, camera angle, and lighting conditions. Image sources included photographs taken in field environments, contributions from community volunteers, and images obtained from the Roboflow image-sharing platform. All collected images underwent manual inspection to ensure clarity, relevance, and annotation suitability, thereby guaranteeing their value for subsequent model training and evaluation.

2.1.1. Roboflow public dataset

As depicted in Fig. 2, the workflow for acquiring images from the Roboflow public dataset is illustrated. A subset of the dataset employed in this study was sourced from the Roboflow platform, which provides extensive, pre-annotated datasets for computer vision applications. Roboflow hosts a large repository of high-quality images covering diverse domains, including animal species, postures, and environmental contexts. In this study, targeted keyword searches related to stray cats and dogs were conducted, and high-resolution, accurately annotated, and diverse image samples were systematically selected. The collected images span multiple breeds, body postures, viewing angles, and illumination conditions, collectively ensuring a robust and heterogeneous dataset to support the subsequent training and validation of the YOLOv8-based recognition model.

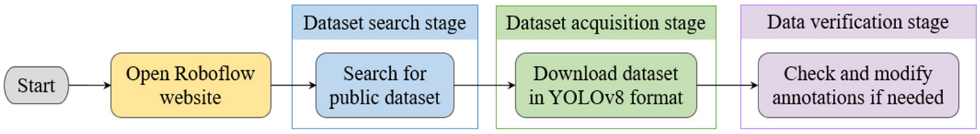


Fig. 2 Workflow for acquiring images from the Roboflow public datasets

2.1.2. Field and on-site photography at stray animal associations

To capture image data that more accurately reflects the local environmental conditions in Taiwan, this study conducted field visits to multiple animal protection associations and shelters actively engaged in TNVR programs. With the assistance of association volunteers, numerous on-site photographs of stray animals were collected [16]. The dataset includes street-level images of stray cats and dogs, pre- and post-surgery comparisons of ear-clipping procedures, and surveillance footage obtained from shelter facilities, all utilized with appropriate authorization. These context-rich and authentic images substantially enhance the dataset's representativeness, thereby improving the robustness and accuracy of the recognition model under real-world application scenarios.

2.2. Data augmentation and annotation

As illustrated in Fig. 3, the data augmentation process was implemented using the Roboflow platform [17]. Roboflow is a cloud-based tool tailored for computer vision applications, providing automated augmentation, format conversion, and annotation management. The platform enables the configuration of multiple augmentation pipelines with real-time preview capabilities, which significantly reduces manual processing time and enhances annotation efficiency.

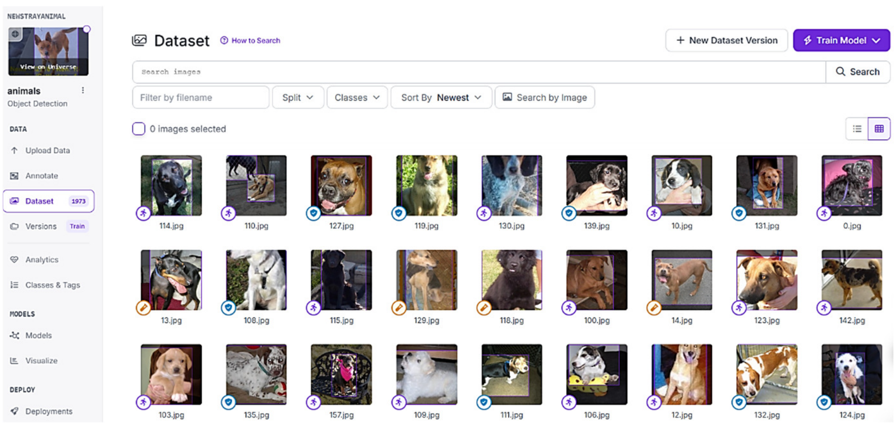


Fig. 3 Dataset augmentation workflow performed using the Roboflow platform

During the annotation phase, manual labeling was also conducted through the Roboflow interface. Each image was annotated into three categories: Cat, Dog, and Ear-tipped. Among these categories, the annotation of Ear-tipped features posed particular challenges, as it required precise attention to subtle ear contours, notches, and asymmetries. To ensure data quality, all annotation tasks were carried out by trained personnel following standardized procedures. Upon completion, the Roboflow platform automatically converted the annotated dataset into the YOLOv8-compatible training format, generating the requisite label files and configuration settings for subsequent model training.

