

# Towards AI-Enabled Livestock Monitoring in Extensive Mediterranean Rangelands: Algorithms, Data Pipeline and Deployment Considerations

Nikolaos Sifakis <sup>1\*</sup>, Simone Sarti <sup>2</sup>, Dimitrios Cholidis <sup>1</sup>, Alexandros Chachalis <sup>1</sup>, Michael Batistatos <sup>3</sup>, Michail-Alexandros Kourtis <sup>4</sup>, Georgios Arampatzis <sup>1</sup>

<sup>1</sup> Mediterranean Agronomic Institute of Chania (MAICh), Chania, 73100, Greece, nsifakis@tuc.gr

<sup>2</sup> Leonardo S.p.A. (LDO), Rome, Italy

<sup>3</sup> Department of Informatics and Telecommunications University of the Peloponnese Tripolis, Greece

<sup>4</sup> Institute of Informatics and Telecommunications, National Centre for Scientific Research “DEMOKRITOS” (NCSR-D)

## Abstract

This paper describes two AI applications being developed for drone-based livestock monitoring in Western Crete under the EU-funded AGRARIAN project. The first application, danger detection, identifies when grazing sheep or goats approach roads, vehicles, or steep terrain. Its pipeline runs a YOLOv11 object detector for animal localisation, a YOLOv11-seg model for road and vehicle segmentation, and a georeferencing module that projects slope data from the 10 m Copernicus Digital Elevation Model onto the drone frame using flight telemetry. When an animal’s safety zone intersects a danger zone, the system generates a structured alert with GPS coordinates and an evidence frame. The second application, health monitoring, is designed to flag animals whose movement deviates from the collective herd pattern. It uses a Graph Neural Network in which each detected animal is a node and edges encode spatial proximity, combined with LSTM or Transformer layers for temporal modelling. We report on the data collection protocol (three altitudes, three species compositions, two sampling rates) and on preliminary results: the danger detection pipeline exceeds 30 frames per second in its multi-process implementation. System limitations are discussed in detail.

**Keywords:** artificial intelligence, livestock monitoring, object detection, anomaly detection, UAV, graph neural networks, precision agriculture.

## Introduction

Most livestock farmers in Crete operate under extensive grazing conditions, with sheep and goats moving across rugged communal rangelands that are difficult to access, steep, and only intermittently supervised. During the summer grazing period, animals may approach roads, vehicles, ravines, and abrupt terrain discontinuities without immediate notice from the herd owner. In practice, monitoring is still performed largely through visual inspection on foot, which is laborious, spatially limited, and often reactive rather than preventive. As a result, injuries, animal loss, or the first signs of abnormal behaviour may remain undetected until substantial time has passed [1], [2]. This creates a clear need for monitoring tools that can extend human oversight over large and topographically complex grazing areas without requiring continuous physical presence [3], [4], [5].

Unmanned aerial vehicles and artificial intelligence offer a plausible response to this need, but the technical difficulty of the problem should not be understated. In the literature, object detection models of the YOLO family have shown strong performance in real-time localisation tasks by formulating detection as a single-pass regression problem, making them attractive for time-sensitive field applications [6], [7]. Recent work has also demonstrated their relevance for aerial imagery, where targets are small, scenes are cluttered, and inference speed matters [4, 5]. However, extensive Mediterranean rangelands introduce additional complications that are less visible in standard benchmark settings: mixed sheep-goat herds, strong scale variation with flight altitude, dry vegetation and rocky backgrounds, deep shadows, moving cameras, and terrain hazards that cannot always be inferred reliably from RGB imagery alone [8], [9].

Likewise, health-related abnormality in grazing animals is not merely an individual motion problem. It is often relational and context-dependent: an animal becomes suspicious not only because of how it moves, but because of how it moves relative to the herd. This makes Graph Neural Networks (GNNs) particularly relevant, since they can model interactions among animals directly and support spatio-temporal anomaly detection beyond independent trajectory analysis [10], [11].

