RESEARCH & INNOVATION FOUNDATION

2016-2020 RESTART RESEARCH
ΠΡΟΓΡΑΜΜΑΤΑ ΕΡΕΥΝΑΣ, ΤΕΧΝΟΛΟΓΙΚΗΣ ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΚΑΙΝΟΤΟΜΙΑΣ

THE RESEARCH & INNOVATION FOUNDATION PROGRAMMES

FOR RESEARCH, TECHNOLOGICAL DEVELOPMENT, AND INNOVATION

RESTART 2016 – 2020

Pillar	I. Smart Growth
Programme	CO-DEVELOP
Project Acronym	Green-HIT
RIF Project Number	CODEVELOP-ICT-HEALTH/0322/0135
Proposal Title	A Green - Holistic IoT platform for Forest Management and Monitoring
Project Coordinator	Frederick Research Center (FRC)
Work Package Number	WP6
Work Package Title	Green-HIT Intelligence Modules
Deliverable Number	D6.4
Deliverable Title	Other Incident and Warning Systems (Prototype & Report)

Dissemination level		
PU	Public	X
CO	Confidential, only for members of the consortium (including RIF)	

Funded by the European Union
NextGenerationEU

Cyprus tomorrow
RECOVERY AND RESILIENCE PLAN

Republic of Cyprus

AUTHORS

Author	Institution	Contact (e-mail, phone)
Andreas Pamboris	**FRC**	res.ap@frederick.ac.cy
Andreas Constantinides	**FRC**	com.ca@frederick.ac.cy

DOCUMENT CONTROL

Document version	Date	Change
v0.1	30/04/2025	Draft
v1.0	26/05/2025	Final Version

Executive Summary

This deliverable presents the extended suite of incident and warning systems deployed under the Green-HIT project to detect and respond to illegal activities that threaten biodiversity in protected areas. These systems complement the audio detection approaches developed under D6.2 for illegal logging and hunting and expand into detection of unauthorized vehicle access (trespassing) using both camera and acoustic-based technologies. The deliverable outlines the full methodology, hardware design, AI models, training approaches, and real-world deployment results. The modules described in D2.4 support real-time alerts and integrate into a three-tier verification framework involving sensor alerting, UAV inspection, and human confirmation.

Table of Contents

1. Introduction

Conservation zones are increasingly vulnerable to unauthorized entry, often occurring in remote and unmonitored terrain. D6.4 introduces and documents in situ and cloud-based solutions for detecting such events through embedded AI systems that interpret audio and visual data. These solutions build on and expand the acoustic detection framework presented in deliverable D6.2, while adding modular camera surveillance and differentiated handling of trespassing events. The solutions were developed and field-tested using hardware adapted to constrained environments, such as embedded systems with LoRaWAN communication and microcontrollers. This deliverable is structured into three (3) core modules that address different aspects of the detection challenge.

The remainder of this deliverable is structured as follows:

- Section 2 describes the three-way verification protocol implemented in Green-HIT, outlining how sensor-based detection, platform alerting, and drone-assisted validation are integrated to ensure accurate incident confirmation and real-time situational awareness.
- Section 3 describes the in-situ camera-based detection modules developed for detecting illegal trespassing in protected biodiversity zones.
- Section 4 details the in-situ acoustic-based modules used for the detection of illegal vehicle presence (trespassing) in ecologically sensitive areas.
- Section 5 outlines the Cloud-based detection pipeline, including real-time data aggregation, signal processing, and alert generation mechanisms.
- Section 6 highlights the broader environmental and infrastructure data collected by Green-HIT's components, emphasizing its potential use for future detection services beyond the project's current scope.
- Section 7 explains the rationale behind key design and implementation choices.
- Section 8 concludes the deliverable.

2. The Green-HIT Three-Way Verification Protocol

The Green-HIT project implements a robust three-way verification protocol to ensure accurate detection, rapid validation, and effective response to incidents identified by its intelligent monitoring modules. This protocol combines sensor- and camera-based detection, platform-based alerting, and aerial visual confirmation via drone deployment.