2.3. YOLOv8 architecture and model selection

The YOLO series has been extensively adopted in the field of computer vision owing to its high detection accuracy, low latency, and lightweight architecture, which make it particularly well-suited for real-time object detection and deployment on edge devices. YOLOv8, developed by the Ultralytics team and released in 2023, represents the latest evolution of this series. Building upon YOLOv5, YOLOv8 introduces enhancements in both network architecture and training workflows, while also extending support to multiple tasks, including object detection, image classification, and image segmentation.

A key architectural innovation in YOLOv8 is its transition from an anchor-based to an anchor-free detection paradigm. This design directly predicts object center points and dimensions, thereby streamlining the prediction process, improving computational efficiency, and enhancing detection accuracy. Moreover, YOLOv8 incorporates an advanced non-maximum suppression (NMS) strategy that reduces redundant detections, a feature that is especially critical in scenarios involving multiple targets or small-object recognition. These characteristics render YOLOv8 particularly advantageous for complex-background tasks requiring fine-grained detection, such as the recognition of ear-clipped regions in stray animals.

In this study, the YOLOv8m variant was selected for model training. Compared with smaller versions, YOLOv8m yields a higher parameter count and stronger feature representation capacity, resulting in improved recognition accuracy. This balance between computational efficiency and discriminative power makes YOLOv8m well-suited for the present task, which requires precise differentiation between ear-clipped and non–ear-clipped stray cats and dogs.

2.4. *YOLOv8 training environment and model training*

The training and validation of the YOLOv8 model were conducted in a local computing environment equipped with an NVIDIA GPU to accelerate computational processes. By exploiting the parallel processing capabilities of the GPU, the system achieved substantial improvements in training efficiency, significantly reducing computational time and enhancing the convergence speed of the neural network. Compared with traditional CPU-based environments, GPUs provide notable advantages in matrix operations and convolutional processing, which are fundamental to deep learning, thereby effectively supporting the training requirements of a medium-scale dataset comprising thousands of images.

Before model training, the annotated and augmented datasets were reformatted into a YOLOv8-compatible structure. A corresponding .yaml configuration file was created to specify three label classes—Cat, Dog, and Ear-tipped—as well as dataset paths and ratio settings for the training and validation subsets. Model training was executed within the Ultralytics YOLOv8 framework with the following parameters: 200 training epochs, a batch size of 16, and an input image resolution of 640 × 640 pixels. In this study, the YOLOv8m variant was adopted, as its deeper architecture and larger parameter space enable superior feature extraction and enhanced recognition accuracy compared with smaller variants.

Recent studies have shown that YOLO-based models are effective for real-time object detection in complex and resource-constrained environments. Previous work has demonstrated reliable real-time performance in dynamic scenarios and low-latency inference for edge deployment [17-18]. These findings support the adoption of YOLOv8 in this study, which further extends YOLO-based detection to domain-specific TNVR ear-tag recognition and operational monitoring.

3. System Design

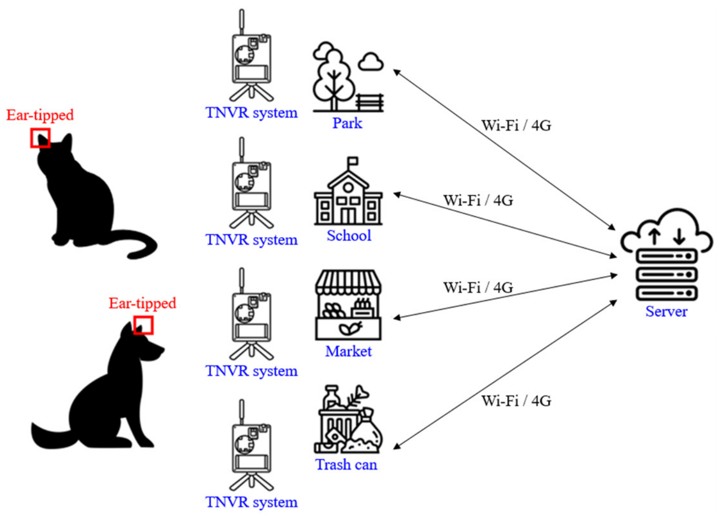


Fig. 4 Application scenario of the TNVR ear-tag recognition system

As illustrated in Fig. 4, the proposed system is designed for deployment in environments with high stray animal activity, such as markets, schools, parks, and waste disposal sites. The system is capable of continuous, 24-hour operation, enabling real-time monitoring of stray animals in surrounding areas and facilitating automated notifications to relevant authorities. By supporting timely detection and reporting, the system curbs the growth of stray animal populations and improves local environmental hygiene. Furthermore, the system exemplifies an application of IoT technologies, integrating sensing, communication, and data management to achieve intelligent monitoring. The following subsections delineate the hardware and software design components of the proposed system, highlighting their integration into a unified framework for stray animal identification, data collection, and decision support, and providing a structured description of how each component contributes to the overall system architecture and functionality.

