DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

A Study on Acoustic Bird Detection in the Context of Smart Agriculture

Fábio Cardoso
Polytechnic Institute of Castelo Branco, Av. Pedro Álvares Cabral, nº 12, 6000-084 Castelo Branco, Portugal
c.fabio@ipcbcampus.pt

Daniel Carvalho
Polytechnic Institute of Castelo Branco, Av. Pedro Álvares Cabral, nº 12, 6000-084 Castelo Branco, Portugal
daniel.carvalho@ipcbcampus.pt

Pedro D. Gaspar
Department of Electromechanical Engineering, University of Beira Interior; Centre for Mechanical and Aerospace Science and Technologies (C-MAST), Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal
dinis@ubi.pt

Vasco N. G. J. Soares
Polytechnic Institute of Castelo Branco, Av. Pedro Álvares Cabral, nº 12, 6000-084 Castelo Branco, Portugal
Instituto de Telecomunicações, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal
AMA – Agência para a Modernização Administrativa, Rua de Santa Marta, n° 55, 1150-294 Lisboa, Portugal
vasco.g.soares@ipcb.pt

João M. L. P. Caldeira
Polytechnic Institute of Castelo Branco, Av. Pedro Álvares Cabral, nº 12, 6000-084 Castelo Branco, Portugal
Instituto de Telecomunicações, Rua Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal
jcaldeira@ipcb.pt

ABSTRACT
Bird attacks on crops represent one of the main challenges faced by farmers to make a farm profitable and sustainable. This challenge requires a paradigm shift from old, traditional, ineffective methods, towards the incorporation of smart farming technologies. Intelligent or precision agriculture contributes to the effective and efficient management of resources and, consequently, an increase in production. This work first reviews projects, techniques and methods that can be used for bird species detection and identification based on their sound. Then, a performance evaluation study is conducted on two different approaches that can be employed for the development of a smart bird deterrence solution. Their strengths and limitations are highlighted. The findings can be used as a foundation for future research in this area.

Keywords
Smart farming; Bird deterrence; Acoustic bird species detection; BirdNET; Crop protection; Sensors; IoT.

1. INTRODUCTION
Throughout human history, agriculture has proved to be a fundamental pillar for ensuring survival and financial sustainability. This ancient practice not only provides essential food, but also economically sustains communities around the world.

Every year, farmers face considerable crop losses due to birds that compromise the development, the quality and quantity of crops. The amounts associated with this problem can range from hundreds to thousands of euros a year in some cases [1], [2], [3]. The problem lies in the nature of the birds, which often feed on crops at crucial stages of growth and is aggravated by the fact that there is a shortage of efficient low-cost bird dispersal and deterrent methods [4]. Traditional methods, such as the use of scarecrows or sound devices, often prove unsuccessful in the face of the birds' ability to adapt to them. Lethal methods, on the other hand, may not be an option due to laws that may exist to prevent the extinction of certain species [3].

Contemporary agriculture is amid a remarkable transformation, driven by the rise of the Internet of Things (IoT). Innovative initiatives such as the implementation of IoT sensors to monitor weather conditions in real time [5] and the use of drones equipped with cameras to inspect crops [6] are two examples of the versatility and effectiveness of these approaches. This evolution shows not only a change in agricultural practice, but also a successful adaptation to emerging technological solutions. The practical applications of IoT sensors and drones demonstrate not only the viability of these technologies, but also their potential to improve and optimize agricultural processes.

In the specific case of crop protection against bird attacks, the use of IoT devices can not only help to detect birds, but also to enable agile responses to disperse them, propelling agriculture into a smarter, more interconnected era. Following up on a previous work [7], the work presented in this paper serves as the foundation for the ongoing development of a prototype capable do detect and analyze the sounds emitted by birds in a crop to identify the species that is causing damage. Upon recognizing the invasive bird species, the system triggers a drone, capable of moving autonomously to the identified location to carry out some bird dispersal maneuvers.

This paper presents the results of the first stage of this project. It focuses on developing a system capable of capturing and identifying the various sound patterns emitted by different species of birds. Hence, it studies, identifies, and evaluates techniques for detecting and classifying bird species based on their sound. Different approaches are proposed, and the most promising ones are implemented and evaluated.

This paper is organized as follows. Section 2 presents the related work on the approaches used for detecting and classifying bird species based on their sound and discusses the challenges and opportunities. Section 3 describes the proposed solutions that can be employed for the development of a smart bird deterrence solution. Section 4 presents a performance evaluation of these approaches. Finally, Section 5 presents the conclusions, along with suggestions for future work.

2. RELATED WORK

As part of the review of the state of the art, we will conduct an analysis of projects that are similar or related to the purpose of ours, with the aim of understanding the existing panorama of similar solutions. This research aims to obtain information on past implementations, including details of the hardware and software used, challenges faced and strategies adopted to overcome these challenges. We intend to acquire information on what has already been accomplished, how it was approached and what results were achieved. This analysis will serve as a conceptual basis for our proposal, allowing us to learn from the experiences of others and develop am innovative, robust, effective and cost-efficient solution.

Bird@Edge [8] is a project dedicated to recognizing 82 species of birds' native to Germany through audio recordings. Using a combination of stations and distributed microphones, the project captures audio in file format with a sampling rate of 44.1 Kilohertz (kHz), capable of detecting frequencies of up to 22.05 kHz. Each audio sample lasts 5 seconds, during which time one of 5 background sound files is randomly selected to be added to the capture of the bird's sound. Devices called Bird@Edge Mics [9] transmit the audio over a wireless network to a Bird@Edge Station, which performs the species recognition analysis. The analysis results from the Bird@Edge Stations are sent to a central infrastructure and stored in a time series database (InfluxDB), allowing the data to be visualized using the Grafana platform in a web interface.

BirdNET-Pi [10] is a real-time bird classification system designed specifically for Raspberry Pi 4B, 3B+ and 0W2. It has several features included and offers a comprehensive solution for automatic and continuous audio-based bird identification. One of its main advantages is its ease of implementation and use, being designed to be accessible and intuitive. In addition to continuously capturing audio from the environment, BirdNET Pi also offers automatic extraction and classification of the bird species detected and then clear and organized presentation of the statistics via a web interface. This combination of features makes BirdNET Pi an autonomous and user-friendly solution, suitable for a variety of biodiversity monitoring scenarios.

BirdNET-Go [11] is a solution specializing in the identification of bird species. This system offers the possibility of analyzing specific audio files or capturing audio in real time, automatically taking 3-second samples for continuous analysis. Using the BirdNET model, BirdNET-Go can identify the different species present in the analyzed bird sounds. It also provides an intuitive web interface for data visualization, where it is possible to check the species detected and their frequency of detection at different times of the day. It also allows this data to be recorded in log files and stored in Structured Query Language (SQL) databases for later analysis and reference.