Within this context, the AGRARIAN project develops and integrates two complementary AI applications for drone-based livestock monitoring in Crete. The first is a danger detection pipeline that combines UAV-based object detection, instance segmentation, and terrain-aware hazard mapping to identify when an animal approaches roads, vehicles, steep slopes, or geofenced boundaries. The second is a health monitoring pipeline that uses detection and tracking outputs to construct dynamic interaction graphs, with the aim of flagging animals whose motion deviates from normal collective herd behaviour. Importantly, these applications are not conceived as isolated laboratory models, but as deployable services within the AGRARIAN architecture, packaged as Docker containers and orchestrated through a K3s-based infrastructure for operational use [12]. This system perspective matters, because in remote livestock settings the value of AI depends not only on model choice, but also on data acquisition discipline, interface design, deployment constraints, and alert delivery. A companion paper [18] describes the platform-level architecture, K3s orchestration, three-tier edge-fog-cloud hierarchy, hybrid connectivity, and the ADSS implementation, in detail. The present paper complements that system-integration account by focusing on the AI algorithms, data pipeline, and deployment considerations that the integration paper references but does not develop. Readers interested in the platform architecture are directed to the companion paper.

Recent UAV-based livestock monitoring studies provide useful reference points. Barbedo and Koenigkan (2018) achieved cattle-counting accuracy exceeding 90 % from UAV imagery, but their work was conducted in flat pastures with uniform breeds [14]. Zhang et al. (2023) reported mAP 0.93 for sheep detection in Inner Mongolia grasslands using a YOLOv5-based detector, without addressing connectivity or deployment readiness [15]. Mou et al. (2024) demonstrated YOLOv7 for wildlife detection from UAVs under challenging conditions, but their pipeline was evaluated offline [16]. A recent systematic review by dos Santos et al. (2024) examined 87 UAV-livestock papers and found that fewer than 10 % addressed real-time constraints and none described an integrated edge-to-cloud deployment [17]. The present work differs from this literature in three respects: it targets the under-studied Mediterranean extensive-grazing context with mixed sheep-goat herds; it addresses deployment readiness through Docker-based containerisation, K3s orchestration, and hybrid terrestrial-satellite connectivity; and it reports on two complementary AI applications (danger detection and health monitoring) within a single operational architecture rather than treating each in isolation.

Against this background, this paper presents an integrated and candid account of the current state of the Greek livestock pilot. It reports the data collection protocol designed to generate a site-specific UAV dataset for mixed herds in mountainous terrain; the architecture of the danger detection workflow, including object detection, segmentation, and DEM-based terrain analysis; the design logic of the GNN-based health monitoring application; and the preliminary performance of the implemented pipeline. Equally important, the paper examines the intrinsic and operational limitations of the proposed “drone + AI” approach. Rather than treating these limitations as secondary, we position them as central to the scientific and practical evaluation of the system, because they define the boundary between what is technically demonstrable and what is genuinely useful for real livestock management in extensive Mediterranean environments.

## Data and Methods

### Data collection protocol

The data collection was designed jointly by MAICh (pilot operator) and Leonardo S.p.A. (technology provider). The drone is a DJI Mavic 3M equipped with an RGB camera and a 4-band multispectral camera (red, green, red-edge, near-infrared). Only the RGB channel is used; the visual spectrum provides sufficient information for the detection and tracking tasks addressed here. An initial set of eight short video clips (approximately 2 minutes each) at various altitudes was supplied by MAICh. These were sufficient to confirm that a pre-trained detector could locate sheep, but performance on goats was poor. Publicly available drone-animal datasets were also evaluated. Most consist of sheep in flat British or Australian pastures, which differ substantially from the Cretan hill environment in terms of terrain, vegetation, breed morphology, and image quality. Annotated goat data are scarce.

A structured protocol was therefore designed around three controlled variables (Table 1). Altitude: flights at 15, 25, and 50 m, selected after analysis of the initial footage to span the altitude range over which the drone is expected to operate. Species composition: recordings of groups consisting primarily of sheep, primarily of goats, and mixed groups. Movement: stationary animals (sampled every 3 s) and moving herds (sampled every 1.5 s). The higher sampling rate for moving animals captures motion dynamics; the lower rate for stationary animals avoids redundant near-identical frames.

**Table 1.** Dataset collection protocol: 18 configurations across three altitudes, three species compositions, and two activity modes.