3.1. Hardware design

As illustrated in Fig. 5, the hardware architecture of the proposed system comprises several key modules, each serving a distinct functional role:

(1) BMduino (BM53A367A) control board, manufactured by Holtek Semiconductor

The BMduino (BM53A367A) control board functions as the central embedded platform of the system, responsible for receiving data inputs from peripheral sensors and managing the operation of automatic detection terminals. Specifically, it performs the core tasks of primary data acquisition, signal processing, and control execution, thereby ensuring stable and efficient system operation.

(2) 8 × 8 Array of IR thermal sensors (GY-AMG8833), manufactured by Panasonic Corporation

The AMG8833 infrared thermal sensor array, consisting of 64 thermal pixels, enables non-contact temperature measurement and heat signature detection. Owing to its ability to capture low-resolution thermal images, the AMG8833 is particularly suitable for identifying moving objects through temperature differentials. In this study, the sensor is employed to support the detection of stray animals in varying environmental conditions, enhancing the robustness of the overall recognition system.

(3) Microwave sensor module (HB100), manufactured by ST Electronics

The HB100 microwave sensor module operates based on the Doppler radar principle, detecting object motion by analyzing variations in reflected microwave signals. In this system, it is integrated with the GY-AMG8833 infrared thermal sensor to enable multimodal sensing. This dual-sensor configuration enhances stability and accuracy in biological target detection, effectively reducing false positives caused by environmental interference.

(4) ESP32-CAM, manufactured by Espressif Systems

The ESP32-CAM module, equipped with an OV2640 camera supporting resolutions up to 1600 × 1200 pixels, provides real-time image transmission and high-quality photo capture capabilities. With built-in Wi-Fi and Bluetooth connectivity, the module is well-suited for IoT applications. In this system, when both the microwave sensor and the infrared thermal array confirm the presence of a valid biological target, the ESP32-CAM is automatically triggered to capture images. Captured data are stored on an onboard SD card for subsequent processing and analysis.

(5) e-paper module, manufactured by Waveshare

An e-paper display module is incorporated to present system location information and contact details, thereby facilitating on-site system management and enabling user feedback collection. Moreover, e-paper technology is particularly energy-efficient for continuous operation due to its low power consumption.

(6) SIM7600X CE module, manufactured by Waveshare

The SIM7600X CE module is a versatile communication unit supporting 4G and LTE Cat-1 networks, offering high-speed data transmission and integrated GPS positioning. Compatible with communication protocols such as TCP, UDP, and HTTP, the module enables reliable real-time positioning and tracking. Furthermore, the module features low power consumption, making it suitable for long-term deployment and battery-powered applications.

In the proposed hardware architecture, the BMduino (BM53A367A) control board functions as the central hub, interfacing with both the HB100 microwave sensor and the 8 × 8 infrared thermal sensor array. The detection process follows a sequential validation mechanism: the microwave sensor first identifies motion within the monitored area, and the infrared thermal sensor then verifies whether the detected object exhibits biological heat signatures. Upon confirmation of an animal target, the ESP32-CAM is triggered to capture images, which are transmitted to the server through either the SIM7600X CE 4G communication module or the ESP32-CAM's integrated Wi-Fi connection. Processed data are uploaded to a web-based interface, where users can access monitoring records and perform query operations. This integrated workflow ensures accurate, efficient, and real-time monitoring of stray animals in diverse urban environments.

