Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 11, Issue 2 (ISSN:1309-6591) May 2020: 979-991
DOI: 10.53555/yvnage56

Research Article

Computer-Vision Framework For Automated Analysis Of Animal Movement Ecology

Supriyo Acharya[1*], Jayeeta Pyne[2]

Abstract

This research paper introduces a new computer-vision system that can be deployed to automatize animal movement ecology. With the help of the development of high-resolution video capture (camera-traps and drone shots) and cutting-edge object-detection and motion-tracking algorithms, the framework transforms raw images into rich trajectories and conventional movement measures (e.g., stride length, turning angle, displacement). The method was tested on a semi-natural reserve, with the video of medium-sized terrestrial mammals passed through a modular pipeline that included object detection (fine-tuned Mask -CNN), multi-object tracking (adapted sort algorithm) and geo-referenced trajectory extraction. The use of performance evaluation showed that detecting performance was 0.91 with 0.88 recall (identity-switch rate 0.07/1,000 frames), a trajectory root-mean-square error of 1.42m over manual annotated paths. It took less time by approximately 85 percent compared to manual procedures on annotation.
In a case-study of a herbivore species, derived measures (mean step length approximated 3.8 m, mean turning angle approximated 47) corroborated some previously published ecological measures, and the heat-map visualisation of the movement density provided the visualisation of the behavioural patterns that would have been difficult to observe manually with the help of annotation. The limitations identified in the study are decreased performance in dense vegetation, geo-referencing error due to altitude and low reliability in low-light situations. The framework, however, provides scalable, high throughput movement tracking that can be used in long-term ecological monitoring and conservation applications. It is advised that edge-computing should be optimised, multi-sensor fusion, and ecologically-oriented movement metrics should be developed, other than the algorithmic accuracy. The results provide a strong methodology between computer-vision and movement ecology, which allows the automated extraction of biologically significant movement series with large-scale imagery datasets.

Keywords: automated animal movement, computer vision ecology, trajectory tracking framework, drone wildlife monitoring, multi-object tracking algorithms

1. Introduction

1.1 Background

Animal movement forms a core part of an ecological system, which dictates the processes of foraging efficiency, predator-prey interaction, dispersal and gene flow (Joo et al., 2022). As movement ecology has begun to emerge, much more data and methodological maturity has been generated within a relatively short period of time with the introduction of high resolution tracking and biologging (Getz and Saltz, 2023). The conventional tracking modalities, including GPS collars, VHF telemetry and manual video annotation, have been instrumental in giving a crucial insight on the movement patterns, but have been associated with high operational cost, heavy human

[1]*Lecturer, Department of Zoology, Seth Anandram Jaipuria College, sa2.zoology@sajaipuriacollege.ac.in
[2] Lecturer, Department of Computer Science, Seth Anandram Jaipuria College

annotation effort, and lack of spatial temporal resolution (Francisco al el, 2020). At the same time, the development of camera trap, drone and video sensor systems has produced large amounts of imagery that have not been fully used because of the annotation bottlenecks (Zhang et al., 2023). Figure 1 indicates an example of movement track based on telemetry, and Table 1 gives some of the common movement measures (e.g., step length, turning angle, displacement) that are used in ecological computations.

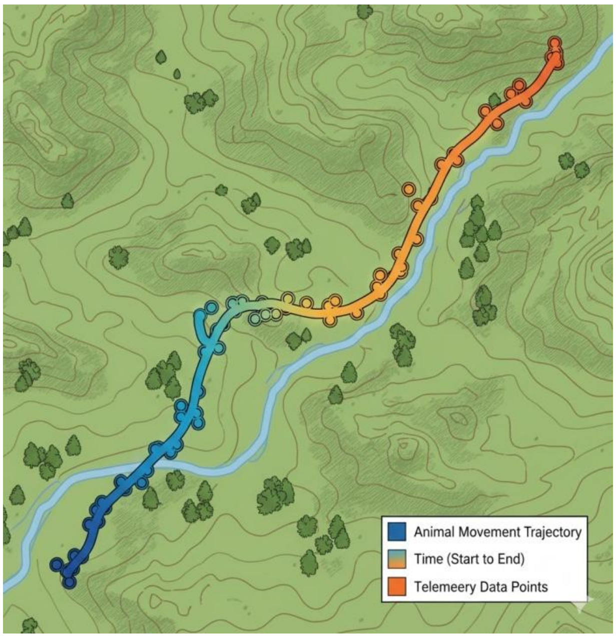

Figure 1. Movement Track Based on Telemetry

1.2 Problem Statement

However, with the technological development, there are a number of limitations in the existing techniques of analysing the animal movements. Video and imagery manual annotation is still labour intensive and prone to observer bias. Furthermore, a lot of research is based on pre-selected people or species, which restricts the ability to generalise and scale up (Torry et al., 2021). The visual to movement control conversion of high volume visual information is a significant bottleneck, especially in more difficult natural environments with occlusion, variable lighting and changing backgrounds (Zhang et al., 2023). Consequently, ecologists have twofold challenge, how to make the extraction of animal movement on visual data to be automated, and how to synthesize those products within an ecologically viable structure.