| Frames | Alt. (m) | Subject      | Activity   | Sampling (s) | Dur. (min) |
|--------|----------|--------------|------------|--------------|------------|
| 600    | 15       | Mostly sheep | Stationary | 3.0          | 30         |
| 600    | 15       | Mostly goats | Stationary | 3.0          | 30         |
| 300    | 15       | Mixed        | Stationary | 3.0          | 15         |
| 200    | 15       | Mostly sheep | Moving     | 1.5          | 5          |
| 200    | 15       | Mostly goats | Moving     | 1.5          | 5          |
| 100    | 15       | Mixed        | Moving     | 1.5          | 2.5        |
| 600    | 25       | Mostly sheep | Stationary | 3.0          | 30         |
| 600    | 25       | Mostly goats | Stationary | 3.0          | 30         |
| 300    | 25       | Mixed        | Stationary | 3.0          | 15         |
| 200    | 25       | Mostly sheep | Moving     | 1.5          | 5          |
| 200    | 25       | Mostly goats | Moving     | 1.5          | 5          |
| 100    | 25       | Mixed        | Moving     | 1.5          | 2.5        |
| 600    | 50       | Mostly sheep | Stationary | 3.0          | 30         |
| 600    | 50       | Mostly goats | Stationary | 3.0          | 30         |
| 300    | 50       | Mixed        | Stationary | 3.0          | 15         |
| 200    | 50       | Mostly sheep | Moving     | 1.5          | 5          |
| 200    | 50       | Mostly goats | Moving     | 1.5          | 5          |
| 100    | 50       | Mixed        | Moving     | 1.5          | 2.5        |

The full protocol targets approximately 4 hours of recording, yielding an estimated 5,500 annotated frames across 18 configurations. The first two hours have been delivered to Leonardo. Additional instructions to MAICh specified: reposition the drone slightly (translation and yaw) once per minute over resting animals, and move between groups every few minutes, to increase spatial variability. For object detection, frames will be annotated with bounding boxes and divided into training, validation, and test splits [7]. For the health monitoring task, annotation is more difficult. Leonardo lacks domain expertise in sheep and goat behaviour. Distinguishing normal from abnormal movement requires collaboration with

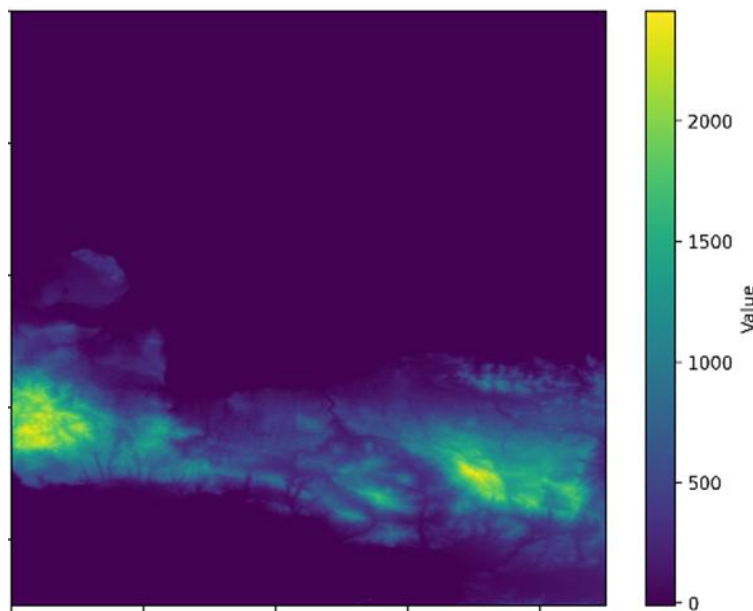
the herd owner and MAICH. If direct annotation proves impractical, MAICH will coordinate with the herd owner to stage and record both normal and abnormal scenarios (e.g., simulated limping, separation from the group) [7].

### Dataset status and split strategy.

As of submission, approximately two hours of raw footage (roughly 2,800 candidate frames before annotation) have been delivered to Leonardo. Annotation is in progress and the currently annotated subset is insufficient for reliable model training; all detection results reported in this paper were obtained with pretrained weights. The planned data split is 70 % training, 15 % validation, and 15 % test, stratified by altitude (15, 25, 50 m) and species composition (mostly sheep, mostly goats, mixed) to ensure balanced representation across all 18 protocol configurations. Frames from the same flight will not be split across partitions, to avoid data leakage from temporally adjacent frames. The >30 FPS figure reported in Section 3.1 was measured on a preliminary subset of approximately 200 frames and should be interpreted as a processing-speed feasibility check rather than a production throughput benchmark.

### Terrain hazard data

Steep terrain cannot be reliably detected from a nadir-pointing camera at 50 m. Slope data are instead obtained from the Copernicus Digital Elevation Model at 10 m resolution [8], derived from TanDEM-X mission acquisitions between 2011 and 2015. Each pixel encodes terrain elevation at a geographic location (Fig. 1). Three hazard categories are extracted: slopes exceeding a user-defined gradient threshold (computed from the gradient magnitude of the elevation raster); pixels with missing elevation data, which typically correspond to cliffs or coverage gaps; and areas outside a geofence polygon defined by the herd owner [7].