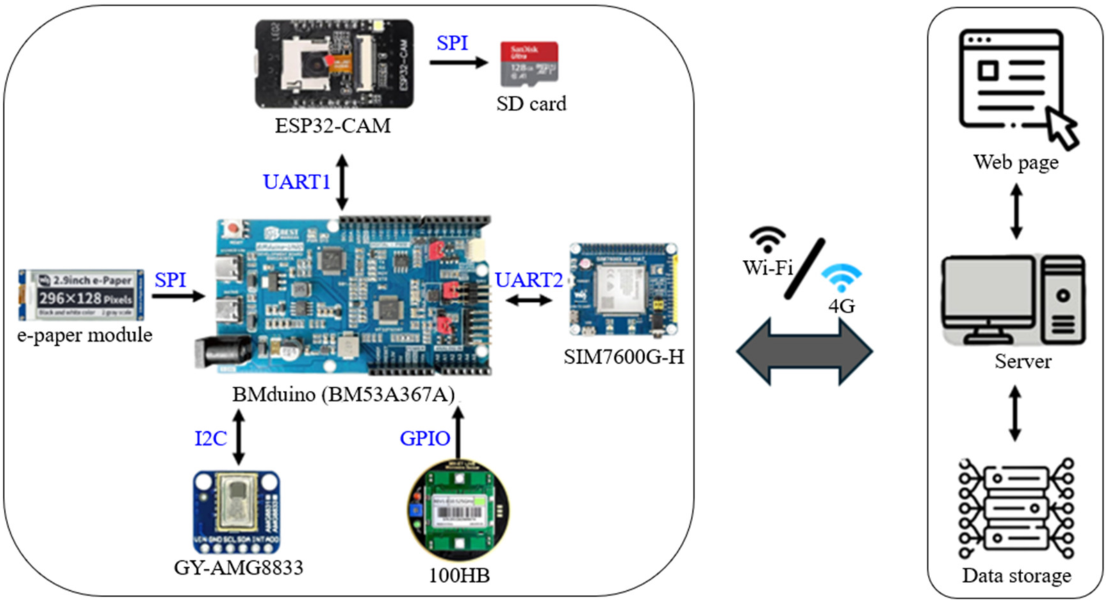


Fig. 5 Hardware architecture of the TNVR ear-tag system

3.2. Software design

As illustrated in Fig. 6, the operational workflow of the proposed system integrates multiple sensing and communication technologies to enable intelligent monitoring of stray animal activities. The system leverages a matrix-type infrared thermal sensing array and microwave radar modules for initial detection, which are subsequently combined with visual recognition, 4G wireless transmission, and cloud-based analytical capabilities. This multimodal approach ensures accurate detection, efficient data transmission, and automated processing of stray animal behaviors in various urban environments.

The integrated framework encompasses animal detection, visual recognition, data analysis, and cloud-based storage, thereby offering an efficient solution for stray animal monitoring and management. By facilitating reliable and automated identification, the system significantly contributes to enhancing the sterilization rate of stray animal populations under TNVR programs. Captured images are transmitted via Wi-Fi to a server, where they undergo processing with computer vision algorithms for recognition and filtering. Only validated images are retained and subsequently uploaded to a cloud database for deep analysis and long-term storage. Finally, the processed results are delivered to end-users through a web-based portal, enabling real-time monitoring, historical data inquiry, and decision support for policy implementation.