The Merlin Bird ID mobile application [12] is available for Android and iOS devices, offering an intuitive and effective way of identifying bird species. With Merlin Bird ID, users can capture photographs or audio recordings to identify the birds they find. The application is not limited to species identification; users can submit their sightings and explore which species are most likely to be found near their location. An additional feature is the maintenance of a list of submitted species, encouraging continued user participation and contributing to the birdwatching community. For those who aren't very knowledgeable or unsure about identification, Merlin Bird ID also offers the option of species identification through three simple steps, making the process easy for all users.

Many of the projects observed so far rely on a database called eBird [13], which stands out for containing a vast collection of samples of birds and their associated species. This database is widely used in projects related to bird identification due to its comprehensiveness and reliability. One of eBird's main features is its openness to the public, allowing bird enthusiasts to contribute their own sightings and audio recordings. In addition, eBird provides an accessible Application Programming Interface (API), allowing developers to use the data from the database in their own projects and applications, thus extending its reach and the amount of data available.

A crucial component in most of the projects mentioned above is BirdNET Analyzer [14]. BirdNET Analyzer analyzes the audio provided to determine the species of birds present. In addition to identifying the species, it also provides additional information, such as the location in the audio where the birds were detected, and the confidence level associated with each detection. BirdNET Analyzer works with files with a sampling rate of 48 kHz and can detect frequencies between 0 Hz and 15 kHz.

2.1 Critical Analysis

After a thoughtful analysis of the different projects available, we highlighted the BirdNET Analyzer algorithms and the eBird database as the most widely used and reliable because they are part of most of the other solutions that exist. The BirdNET-PI solution aroused interest because of its ability to analyze audio in real time and because it incorporates automatic statistical analysis and graphical demonstration of the data while keeping the necessary computing power to a minimum. Bird@Edge, although it seemed a promising option compared to BirdNET, turned out to be an extremely limited solution due to its exclusive availability for NVIDIA Jetsons platforms [15] and the high costs of the components required for its implementation.

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

The next section will present the proposal for building a prototype capable of detecting and identifying bird sounds based on the use of the BirdNET-Pi and BirdNET Analyzer solutions and the eBird database.

3. PROPOSAL

Our project aims to help farmers to protect against the damage to crops caused by birds. To achieve this, we sought a solution that would be effective and able to deal with the complexities of different agricultural scenarios while remaining affordable. Our aim is to develop an integrated system that combines the detection and identification of bird species through sound processing with dispersal strategies using a drone equipped for this purpose.

At this stage of the project, we only intend to focus on bird species detection and recognition, leaving their dispersal for a second stage. We will use an architecture made up of several nodes distributed among different trees in the orchard that will forward the information collected to a central node. This node will then process it and send an order to a drone that will fly to the location where the bird detection occurred and perform a dispersal maneuver. This concept is illustrated in Figure 1.

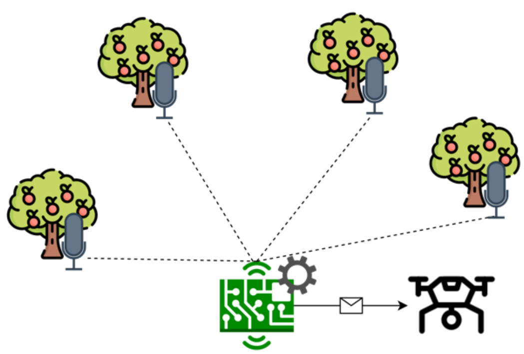

Figure 1. Diagram of the infrastructure for detecting bird species in an orchard.

3.1 Testbed

To implement the proposed architecture, it is important to carefully select the appropriate hardware for the nodes so that they can perform their function successfully.

Although other options could have been considered, given the budget and the equipment we had available, a Raspberry Pi 4B [16] shown in Figure 2, was used as the central node. This equipment proved to be a solution that met the requirements initially defined, due to its computing capacity and affordable cost. Equipped with a Quad-core 64-bit Advanced Risc Machine (ARM)-Cortex A72 1.5Gigahertz (GHz) processor and 4 Gigabytes (GB) of Low Power Double Data Rate 4 (LPDDR4), along with an integrated Wireless Fidelity (Wi-Fi) antenna operating at a frequency of 2.4GHz, this device offers performance and connectivity for the central function of the system. To add the ability to capture audio for testing, a HyperX Cloud II headset was connected to access its microphone.

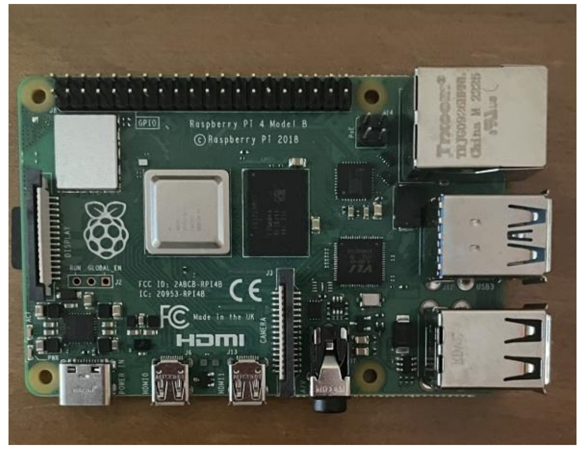

Figure 2. Raspberry Pi 4B 4GB.

For the same reasons stated in the previous case, the Raspberry Pi 02W models shown in Figure 3, were selected for the distributed nodes, due to their versatility and variety of connections. In addition to the integrated Wi-Fi antenna and Bluetooth 4.2 support, including Bluetooth Low Energy (BLE), this device offers a Universal Serial Bus (USB) On The Go (OTG) port and a 40-pin General Purpose Input/Output (GPIO). These features provide a wide range of connectivity options, allowing not only the integration of high-quality microphones, but also more economical and accessible alternatives [17].

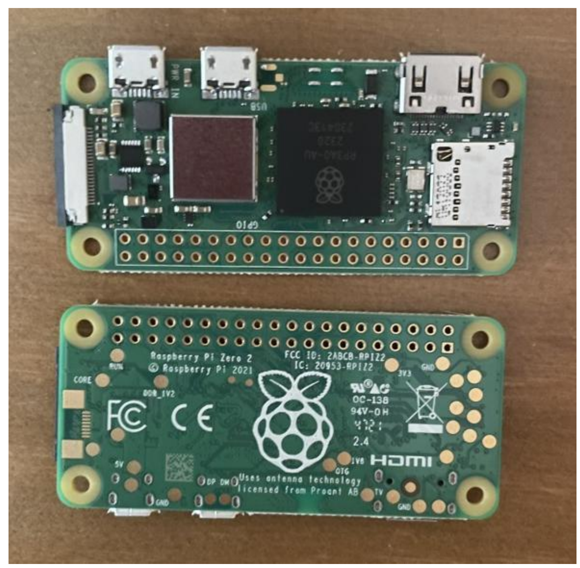

Figure 3. Raspberry Pi Zero 2 W.

3.2 Development

Based on the proposed architecture and the available hardware solutions, two approaches were considered. Each of them will be described below, along with the reasons why they were considered in the construction and operation of the proposed solution.