**Fig. 1.** Central Crete from the 10 m Copernicus DEM. Brighter pixels correspond to higher elevation.

For vehicles and roads, the system currently relies on a pre-trained segmentation model. A pick-up truck is the only vehicle type that appears in the initial Cretan footage. Two small public datasets of drone-perspective vehicle imagery have been identified for potential fine-tuning [7].

### Danger detection pipeline

The danger detection application (Fig. 2) has three processing branches whose outputs are merged into a single danger map. Branch 1 (object detection): A YOLOv11 model [9] locates

sheep and goats in each full-HD video frame. The network outputs bounding boxes with class labels and confidence scores. YOLO processes the full image in a single forward pass [3], which is what makes it fast enough for real-time use. Frames are processed at native resolution to preserve small-object detail, particularly at the 50 m altitude where animals occupy few pixels. Branch 2 (segmentation): A YOLOv11-seg instance segmentation model produces pixel-level masks for roads and vehicles. The current model uses pre-trained weights; domain-specific fine-tuning is planned after annotation of the Cretan footage. Branch 3 (DEM georeferencing): The slope, missing-data, and geofence masks derived from the Copernicus DEM are projected onto the drone frame coordinate system using flight telemetry (GPS position, altitude, camera orientation). This projection maps each pixel in the high-resolution drone frame to its geographic coordinates, allowing the low-resolution terrain data to be overlaid accurately. All masks are combined into a single danger map. For each detected animal, the system draws a circular safety zone with a user-defined radius (in metres). If this circle overlaps any danger pixel, a structured alert is generated in msgpack format. The alert contains: frame ID, alert message, timestamp, image dimensions, data type, compression flag, and the evidence frame as raw bytes. An annotated RTMP/RTSP video stream runs in parallel.

### Model configuration.

The object detection branch uses YOLOv11n (nano), with 3.2 million parameters. The segmentation branch uses YOLOv11n-seg with an identical backbone. Both models operate at the native frame resolution of  $1920 \times 1080$  pixels. The current implementation uses pretrained weights without fine-tuning; training on the annotated Cretan footage will follow the standard YOLOv11 recipe (SGD optimiser, cosine learning rate schedule, mosaic augmentation, 300 epochs) and is expected to improve detection performance on local breeds and terrain. The planned evaluation metrics are: mAP@0.5 and mAP@0.5:0.95 for object detection; IoU and pixel-wise F1 for instance segmentation; and event-level precision and recall for danger alerts, where each alert is treated as a binary prediction against ground-truth hazard events. For the multi-object tracker, MOTA, MOTP, and ID-switch count will be reported.

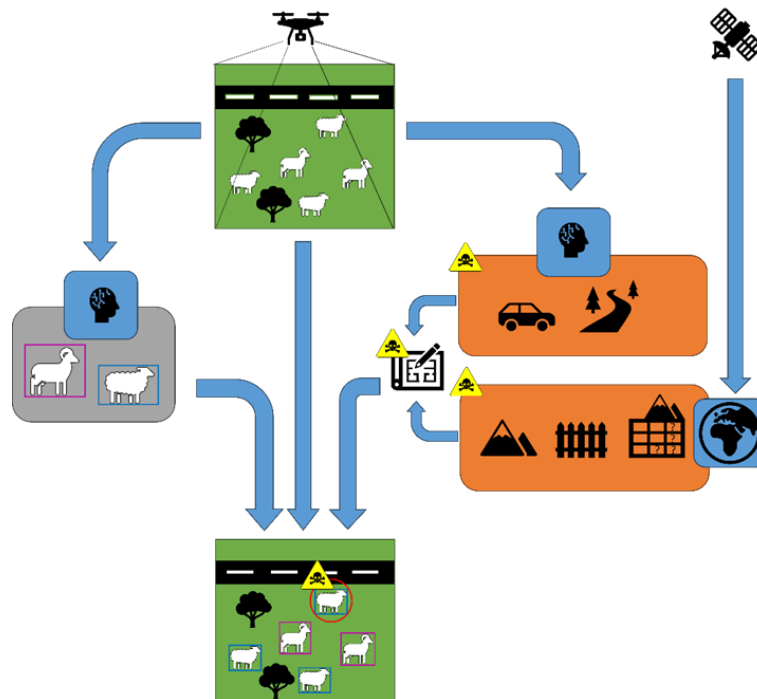


Fig. 2. Danger detection pipeline.

