Real-Time Multi-Camera Dog Track and Alert System

Xiaoquan Zhang
Mahidol University
Henan, China
xiaoquan.zha@student.mahidol.ac.th

Thanaporn Pichaisongkram
Mahidol University
Bangkok, Thailand
thanapornpichai@gmail.com

Chatchai Wangwiwattana
University of the Thai Chamber of Commence
Bangkok, Thailand
chatchai_wan@utcc.ac.th

***Abstract*—This paper presents a system titled Real-Time Multi-Camera Dog Track and Alert, which leverages the YOLOv11m object detection algorithm to detect dogs in real-time video streams and recognize individual dogs by name. It addresses challenges in pet monitoring and behavior research. YOLO-based detection is combined with identity-aware annotation and a timing system for absence alerts. Experiments with dual video feeds show robust real-time detection and alert capabilities for dogs "Minney" and "Lucky."**

***Index Terms*—YOLO, dog monitoring, object detection, pet tracking, real-time alert.**

I. Introduction

Pet monitoring has become a growing concern for modern households, especially for pet owners who need to ensure the well-being and safety of their animals while away. With the increasing use of smart surveillance systems and affordable home cameras, pet tracking solutions have become more accessible [1], [2]. However, most existing systems focus solely on general object detection and motion monitoring. They can detect the presence of a dog but lack the ability to recognize and differentiate between individual animals. This limitation is critical in multi-pet homes or in shelters where identity-based tracking is essential for behavior analysis, health management, and security [3], [4].

Recent work has addressed various aspects of pet surveillance. Singh et al. [5] proposed a YOLO-LSTM hybrid for pet behavior recognition. Kim and Moon [6] integrated video and wearable sensors to improve recognition accuracy. While these solutions demonstrate strong technical performance, they often rely on additional hardware such as wearable devices, which are not always feasible for daily use or scalable for shelters. In contrast, Smith and Johnson [7] explored face-based pet ID systems, but their approach requires biometric-level facial capture, which may not perform well with low-angle, full-body views from typical CCTV systems.

Other biometric-based recognition systems, such as nose-print identification [8] or breed/face classification [9], require specialized imaging, which is not available in most homes. Additionally, multi-camera tracking frameworks [10] or embedded kennel monitoring devices [11] may be too complex or costly for everyday use.

To address these limitations, we propose the "Real-Time Multi-Camera Dog Track and Alert System," a vision-only, identity-aware solution that can detect and recognize household dogs in real time. The core goal of this project is to develop a lightweight, deployable program that alerts the owner if a specific dog has not been detected on any available camera for over one hour. The system is based on YOLOv11m, a deep-learning model known for balancing real-time speed and high detection accuracy. Our model is trained to recognize individual dogs—specifically "Minney" and "Lucky"—using only standard video input without any biometric preprocessing.

To simulate a realistic household scenario, we constructed a custom dataset using 4 videos captured under different conditions. These videos were recorded with assistance from a family member located remotely in Chiang Mai, Thailand. The remote recording was facilitated using mobile phone video calls and a CCTV application that allowed us to capture footage from both indoor and outdoor environments. Each video ranged from 1 to 5 minutes in length, recorded at 30 frames per second (FPS) and 1080p resolution to match the quality of typical home surveillance systems. This approach allowed us to replicate a real-life smart home use case without the need for special equipment.

Once the footage was collected, individual frames were extracted and manually annotated in YOLO format. Each dog was treated as a unique class. The detection pipeline processes each incoming frame, assigns labels with confidence scores, and tracks the temporal presence of each dog. When a dog remains undetected for a configurable duration (default: one hour), the system automatically triggers an alert. Our current prototype supports alert delivery via mobile notification, sound alarm, or log entry, depending on deployment setup [12], [13].

This system is inspired by emerging work in animal behavior surveillance [14], facial recognition for pets [7], and smart home integration for pet care [1], but we focus on keeping the system low-cost, non-invasive, and easy to implement. Furthermore, we explore the potential of extending this method to multiple camera streams simultaneously, offering robust spatial coverage of the household environment.

The remainder of this paper is organized as follows: Section II presents our methodology, including dataset creation and model training. Section III showcases experimental results and real-time tracking performance. Section IV discusses the results, limitations, and opportunities for further development. Finally, Section V concludes with future enhancements.

II. Methodology

A. Dataset Creation

Video data of dogs “Minney” and “Lucky” was collected in indoor and outdoor environments. Frames were extracted and annotated using YOLO-format bounding boxes. Each dog’s name was treated as a separate class.