Each intelligent module is triggered based on a distinct set of signals corresponding to its domain:

- Fire Detection
- Illegal Logging
- Illegal Hunting
- Trespassing

The protocol is depicted in Figure 1. Upon detection, the Web platform automatically highlights the event on the interactive map interface and issues a notification in the designated alert area. These notifications are directly linked to the precise location of the triggered sensor, allowing operators to immediately understand the context and location of the alert.

Simultaneously, the location data of the triggering sensor is stored and relayed to the drone operator's controller, provided the drone is powered on and in range. This allows for seamless coordination between ground detection systems and aerial verification.

Once activated, the drone navigates to the incident site and initiates a real-time video stream, which is made available to authorized platform users. This live feed enables immediate situational awareness, helping stakeholders distinguish between false positives and actual threats, thereby enhancing the system's reliability, response time, and operational efficiency.

The three-way verification protocol serves as a critical layer of redundancy and trust in the Green-HIT monitoring framework, ensuring that all environmental and security alerts are verified swiftly and accurately before action is taken.

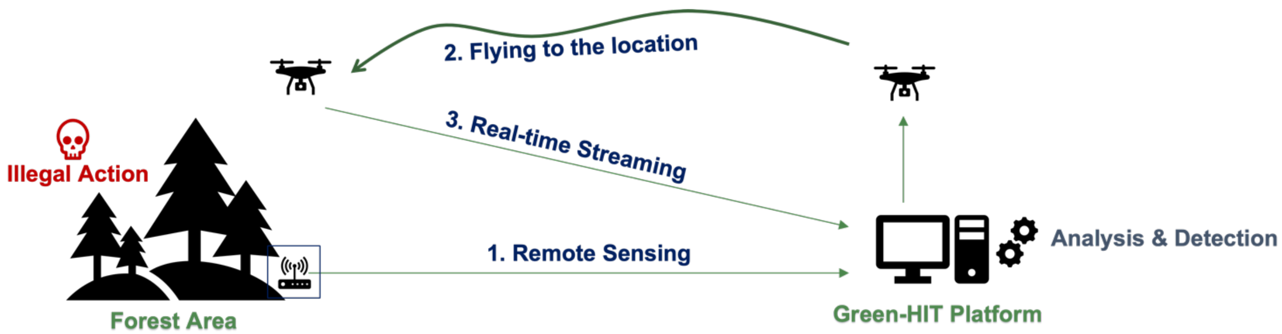

Figure 1: Three-way verification protocol.

Figure 2 showcases all stages involved through photos taken during real-life scenarios tested as part of the project’s pilot trials.

Gunshot sound and audio source detection and alert

UAV flies to the area to verify

UAV live feed – streamed to Web platform

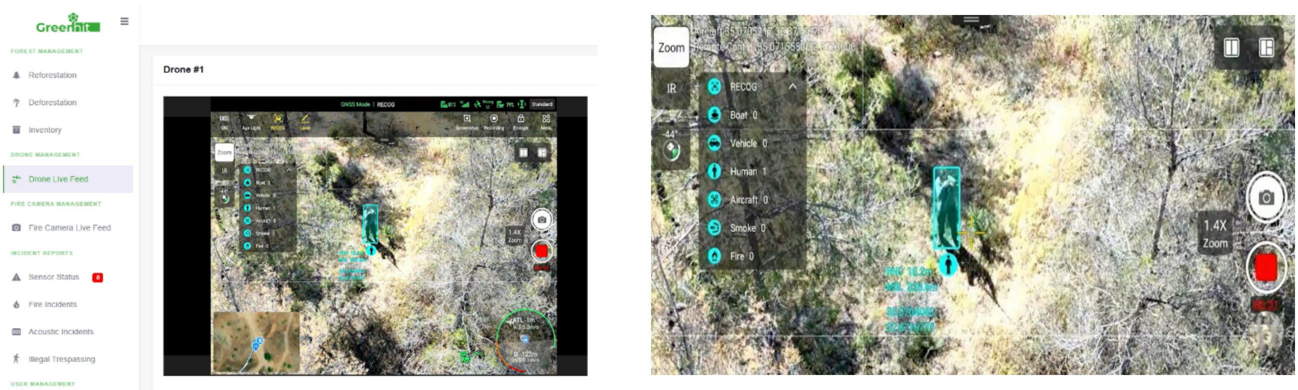

Figure 2: Three-way verification protocol in action – illegal hunting scenario.