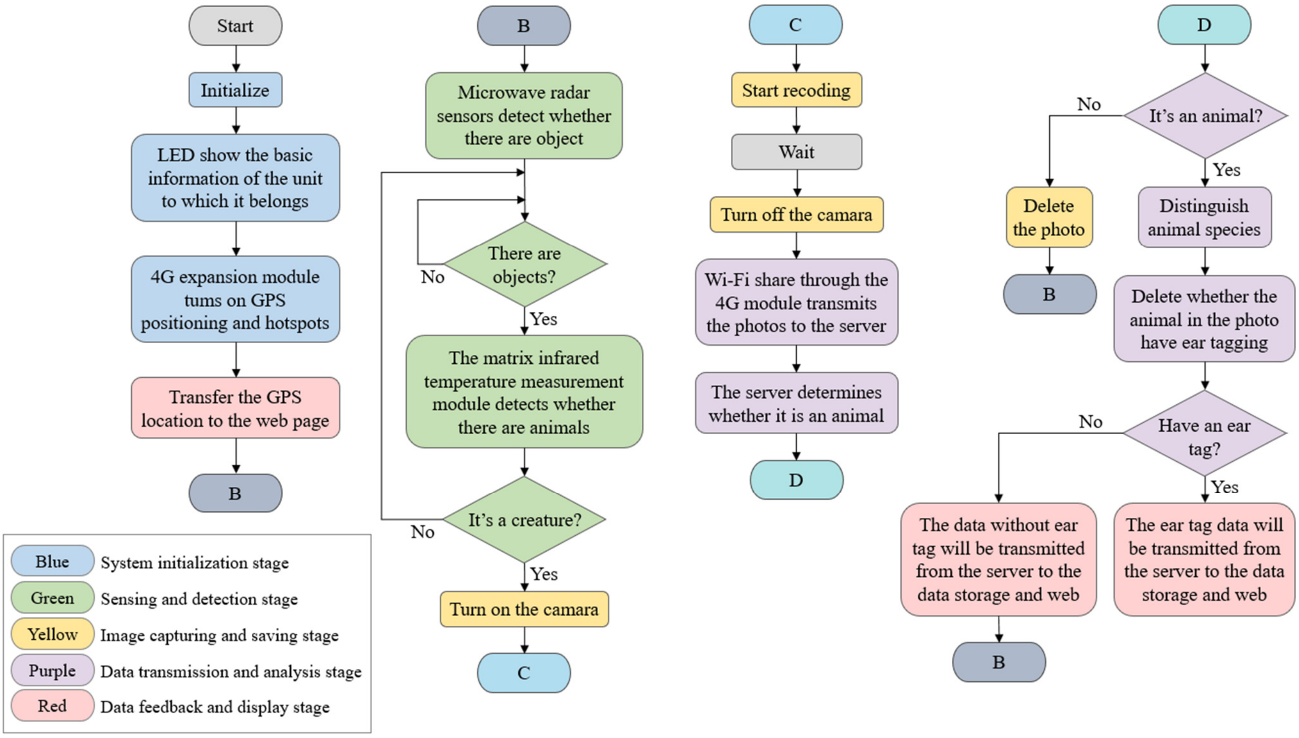


Fig. 6 Operational flow chart of the proposed system

The system integrates advanced computer vision technologies, including OpenCV and YOLO, to preprocess images prior to transmission. Specifically, the preprocessing pipeline crops images to retain only the regions containing cats and dogs, thereby reducing data volume and alleviating computational load. By transmitting cropped images rather than raw inputs, the system evinces notable improvements in both transmission efficiency and recognition accuracy. Within the visual recognition workflow, all images are annotated and categorized using the Roboflow platform. The labeled datasets are then exported in YOLO-compatible format and subsequently loaded into the program to train a YOLOv8-based model specifically designed to identify cats, dogs, and ear-clipping features. In this way, this tailored training strategy enhances model precision and robustness under diverse environmental conditions.

On the server side, the system is implemented using Node.js, with the Express framework employed for the development of web-based applications. This configuration provides an efficient and flexible runtime environment capable of handling diverse HTTP requests while seamlessly interacting with the backend database. The frontend web interface extends multiple user-oriented functionalities, including real-time data queries, monitoring dashboards, and automated report generation. These features empower users to filter, retrieve, and export specific records according to operational needs, thereby supporting both day-to-day management tasks and long-term policy evaluation.

4. System Testing and Results

As illustrated in Fig. 7, the physical prototype and deployment setup of the proposed system are presented. The overall device height can be flexibly adjusted according to the specific requirements of the installation site, thereby accommodating diverse terrain conditions. This adaptability ensures that the camera is positioned at approximately the head level of stray cats and dogs, thereby providing an optimal angle for image acquisition and recognition.

In the experimental phase, comparative evaluations were conducted between different YOLOv8 variants. As depicted in Fig. 8, the performance metrics during the training of the YOLOv8m model are presented. Both the bounding box loss and classification loss steadily decreased and stabilized within the first 50 epochs, while the validation loss exhibited consistent convergence, thereby indicating the model's effective learning and generalization capacity. Furthermore, the mAP50–95 metric consistently exceeded 0.98 throughout training, with both precision and recall approaching 100%. These results demonstrate the high recognition accuracy of the YOLOv8m model in identifying stray animals and distinguishing ear-clipped individuals, confirming its robustness and suitability for real-world TNVR monitoring applications.