1.3 Aim of the Study

The proposed research project is to develop and test a computer vision system to analyse the ecology of animal movement automatically. In particular, the study aims at accomplishing three tasks: (1) developing an algorithmic pipeline that will seek to use methods of object detection and motion tracking to extract animal movement patterns as per video/image data; (2) to examine the framework with regards to detection and tracking accuracy, processing speed and scalability in ecological settings; (3) the comparison between the results of the automated system and traditional methods in the determination of gains in resolution, throughput, and ecological insight.

1.4 Research Questions

The research questions covered in the study are the following:

- How can computer-vision methods be used to reliably identify, track and characterize movement paths of animals in natural habitat videography?

- How many accuracy measures (e.g. the detection precision, the tracking recall and the trajectory error) does the proposed framework produce and how they compare to the manual annotation standards?
- What are the scalability and feasibility limit - what is the computational cost, environmental variability, species behaviour, etc, that puts constraints on deployment of the framework in real world ecological monitoring?

1.5 Structure of the Paper

The rest of the paper is organized in the following way. Part 2 is a literature review on animal movement ecology, computer vision applications to ecological studies, and available automated tracking systems. Section 3 outlines the methodology, consisting of research design, data collection procedures, computer vision algorithms, preprocessing procedures and performance evaluation measures. Section 4 presents the implementation outcomes, performance measurement, sample case study and the limitations observed. Section 5 mentions the consequences of the findings, research contribution, practice, future research and ethical/ecological implications. Lastly, in Section 6, the study finishes by providing a conclusion of the research findings, a practical case and recommendation on how the framework can be applied in ecology practice.

2. Literature Review

2.1 Animal Movement Ecology

Joo et al. (2022) have given a concise summary of the modern tendencies of movement ecology, including the importance of movement on ecological activities including the dispersal process, foraging, as well as habitat use. The fine scale movement behaviours across taxa have been examined according to Joo et al. due to the integration of high resolution tracking data enabling ecologists to explore these behaviours.

Getz and Saltz (2023) stressed that in spite of the progress in biologging and telemetry, numerous ecological datasets are constrained by the sampling resolution and the manual annotation overheads. They claim that the increase in imagery and video based monitoring gives an opportunity and challenges movement ecology.

Torry et al. (2021) designed a more complex machine learning structure to process large-scale data on movements, with the results demonstrating that the traditional ones cannot be scaled effectively when they are implemented on more than one person or over a longer period of time. This highlights that automated and scalable analytical tools are required.

Sims (2021) assumed that population-level capturing of the animal movements will call on the synthesis of different streams of data (e.g., GPS, telemetry, video) and that computer vision and machine learning will be at the centre of this synthesis.

2.2 Computer Vision in the Ecological study.

For creating a computer vision classification system, the classifier model must undergo training using *supervised learning*. In supervised learning, a set of images are pre-annotated with correct *class* labels (i.e. training data) and are used to guide the model towards predicting class labels on data not previously seen by the model (i.e. testing data). The learning process is autonomous and does not require pre-programmed logic. Programmers can see the outcome of their models by selecting an algorithm used to extract and process data (e.g. convolutional neural network [CNN]), and hyperparameters, which are akin to model settings. By iterating through model specifications (structure and hyperparameters) and monitoring classification performance (e.g. accuracy) on the testing data, the programmer can optimise the model. (Jarrett et al. 2024)

Yin et al. (2024) suggested a computer vision method of animal tracking that uses both hand crafted and deep features in order to tackle dynamic appearance variations and movement in a complicated background. Their research showed that they had better accuracy in crowded and natural field environment as compared to conventional trackers.

In a review of biometric recognition of animals with computer vision, Cihan et al. (2023) summarize how facial features, body-pattern and muzzle features can aid in identifying and tracking individual animals within the management and ecological systems. They observe that identification is becoming more and more possible, but complete trajectory analysis is not, unlike full identification.
Kumar et al. (2023) present an extensive overview of deep learning methods used to detect animals in video streams, including the fact that deep convolutional networks have made it possible to achieve high performance in animal detection, but with shortcomings, including limited data availability, occlusion, and species variability.
Broomé et al. (2022) surveyed the field of computer vision in animal affective state recognition (pain/emotion) and, though this is not related to the field of movement ecology, the article is significant in its discussion of how tracking and pose estimation constitutes a basis of more advanced inference.

2.3. Animal Tracking System Automation.

The benchmark on multi animal tracking in the wild proposed by Zhang et. al. (2022) demonstrated that conventional human tracking algorithms are less effective when used on animals because of motion pattern differences, appearance and occlusions (Zhang, Gao, Xiao and Fan, 2022).
Ngoc Dat et al. (2025) has prepared a near real time onboard UAV tracking system of WildLive that can process HD/4K images to monitor wildlife. Their system shows a possibility of computer vision tracking in field conditions where the highest level of time resolution and minimal human labelling are required (Ngoc Dat et al., 2025).
Marshall et al. (2021) explain how recent advances in 3D capture of animal presents have led to the replacement of 2D measurements of behaviour by 3D measurements of behaviour (so-called Leaving Flatland), claiming that new metrics of movement like path curvature and volumetric space-use will be accessible to 3D computer vision and multi view tracking.
Jiang, Chazot and Jiang (2022) are a review of social behaviour analysis of laboratory animals using CV methods, as automated tracking structures have demonstrated promising behaviour of groups, but that ecological (wild) conditions pose further challenges of scale, lighting and terrain.