The task of detecting and identifying bird species, using the techniques identified in Section 2, is demanding from a computational point of view. The computing power on the

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

distributed nodes (Raspberry Pi Zero 2 W) is more limited, which can have an impact on achieving the desired results. On the other hand, the time needed to obtain these results may not be appropriate for the drone to receive an order in appropriate time. An alternative to consider would be to use the additional resources present on the central node (Raspberry Pi 4B) and perform the most demanding computational tasks on that node. In this case, the impact of transporting the data on the network and the work overload on this node could have an impact on obtaining results in appropriate time.

With these considerations in mind, two possible approaches were defined. The first involved processing the captured sound data to identify the bird species directly on the distributed nodes and only sending the result to the central node. In the second, the distributed nodes send the captured sound data to the central node and the latter carries out the necessary tasks to identify the bird species. The developments carried out to implement these two approaches will be detailed below. This will be followed by a performance evaluation of the two approaches and a critical analysis of the results obtained.

Before moving on to the specific details of each approach, it is important to first present the general operation of the solution. Initially, the system is prepared to listen constantly for any sound changes. When a change is detected, an audio file is created with the captured sound. Then, using the BirdNET-Pi and BirdNET Analyzer algorithms, the file is analyzed to identify the presence of any bird species. If any species is identified, the results are recorded in a text file which is then sent to the central node. This file includes the following data: the common name of the species detected, the date and time of detection and the coordinates of the node.

The central node also remains in a constant state of waiting for data. When it receives data from one of the distributed nodes, it searches the received file to find the location of the sending node. Based on this search, an order is sent for the drone to travel to the indicated location to disperse the birds. This will be explored in a later stage of this project.

3.2.1 Software

At the core of both approaches is the BirdNET algorithm, which is a robust neural network that is highly specialized in detecting and identifying birds from audio recordings. This algorithm is an ideal choice for our project as it offers a unique combination of accuracy and efficiency [18].

We will highlight two specific variations of BirdNET that will be used in the tests: BirdNET-Pi and BirdNET Analyzer. BirdNET-Pi is optimized for Raspberry Pi devices, offering the ability to capture and analyze audio in real time. This functionality is crucial for our application, as it allows birds to be detected immediately and preventative measures to be taken quickly. The installation process simply consists of downloading the project from the project's GitHub via the link provided on the project's main page, visible in Figure 4.

Once the provided command has been executed and everything has been installed, simply open the website with the hostname you defined when creating the system image. If no link has been defined, *http://birdnetpi.local* is used by default. Access can also be made via the IP address of the machine where the installation took place. Figure 5 shows the BirdNET-Pi dashboard. To complete the installation, you just need to update the latitude and longitude so that the system knows which geographical zone you are in.

In addition, we will be using BirdNET Analyzer, an easy-to-use tool that offers a simplified approach to bird detection. BirdNET Analyzer is a convenient option for offline analysis of audio recordings, providing an effective way to identify bird species quickly and accurately. The installation process is very similar to the previous one, with the only requirement being to install the following python dependency packages *tflite-runtime*, *librosa*, *resampy* and *ffmpeg*, using the *pip* tool. After installing all the necessary dependencies, all that's left to do is clone the project's Github repository, which can be seen in Figure 6. The files are analyzed by running the *analyze.py* script located inside the folder with the parameters "*--i*" to provide the file or directory with files to be analyzed and "*--o*" to specify the output path of the results.

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

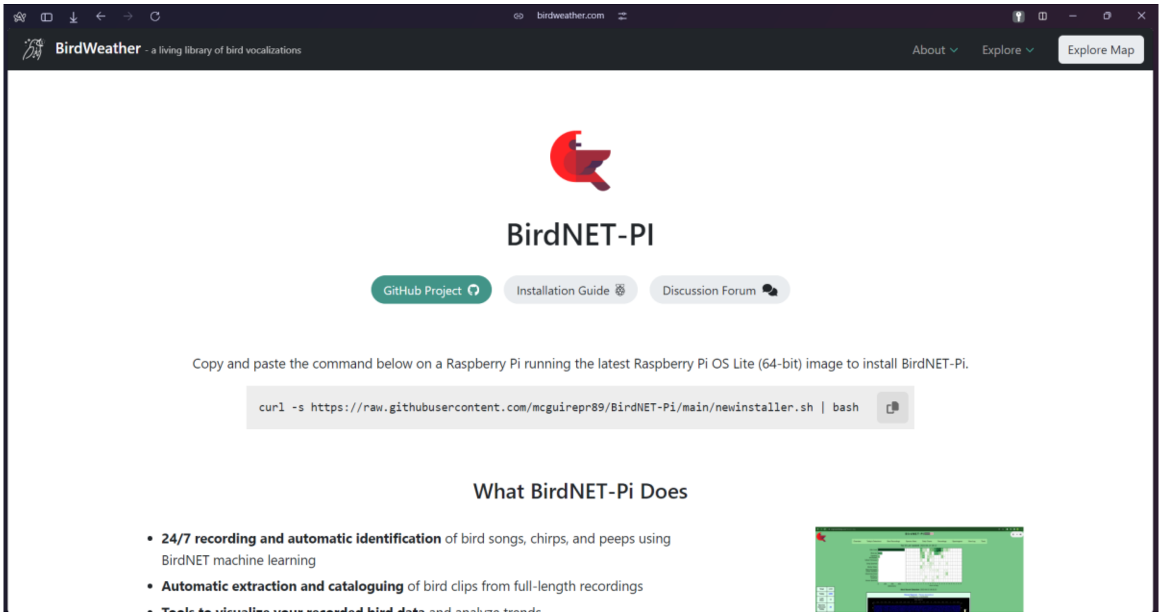

Figure 4. BirdNET-Pi's homepage.

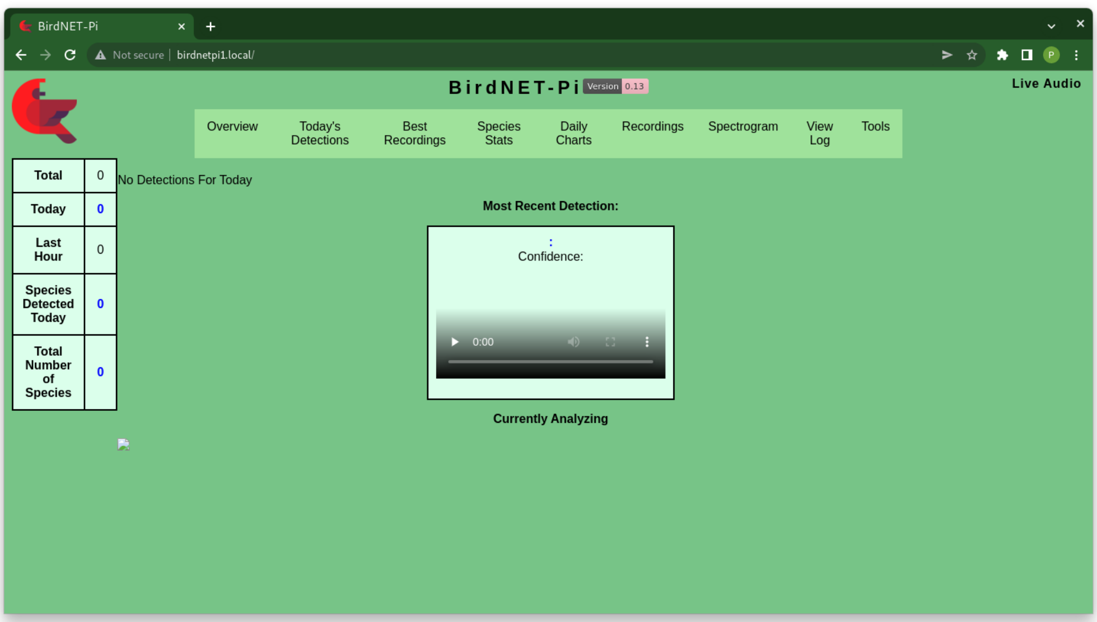

Figure 5. BirdNET-Pi's Dashboard.