Fig. 7 Actual placement of the system equipment and physical view of the device

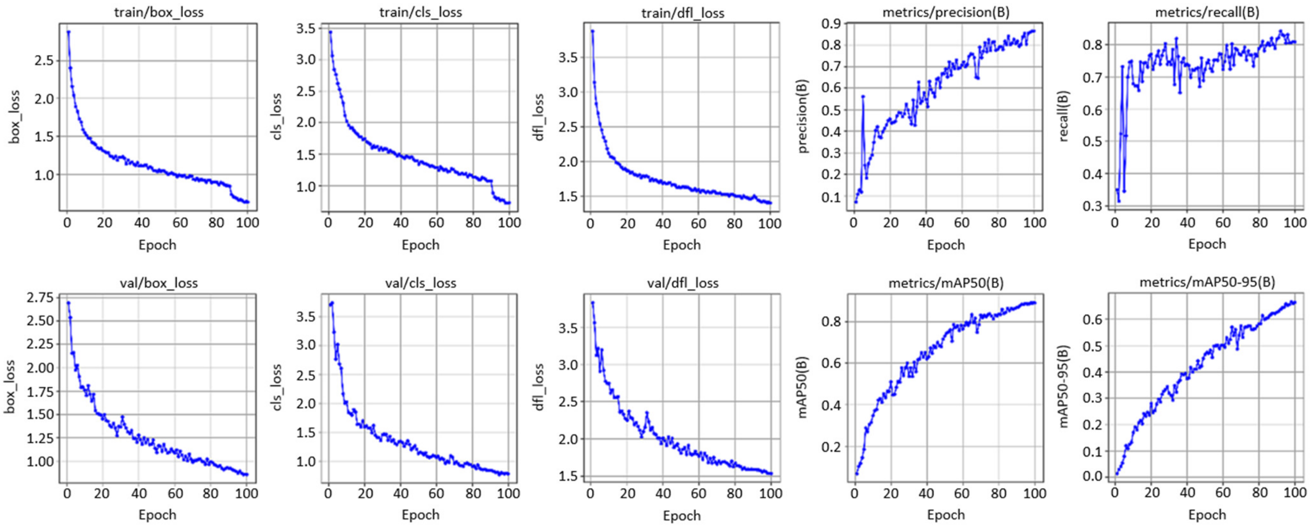
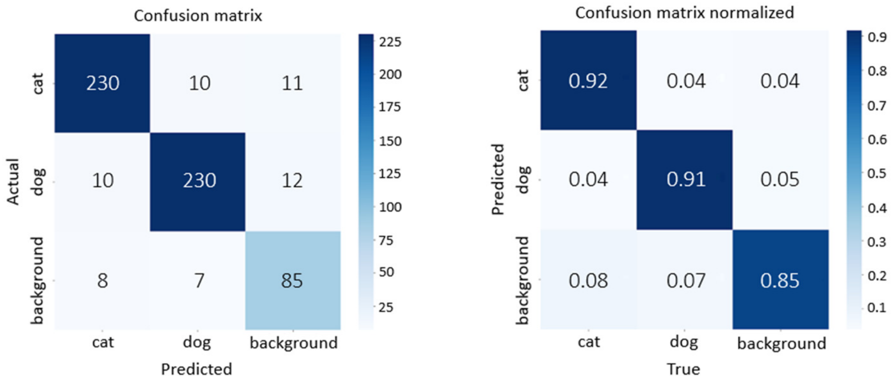


Fig. 8 Performance metrics during the YOLOv8m model training process

4.1. Confusion matrix

The confusion matrix is a fundamental tool for evaluating the performance of classification models, as it provides a detailed representation of both correct predictions and misclassifications across different categories. As illustrated in Fig. 9, two forms of confusion matrices are presented for the YOLOv8m model evaluated on the test dataset, offering a comprehensive view of its classification capabilities. Fig. 9(a) displays the original confusion matrix, which records the absolute number of classifications per category, thereby revealing the distribution of correct predictions as well as category-specific misclassifications. In contrast, Fig. 9(b) presents the normalized confusion matrix, where classification outcomes are expressed as proportions, allowing for a clearer comparison of classification accuracy and confusion levels across categories.

Overall, the system achieved a visual recognition accuracy of 91%, successfully identifying the ear-clipping status of stray cats and dogs. These results highlight the robustness of the YOLOv8m model in handling complex classification tasks and demonstrate its practical applicability for automated monitoring within TNVR management systems. Through the two forms of confusion matrices, the prediction accuracy and error distribution of the YOLOv8m model across the categories of Cat, Dog, and Background can be more intuitively examined. These visualizations provide insights into the model's strengths and weaknesses while serving as critical references for guiding future optimization of classification performance and enhancing system robustness.



