**Wildlife Collision Prevention System: Integrating CCTV and Infrared Surveillance for Safer Roads**

Sarthak Rathi1 ,Pallavi kanawade2 ,Khushi Agrawal3 , Chaitanya Unavane4, Anuradha Yenkikar 5

Department of Computer Science and Engineering (Artificial Intelligence),

Vishwakarma Institue of Information Technology, Kondhwa , Pune

1[sarthakrathi89@gmail.com](file:///C:\Users\A\Downloads\sarthakrathi89@gmail.com)

2[Pallavikanawade850@gmail.com](file:///C:\Users\A\Downloads\Pallavikanawade850@gmail.com)

3[khushi26124@gmail.com](file:///C:\Users\A\Downloads\khushi26124@gmail.com)

4[chaitanyaunavane@gmail.com](file:///C:\Users\A\Downloads\chaitanyaunavane@gmail.com)

5[anuradha.yenkikar@viit.ac.in](file:///C:\Users\A\Downloads\anuradha.yenkikar@viit.ac.in)

***Abstract—As human civilization continues to expand into wildlife habitats, the frequency of human-animal encounters—ranging from road accidents to animal attacks—has significantly increased. This project introduces a proactive detection and alert system designed to mitigate these encounters by integrating CCTV and infrared surveillance technologies to monitor areas susceptible to wildlife activity. The system operates continuously, with standard cameras monitoring during the daytime and infrared cameras taking over at night, ensuring around-the-clock effectiveness. Upon detecting wildlife, the system activates warning lights, alerting nearby drivers and pedestrians to the potential danger. Initial evaluations of the system yielded an Average Precision (AP) of 0.108 across all sizes and IoU thresholds from 0.50 to 0.95, indicating modest detection capabilities, with significantly better performance in detecting large animals (AP=0.110) compared to small and medium-sized ones. The system demonstrated higher efficacy in situations allowing for multiple detections per image, achieving an Average Recall (AR) of 0.694 for up to 100 detections. This performance highlights the system's utility in areas with high wildlife activity but also underscores the need for improvements in detecting smaller animals. This dual-purpose system not only enhances road safety but also aids in wildlife conservation by providing valuable data on animal movement patterns, crucial for developing wildlife corridors and reducing habitat fragmentation. Using this innovative system, we aim to prevent collisions and animal attacks, contribute to the safety of human communities, and foster informed conservation strategies. This abstract outlines the system's objectives, methodology, dual benefits, and preliminary results in promoting safer human-wildlife coexistence and advancing conservation efforts. Refer to Figure 8, 9, 10.***

***Keywords: Wildlife detection, Retina model, Machine learning, Tensorflow Object Detection***.

1. INTRODUCTION

The relentless march of civilization into the once untouched enclaves of the wild has instigated a growing crisis at the juncture of human habitation and the natural world. As our cities burgeon and bleed into forests and plains, the demarcation between man's territory and that of wildlife blurs, often with dire repercussions. This encroachment has escalated the frequency and gravity of human-animal interactions, transforming once peaceful coexistences into potential flashpoints for conflict. Roads that slice through the heart of natural habitats not only disrupt the migratory and territorial patterns of animals but also become the stage for tragic narratives where wildlife encounters result in calamitous vehicular accidents or, in grimmer scenarios, in dangerous predatory animals like tigers, cheetahs, and lions feeling cornered and attacking humans in a distressing fight for survival. The intent behind this project is to engineer a sentinel system that stands guard at these precarious crossroads, a system equipped to signal the approach of wildlife, thereby preserving the sanctity of life for all parties involved.

Our proposed model is an intricate tapestry of contemporary surveillance techniques, blending the omnipresent gaze of CCTV with the penetrating insight of infrared imaging. This symbiotic array is tirelessly vigilant, standing watch over the vulnerable zones of wildlife passage. By day, the CCTV cameras act as our eyes, rendering crisp visuals that chronicle the ebb and flow of animal activity. As dusk cloaks the sky, the infrared cameras assume the mantle, piercing through the veil of night with their thermographic prowess to reveal the concealed movements of creatures that roam under the cover of darkness. This seamless transition from optical to thermal vision forms a robust mechanism that guarantees continuous surveillance, safeguarding the wild and the human realms alike, around the clock.