While previous datasets (e.g., wildlife traps [14]) emphasize background diversity, we focus on visual clarity for reliable identity labeling. Turečková et al. [15] successfully trained YOLO on dog face crops; our approach generalizes this to whole-body CCTV footage.

B. Model Training

We used YOLOv11m due to its speed–accuracy tradeoff. Each dog’s name served as a label. Similar to work on embedding-based recognition [7], but without requiring facial alignment or high-res imagery.

C. Testing

Two different videos were run simultaneously through a custom Python pipeline. Each frame was processed for detection, name assignment, and confidence score logging.

D. Alert System

The alert mechanism is a critical component of the proposed system, designed to notify owners when a dog remains undetected for a predefined duration (one hour). The timer-based module operates as follows:

- Detection Monitoring: After the model identifies and labels a dog in a video frame, the system initializes a timer for that specific dog.
- Timer Reset Logic: If the dog reappears in subsequent frames within the hour, the timer resets to zero. This ensures that transient occlusions (e.g., temporary movement outside of the camera field of view) do not trigger false alarms.
- Alert Trigger: If the dog is continuously undetected for one hour, the timer expires, and the system sends an alert (email, sound, or log).

III. Result

A. Detection Accuracy

Precision reached 1.0 at confidence threshold 0.913. Figure 1 shows the curve from training validation. Our model outperformed generic detectors that don’t distinguish between individuals [9].

B. Recall Curve

The graph 2 reflects the change in recall of the model for different confidence thresholds and is used to evaluate performance under varying sensitivity.

C. Confusion Matrix

Figure 3 shows the model’s classification performance for Minney, Lucky, and background. The model performs especially well with Lucky, achieving 1.00 accuracy.

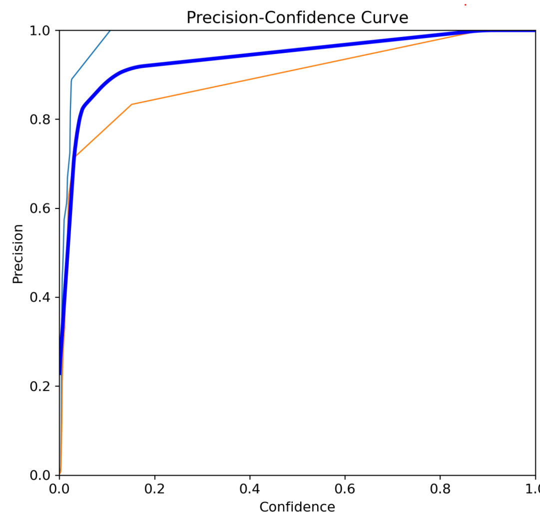

Fig. 1. Precision and confidence curve

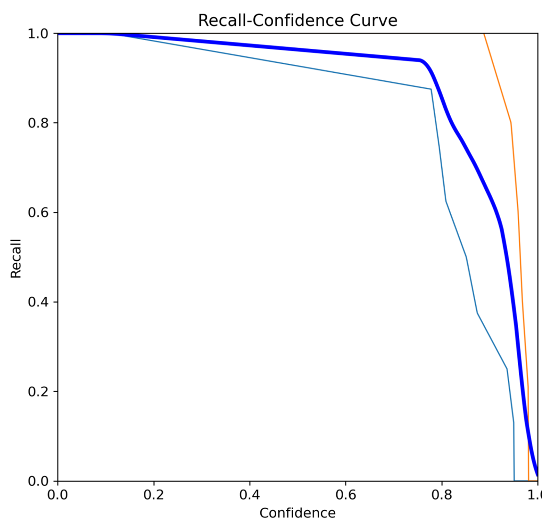

Fig. 2. Recall and confidence curve

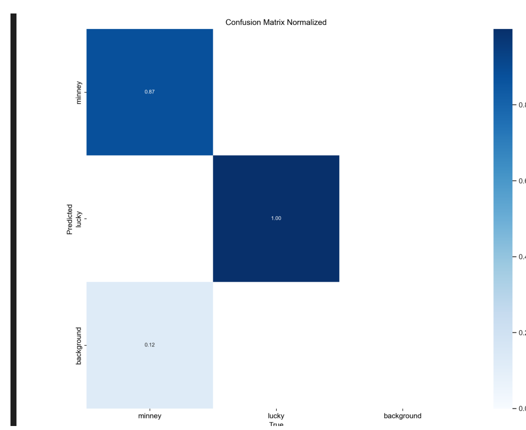

Fig. 3. Confusion Matrix Normalized

D. Real-Time Tracking

Figure 4 shows the successful labeling of two dogs from separate video streams. Label stability was maintained despite occlusion and motion.