3. In-Situ Camera-Based Detection of Illegal Trespassing – Biodiversity Protection

The in-situ camera-based modules are intended for deployment at key entry points of protected areas (e.g., high ecological value sites, nature trails, etc.) to detect unauthorized vehicle entry, offering surveillance for overgrazing and other in-forest human activities, for e.g., camping and picnics, illegal littering, and dumping of wastes.

In the *Platy Valley* hiking trail, the *Milesight X5* camera (Figure 3) is positioned eight (8) meters inside the trail entrance to distinguish trespassing vehicles from passersby on nearby roads. The system operates in continuous mode or on motion detection and has an effective range of six (6) meters. When triggered, the camera sends images to the Green-HIT platform (Figures 4 & 5), logging the incident and initiating the Green-HIT three-way verification protocol. The camera's location ensures minimal false positives while maintaining effective detection.

Figure 3: Milesight X5 camera Installed at the entrance of the hiking trail.

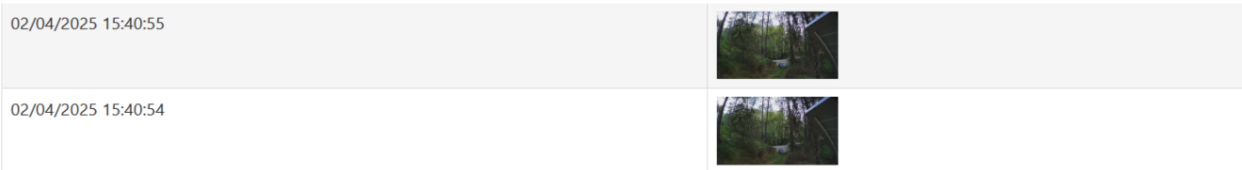

Figure 4: Logs of illegal trespassing incidents (accessed from within the Green-HIT platform).