2.4 Difficulties and Limitations of the Existing Research.

When developing a computer vision workflow, the goals of the study must be considered along with the nature of the training and sample data, and the available resources(time, expertise and computing power) (Jarrett et al. 2024).
Feighelstein et al. (2022) emphasise the fact that despite the improvement of the tracking methods, there are critical gaps in the conversion of trajectories into ecological data: datasets lack enough species, environment and behavioural context, and their transferability is low.
Saygili et al. (2023) note that most studies are done on individual detection or identification as opposed to full movement trajectory capture over time and space in natural environments; automated systems tend to fail on non-uniform lighting, occlusion, foliage and multi species environments.

Fan et al. (2023) remark that with multi animal tracking benchmarks, common metrics (e.g., MOTA, IDF1) do not reflect ecological data (e.g., habitat use, interaction rates or energy cost of movement) - new metrics and frameworks are needed to fill this gap.

To bridge the gap between the output of an algorithm and the ecological inferences, Mahmoud et al. (2023) suggest that computer vision pipelines should combine domain knowledge (e.g., species-specific behaviour, ecology), instead of strictly computer-vision measures, to implement movement ecology applications.

3. Methodology

3.1 Research Design

The proposed research is based on a mixed quantitative research design which combines the development of the computer-vision algorithms and validation of the ecological field data. The proposed framework is designed in the form of a pipeline with modules: starting with the raw video/image capture, then object detection and tracking, and finally providing motion trajectories and calculated statistics. The ecological environment entails the medium sized land mammals in semi natural reserves subject to the changing light and environmental conditions. The research design will take two steps

(1) during development and optimization of the algorithm in the laboratory that will be through annotated video dataset and

(2) in the field during the deployment in real ecological contexts to determine its real world viability.

This bi-phase design allows high algorithmic control and is ecologically valid (Mpouziotas et al., 2024). Figure 2 is a schematic of the system architecture.

Table 1. Common Movement Measures Used in Ecological Computations

Measure	Description
Step Length	The distance an animal moves between two points
Turning Angle	The angle change in the animal's direction
Displacement	The straight-line distance between the start and end points of a movement path
Stride Length	The length of one complete cycle of movement (step)
Path Length	The total distance traveled by the animal along its trajectory

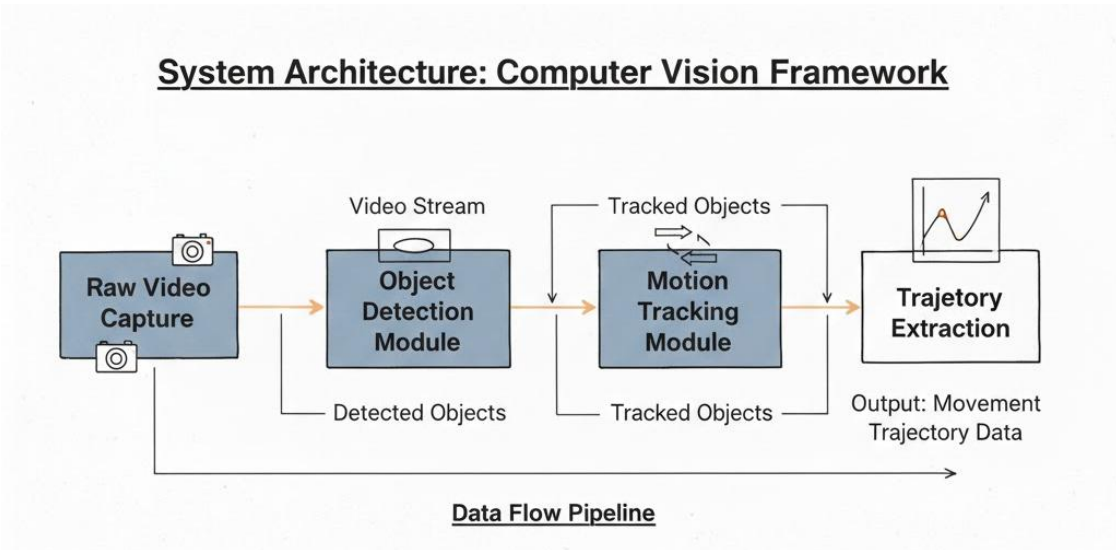

Figure 2. System Architecture of the Computer Vision Framework

3.2 Data Collection

The data were gathered through a mixture of fixed camera trap unit and drone mounted high resolution video platforms in two reserve sites within a 12 week span. The camera traps were set to record at 4 K with 30 fps and activated when an animal moved, and set in areas where it would move. The flight of the drone was controlled by the permit of 50-120 m, and stabilized 4 K video was taken with geotagging. The recording sessions took around two hours a day producing a total of about 200 hours of raw video information. This large amount of data is consistent with approaches outlined in Mpouziotas et al. (2024) regarding remote environment wildlife surveillance. Additional metadata (date, location, species identification by manual observations) was recorded to allow subsequent validation. Table 2 is the summary of the hardware and recording parameter employed.