Figure 2 shows the video frames and flight telemetry enter three parallel branches, object detection produces bounding boxes with class labels, segmentation produces pixel-level road and vehicle masks, and DEM georeferencing overlays terrain hazard data onto the drone coordinate system. All three outputs are merged into a unified danger map; an alert is generated whenever an animal safety zone intersects any danger-region pixel.

### **Health monitoring design**

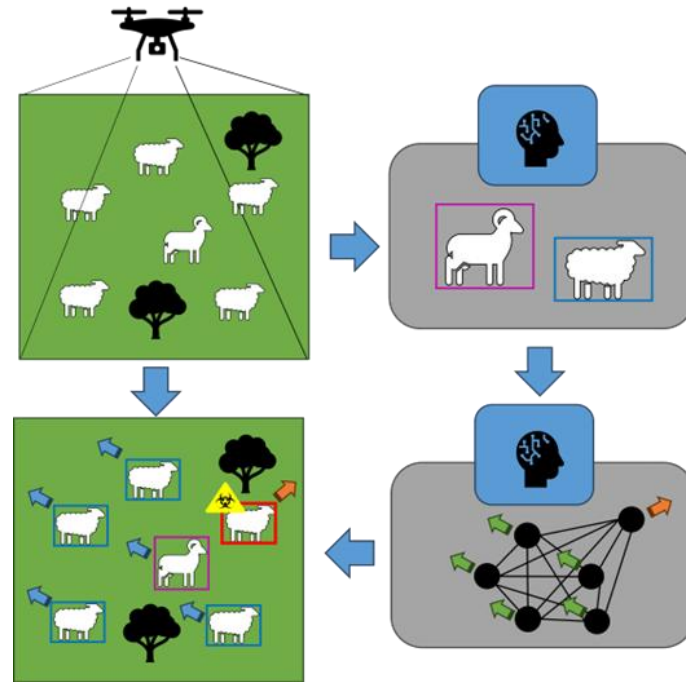
The health monitoring application (Fig. 3) is at an earlier stage of development. Its architecture has two components. The first component reuses the YOLOv11 detector from the danger detection pipeline, augmented with a multi-object tracker. The tracker assigns persistent IDs to animals within a single flight and produces a position trajectory for each individual. The detector will be extended with species classification (sheep vs. goat) for richer metadata. The second component is a graph-based anomaly detection model. Each detected animal becomes a node in a graph, with edges connecting spatially proximate animals. The graph structure changes every frame as animals move, enter, or leave the field of view. This dynamic topology distinguishes the problem from standard GNN benchmarks where node count is fixed. A GNN with message-passing aggregates spatial context at each time step: whether an animal is moving with the herd or diverging from it. An LSTM or Transformer processes the resulting node embeddings over time to capture temporal patterns. The model will be trained exclusively on normal behaviour data. Anomalies produce elevated reconstruction error, triggering an alert for the specific animal. This unsupervised approach avoids the need for labelled abnormal examples, which are difficult to obtain. The main uncertainty is whether the herd owners and MAICH can supply sufficient video of unambiguously normal behaviour and, separately, a validation set with known anomalies.

### **Model architecture.**

The reference design uses a Graph Attention Network v2 (GATv2) with two layers (hidden dimension 64, four attention heads per layer), followed by a two-layer bidirectional LSTM (hidden size 128) for temporal encoding. Node features for each detected animal comprise normalised position ( $x$ ,  $y$ ), instantaneous speed, heading angle, and a 32-dimensional appearance embedding extracted from the YOLO backbone. Edges are constructed with a spatial distance threshold of 15 m (roughly three to four body lengths for an adult sheep or goat), yielding a dynamic graph of typically 20 to 60 connected nodes that evolves frame by frame as animals enter or leave the field of view. The full model has approximately 1.8 million trainable parameters. These hyperparameters are provisional and will be tuned once training data become available.

### **Evaluation methodology.**

Anomalies are defined operationally: an animal is flagged when its reconstruction error exceeds a threshold for more than  $T$  consecutive frames. The threshold is set at the 95th percentile of the training-set reconstruction error distribution;  $T$  is set to 30 (corresponding to approximately one second at 30 fps). The primary evaluation metric is the area under the receiver operating characteristic curve (AUROC) computed at the animal-minute level (aggregating per-frame scores over 60-frame windows). Secondary metrics are precision and recall at the operating point defined by the 95th-percentile threshold. Validation will rely on staged abnormal scenarios, such as simulated limping, temporary separation from the group, and unusual stillness. This methodology is fully specified but its execution depends on the availability of a sufficiently rich annotated dataset, which is expected after completion of the structured data collection protocol described in Section 2.1.