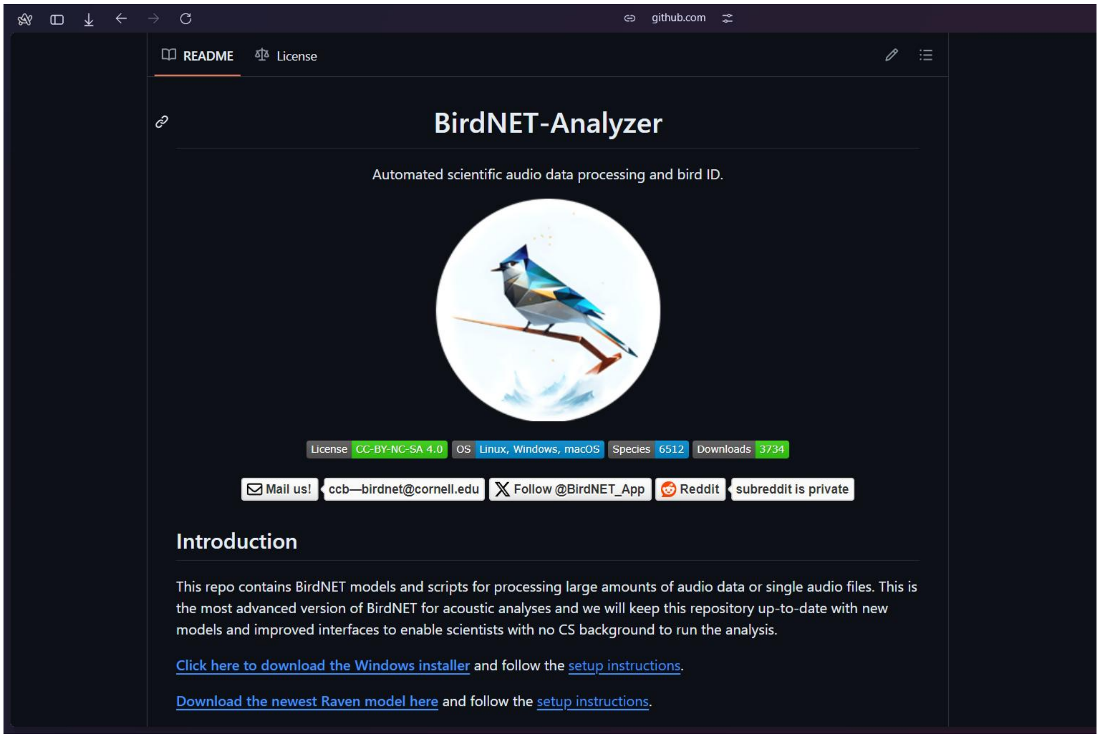

Figure 6. BirdNET-Analyzer's GitHub repository.

3.2.2 Approach 1

The architecture shown in Figure 7 corresponds to the first approach described above. The distributed nodes consist of a Raspberry Pi 02W equipped with a microphone, storage and Wi-Fi and Bluetooth communication capabilities. These nodes play a crucial role in the system's infrastructure, being responsible for the essential functions of capturing and analyzing sound, as well as communicating with the central node.

The audio is analyzed using the algorithms described above and if any bird species is detected in the results, a text file is created with the species data, date and time and location. The results are then sent to the central node which, upon receiving this data, searches the file for the location of the sending node and sends orders for the drone to be activated and move to the specific location to disperse the birds as quickly as possible.

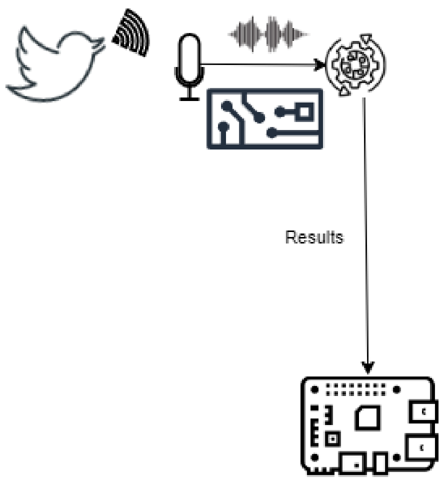

Figure 7. Infrastructure diagram for Approach 1.

As we can see in Figure 8, each distributed node follows a cycle in which it captures the sound and immediately analyzes it using the algorithms. It then checks whether a bird has been detected and, if so, sends the file in the format indicated above to the central node. The central node waits for some data reception. When it receives the data, it determines the location of the sending node and sends this information to the drone, which flies there to disperse the birds present at the location, as illustrated in Figure 9.

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

3.2.3 Approach 2

In the second approach, the functioning of the system is slightly altered in the sense that audio analysis becomes a function of the central node instead of each distributed node, as we can see in Figure 10.

The central node consists of a Raspberry Pi 4B, which has data storage, Wi-Fi and Bluetooth communication capabilities. This central node acts as the convergence point for the distributed nodes and is responsible for processing the collected audio. To do this, it is equipped with the analysis algorithms.

Each distributed node consists of a Raspberry Pi 02W, equipped with a microphone and Wi-Fi and Bluetooth communication capabilities. These nodes have the function of continuously monitoring the environment for sound changes and, if they identify any, capturing the audio and sending it, along with its location, to the central node.

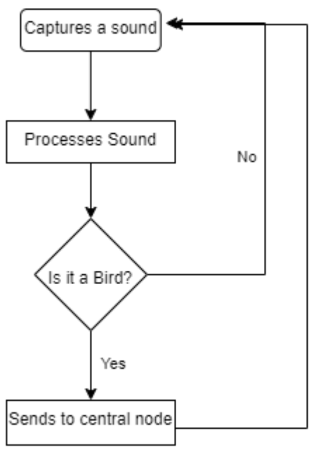

Figure 8. Flowchart of the Distributed Node in Approach 1.

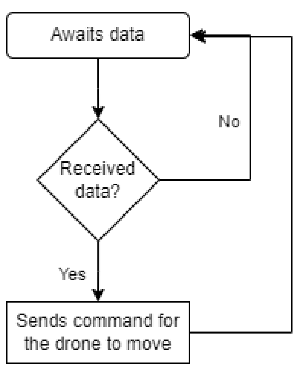

Figure 9. Flowchart of the Central Node in Approach 1.