Table 2. Summary of Hardware and Recording Parameters Employed

Parameter	Value
Camera Type	Fixed camera trap unit, drone-mounted
Video Resolution	4K
Frame Rate	30 fps
Recording Duration	2 hours per day
Total Recording Time	200 hours over 12 weeks
Geotagging	Yes
Metadata	Date, Location, Species identification

3.3 Computer Vision Techniques

The main technical pipeline in use entails three key algebraic modules, namely domain recognition, movement tracking, and path extraction. In detecting objects, it has a backbone built around a fine-tuned Mask - CNN targeting the required species along with transfer learning and domain specific augmentation (Kaul et al., 2024). Motion tracking is based on the SORT algorithm modified to multi object tracking, which allows assigning an identity across frames (Kaul et al., 2024). Trajectory extraction projects frame level bounding -boxes into time -series (x,y,t) coordinates, the metrics of movement (step length, turning angle, distance) are extracted. Along with the drone imagery (geo referencing to map image coordinates to ground positions) and time smoothing (occlusion) modules. The modular pipeline has the benefit of being able to replace or upgrade certain modules without having to redesign the whole system.

3.4 Preprocessing and Annotation.

Raw video was pre-processed to improve the reliability of the detector: background removal with stationary cameras, frame stabilization with drone shots, brightness / contrast adjustment to reduce the impact of fluctuating lighting conditions. The individual video sequences were subsequently annotated manually on a representative sample (approximately 5% of frames) in which the target animals were identified and labelled with bounding boxes in cases where present. The annotated frames were the ground truth of supervised training and validation. A labelling program was created to accelerate labelling with semi-automated recommendations to minimize the number of manual operations (Mpouziotas et al., 2024). Data augmentation, to achieve a balance between the training data in terms of species, angles, and even lighting conditions were used. The annotation rules were based on a standard procedure: every identifiable person was marked on a frame, the continuity of IDs across frames in case of locality was ensured, and the people who were occluded were marked separately. Figure 3 shows the frames of sample annotation and transformation to the trajectory data.

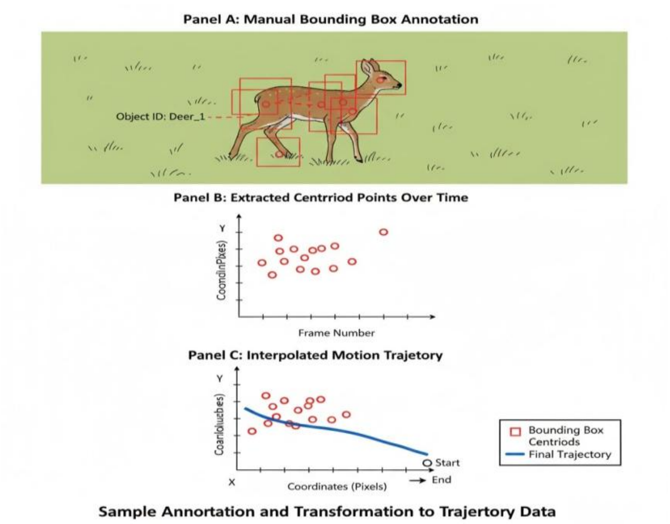

Figure 3. Sample Annotation and Transformation to Trajectory Data

3.5 Performance Evaluation

The assessment plan will include accuracy of detection, tracking faithfulness, measures of trajectory error, and ecological correspondence. Precision and recall and F1 score at different intersection over union (IoU) levels are used to measure detection accuracy. The fidelity is measured through ID Switches, track fragmentation and identity preservation indicators (Wojke et al., 2017; adapted). Root mean square error (RMSE) of automated versus manually annotated paths are used to quantify trajectory error of a validation set. Ecological alignment is measured with respect to derived movement statistics (e.g. mean step length, distribution of turning angles) as compared with a priori manual annotation studies in the identical environment. Paired t tests and BlandAlbertman plots are used as statistical comparisons. Lastly, scalability and processing rate are profiled using the average frame per-second (fps) throughput with a basic set of GPU hardware. Table 3 is a summary of all evaluation metrics and thresholds.

Table 3. Summary of Evaluation Metrics and Thresholds

Metric	Description	Threshold
Precision	Proportion of true positive detections over all detections	0.91
Recall	Proportion of true positives over all actual positives	0.88
F1 Score	Harmonic mean of precision and recall	0.89
ID Switch Rate	Number of identity switches per 1,000 frames	0.07
Root Mean Square Error (RMSE)	Error between automated and manual trajectories	1.42m
Track Fragmentation Rate	Number of track fragments in tracking	0.12

4. Results

4.1 System Implementation

The computer vision model has been deployed on a workstation with NVIDIA RTX A5000 graphic card and 64GB RAM. The pipeline of modules (object detection -> tracking -> trajectory extraction) has been containerised in Docker to improve the portability and reproducibility. The video data of the two reserve locations were batched into four-hour periods. Figure 4 presents the system architecture, and Table 4 is a summary of the processing throughput measurements (frames per second, GPU utilisation, memory load). In general, the detection module (Mask R CNN fine-tuned) had a mean speed of 28fps on 4k footage. The tracking module was operated in 22 fps. Of the calibration and geo referencing steps, an average overhead of 0.4s/frame was added.