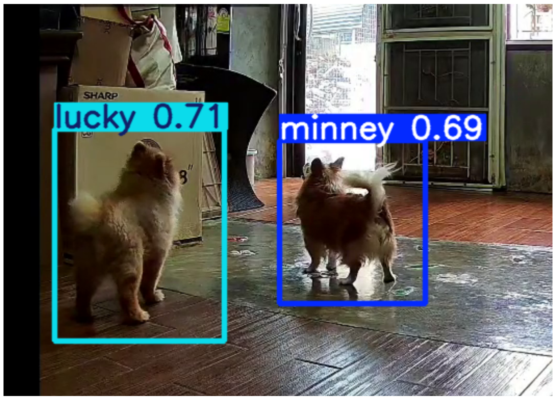

Fig. 4. Named detection output ("Minney" and "Lucky")

IV. Discussion

This system shows real-time pet recognition is feasible without biometrics or wearables. Compared to Kim and Moon's wearable-camera fusion [6], our method is more deployable in average homes. While Smith and Johnson's embeddings achieve higher precision in lab conditions [7], our YOLO model balances speed and generalizability. Future improvements could include embedded deployment or voice-activated control.

V. Conclusion

We propose a real-time system that detects, identifies, and monitors household dogs using YOLOv11m. Combining video input with dog identity tracking and alert generation makes this a useful tool for smart homes and shelters.

References

[1] A. Johnson and E. Lee, "Smartpet: Real-time pet tracking in smart homes," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4555–4564, 2020.

[2] S. Akhtar and H. Raza, "Smart pet care: An iot-based solution for monitoring and caring for domestic animals," *International Journal of Distributed Sensor Networks*, vol. 16, no. 7, 2020.

[3] X. Liu and J. Wang, "Petcam: A lightweight framework for pet detection and activity monitoring," *Sensors*, vol. 21, no. 3, p. 698, 2021.

[4] L. Brown, M. Hart, and J. Kim, "Deeppets: Deep learning for pet monitoring using surveillance data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019, pp. 1050–1055.

[5] R. Singh, S. Bansal, and M. Bhatnagar, "Deep learning-based surveillance system for pet activity monitoring using yolo and lstm," in *2021 IEEE International Conference on Smart Technologies (ICST)*, 2021, pp. 200–205.

[6] J. Kim and N. Moon, "Dog behavior recognition based on multimodal data from a camera and wearable device," *Applied Sciences*, vol. 12, no. 6, p. 3199, 2022.

[7] M. Smith and L. Johnson, "Dog face recognition using deep feature embeddings," in *Proceedings of the 2023 International Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, 2023, pp. 1234–1240.

[8] C. Caya, J. Suchanek, and A. Razi, "Dog identification system using nose print biometrics," in *International Conference on Biometrics*, 2022.

[9] Y. Boulaouane, M. Garifulla, J. Lim, and J. Lim, "Advancing pet biometric identification: A state-of-the-art unified framework for dogs and cats," *arXiv preprint arXiv:2403.12415*, 2024.

[10] M. Garcia and T. Perez, "Multi-camera object tracking using yolo and deep sort," *IEEE Access*, vol. 8, pp. 29 510–29 520, 2020.

[11] J. Lee and Y. Choi, "Smartkennel: Embedded dog monitoring and feeding system," *Electronics*, vol. 11, no. 6, p. 884, 2022.

[12] C. Fernandez and P. Nunez, "Petwatch: Real-time video-based pet surveillance with anomaly detection," in *2021 IEEE International Conference on Consumer Electronics (ICCE)*, 2021, pp. 1–4.

[13] Z. Wang and Y. Chen, "Tracking pets in smart homes: Deep learning and iot integration," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 2501–2514, 2023.

[14] Y. Zhang, H. Wang, and X. Li, "Animal detection and classification from camera trap images using deep learning," *Sensors*, vol. 21, no. 12, pp. 1–15, 2021.

[15] A. Turečková, T. Holík, and Z. K. Oplatková, "Dog face detection using yolo network," in *Proceedings of the Mendel Conference on Soft Computing*, Brno, Czech Republic, 2020, pp. 18–23.