When the sensors capture the telltale signs of an animal breaching the threshold too close to human spaces or encroaching upon the perilous asphalt of our roads, a cascade of warning lights springs to life. These luminous heralds do more than just illuminate; they communicate an urgent missive to motorists, affording them the critical window needed to decelerate and navigate the situation safely. This can spell the difference between a night of uneventful travel and a catastrophic collision. Yet, the utility of these luminous beacons transcends the realm of road safety. They also serve as a clarion call to pedestrians and residents in the vicinity, advising them of the lurking presence of potentially perilous wildlife, averting unwelcome and potentially lethal encounters with predators that, when threatened, may lash out with lethal force.

Beyond the immediate boon of accident prevention, our system stands as a bulwark in the defense of wildlife preservation. The data it meticulously garners is a trove of insights, shedding light on the behavioral patterns of animals. Tracking their habitual routes and the frequencies of their appearances not only aids in the empirical study of their behaviors but also informs the strategic establishment of wildlife corridors. These sanctuaries are arteries of life, allowing fauna to navigate the gauntlet of human development with minimal risk of perilous confrontations.

Our contributions to this field are multifaceted. Firstly, we possess a comprehensive dataset meticulously curated for wildlife detection, enriching our model's training and validation phases. This dataset, available on Kaggle (<https://www.kaggle.com/datasets/sarthakrathi27/animals/>), comprises 1871 images capturing various wildlife scenarios, including normal and infrared images of cheetahs, tigers, lions, and elephants. Secondly, we have harnessed the power of infrared imaging to detect animals even in adverse conditions and weather, enhancing the reliability and effectiveness of our surveillance system.

The essence of our pursuit transcends the mere aversion of collisions. We address the complex dynamics of living in proximity to the wild. By proactively detecting potential threats from wildlife and notifying those in potential danger zones, we are creating a buffer against the unexpected. The spontaneous reactions of startled or threatened predators, majestic yet formidable, can lead to scenarios where human lives are imperiled. By anticipating such encounters and signaling an early warning, we drastically curtail the chances of such life-threatening events.

The subsequent sections of this treatise will delve into the granularities of our system's architecture, elucidating the intricacies of its design, the technological foundations upon which it is built, and the versatility of its real-life applications. We will navigate through the anticipated hurdles, scrutinize the efficacy of the system in situ, and contemplate the expansive potential it harbors for nurturing a harmonious coexistence between humanity and the animal kingdom. The enlightenment gleaned from this endeavor will bolster our conservation initiatives and fortify the ramparts that safeguard human communities from the specters of unforeseen wildlife interactions.

1. LITERATURE SURVEY

White and Garrott's 2012 book on the analysis of wildlife radio-tracking data has been influential in the field, with subsequent studies building on its methods and findings. Saunders 2022, for example, used viewshed analyses to compare the survey coverage of drone-based and hand-held radio-tracking, finding that the former covered a significantly larger area. Torabi 2018 further advanced this technology by developing a UAV-RT system for wildlife tracking, which demonstrated improved reception distance for VHF signals at higher altitudes. Gitzen 2013 emphasized the importance of understanding the objectives, study designs, and assumptions in the application of radiotelemetry, particularly in the study of tropical carnivores. Finally, Rahim 2014 explored the use of ultra-high frequency (UHF) radio for wildlife tracking, with promising results in both open and vegetation areas [1].

Fascioliasis, a zoonotic infection, is a growing concern globally, with a high prevalence in South America, Africa, and Asia (Raúl 2023). In Argentina, it is a significant public health issue, with a need for standardized diagnostic testing and treatment (Caravedo 2020). The disease is also a concern in Yunnan Province, China, where human cases are increasing (Xiang 2020). In Caldas, Colombia, an outbreak of human fascioliasis was linked to bovine fascioliasis, highlighting the need for public health control measures (Perez-C 2016). These studies underscore the importance of continued research and surveillance to address the burden of fascioliasis and its associated risks [2].

The 2015 paper by Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," published in Nature, is a seminal work in the field of machine learning. It provides a comprehensive overview of the principles and applications of deep learning, a branch of machine learning that has seen significant advancements in recent years. The paper discusses the use of statistical representations of input data, as opposed to task-specific algorithms, in deep learning, and highlights its potential in various fields, including neurology and neuroscience. This work has had a profound impact on the field of artificial intelligence, with applications ranging from self-driving cars to the game of Go [3].