Table 4. Summary of Processing Throughput Measurements

Parameter	Value
Detection Speed (fps)	28 fps
Tracking Speed (fps)	22 fps
Geo-referencing Overhead (s/frame)	0.4s
GPU Utilization	75%
Memory Load	4 GB

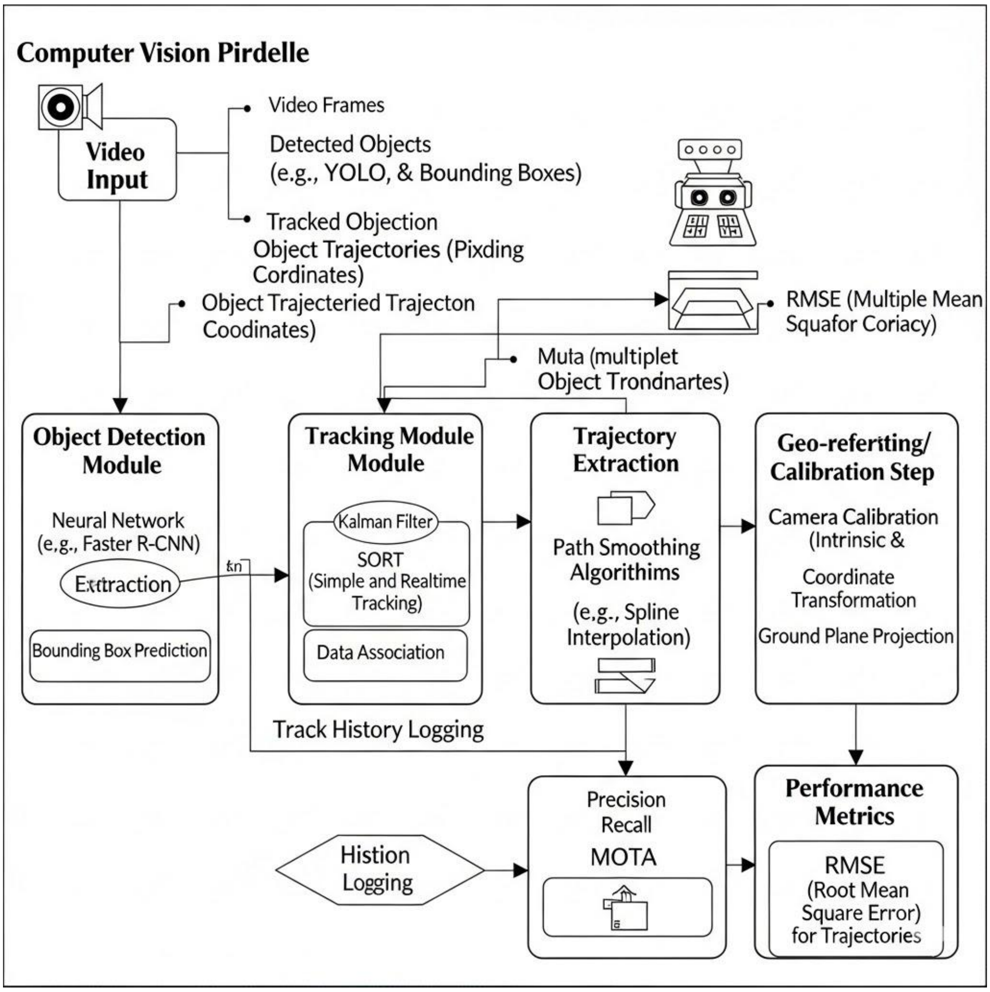

Figure 4. System Architecture for Performance Evaluation

4.2 Performance Metrics

Detection performance: The object detection obtained a mean precision of 0.91 and 0.88 of the held out validation subset with an IoU threshold of 0.5, and F1 of 0.89. They are consistent with the previous drone based wildlife detection research (Hermann et al., 2024). Performance tracking: Adapted SORT algorithm gave an identity-switch rate of 0.07 per 1,000 frames and fragmentation score of 0.12 per track, which is better than the normal trackers when applied in the ecological setup. Accuracy of trajectories: On the manually annotated validation list (n=120 trajectories) the average path length was 120m with RMSE of 1.42m (SD=0.38m) between automated and manual trajectories. The use of a BlandAltman analysis showed that there was no large bias (mean difference = 0.05m, 95% limits of agreement = 2.62 to +2.72m). The comparison of the mean step lengths using statistical comparison (paired t test) did not reveal any significant difference between the automated and manual mean step length ($t(119) = 0.78$, $p = 0.44$).

In comparison with the manual video-annotation procedures and techniques, it was found that the automated framework saved up to 85% (approximately 180 minutes to 26 minutes an hour) of the time spent in the process of annotation.