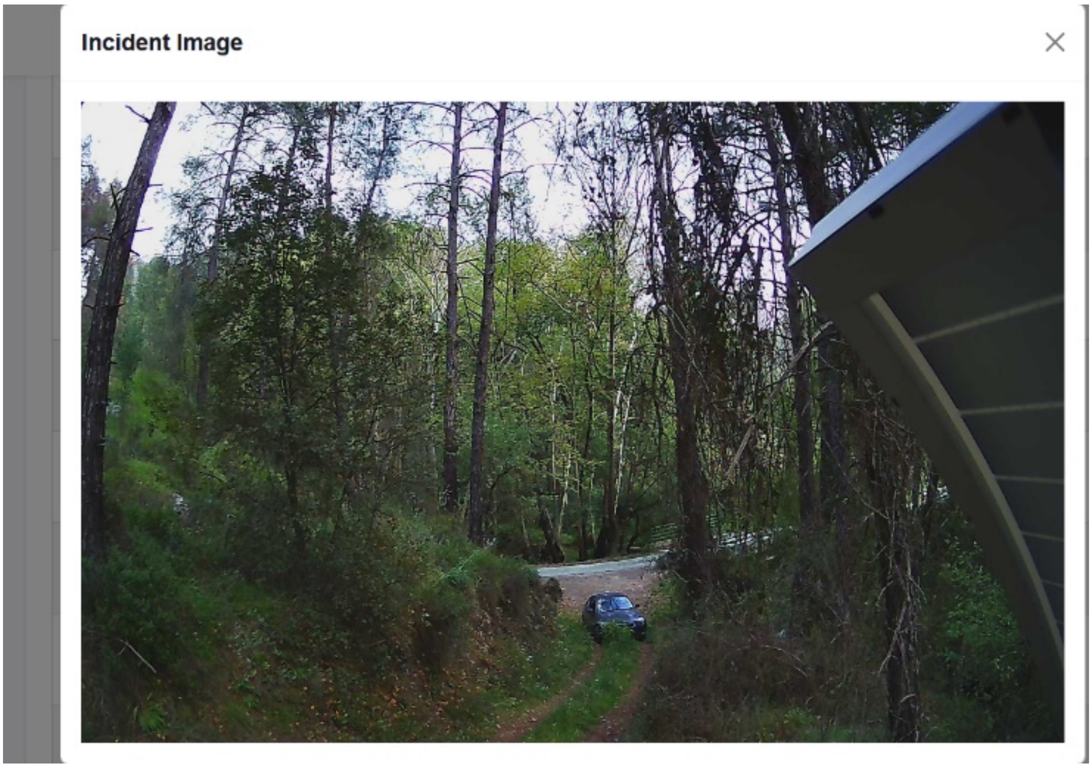

Figure 5: Image capture of trespassing vehicle (accessed from within the Green-HIT platform).

4. In-Situ Acoustic-Based Detection of Illegal Trespassing – Biodiversity Protection

The in-situ acoustic module (Figure 6) detects trespassing sounds, namely vehicle engine sounds indicating unauthorized entry into nature trails or restricted zones. This module was installed In the *Platy Valley* area, deeper into the trail, to complement the camera-based approach deployed at the entrance. It uses *Edge Impulse* and follows the same signal processing and training pipeline as the illegal hunting and logging detection model (see deliverable D6.2), using engine sounds as the target class 'Engine' and environmental audio as 'Other'.

The engine detection model achieved 100% accuracy for detecting engine sounds, with some false positives (8%). These models were tested under both lab and real-life conditions. The latter entailed detection of live vehicle entries. When detection surpasses the confidence threshold, an alert is sent over LoRa to the Green-HIT platform. This system ensures that even if a vehicle passes undetected by the entrance camera, it can still be detected acoustically further down the trail.

Figure 6: In-situ audio recognition module installed deeper inside the forested areas.

4.1 Signal Processing Approach

Audio preprocessing was performed using Mel-filterbank Energy (MFE), which is well-suited for non-speech data. The MFE extracts a spectrogram from the audio signals provided in the dataset using time and frequency features in a non-linear scale called the Mel-scale. The MFE was set with the following parameters:

- Frame Length: 0.02 seconds
- Frame Stride: 0.01 seconds
- Filter Number: 40
- Fast Fourier Transform (FFT) Points: 512
- Low Frequency Band: 300Hz
- Noise Floor: -75dB

First a spectrogram was created using the Frame Length, Frame Stride and FFT points provided. It divides the signal window of the data into multiple overlapping frames based on the Frame Length and Stride provided. For example, a sample with a window of 1 second (using the above parameters) would create 99 timeframes. An FFT is then calculated for each frame. The number of frequency features is equal to the FFT points divided by two (2) plus one (1). The Noise Floor is then applied to the spectrum.

After the spectrogram is computed, the triangular filters are applied on a Mel-scale to extract frequency bands, using the Low Frequency Band parameter as low and zero as high. The number of frequency features extracted is determined by the Filter Number parameter.

The FFT Bin Weighting graph (Figure 7) shows how the FFT bins are scaled and summed into the output columns based on the parameters above.