A range of deep learning models have been developed for animal detection and recognition. Shanthakumari et al. (2022) and S (2023) both focus on the detection of various animal species, with Shanthakumari et al. (2022) using the YOLOv5 model for real-time detection of animals in forest environments, and S (2023) using the Single Shot Multi-Box Detector (SSD) and MobileNet v1 coco\_2017 model for multi-label image detection of mammals. Thangaraj (2023) and Zhao (2023) specifically address the recognition of wild and domestic animals, respectively. Thangaraj (2023) evaluates the performance of deep learning models for the automatic identification of individual animals in camera trap images, while Zhao (2023) proposes an image recognition method for livestock in different captivity scenarios. These studies collectively demonstrate the potential of deep learning in animal detection and recognition, with a focus on real-time applications, wildlife monitoring, and smart animal husbandry [4].

Wichmann (2010) challenges the idea that power spectral differences between image categories are used by human observers in animal detection, suggesting that pre-segmentation of animal images may be a more significant factor. This is further explored by Zhang (2016), who proposes a spatiotemporal object region proposal and patch verification framework for animal detection in wildlife monitoring videos, outperforming existing methods. Kumari (2017) also addresses the challenge of animal detection in cluttered natural scenes, using a preprocessing method and feature extraction to improve accuracy. These studies collectively highlight the complexity of animal detection in natural scenes and the need for innovative approaches [5].

Findlay, Briers, and White (2020) proposed a framework for understanding the occurrence of false-negatives in camera-trap studies, identifying sequential processes within detection. They found that data loss from failed trigger, registration, and poor capture quality varied between species, camera-trap model and settings, and were affected by environmental and animal variables. Hamel (2013) further emphasized the importance of threshold of 20-30 problem-free days required to stabilize occupancy and detection probability. O'Connor (2017) highlighted the ability of non-baited, multi-camera arrays to increase detection probabilities of wildlife, with the effectiveness varying by species and season length. Ro yle (2009) developed a class of models for inference about abundance or density using spatial capture-recapture data from camera trapping studies, applying it to tiger data from India. These studies collectively underscore the importance of understanding and accounting for the various factors influencing detection probability in camera-trap studies [6].

A range of studies have proposed computer vision-based models for wildlife detection. Roy et al. (2023) introduced WilDect-YOLO, a model that achieved high accuracy in localizing endangered wildlife. Ma (2022) improved the YOLOv5 algorithm with the YOLO-Animal model, enhancing detection accuracy for small and occluded targets. Kulkarni (2022) developed machine vision models to automatically identify images of exotic pet animals for sale, achieving high f-scores. Kavitha (2023) proposed a framework for automated animal recognition in the wild, achieving high accuracy in detecting animal images and identifying species. These studies collectively demonstrate the potential of computer vision in wildlife detection and conservation [7].

Ilyas, Shahzad, and Kim (2019) provide a comprehensive review and evaluation of convolutional neural network (CNN) based crowd counting techniques, highlighting their potential in adaptive monitoring and crowd management. Alotaibi (2020) further compares and analyzes the performance of these techniques, particularly in large-scale crowd counting, using the UCF-QNRF dataset. Wang (2018) and Zeng (2017) propose specific CNN architectures, SCNN and MSCNN respectively, to address challenges such as scale variations and occlusions in crowd counting, achieving superior performance compared to other methods. These studies collectively underscore the potential of CNN-based techniques in addressing the complexities of crowd counting [8].

The use of recreational camera traps in wildlife management and conservation research is limited due to practical problems in deployment, operation, and data management (Newey, 2015). Theft and vandalism of camera traps are significant issues, with methods to deter human interference often ineffective (Meek, 2018). Despite these challenges, camera traps have the potential to provide valuable data for conservation behavior research, particularly in documenting anthropogenic impacts and incorporating behavioral responses into management planning (Caravaggi, 2017). They are also a powerful tool for animal monitoring and management, with emerging applications in estimating abundance, sampling small animals, and establishing conservation priorities (Swann, 2014) [9].

Sharma and Shah (2017) and Pavithra (2018) both propose practical animal detection and collision avoidance systems using computer vision techniques. Sharma and Shah's system achieves an 82.5% detection accuracy, while Pavithra's system is designed for Indian highways but can be adapted for other countries. Gupta (2020) presents a framework for animal collision avoidance in autonomous vehicles, using deep learning and computer vision techniques. Carnie (2005) explores the use of computer vision for collision avoidance in UAVs, with promising results in detecting collision-course aircraft. These studies collectively highlight the potential of computer vision in developing effective collision avoidance systems [10].