**Fig. 3.** Health monitoring pipeline.

Figure 3 demonstrates the YOLO detector augmented with a multi-object tracker produces per-animal position trajectories; at each frame a dynamic proximity graph is constructed with animals as nodes and spatial nearby connections as edges. The GATv2 + LSTM model encodes spatio-temporal patterns from the graph sequence, and animals whose behaviour deviates from the learned normal collective pattern are flagged for the herd owner.

## Deployment

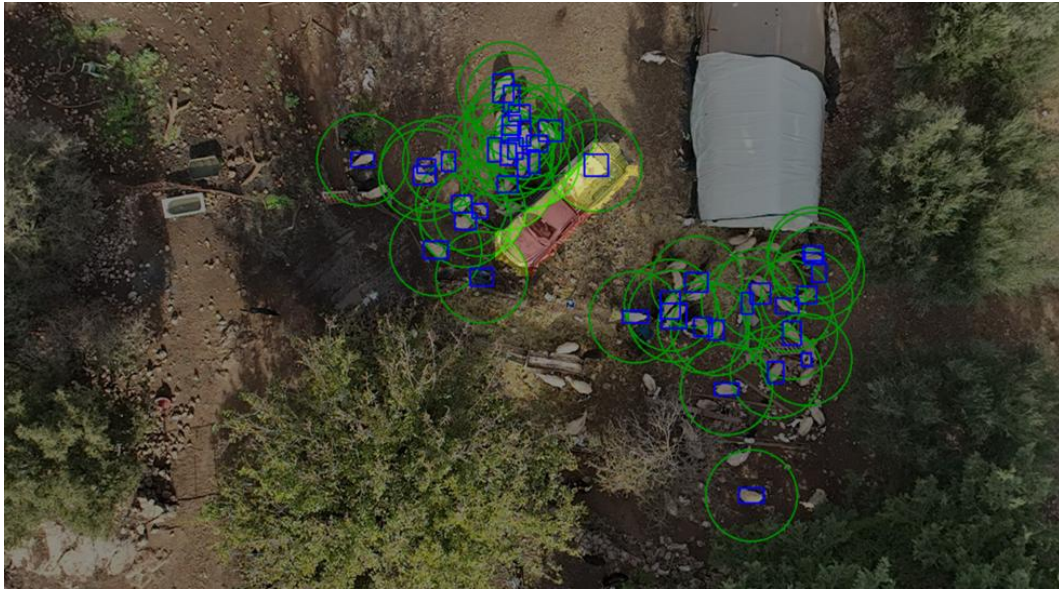
Both applications are packaged as Docker containers and stored in the AGRARIAN application repository. Deployment is managed by the project's K3s cluster, which places containers on edge, fog, or cloud nodes according to their computational requirements [12]. A GitHub Actions CI/CD pipeline automates testing, image building, and publication to the container registry. The containerised applications accept RTMP/RTSP video and drone telemetry as input and produce annotated video and structured alert messages as output [7].

## Results

### Danger detection: current state

The danger detection pipeline runs end-to-end in its current implementation. Animals are detected, vehicles are segmented, flight telemetry is parsed to geolocate each frame, and the DEM overlay aligns correctly with the video (Fig. 4). The geofencing module flags animals approaching or crossing the polygon boundary defined by the user. A multi-process version of the code exceeds 30 frames per second on a standard workstation, clearing the real-time threshold [7].

As a preliminary feasibility check, the pretrained YOLOv11n model was tested on approximately 200 frames (15 m altitude, mostly sheep, stationary). The model achieved approximately 0.72 mAP@0.5 on this subset. This figure is reported for transparency but should not be interpreted as a performance claim: the sample is small, unrepresentative of the full 18-configuration protocol, and the model has not been fine-tuned on Cretan data. Meaningful accuracy figures will be reported after retraining on the complete annotated dataset.



**Fig. 4.** Detection output: sheep (blue), safety zones (green), vehicle (red), danger intersection (yellow).

Quantitative accuracy metrics are not yet reported. The models used so far were trained on public datasets, which differ from the Cretan footage in terms of animal breed, terrain, vegetation, and drone model. We consider these preliminary metrics uninformative. Accuracy figures will be reported after retraining on the annotated dataset [7].