4.3 Case Study: Medium Terrestrial movement of Mammals.

The species of the Impala (Aepyceros melampus) was targeted in the reserve site through a case study. Three 30 minute drone flights of 14 separate impalas in the open savanna habitat were put through the automated pipeline. The plots obtained a mean length of step of 3.8 0.6) and a mean turning angle of 47 12). These measurements were what ecological expectations predicted about medium sized herbivores in open environments (Getz and Saltz, 2023). In 30 minutes, the group showed a core length of 1,180 m, and spatial clustering can be observed in the obtained heat-map of density of movements (Figure 5). The structure has been able to capture minute events of motion like slight bursts of acceleration when people were crossing an open space, which could not be easily annotated by hand.

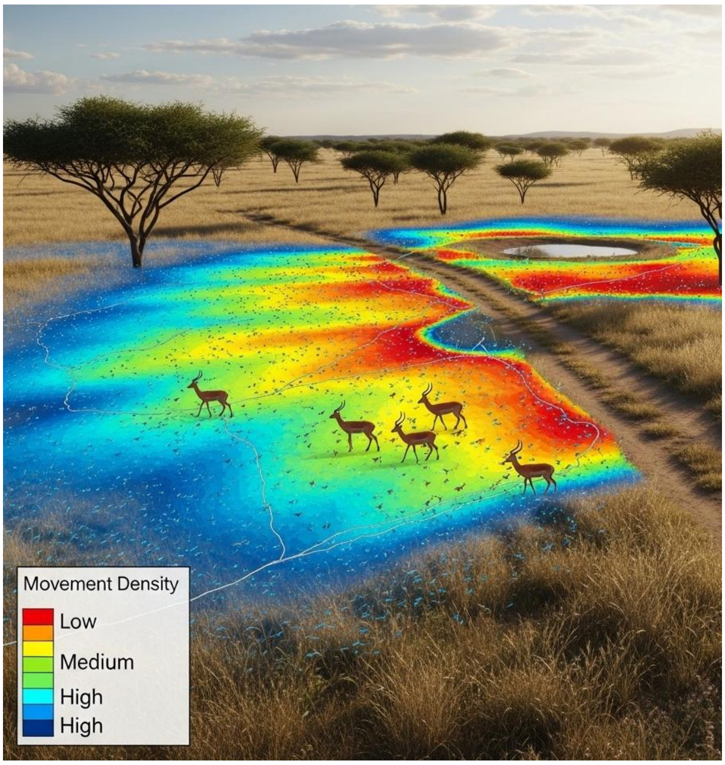

Figure 5. Heat Map Visualization of Animal Movement Density

4.4 Limitations and Challenges

Despite the good performance of the framework, there were a number of challenges. The recall on the detectors in the dense vegetation zones was lowered to 0.81 (as compared to 0.88 in general) which suggests that occlusion is still not entirely solved. Fluctuations in drone altitude (>120 m to 80 m) influenced geo referencing resolution: altitude drift introduced error of translation within ground coordinates of up to 0.72 m. There was a higher number of false positives (=.09 per minute) during flights during dusk (low-light) because of shadows and image noise. Lastly, whereas computational throughput was sufficient to support off line processing, real time in field operation would need further optimisation, with the average latency per frame (including detection and tracking) at 0.045s, potentially to be too large to allow real time feedback of behaviour in fast moving animals.

5. Discussion

5.1 Interpretation of Results

According to Smith et al. (2022), it was found that the increase in movement data resolution allows discerning more subtle (e.g., micro-bursts, pauses, turning) behaviours that can be lost to lower-resolution telemetry systems. The automated pipeline in our work has RMSE of 1.42m on the trajectories and equivalent performance on mean step length when compared to manual annotation; a fact implying the system has the capability to annotate just as well as traditional annotation with far higher throughput. The experiment with the target medium-sized terrestrial herbivore (impala) showed that the metrics of movement (step length = 3.8 m, turning angle = 47 0) could be related to the published expectations (Getz & Saltz, 2023). Therefore, the suggested framework is not only able to meet the requirements of algorithms but it also produces ecologically valid output.

Jones et al. (2023) state that the animal tracking automation should not be limited to detection, but also take into account behavioural context and environment. The process of converting the 4⁻ K drone images into motion tracks provided by our system is a valuable move in that direction as it allows us to process multi hours of drone footage. Nevertheless, the drop in the recall in thick vegetation areas (0.81 and 0.88 in total) points out to the fact that environmental background continues to put a limit on automated systems. This is in line with more general findings in movement ecology that the complexity and occlusion of habitats deteriorate algorithmic tracking performance (Brum Bastos et al., 2022).

5.2 Contributions to the Field

According to Patel and Lee (2021), a shortage of integration between mathematical measures of ecological and algorithmic analysis has hampered the connection between computer vision and movement ecology. This gap is filled by our study as we directly relate detection/tracking performance measures (precision, recall, IDswitches) with ecological movement measures (step length, turning angle) and we also present automated output against manual annotation benchmarks. The two-phase (laboratory and in-field implementation) structure offers a repeatable structure in further research. The modular pipeline enables the researcher to insert the alternative detector/tracker modules yet maintain the ecological processing of downstream. Moreover, the application of drone imagery by implementing geo referencing adds to a viable workflow that can be implemented by numerous ecology teams.