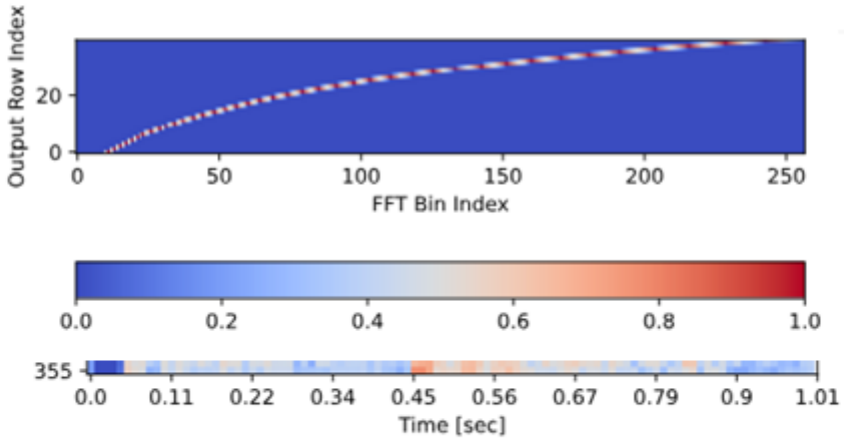

Figure 7: FFT Bin weighting graph.

4.2 Model Training and Neural Network Architecture

Model training was done using the Classification (Keras) learning block while utilizing a One-Direction Neural Network architecture. This was chosen because it is suitable for two-dimensional data like audio. The following settings were used to train the model:

- Training Cycles: 200
- Learning Rate: 0.0005

The training algorithm passed through the training data 200 times (200 training cycles) and adapted the model's parameters at the set learning rate. Training cycles and learning rates were determined by considering several models with different training settings to find those exhibiting the best accuracy while avoiding overfitting. The network used the following neural network architecture.

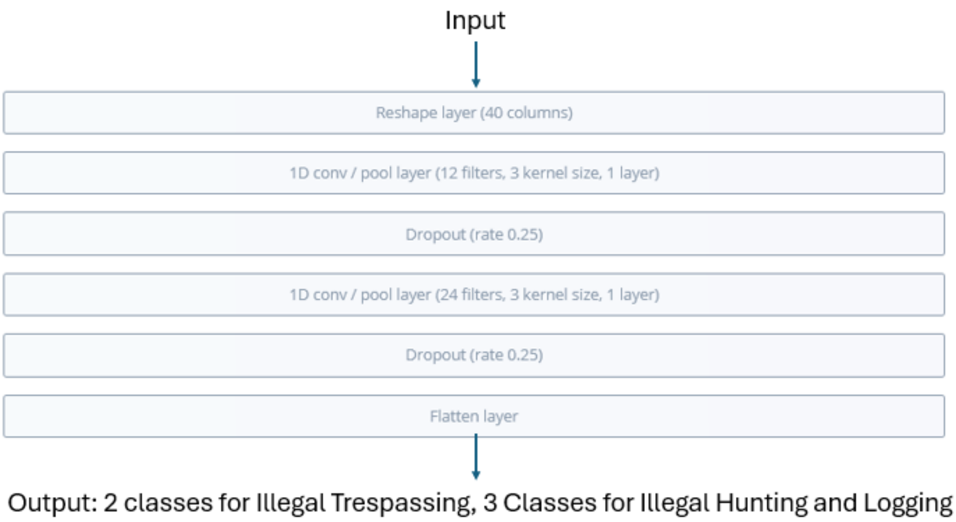

Figure 8: Neural Network architecture.

1. The inputs are the extracted features taken during signal processing; these features pass through each layer of the above architecture.
2. The reshape layer turns the one-dimensional data from the feature into multi-dimensional data to feed into the convolutional layer.
3. The data is then passed through two (2) convolutional layers:
   - the first slides 12 filters across the sequence with three (3) kernel size that moves at one (1) step at a time.
   - the second slides 24 filters.
4. After each convolutional layer, several network connections are cut from the model to reduce overfitting with a dropout probability set at 25%.
5. The features are then flattened back into a single dimension to provide the output, which is separated in two (2) classes, i.e., *Engine* and *Other*.

4.3 Model Accuracy and Performance

This section discusses the detection accuracy of the developed models.

For the illegal trespassing detection module (Figures 9 and 10), *Engine* was correctly classified at 100%, while 8% of the *Other* samples were classified as *Engine*.