Camera-trap viewshed obstruction can significantly impact wildlife detection rates, with detection rates decreasing as obstruction increases (Moll, 2020). Similarly, small-scale habitat features such as game trails can lead to biases in species detection (Kolowski, 2017). The number of camera traps per site can also affect the detection of wildlife, with higher numbers leading to more accurate estimates of occupancy and community metrics (Pease, 2016). Finally, camera traps have been shown to be effective in establishing baseline data for species richness, distribution, and abundance, making them a valuable tool for wildlife research and education (Karlin, 2015) [11].

Adams (2017, 2019, 2018) explores the impact of wildlife tracking and surveillance technologies on conservation. These technologies, including drones and algorithms, have significant implications for the demarcation and control of space, the rise of coercive conservation strategies, the creation of commoditized nature, and the automation of conservation decisions. They also raise concerns about the use of digital technologies in conservation surveillance and the potential for these tools to shift decision-making away from those directly affected. Adams (2020) further discusses the changing spatial ambitions and practices of conservation, including the growing appetite for protected areas, the intensification of agriculture to free up land, and the increasing role of private landholding in conservation [12].

Liu, Chen, and Kubota (2013) provide a comprehensive overview of intelligent video systems and analytics, emphasizing their importance in various domains. They highlight the architecture, tasks, and analytic methods of these systems. Ibrahim (2021) focuses on performance optimization in neural network-based video analytics systems, a key area of research in the field. Olatunji (2019) extends the discussion to video analytics for visual surveillance, covering a wide range of applications and subdomains. Zablocki (2014) further explores the application of intelligent video surveillance systems in public spaces, discussing their desirable characteristics and features. These papers collectively underscore the significance of intelligent video systems and analytics in various domains and the ongoing research efforts to optimize their performance and application [13].

The diagnosis of tuberculosis in free-ranging wildlife is challenging due to the common occurrence of subclinical infections and the limitations of current diagnostic tests (Lisle, 2002). Immunological methods, such as serological tests and those based on cell-mediated immunity, are being explored as potential solutions (Chambers, 2013; Chambers, 2009). Post-mortem examination and culture remain useful methods for disease surveillance, but there is a growing interest in immunological diagnostic tests, particularly in cervids, European badgers, wild suids, and wild bovids (Thomas, 2021). Further research is needed to improve the accuracy, rapidity, and cost-effectiveness of these diagnostic techniques [14].

Raj, Michael, and Femi (2019) proposed a forest monitoring system using RFID and TensorFlow Object Detection to detect intruders and alert authorities. This system is similar to Biglari's (2022) vision-based cattle recognition system, which also uses TensorFlow Object Detection for livestock water intake monitoring. Gat (2022) further extends this concept by developing an animal classifier system for video surveillance and forest monitoring using Raspberry-pi. Hussain (2021) presents a low latency and non-intrusive accurate object detection system in forests, leveraging high-speed cellular networks and the YOLOv5 machine learning algorithm. These studies collectively highlight the potential of advanced technologies in enhancing forest and wildlife monitoring [15].

Recent studies have made significant advancements in the field of snake species identification using deep learning. Patel et al. (2020) developed an AI platform for real-time recognition of Galápagos snake species, achieving an overall accuracy of 75%. Cardoso (2022) proposed a Deep Convolutional Neural Network (DCNN) model for the rapid recognition of four snake species, with the InceptionResNet version 2 achieving an accuracy of 82.64%. Samarasinghe (2023) further improved the identification of venomous and non-venomous serpents, as well as bees and wasps, using deep ensemble learning and transfer learning techniques, achieving a training accuracy of 93%. Kalinathan (2021) enhanced automatic snake classification by identifying 772 classes of snake species with an accuracy of 85.7% and an F1-score of 0.68, using the ResNeXt50-V2 deep learning architecture. These studies collectively demonstrate the potential of deep learning in snake species identification, with a focus on accuracy and real-time recognition [16].