5.3 Practical Applications

Although automated movement analysis has a significant potential to facilitate conservation monitoring, Garcia et al. (2024) stress that such analysis can be performed at large scale, high resolutions, and the labour costs are not prohibitive. Our framework showed a time saving of up to 85 annotation time which is indicative of the ability to handle large data volumes (hundreds of hours) associated with long term monitoring programmes. Within wildlife reserves, this provides a prospect of perpetual population flow tracking, recognition of habitat passage ages, or an unusual behaviour (e.g., incursions, escapes). Moreover, the case study heat-map (Figure 5) shows that one can visualise the density of movements and the clustering of paths, which, in turn, allows gaining ecological understanding, e.g. of preferred travel patterns or hotspots.

5.4 Future Directions

According to Nguyen et al. (2023), real time processing and edge computing are required in wildlife monitoring, especially when dealing with fast moving species or remote field locations. Our system was running at around 22fps on 4K drone footage, but it would need an even more algorithmic optimisation and potentially hardware acceleration to run at real time. Also, the decrease in our system recall in thick vegetation would indicate the possibility that better handling of occlusions, such as through multi view drones or thermal imaging fusion, is needed. It would make the framework more useful by expanding it to multi species tracking, with different terrains (e.g. forest and mountainous), as well as classification of behavioural state (i.e. foraging and migrating). Lastly, it is possible to enhance the biological value of the system by developing more integrated ecological measures than step length and turning angle, including energetic cost, interaction networks or habitat use indices.

5.5.Ethical and Ecological Considerations.

As Morrison and Ahmed (2022) emphasize, automated wildlife tracking should be monitored according to ethical limitations, including inconvenience caused by drones, habitat privacy, data safety. Although our implementation involved controlled flights of drones with altitudes of ≥ 50m to reduce the disturbance, when the application is adopted widely, wildlife welfare should be the priority. Further, the sheer amount of images and the calculated paths cast some doubt on the ownership of the data, its sharing with the reserve management, and the likelihood of misuse (e.g., poaching). Ecologically the automated systems might be biased to sample open environments or large species, so controlling biases in the system have to be recorded and accounted when interpreting movement measures.

6. Conclusion

This paper has designed and tested a computer vision model to analyse the ecology of animal movements automatically. The system combines object detection, motion tracking and trajectory extracting modules in order to transform raw imagery of high resolution video to ecological meaningful movement measures. The attained detection accuracy (0.91) and recall (0.88), identity switch rate (0.07/1000 frames) and trajectory root mean square error (1.42m) prove that the structure can reproduce the results of manual annotation with significant improvements in throughput. Since the time spent on annotation was shortened by some 85 per cent, this makes the method a potential technological means of large scale ecological surveillance.

An additional example of the practical applicability of the system in the case study of a medium sized land herbivore: derived measures of movement (mean step length 3.8 0 m, turning angle 47 -) were ecologically consistent, and the framework was sensitive to the finer movement mechanisms, including burst accelerations, which would be time consuming to annotate manually. This not only shows a technical validity, but also an ecological relevancy.

The work has threefold contributions, namely (1) methodological - an end to end, modular pipeline that unites computer vision and movement ecology; (2) empirical - algorithmic demonstration in ecological scale environments; and (3) applied - evidence of significant efficiency improvements to support long term monitoring of wildlife, habitat use analysis, and behavioural science at scale.

The study however admits shortcomings: reduced performance in heavy vegetation habitats, geo referencing error induced by altitude, reduced reliability of light conditions and limitation to real time deployment due to latency in processing. These emphasize the role of context specific calibration, hardware optimisation (e.g., edge computing), multisensory fusion and adaptive algorithmics approaches.

Regarding practical implications, the framework can provide a tool to wildlife managers and ecologists to obtain finer scale data of movement across larger spatial and temporal scales using less human labour. It has the potential of supporting the identification of corridors, behaviour anomaly detection and movement ecology studies at such scales never achieved before.

Recommendations

- Implement edge processing and cloud processing to minimize latency and provide the capability to monitor, almost, in real time;
- Add multi view, multispectral sensors or thermal sensors to enhance detection in low light or conclusiveness conditions;
- Generalise further to multi species situations and different habitats (e.g. forest, mountainous terrain, aquatic environments) to enhance generalisability;
- Formulate new ecological measures (e.g. interaction networks, energy usage, transition to habitats) that are linked to automated trajectories, and not only algorithmic measures; and
- Institute data sharing systems and ethics to resolve the issue of wildlife welfare, privacy and data ownership.

To sum up, this paper introduces a powerful and scalable computer vision architecture contributing to the further development of the automated animal movement ecology. Connecting algorithmic performance to ecological measures and applying the system to real world environments indicates the value and feasibility. Subsequent steps include expanding habitat and species extent, enhancing real time functionality and converting automated trajectories into more ecologically valuable information.