ACCURACY
97.2%

LOSS
0.04

Confusion matrix (validation set)

	ENGINE	OTHER
ENGINE	100%	0%
OTHER	8%	92%
F1 SCORE	0.98	0.96

Figure 9: Illegal trespassing detection model - confusion matrix.

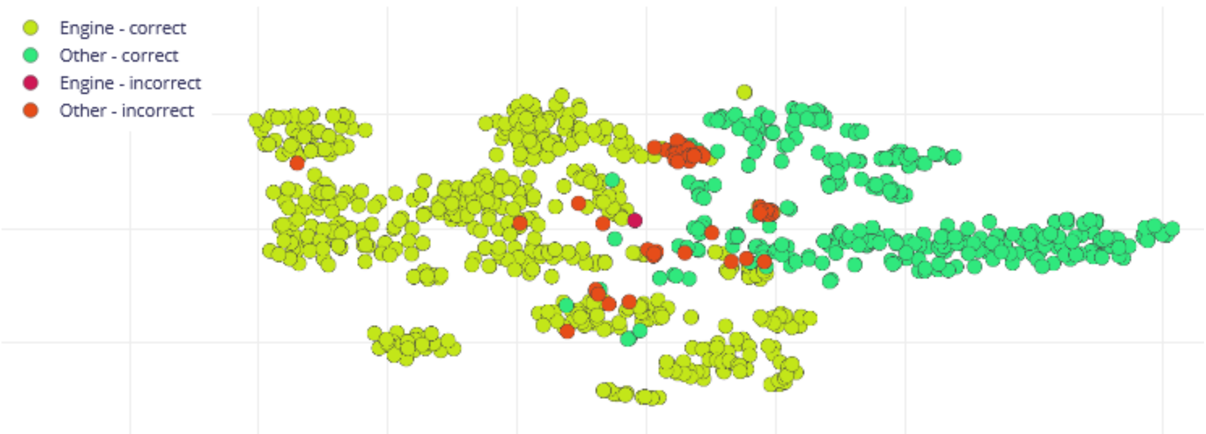

Figure 10: Illegal trespassing detection model - accuracy visualiser.

The performance metrics of the model are the following:

- Inferencing Time: 7ms
- Peak Ram Usage: 20.8k
- Flash Usage: 51.8k

Considering the Nordic nRF52840 module inside the Audio Recognition Module has 1MB flash and 256KB RAM, the model is well within the acceptable performance requirements to function.

4.4 Model Implementation

The developed models were deployed as an Arduino library to perform inference continuously in the field. The models were tested with the audio samples used in the dataset and real-life scenarios, such as real vehicle engines.

5. Cloud-Based Acoustic Detection of Illegal Trespassing – Biodiversity Protection

In addition to the embedded, in-situ illegal trespassing detection module, a Cloud-based pipeline was also implemented using YAMNet, in accordance with the approach described in deliverable D6.2. This pipeline allows offline or batch analysis of uploaded audio clips to detect signs of human activity in protected environments. The system processes audio using standardized preprocessing (resampling to 16kHz mono, STFT, log Mel spectrograms), and performs frame-level inference using YAMNet. Outputs include 521-class probabilities and 1024-dimensional embeddings per frame, which are aggregated and passed through a downstream classifier.

In the context of trespassing detection, the embeddings are used to train a simple logistic regression model to recognize acoustic signatures such as speech, vehicles, or other human-related activity. This model complements the in-situ modules by enabling retrospective analysis of collected audio data and continuous model improvement with new field data.

This module was designed for seamless integration into broader environmental monitoring frameworks. Depending on deployment needs and available infrastructure, it can operate either on edge devices (e.g., Raspberry Pi, microcontrollers with AI accelerators) or be cloud hosted. Audio sensors in forested or protected areas can record data continuously, or on a scheduled basis, which is then streamed or uploaded for analysis. The modular design ensures that updates to the model or preprocessing pipeline can be implemented without disrupting existing installations. Moreover, the module can be expanded to include other modalities, such as image or video detection, to provide a multi-sensory fusion-based monitoring capability.