(a) Confusion matrix (b) Normalized confusion matrix

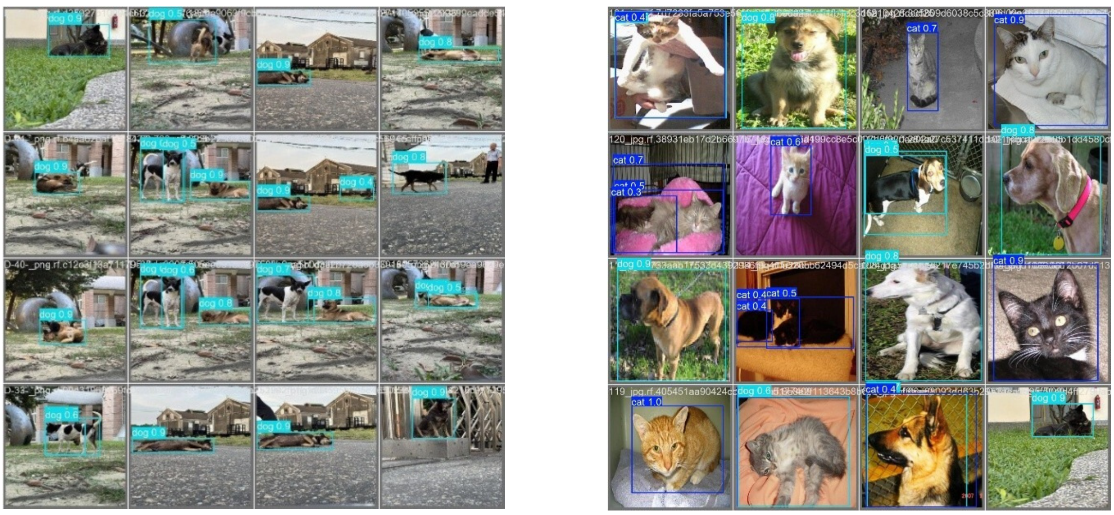
Fig. 9 Confusion matrices of the YOLOv8m model on the evaluation dataset

4.2. Prediction result

As illustrated in Fig. 10, the inference outputs of the YOLOv8m model are presented, demonstrating its recognition capabilities across 32 diverse images. These results cover multiple categories, including cats and dogs, and reflect the model's robustness across varied environments, postures, and contextual conditions. Figs. 10(a) and 10(b) display detection results on the first batch of images, which include scenarios such as multiple cats within a single frame, interactions among multiple dogs, and cases involving motion blur.

More specifically, Fig. 10(a) illustrates that the model can accurately detect and annotate multiple dogs in varying postures, including lying, running, and interacting, with consistently high confidence scores. Fig. 10(b) demonstrates that even under outdoor environments with substantial lighting variation and ground surfaces resembling the animals' fur color, the YOLOv8m model maintains effective identification performance, highlighting its adaptability to scene variability. Furthermore, Fig. 10 illustrates the diversity of input image sources used for evaluation. Fig. 10(a) presents images captured by the customized device developed in this study, representing real-world application scenarios, while Fig. 10(b) presents samples from publicly available datasets covering diverse environments and animal conditions. The combination of field-collected and public images supports model generalization and provides a basis for further optimization.

As shown in Table 1, the proposed YOLOv8m-based model achieves consistently higher performance across all evaluation metrics, particularly under stricter IoU thresholds, indicating enhanced localization robustness for small and partially occluded ear-tag regions.



(a) Detection results for multiple cats and dogs (b) Detection results under outdoor conditions with lighting variation

Fig. 10 Inference bounding boxes and classification results for the first batch of images

Table 1 Performance comparison of different detection models

Model	Precision	Recall	F1-score	mAP@0.5	mAP@0.5:0.95
Yolov8n	0.93	0.92	0.92	0.94	0.86
Yolov8s	0.95	0.94	0.94	0.96	0.91
Yolov8m	0.98	0.97	0.97	0.99	0.95

5. Conclusions

This study developed an intelligent TNVR monitoring and management system to enhance the efficiency, accuracy, and sustainability of stray animal control. The proposed framework integrates radar sensing and image recognition to enable real-time detection, automated tracking, and cloud-based data management. A domain-specific ear-tag recognition model, combining OpenCV-based preprocessing with YOLOv8 detection, achieved an accuracy of 91% under diverse lighting and environmental conditions, demonstrating robust performance for practical deployment. The main conclusions are summarized as follows:

(1) The proposed TNVR system effectively integrates radar and vision sensors to enable continuous detection and identification of stray animals in high-density environments.

(2) The YOLOv8-based ear-tag recognition model enhances detection precision and operational stability compared with conventional image-based approaches.

(3) Automated data transmission via 4G communication supports real-time monitoring and mitigates manual workload, facilitating timely and informed decision-making.

(4) The integrated platform contributes to humane and sustainable TNVR implementation by lowering labor demands and supporting long-term animal welfare management.