As a result of the processing now being done at the central node, the distributed nodes now only have the function of continuously detecting sound changes and, when a change is detected, capturing the audio and sending the file, along with the location, to the central node, as shown in Figure 11.

Figure 12 shows the flowchart of the central node. This node is in a continuous state of waiting for files sent by the distributed nodes. When an audio file is received, the central node analyzes it using the BirdNET-PI and BirdNET Analyzer algorithms to identify possible bird species. If any species is detected, the central node sends a signal for the drone to fly to the location.

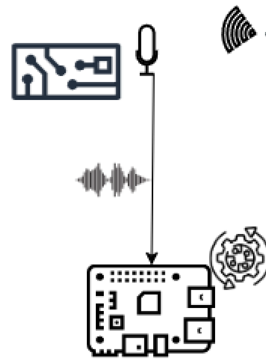

Figure 10. Infrastructure diagram in Approach 2.

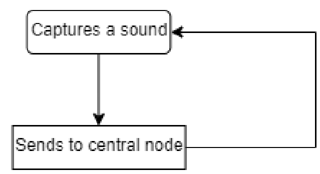

Figure 11. Flowchart of the Distributed Node in Approach 2.

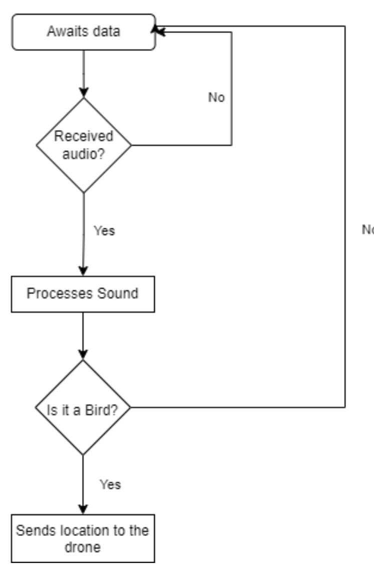

Figure 12. Flowchart of the Central Node in Approach 2.

4. PERFORMANCE EVALUATION

This section presents the performance evaluation of the two proposed approaches described in the previous section. First, we are interested in assessing the times required to obtain the bird species classification. Then, the impact of using different types of audio compression and different levels of audio quality will be evaluated. This section describes the performance metrics considered and presents a critical analysis of the results obtained.

4.1 Performance Metrics

We evaluated the performance in terms of analysis time, level of confidence obtained by the BirdNET algorithm, and success rate. The analysis time refers to the time it takes to carry out the analysis (i.e., determine the bird species) at a node. BirdNET assigns classification confidence score that ranges from 0 to 1 to each 3-s clip, with scores closer to 1 meaning greater confidence that the clip contains a sound of a specific bird species. The success rate was calculated to indicate the accuracy in identifying the bird species captured. This metric is determined by dividing the number of correct detections by the total number of tests carried out on each audio, as shown in Eq 1. A detection is considered correct when the species identified in the result returned by the algorithm has the highest percentage of confidence.

$$\text{Success rate} = \frac{\text{Number of correct species detected}}{\textit{Total number of tests}} \qquad \text{(Eq.1)}$$

The selection of the type of audio file compression was also considered. Files compressed in Waveform Audio File Format (WAV) and MPEG 1 Audio Layer 3 (MP3) were analyzed to assess the impact on the analysis and the reliability of the algorithm.

Another parameter tested was audio quality, classified from A to E, where A represents files with clear audio and E refers to files with more background noise. The aim was to examine the impact of audio quality on the success rate.

Each test was repeated 30 times for each minimum confidence level established, totaling 330 results for each file analyzed. The measurements were repeated to minimize the influence of atypical results or outliers, thus increasing the consistency and representativeness of the data.

All the audio files tested contained audible sounds and had a fixed duration of 15 seconds. The variations between the files lay in the type of compression (MP3 or WAV), size and quality of the audio, graded according to a scale from A to E.

4.2 Results

Various tests were carried out using the approaches and configurations described above, with the aim of evaluating the system's performance. The results reported are the average of 30 runs. The following results were obtained.

4.2.1 Software Performance Analysis

We selected two audios to use as a test base where you can clearly hear the sound of both birds and can identify their species without any problems: in the first audio is from a common ground dove and the second is the sound of crows.

We started with the options that had the most functionality, which were BirdNET-Pi and BirdNET Analyzer. The sounds were played on an external device near the microphone and when we entered the BirdNET-Pi web interface, shown in Figure 13, we saw that it was detecting a sound and showing the spectrogram of the last 10 seconds. After several detection attempts, we concluded that this would not be a good solution because two detections were made and, of those two, only one contained the sound of a bird, as displayed in Figure 14. What's more, the species identified in it didn't match the one in the audio.

As this option did not provide satisfactory results, we turned to the next option, focusing on BirdNET Analyzer. After installing and analyzing the same two audios tested previously, we noticed that not only did it correctly identify the species, but it also provided more detailed results, as we can see in Figures 15 and 16.

In the results, we can see that besides showing the species detected, it also indicates where in the audio the detection was made and the confidence of all detections, even if it is low. In Figure 15, we can see that it detected the species common ground dove several times throughout the audio. In Figure 16, we can see that the same happened with the species common raven and it also detected several varieties of crows.

Considering the results obtained by these two solutions, we decided to carry out the remaining tests using BirdNET Analyzer because it gave better results and identified the correct species in both audios.

4.2.2 Analysis Times in Approach 1

The first test consists of a local analysis of the audio files with a minimum level of confidence, which means eliminating results that have a lower percentage of confidence than that configured. In this test, we register for each bird species the type of file compression, the analysis time, whether the species was correctly identified and, where applicable, the percentage of certainty with which it was detected. We then calculated the overall average of the analysis times.

As we can see in the chart of Figure 17, there is a variation in the values of analysis times between the different species, although all the audios have the same duration. At the minimum end, the average analysis time is 36,51 seconds, while at the maximum end it reaches 61,13 seconds, almost double the minimum time. These values provide an important benchmark for establishing a base value of expected time for each bird species, allowing improvements to be assessed later.

The bar chart of Figure 18 shows the variation in analysis times. The aim of this analysis is to determine whether the confidence level has an impact on the analysis time. However, the results of this chart do not allow for clear conclusions, as the patterns are not consistent. In some cases, analysis times decrease or remain stable as the minimum confidence level increases, but in other cases, these same times increase again halfway through the confidence interval. Therefore, it is not possible to conclude that the minimum confidence level has a significant effect on the analysis time.

This approach, while offering a decentralized and efficient approach to bird detection, also presents significant challenges related to costs and processing time. The most significant cost is associated with the requirement for computing power on each node. To ensure local analysis of sounds, each node must be powerful enough to process the audio data. In larger plantations, this computational demand can become substantial, leading to a considerable increase in implementation costs.