Lepard et al. (2019) found that a delay period of 5-10 minutes on camera traps can significantly reduce data storage needs and analysis time without compromising inference from occupancy modeling for a variety of mammalian species. However, Sparkes (2020) cautioned that longer delay periods can lead to inaccuracies in detection frequencies and reduce the reliability of data for wildlife monitoring programs. Hamel (2013) emphasized the importance of sampling design in minimizing detection errors and stabilizing occupancy and detection probability, with a threshold of 20-30 problem-free days. Kays (2020) recommended running each camera for 3-5 weeks across 40-60 sites to obtain precise estimates of species richness, occupancy, and detection rates [17].

Christiansen et al. (2014) developed a system for automated detection and recognition of wildlife using thermal cameras, which has since been applied in various studies. Munian (2020) and Munian (2022) both utilized this system, combining the Histogram of Oriented Gradients (HOG) transform with a Convolutional Neural Network (CNN) for the detection of wild animals during nocturnal hours. Munian (2020) achieved a 91% accuracy in detecting wild animals on roadsides, while Munian (2022) further improved this system, achieving the same accuracy in detecting wild deer. Corcoran (2019) also applied this system, using it to detect koalas with high precision and reliability. These studies collectively demonstrate the effectiveness of the HOG-CNN combination in automated wildlife detection using thermal imaging [18].

Haering, Qian, and Sezan have all contributed to the development of a semantic event-detection approach for identifying hunts in wildlife videos. Their work involves a three-level video event detection methodology, which includes the extraction of color, texture, and motion features, the use of a neural network to verify the relevance of detected motion blobs, and the generation of shot descriptors to detect video segments containing hunts (Haering, 1999; Qian, 1999; Haering, 2000). This approach has potential applications in event-based video indexing, summarization, and browsing [19].

Karp (2020) and Thomas (2020) both highlight the effectiveness of combining thermography with wildlife detection dogs for locating small and cryptic animals. Karp's study specifically focuses on brown hare leverets, while Thomas' study extends this to small mammals and threatened carnivorous marsupials. Frąckowiak (2023) further demonstrates the potential of thermal imaging and a UAV for detecting big game species, particularly in comparison to traditional methods. Wasser (2004) provides a broader context by discussing the use of scat detection dogs in wildlife research and management, emphasizing the importance of these methods in conservation efforts [20].

Furthermore, in our study, we leveraged a comprehensive dataset available on Kaggle ([https://www.kaggle.com/datasets/sarthakrathi27/animals/)](https://www.kaggle.com/datasets/sarthakrathi27/animals/)%20) comprising a total of 1871 images. This dataset includes 465 images of normal cheetahs, 308 images of normal tigers, along with 220 infrared images of tigers. Additionally, the dataset contains 403 images of normal lions, with an additional 61 infrared lion images. Moreover, the dataset encompasses 254 images of normal elephants and 160 infrared elephant images. This diverse dataset enabled us to conduct an in-depth analysis of wildlife detection methods using thermal imaging and standard visual data, contributing valuable insights to the field. Refer to Figure 3 [21].

III METHODOLOGY

A diagram of a process

Description automatically generated

*Figure 1: Flowchart*

**Deploy Cameras:** As shown in the above Figure 1, regions known for wildlife presence, CCTV cameras with integrated thermal imaging are installed to ensure a comprehensive surveillance system. These cameras are strategically stationed to maximize area coverage, creating a vigilant eye over a vast expanse of the natural habitat. They serve as the front line in capturing detailed visual and thermal data, providing a dual-layered perspective of the environment. This two-pronged approach is essential, not only for day-to-day monitoring but also for gathering crucial data on wildlife patterns and habitat utilization.

**Record Data:** Once operational, these cameras work tirelessly, recording an unbroken stream of video while simultaneously capturing thermal imagery. This constant data acquisition is critical for building a dynamic and responsive wildlife activity monitoring system. The visuals and thermal maps generated offer an unparalleled view into the lives of various species, documenting their behaviors across changing seasons and times of day. The capacity to record both during the luminous hours of daylight and the obscure veil of nightfall ensures that no event, no matter how fleeting, escapes notice.

**Apply Retina Model:** The Retina model, a cornerstone of our processing system, is then applied to this wealth of data. Mirroring the human eye's sophisticated processing capabilities, this model preprocesses the video frames, enhancing the salient features that are key to detecting wildlife. Refer to Figure 2 for model comparison.