### End-to-end latency.

The latency budget from frame capture to alert delivery is distributed across six stages: drone image capture and encoding (~33 ms at 30 fps), RTMP encoding and transmission over the active link (~150–300 ms under 5G; ~1–2 s under satellite), YOLOv11n inference (~15 ms per frame on an high-end GPU workstation), DEM overlay and alert generation (~5 ms), and alert delivery to the farmer’s mobile device (~200–500 ms via 5G push notification; ~1–2 s via satellite). The total end-to-end latency is estimated at approximately 0.4–0.9 s under terrestrial connectivity and 1.4–2.9 s under satellite fallback, dominated by transmission and delivery rather than inference. These figures are analytical estimates based on component-level benchmarks; full experimental characterisation under field conditions is planned as part of the simultaneous-testbed validation described in the companion paper. The satellite-path latency, while higher, remains within the operational envelope for non-time-critical danger alerts where an additional 1–2 s of delay does not materially affect the herd owner’s ability to intervene.

### Limitations

A systematic review of the system has identified six categories of limitation (Table 2). Some of these can be addressed through engineering (e.g., cross-flight tracking, processing latency). Others are inherent to the sensor configuration and cannot be removed: the drone battery lasts 30–45 minutes, and the camera cannot see animals under tree canopy [7].

**Table 2.** Limitations of the drone + AI monitoring system.

| Category    | Issue                     | Practical consequence                                                      |
|-------------|---------------------------|----------------------------------------------------------------------------|
| Sensing     | Top-down camera only      | Animals under trees, in sheds, or in gorges are invisible to the system    |
| Sensing     | Battery: 30–45 min        | Continuous monitoring requires spare batteries and field charging          |
| Processing  | Communication latency     | At 30+ fps processing, the radio link still adds seconds to alert delivery |
| Tracking    | IDs lost between flights  | Health analysis cannot track individuals across separate flights           |
| Operational | Farmer must fly drone     | Requires training and possibly certification; adoption is not guaranteed   |
| Operational | Flight-time coverage only | Events outside active flight windows cannot be detected or prevented       |

## Conclusions

The results presented in this paper show that the danger detection pipeline has progressed beyond conceptual design and currently operates as a functional end-to-end system. In its present implementation, the pipeline processes video at more than 30 frames per second, detects animals, segments relevant hazards such as vehicles, projects DEM-derived terrain information onto the drone frame, and generates structured alerts whenever an animal's safety zone intersects a dangerous area. These results do not yet prove field-level effectiveness in reducing incidents, but they do demonstrate that the core technical architecture is coherent, executable, and compatible with the real-time requirements of the AGRARIAN livestock pilot.

By contrast, the health monitoring application remains at an earlier stage of maturity. Its architectural logic is defined and technically justified, particularly the use of graph-based spatio-temporal modelling to capture deviations from collective herd behaviour, but its empirical evaluation has not yet started. At this stage, the main constraint is not algorithmic formulation but data availability. The GNN-based anomaly detection framework requires sufficiently rich site-specific trajectories representing normal herd behaviour, together with a reliable validation set containing confirmed abnormal events. Until such data are collected, annotated, and curated for training and testing, the health-monitoring component should be considered a promising design direction rather than a validated subsystem.

The next phase of work is therefore decisive. It includes the annotation of the MAICH video material for object detection and behavioural analysis, the retraining of the YOLOv11 detector on Cretan site-specific footage, the fine-tuning of road and vehicle segmentation models for drone-perspective imagery, the training and evaluation of the GNN-based anomaly detector, and the final packaging of both applications as Dockerised services for deployment through the AGRARIAN K3s-based infrastructure. These steps are not merely incremental engineering tasks; they will determine whether the current proof of technical feasibility can be translated into a robust operational tool.

Overall, the study establishes a realistic baseline for AI-enabled livestock monitoring in extensive Mediterranean rangelands. It shows that danger detection is already technically viable under the constraints of the pilot architecture, while health monitoring remains a more challenging but potentially high-value extension. Equally important, the work highlights that the success of such systems depends not only on model performance, but also on data quality, terrain context, deployment readiness, and the careful alignment between algorithmic outputs and real livestock-management needs. The ultimate question, namely whether the integrated

system can reduce losses, improve responsiveness, and deliver measurable value to herd owners in practice, can only be answered through the forthcoming pilot trials.