Future work will focus on extending system validation to additional urban environments and exploring lightweight models and cloud-based analytics to further enhance adaptability and performance. From a public administration perspective, the proposed system functions as a practical decision-support tool for local governments by enabling continuous monitoring of sterilized and non-sterilized stray animals. Real-time and historical data support evidence-based prioritization of TNVR interventions and resource allocation, while the web-based interface minimizes operational complexity. Although a detailed cost–benefit analysis is beyond the scope of this study, automated detection and remote monitoring are expected to reduce long-term labor costs. The modular architecture further supports scalability, institutional integration, and potential extension to other urban animal or wildlife monitoring applications.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] J. K. Levy, D. W. Gale, and L. A. Gale, "Evaluation of the Effect of a Long-Term Trap-Neuter-Return and Adoption Program on a Free-Roaming Cat Population," Journal of the American Veterinary Medical Association, vol. 222, no. 1, pp. 42-46, 2003.

[2] S. Zito, G. Aguilar, S. Vigeant, and A. Dale, "Assessment of a Targeted Trap-Neuter-Return Pilot Study in Auckland, New Zealand," Animals, vol. 8, no. 5, article no. 73, 2018.

[3] K. L. Hughes and M. R. Slater, "Integrated Return-to-Field and Targeted Trap-Neuter-Vaccinate-Return Programs Result in Reductions of Feline Intake and Euthanasia at Six Municipal Animal Shelters," Frontiers in Veterinary Science, vol. 6, article no. 77, 2019.

[4] A. M. Johnston and D. S. Edwards, "Welfare Implications of Identification of Cattle by Ear Tags," Veterinary Record, vol. 138, no. 25, pp. 612-614, 1996.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017.
[6] Visual Geometry Group, "The Oxford-IIIT Pet Dataset," https://www.robots.ox.ac.uk/~vgg/data/pets/, 2025.
[7] J. LeBien, M. Zhong, M. Campos-Cerqueira, J. P. Velev, R. Dodhia, J. L. Ferres, et al., "A Pipeline for Identification of Bird and Frog Species in Tropical Soundscape Recordings Using a Convolutional Neural Network," Ecological Informatics, vol. 59, article no. 101113, 2020.
[8] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," https://doi.org/10.48550/arXiv.2004.10934, 2020.
[9] U. P. S. Lundquist, S. Afridi, C. Berthelot, N. Ngoc Dat, K. Hlebowicz, E. Iannino, et al., "WildDrone: Autonomous Drone Technology for Monitoring Wildlife Populations," Frontiers in Robotics and AI, vol. 12, article no. 1695319, 2025.
[10] E. Williams, A. Carter, and J. Boyd, "Kinetics and Kinematics of Working Trials Dogs: The Impact of Long Jump Length on Peak Vertical Landing Force and Joint Angulation," Animals, vol. 11, no. 10, article no. 2804, 2022.
[11] A. Sudarsono, S. Huda, N. Fahmi, M. U. H. Al-Rasyid, and P. Kristalina, "Secure Data Exchange in Environmental Health Monitoring System through Wireless Sensor Network," International Journal of Engineering and Technology Innovation, vol. 6, no. 2, pp. 103-122, 2016.
[12] M. Boukabous and M. Azizi, "Image and Video-Based Crime Prediction Using Object Detection and Deep Learning," Bulletin of Electrical Engineering and Informatics, vol. 12, no. 3, pp. 1630-1638, 2023.
[13] J. Garcia-Pajuelo and E. Paiva-Peredo, "Comparison and Evaluation of YOLO Models for Vehicle Detection on Bicycle Paths," IAES International Journal of Artificial Intelligence, vol. 13, no. 3, pp. 3634-3643, 2024.
[14] H. Gao, "A YOLO-Based Violence Detection Method in IoT Surveillance Systems," International Journal of Advanced Computer Science and Applications, vol. 14, no. 8, pp. 143-149, 2023.
[15] Faith for Animals, https://www.faithforanimals.org.tw/, 2025.
[16] Roboflow, https://roboflow.com/, 2025.
[17] S. S. More, and R. Bansode, "FCN-YOLOS: An Effective Deep-Learning Model for Real-Time Object Detection," Journal of Field Robotics, vol. 42, no. 8, pp. 4053-4074, 2025.
[18] S. S. More, N. Patil, V. B. Lobo, N. Shet, D. Goswami, and P. Rane, "Empowering the Visually Impaired: YOLOv8-Based Object Detection in Android Applications," Procedia Computer Science, vol. 252, pp. 457-469, 2025.