In addition, the effectiveness of this approach is limited by the processing time of the local nodes. The time needed to analyze the audio data can vary from 30 seconds to just over a minute. Considering the additional time needed to send the data to the central node, coordinate with the drone and fly it to the identified area, the total time to disperse the birds can reach an order of magnitude of several minutes. These delays can be critical, especially in situations where an immediate response is essential to avoid significant damage to crops. Thus, the impact on response time can affect the effectiveness of the system, especially in large plantations where efficient coordination is vital.

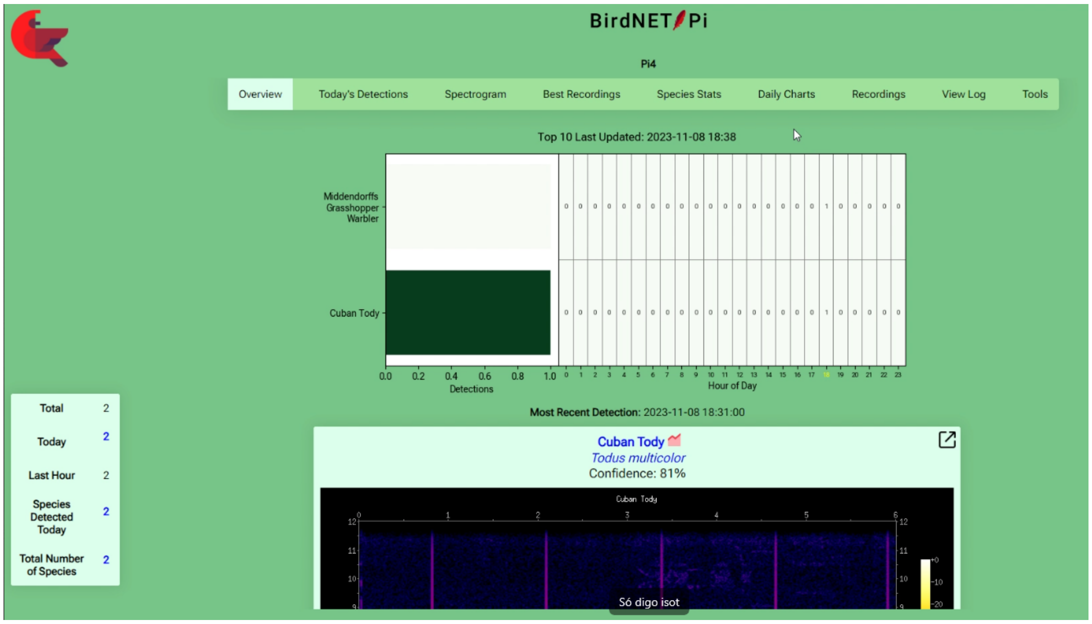

Figure 13. BirdNET Pi's web interface.

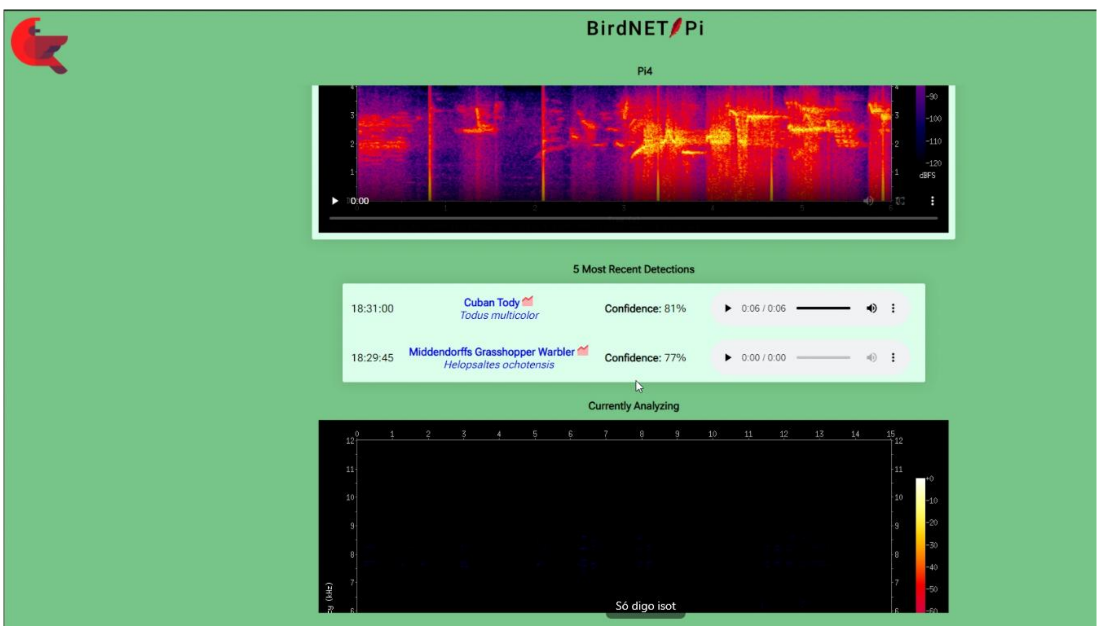

Figure 14. BirdNET Pi's detections.

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

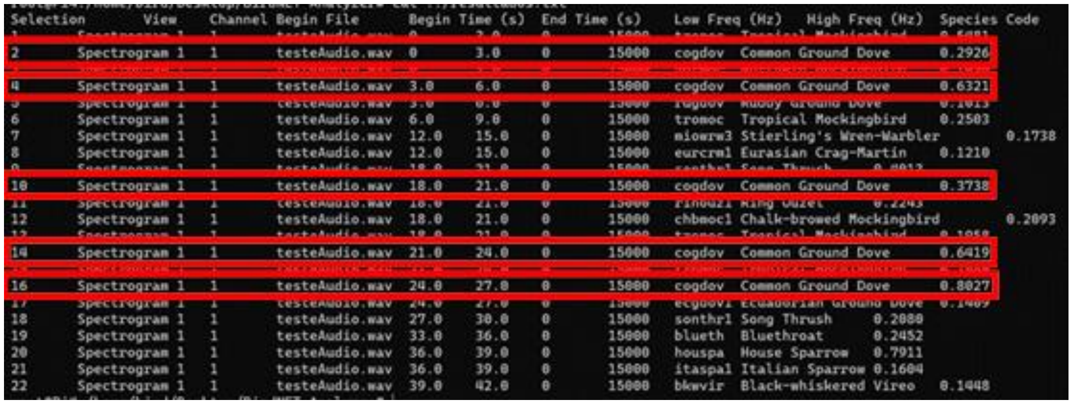

Figure 15 BirdNET Analyzer's common ground dove audio detection results.