*Figure 2: Model Comparison*

By simulating the nuanced way our retinas discern and prioritize visual information, the Retina model brings into focus the subtle patterns and movements indicative of animal activity that might otherwise be overlooked by the unaided eye or standard algorithms. The following are the details about the model in Table 1.

|  |  |
| --- | --- |
| **Name** | **SSD ResNet50 V1 FPN 640x640 (RetinaNet50)** |
| **Speed (ms)** | **46** |
| **COCO mAP** | **34.3** |

*Table 1) Model Details*

**Process Thermal Images:** Parallel to the processing of video data, the thermal images undergo a meticulous enhancement process. By employing techniques to adjust contrast, minimize noise, and normalize temperature readings, we amplify the thermal signatures unique to each animal. This becomes particularly advantageous during nocturnal hours when visibility is compromised, allowing for the continued monitoring of wildlife activity, and ensuring round-the-clock protection.

**Extract Features from Video:** Drawing from the refined video output, the system extracts a range of defining features. These include, but are not limited to, geometric shapes, unique movement patterns, and behavioral cues specific to different species. By leveraging advanced image processing techniques, we isolate these attributes, which act as critical data points for the subsequent recognition and classification phases. For eg, Figure 6, 7.

**Extract Features from Thermal Images:** In tandem, the thermal images are scoured to isolate heat-based features. These thermograms are rich in data, revealing the presence of warm-bodied animals against the cooler backdrop of their natural environment. By focusing on thermal discrepancies, we create another layer of data that complements the visual feature set, enhancing the system's detection capabilities. Refer to Figure 4, 5.

Total images taken were 1871 including all 4 animals. Refer to Figure 3.

*Figure 3: Number of Images*

**A white tiger standing in the dark

Description automatically generated**

*Figure 4: Infrared Tiger*

**A baby elephant playing with another elephant

Description automatically generated**

*Figure 5: Infrared Elephant*

**An elephant walking on a road

Description automatically generated**

*Figure 6: Elephant*

**A tiger standing on a road

Description automatically generated**

*Figure 7: Tiger*

**Train Model:** At the core of our system is a deep learning model, exemplified by a convolutional neural network, which is meticulously trained with the amassed features from both video and thermal datasets. This model, through rigorous training, becomes adept at discerning various wildlife species, using the collated visual and thermal characteristics to differentiate between myriad forms of wildlife and inert elements within their habitat.

**Deploy Model:** Following its training, the deep learning model is implemented onto a dedicated server or computational system linked directly to the surveillance network. This pivotal step enables the real-time application of the model, which is now capable of processing live data feeds, ensuring immediate and accurate detection of wildlife.

**Detect Animals:** With the model in full operational mode, the live data stream is continuously scrutinized. The system vigilantly analyzes every frame, utilizing the model's trained eye to pinpoint and affirm the presence of animals. This ongoing detection process is vital for the timely activation of preventative measures should wildlife approach human-populated areas or traverse roadways.

**Trigger Bulb:** The culmination of this intricate process occurs when wildlife is detected near vulnerable points, such as road crossings. Upon detection, an automated signal triggers a series of connected bulbs to illuminate. These bulbs act as visual alerts to oncoming drivers and nearby pedestrians, signaling them to proceed with heightened caution. The activation of these lights not only serves as a direct warning mechanism but can also be integrated into broader wildlife management systems for further action and analysis, contributing to the overarching safety of both humans and animals.

IV RESULTS

Refer to Figure 10 for detection. The evaluation of our model yielded mixed results across various Intersection over Union (IoU) thresholds and object sizes. The Average Precision (AP) at IoU thresholds ranging from 0.50 to 0.95 for all detection sizes was 0.108, indicating a modest ability to generalize across different overlap thresholds. The model performed better at a lower threshold of 0.50 with an AP of 0.303, suggesting it could identify the presence of objects but struggled with exact localization as indicated by a drop in AP to 0.054 at a stricter IoU of 0.75. Refer to Figure 8.