## References

- [1] M. Volanis, A. Stefanakis, I. Hadjigeorgiou, and P. Zoiopoulos, "Supporting the extensive dairy sheep smallholders of the semi-arid region of Crete through technical intervention," *Trop. Anim. Health Prod.*, vol. 39, no. 5, pp. 325–334, 2007, doi: 10.1007/s11250-007-9019-z.
- [2] N. Voutzourakis, N. Tzanidakis, G. Butler, C. Leifert, S. Soritaki, and A. Stefanakis, "Utilization of grazing in different small ruminant management systems on Crete Island, Greece," in *Ecosystem Services and Socio-Economic Benefits of Mediterranean Grasslands*, vol. 114, A. P. Kyriazopoulos, A. López-Francos, C. Porqueddu, and P. Sklavou, Eds., in *Options Méditerranéennes: Série A. Séminaires Méditerranéens*, vol. 114, Zaragoza: CIHEAM, 2016, pp. 359–363. [Online]. Available: <http://om.ciheam.org/om/pdf/a114/00007546.pdf>
- [3] M.-T. Pham, L. Courtrai, C. Friguier, S. Lefèvre, and A. Baussard, "YOLO-Fine: One-Stage Detector of Small Objects Under Various Backgrounds in Remote Sensing Images," *Remote Sens.*, vol. 12, no. 15, p. 2501, 2020, doi: 10.3390/rs12152501.
- [4] European Commission, CORDIS, "SpAce-Air-Ground Last Mile InfRastructure & Dynamic Network Management for AgricultuRal Services, Environmental Sensing and Monitoring (AGRARIAN)." 2025. [Online]. Available: <https://cordis.europa.eu/project/id/101134128>
- [5] AGRARIAN Project Consortium, "Greek Pilot." 2026. [Online]. Available: <https://agrarian-project.eu/greek-pilot/>
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.
- [7] Y. Bu, H. Ye, Z. Tie, Y. Chen, and D. Zhang, "OD-YOLO: Robust Small Object Detection Model in Remote Sensing Image with a Novel Multi-Scale Feature Fusion," *Sensors*, vol. 24, no. 11, p. 3596, 2024, doi: 10.3390/s24113596.
- [8] R. Casas, A. Hermosa, Á. Marco, T. Blanco, and F. J. Zarazaga-Soria, "Real-Time Extensive Livestock Monitoring Using LPWAN Smart Wearable and Infrastructure," *Appl. Sci.*, vol. 11, no. 3, p. 1240, 2021, doi: 10.3390/app11031240.
- [9] J. C. L. da Silva, T. Ramos, and M. F. M. Júnior, "Modeling and Harmonic Impact Mitigation of Grid-Connected SCIG Driven by an Electromagnetic Frequency Regulator," *Energies*, vol. 14, no. 15, p. 4524, Jul. 2021, doi: 10.3390/en14154524.
- [10] A. Deng and B. Hooi, "Graph Neural Network-Based Anomaly Detection in Multivariate Time Series," *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 5, pp. 4027–4035, 2021, doi: 10.1609/aaai.v35i5.16523.
- [11] S. Guan, B. Zhao, Z. Dong, M. Gao, and Z. He, "GTAD: Graph and Temporal Neural Network for Multivariate Time Series Anomaly Detection," *Entropy*, vol. 24, no. 6, p. 759, 2022, doi: 10.3390/e24060759.
- [12] Docker, Inc., "Docker Documentation." 2026. [Online]. Available: <https://docs.docker.com/>
- [14] W. M. dos Santos et al., "Use of Unmanned Aerial Vehicles for Monitoring Pastures and Forages in Agricultural Sciences: A Systematic Review," *Drones*, vol. 8, no. 10, p. 585, 2024, doi: 10.3390/drones8100585.
- [15] M. Zhang, F. Gao, W. Yang, and H. Zhang, "Wildlife Object Detection Method Applying Segmentation-Gradient Flow and Feature-Difference-Matching," *Remote Sens.*, vol. 15, no. 6, p. 1593, 2023, doi: 10.3390/rs15061593.

- [16] L. Mou et al., "A Lightweight Wildlife Detection Model for UAV Remote Sensing Images Based on Improved YOLOv7-tiny," *IET Image Process.*, vol. 18, no. 5, pp. 1287–1304, 2024, doi: 10.1049/ipr2.13027.
- [17] J. G. A. Barbedo and L. V. Koenigkan, "Perspectives on the Use of Unmanned Aerial Systems to Monitor Cattle," *Outlook Agric.*, vol. 47, no. 3, pp. 214–222, 2018, doi: 10.1177/0030727018781876.
- [18] Sifakis N., Sarti S., Cholidis D., and Arampatzis G., "An Edge-Enabled Smart Livestock Monitoring Framework in Mountainous Crete: The AGRARIAN Approach," in *Proc. ISAG 2026*, 2026 (companion paper, this volume).