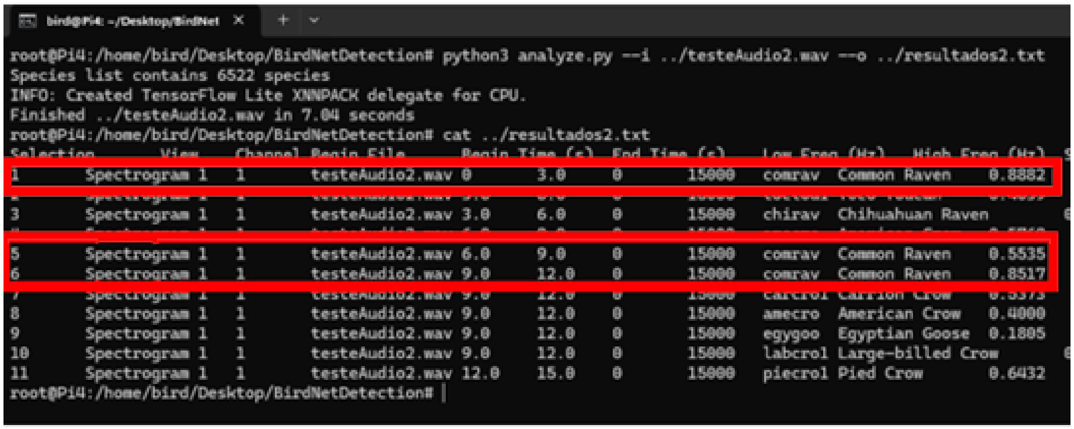

Figure 16. BirdNET Analyzer raven audio detection results.

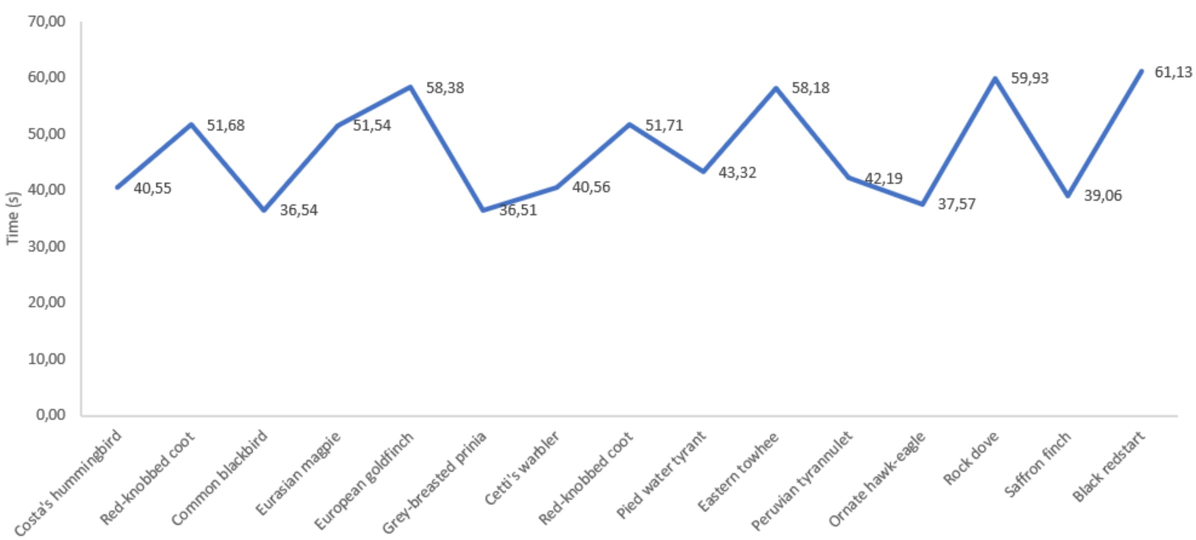

Figure 17. Chart with the average analysis times at the distributed node per bird species.

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

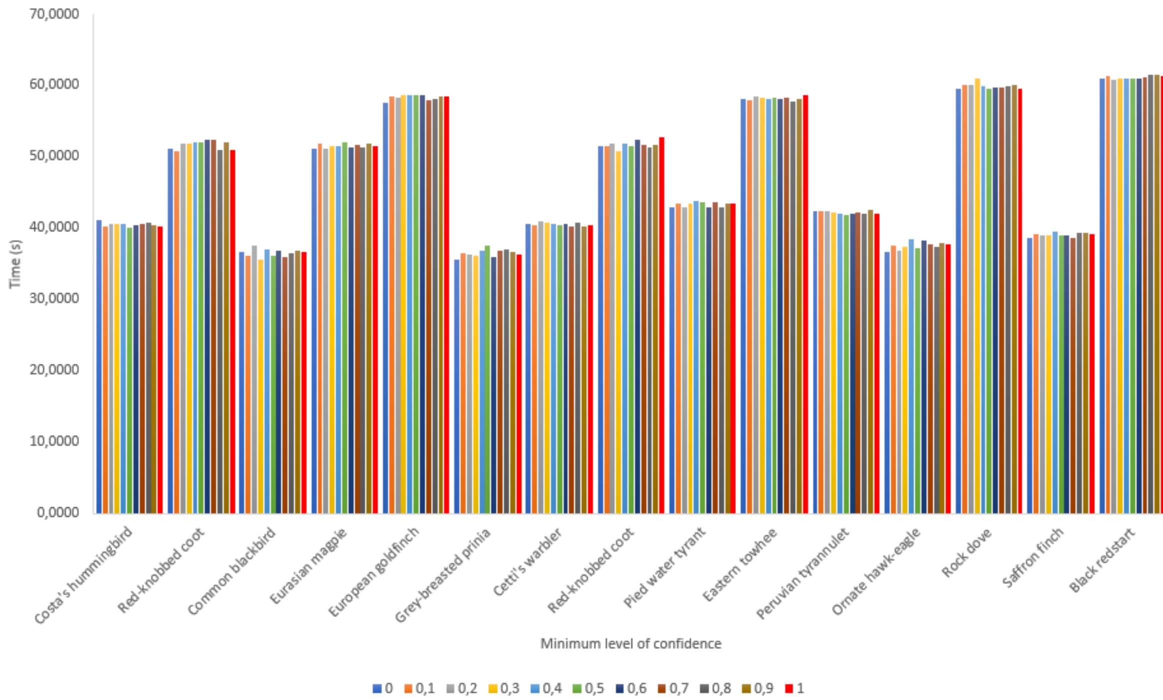

Figure 18. Chart comparing analysis times at the distributed node per bird species for different minimum levels of confidence.

4.2.3 Analysis Times in Approach 2

Like the first test, we will analyze the audio files, except that this time the file will be analyzed on the central node. To try to simulate the real scenario more closely, the file is originally sent from a distributed node using the BirdNET "client" script, to the central node that runs the "server" script.

Now we are interested in measuring not only the analysis time, but also the overall time and the transmission time, illustrated in Figure 19. The analysis time refers to the time it takes to carry out the analysis (i.e., determine the bird species) at the central node. The overall time is the total interval from sending the file from the distributed node to the central node, processing it, and sending the result back to the distributed node. The transmission time is the sum of the time it takes to send the birds sound data to the central node and then send the result with the bird species and confidence levels back to the distributed node. It allows to know how much time is spent exclusively on transmission between the nodes.