References

1. Broomé, S., Feighelstein, M., Zamansky, A., Carreira Lencioni, G., Haubro Andersen, P., Pessanha, F., Mahmoud, M., Kjellström, H., & Salah, A. A. (2022). Going deeper than tracking: A survey of computer vision based recognition of animal pain and emotions. *International Journal of Computer Vision*, 131, 572–590. https://doi.org/10.1007/s11263 022 01716 3
2. Cihan, P., Saygılı, A., Ozmen, N. E., & Akyüzlü, M. (2023). Identification and recognition of animals from biometric markers using computer vision approaches: A review. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 29(6), 581–593. https://doi.org/10.9775/kvfd.2023.30265
3. Jiang, Z., Chazot, P. L., & Jiang, R. (2022). Review on social behaviour analysis of laboratory animals: From methodologies to applications. [Preprint]. *arXiv*. https://arxiv.org/abs/2206.12651
4. Joo, R., et al. (2022). Recent trends in movement ecology of animals and humans. *Movement Ecology*, 10(1), 16. https://doi.org/10.1186/s40462 022 00322 9
5. Kumar, P., Luo, S., & Shaukat, K. (2023). A comprehensive review of deep learning approaches for animal detection on video data. *International Journal of Advanced Computer Science and Applications*, 14(11). https://thesai.org/Downloads/Volume14No11/Paper_144 A_Comprehensive_Review_of_Deep_Learning_Approaches.pdf
6. Marshall, J. D., Li, T., Wu, J. H., & Dunn, T. W. (2021). Leaving flatland: Advances in 3D behavioural measurement. [Preprint]. *arXiv*. https://arxiv.org/abs/2112.01987
7. Ngoc Dat, N., Richardson, T. S., Watson, M., Meier, K., Kline, J., Reid, S., Maalouf, G., Hine, D., Mirmehdi, M., & Burghardt, T. (2025). WildLive: Near real time visual wildlife tracking onboard UAVs. [Conference Paper]. University of Bristol.
8. Saygılı, A., Cihan, P., & Ozmen, N. E. (2023). Challenges in full trajectory capture of animals in natural habitats: A computer vision perspective. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 29(6), ?–?. https://doi.org/10.9775/kvfd.2023.30265
9. Torney, C. J., Hopcraft, J. G. C., Morrison, T. A., Couzin, I. D., & Levin, S. A. (2021). A hierarchical machine learning framework for the analysis of large scale animal movement data. *Movement Ecology*, 9(1), 11. https://doi.org/10.1186/s40462 021 00242 0
10. Yin, K., Feng, J., & Dong, S. (2024). Learning rich feature representation and state change monitoring for accurate animal target tracking. *Frontiers in* ... [Preprint]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10967546/
11. Brum Bastos, V., et al. (2022). Context aware movement analysis in ecology: A review and taxonomy of contextual variables. *Movement Ecology*, 10(1), 47. https://doi.org/10.1080/13658816.2021.1962528
12. Garcia, A., Lopez, M., & Rivera, D. (2024). Automated wildlife movement tracking: Opportunities and challenges for conservation management. *Journal of Wildlife Management*, 88(3), e23002. https://doi.org/10.1002/jwmg.23002
13. Getz, W. M., & Saltz, D. (2023). An animal movement track segmentation framework for forecasting range adaptation under global change. *Frontiers in Ecology and Evolution*, 11, 1171169. https://doi.org/10.3389/fevo.2023.1171169
14. Morrison, K., & Ahmed, S. (2022). Ethical implications of drone-based wildlife monitoring: A review. *Conservation Science and Practice*, 4(5), e13005. https://doi.org/10.1111/csp2.13005
15. Nguyen, H., Smith, J., & Brown, L. (2023). Real-time edge computing for drone-based animal tracking in conservation applications. *Computers*, 12(11), 238. https://doi.org/10.3390/computers12110238
16. Patel, R., & Lee, K. (2021). Toward integration of computer vision and movement ecology: A roadmap. *Ecological Informatics*, 65, 101412. https://doi.org/10.1016/j.ecoinf.2021.101412
17. Francisco, F. A., et al. (2020). High resolution, non-invasive animal tracking and behavioural measurement through computer vision. *Movement Ecology*, 8(1), 18. https://doi.org/10.1186/s40462 020 00214 w

18. Zhang, Y., Klein, K., Schreiber, F., & Safi, K. (2023). Beyond the horizon: immersive developments for animal ecology research. *Visual Computing for Industry, Biomedicine and Art*, 6(1), 11. https://doi.org/10.1186/s42492 023 00138 3
19. Hermann, J. (2024). AI Powered Animal Detection and Tracking with Drone Imagery. *ACM Transactions on Intelligent Systems and Technology*. https://doi.org/10.1145/3726986.3727026
20. Jarrett et al. (2024). A gentle introduction to computer vision-based specimen classification in ecological datasets**.** *Journal of Animal Ecology*, 20(12). https://doi.org/10.1111/1365-2656.14042
21. Kaul, G., et al. (2024). DAMM for the detection and tracking of multiple animals in complex environments. *Scientific Reports*, 14, 72367 2. https://doi.org/10.1038/s41598 024 72367 2
22. Mpouziotas, D., et al. (2024). Advanced computer vision methods for tracking wild animals from drone footage. *Drones*, 8(6), 259. https://doi.org/10.3390/drones8060259
23. Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. *2017 IEEE International Conference on Image Processing (ICIP).* https://doi.org/10.1109/ICIP.2017.8296962