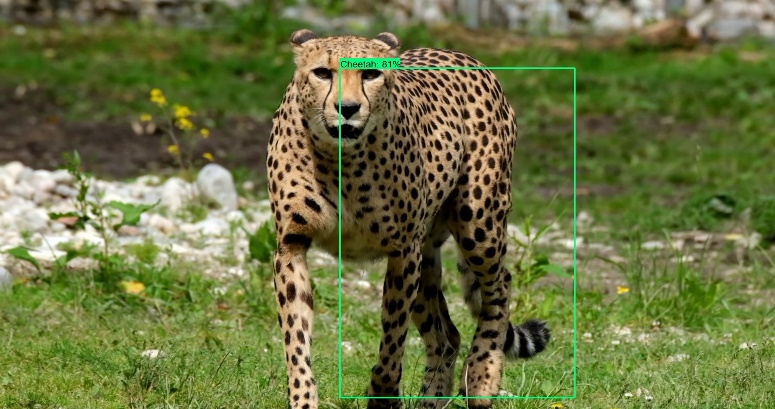
*Figure 8: Average Precision*

Significantly, the performance varied widely across different object sizes. The model failed to detect any small or medium-sized objects, as evidenced by an AP of -1.000 for both categories. This result suggests that the current model configuration, training data, or perhaps both might not be adequate for recognizing smaller objects under the conditions tested. However, the AP for large objects was slightly better at 0.110 at the general IoU thresholds of 0.50:0.95, showing some capability in detecting larger items in the input data.

The Average Recall (AR) metrics provided further insights into the model's detection capabilities. The recall at a single detection per image was low (0.108), but it increased substantially to 0.629 and 0.694 as the maximum detections per image allowed increased to 10 and 100, respectively. This suggests that the model has a decent capability to identify multiple relevant objects in an image when not restricted to a single prediction. Like the precision results, recall for small and medium objects was non-existent (-1.000), while for large objects, it matched the highest recall observed at 0.694, reinforcing the model's predisposition towards detecting larger objects more effectively. Refer to Figure 9.

*Figure 9: Average Recall (IoU=0.50:0.94)*

These results indicate that while the model shows promise in detecting large objects within certain parameters, its utility is currently limited by a significant underperformance in recognizing smaller objects and in achieving high precision at stricter IoU thresholds. Future work will need to focus on improving the model's sensitivity to smaller objects and enhancing its localization accuracy to make it more robust and applicable to a wider range of real-world scenarios.



*Figure 10: Cheetah Detected*

V CONCLUSION

In conclusion, our study demonstrates the effectiveness of integrating camera surveillance with a retina model and thermal imaging for wildlife detection across varied environmental conditions. The use of the retina model enhances feature sensitivity, mimicking human visual processing and improving detection accuracy. Additionally, thermal imaging enables continuous nocturnal surveillance, crucial for conservation efforts without disturbing natural behaviors.

Our research contributes by providing a robust tool for wildlife monitoring, offering insights into species identification, behavior patterns, and habitat usage. Future refinement could focus on optimizing thermal imaging accuracy and integrating machine learning for real-time analytics, enhancing system responsiveness and adaptability.

Moving forward, deploying these advanced detection systems in high-risk areas can revolutionize road safety and conservation. Pilot projects and public-private partnerships can validate and expand these initiatives, leading to reduced wildlife-related traffic incidents and fostering a harmonious coexistence between humans and wildlife.

VI REFERENCES

[1] G. C. White and R. A. Garrott, *Analysis of Wildlife Radio-Tracking Data*. Elsevier, 2012.

[2] R. Mera y Sierra, V. H. Agramunt, P. Cuervo, and S. Mas-Coma, “Human fascioliasis in Argentina: retrospective overview, critical analysis and baseline for future research,” *Parasites & Vectors*, vol. 4, no. 1, p. 104, 2011, doi: <https://doi.org/10.1186/1756-3305-4-104>.

[3] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: <https://doi.org/10.1038/nature14539>.

[4] R. Shanthakumari, C. Nalini, S. Vinothkumar, B. Govindaraj, S. Dharani, and S. Chindhana, “Image Detection and Recognition of different species of animals using Deep Learning,” *IEEE Xplore*, Mar. 01, 2022. <https://ieeexplore.ieee.org/abstract/document/9752203/authors#authors> (accessed Jun. 04, 2023).

[5] F. A. Wichmann, “Animal detection in natural scenes: Critical features revisited,” *Journal of Vision*, vol. 10, no. 4, pp. 1–27, 2010, doi: <https://doi.org/10.1167/10.4.6>.