6. Additional Environmental and Infrastructure Data Captured

Beyond its core focus areas (related to fire, illegal logging and hunting, trespassing, and forest regeneration incidents), Green-HIT collects a wide range of environmental and infrastructure-related data through its network of composite sensors, drone systems, and cameras. While this data is not directly used by the intelligent modules described in previous sections, it represents a valuable resource for expanding the platform's capabilities in future warning and detection scenarios. Key categories of general data collected include:

- Environmental Parameters: Through deployed weather stations (e.g., Seeed SenseCAP S2120), the system captures temperature, humidity, barometric pressure, wind speed and direction at regular intervals. This data is useful not only for fire risk modelling but also for applications in climate monitoring, disease outbreak prediction (e.g., pest-borne), and ecological research.
- Soil and Vegetation Monitoring: The SATIoT multi-sensor hub integrates with soil temperature, humidity, and conductivity sensors. This information can support future services related to drought prediction, precision agriculture, and land degradation assessment.
- Thermal Imagery: Thermal cameras installed on lookout towers and drones provide a continuous feed of temperature distributions across monitored zones. This data could aid in wildlife tracking, energy infrastructure surveillance, or early heat-stress detection in ecosystems.
- LoRaWAN Network Telemetry: Sensor uptime, signal strength, and network performance metrics are also logged and can be analyzed to improve infrastructure resilience or detect potential network interference patterns over time.
- Drone-Based Visual Data: High-resolution RGB imagery captured during drone flights is currently used for incident verification, but may also serve for land-use classification, erosion monitoring, or identifying changes in vegetation density.

All collected data is securely transmitted and stored within the Green-HIT platform's backend infrastructure. Thanks to its modular and scalable design, this data can be readily leveraged by additional detection services or third-party applications, reinforcing the platform's long-term value and extensibility. This broad-spectrum environmental dataset positions Green-HIT as a foundational tool not only for immediate incident response, but also for strategic environmental management and planning.

7. Rationale Behind Choices Made

The Green-HIT other incident and warning system design is guided by the need to balance detection reliability, energy efficiency, hardware limitations, and deployment feasibility in remote or rugged environments. Each technology included was selected based on its ability to satisfy those constraints while delivering accurate real-time detection.

Camera-based modules offer clear visual validation of unauthorized activity but are constrained by visibility, lighting, and line-of-sight requirements. Their placement must be strategic to avoid false positives and to optimize coverage. They are ideal for fixed entry points or narrow passages.

Acoustic sensors, on the other hand, provide round-the-clock monitoring and do not require line-of-sight, making them suitable for wider area surveillance and detection of distant or obscured events. These sensors are less affected by terrain and lighting but may suffer from environmental noise or audio muffling depending on placement. This makes elevation and redundancy essential.

The use of the *Edge Impulse* platform for embedded AI model development was based on its rapid prototyping capabilities, seamless deployment on microcontrollers, and suitability for low-power environments. It enabled the design of efficient neural networks with short inference times and minimal memory usage, making them ideal for Nordic nRF52840 microcontrollers used in the field.

For centralized, post-processed analysis, Google's YAMNet was selected due to its robust performance across 521 audio classes, lightweight MobileNetV1 architecture, and availability of pre-trained weights, allowing rapid deployment and further fine-tuning using domain-specific datasets. Its dual output (interpretable class scores and 1024-d feature embeddings) enables both classification and anomaly detection.

8. Conclusions

Deliverable D6.4 showcases the comprehensive modular incident detection infrastructure deployed in the Green-HIT platform. By combining real-time edge AI modules with cloud-based classification, and fusing acoustic and visual modalities, the system achieves robust, scalable monitoring of remote conservation areas. The field deployment in *Platy Valley* and *Orkonta* demonstrates practical feasibility, high model accuracy, and reliable alerting. Future directions include deeper sensor fusion, anomaly detection techniques, and integration with UAV path planning systems.