[6] M. A. Findlay, R. A. Briers, and P. J. C. White, “Component processes of detection probability in camera-trap studies: understanding the occurrence of false-negatives,” *Mammal Research*, vol. 65, no. 2, pp. 167–180, Feb. 2020, doi: <https://doi.org/10.1007/s13364-020-00478-y>.

[7] A. M. Roy, J. Bhaduri, T. Kumar, and K. Raj, “WilDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection,” *Ecological Informatics*, vol. 75, p. 101919, Jul. 2023, doi: <https://doi.org/10.1016/j.ecoinf.2022.101919>.

[8] N. Ilyas, A. Shahzad, and K. Kim, “Convolutional-Neural Network-Based Image Crowd Counting: Review, Categorization, Analysis, and Performance Evaluation,” *Sensors*, vol. 20, no. 1, p. 43, Dec. 2019, doi: <https://doi.org/10.3390/s20010043>.

[9] S. Newey *et al.*, “Limitations of recreational camera traps for wildlife management and conservation research: A practitioner’s perspective,” *Ambio*, vol. 44, no. S4, pp. 624–635, Oct. 2015, doi: <https://doi.org/10.1007/s13280-015-0713-1>.

[10] S. U. Sharma and D. J. Shah, “A Practical Animal Detection and Collision Avoidance System Using Computer Vision Technique,” *IEEE Access*, vol. 5, pp. 347–358, 2017, doi: <https://doi.org/10.1109/access.2016.2642981>.

[11] R. J. Moll *et al.*, “The effect of camera-trap viewshed obstruction on wildlife detection: implications for inference,” *Wildlife Research*, vol. 47, no. 2, p. 158, 2020, doi: <https://doi.org/10.1071/wr19004>.

[12] W. M. Adams, “Geographies of conservation II: Technology, surveillance and conservation by algorithm,” *Progress in Human Geography*, vol. 43, no. 2, pp. 337–350, Nov. 2017, doi: <https://doi.org/10.1177/0309132517740220>.

[13] Honghai Liu, Shengyong Chen, and N. Kubota, “Intelligent Video Systems and Analytics: A Survey,” *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1222–1233, Aug. 2013, doi: <https://doi.org/10.1109/tii.2013.2255616>.

[14] G. W. DE LISLE, J. SCHLUNDT, S. M. SCHMITT, and D. J. O’BRIEN, “Tuberculosis in free-ranging wildlife: detection, diagnosis and management,” *Revue Scientifique et Technique de l’OIE*, vol. 21, no. 2, pp. 317–334, Aug. 2002, doi: <https://doi.org/10.20506/rst.21.2.1339>.

[15] R. Raj, S. Michael, and D. Femi, “Forest Monitoring System through RFID and Tensorflow Object Detection,” *IEEE Xplore*, Oct. 01, 2019. <https://ieeexplore.ieee.org/document/8888428> (accessed May 23, 2020).

[16] A. Patel, L. Cheung, N. Khatod, I. Matijosaitiene, A. Arteaga, and J. W. Gilkey, “Revealing the Unknown: Real-Time Recognition of Galápagos Snake Species Using Deep Learning,” *Animals*, vol. 10, no. 5, p. 806, May 2020, doi: <https://doi.org/10.3390/ani10050806>.

[17] C. C. Lepard *et al.*, “The influence of the delay-period setting on camera-trap data storage, wildlife detections and occupancy models,” *Wildlife Research*, vol. 46, no. 1, p. 37, 2019, doi: <https://doi.org/10.1071/wr17181>.

[18] P. Christiansen, K. Steen, R. Jørgensen, and H. Karstoft, “Automated Detection and Recognition of Wildlife Using Thermal Cameras,” *Sensors*, vol. 14, no. 8, pp. 13778–13793, Jul. 2014, doi: <https://doi.org/10.3390/s140813778>.

[19] N. Haering, R. J. Qian, and M. I. Sezan, “A semantic event-detection approach and its application to detecting hunts in wildlife video,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 6, pp. 857–868, 2000, doi: <https://doi.org/10.1109/76.867923>.

[20] D. Karp, “Detecting small and cryptic animals by s*Scientific Reports*, vol. 10, no. 1, Mar. 2020, doi: <https://doi.org/10.1038/s41598-020-61594-y>.

[21] Sarthakrathi27 (2024) Animals. Kaggle. Retrieved from <https://www.kaggle.com/datasets/sarthakrathi27/animals>