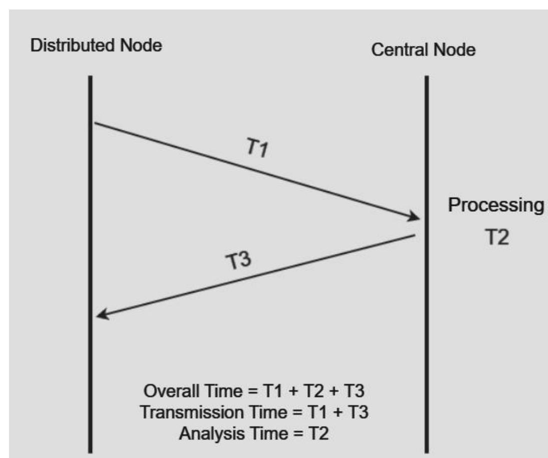

Figure 19. Diagram illustrating the various times.

The chart of Figure 20 shows the same pattern as the previous test. The values tend to be around a certain range. However, two bird species end up showing higher analysis times. As in Figure 17, the same bird species requires the longest analysis time that almost doubles the lowest value. Despite these variations, the times are much lower than in the previously tested approach and are considered more acceptable for a solution that requires a quick response, as is the case with our proposal. In addition, we can see a significant variation in the analysis times for all bird species when the minimum confidence level is set to 0. The difference between the minimum confidence levels of 0.1 and 1 does not show a large variation, but from level 0 to the others, there is a difference of more than double in some cases, as illustrated in Figure 21. The chart displayed in Figure 22 allows us to get an

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

idea of the transmission times we can potentially expect in the final solution. It can be concluded that the transmission time is around 1 second. However, there are higher values, as much as triple the average value. As expected, those higher transmission time values are reflected in the chart presented in Figure 23, when the minimum confidence level is set at 0.9. The strange discrepancies observed in Figure 23 maybe caused by network bandwidth saturation. It is important to note that the times registered refer to a round-trip time between the distributed node and the central node. The closest realistic time would be half of these values, excluding network overhead factors and other possible delays. It is important to note that although this approach, offers very low analysis times, it can result in higher investment costs due to the requirement for significant computing power in the central equipment. This central equipment must be able to efficiently process the quantity of requests coming from all the distributed nodes. Moreover, there are challenges to consider such as the fact that data traffic sent over the network may overload its capacity. In addition, this centralization can result in times when many requests are sent simultaneously, increasing the total analysis time due to the possible overload on the central equipment.

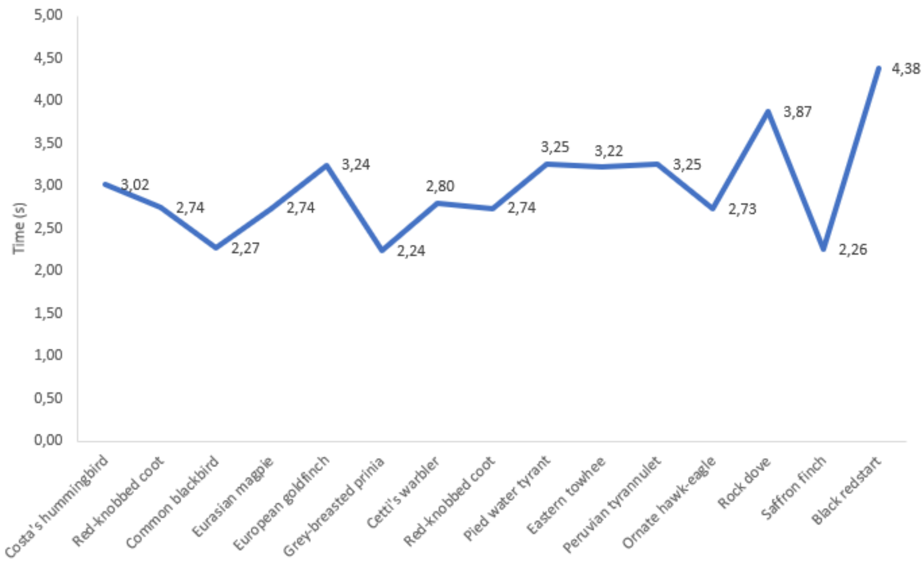

Figure 20. Chart of the average analysis times at the central node per bird species.

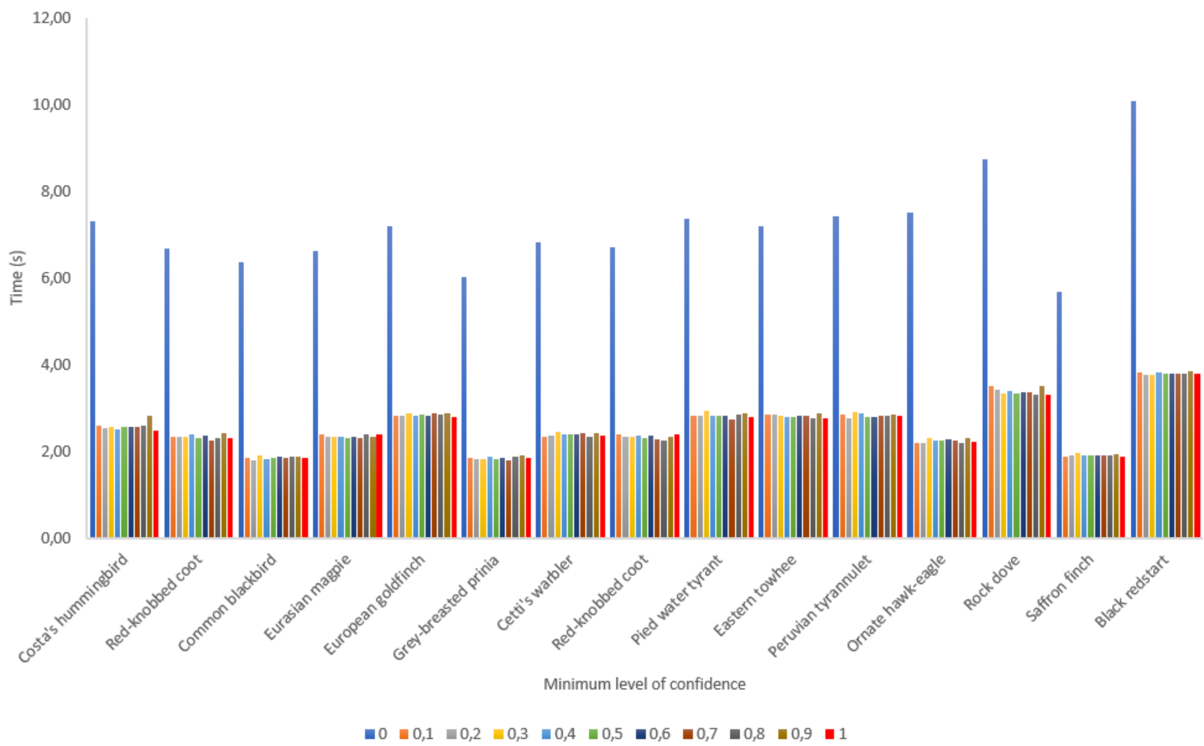

Figure 21. Chart comparing analysis times at the central node per bird species for different minimum levels of confidence.

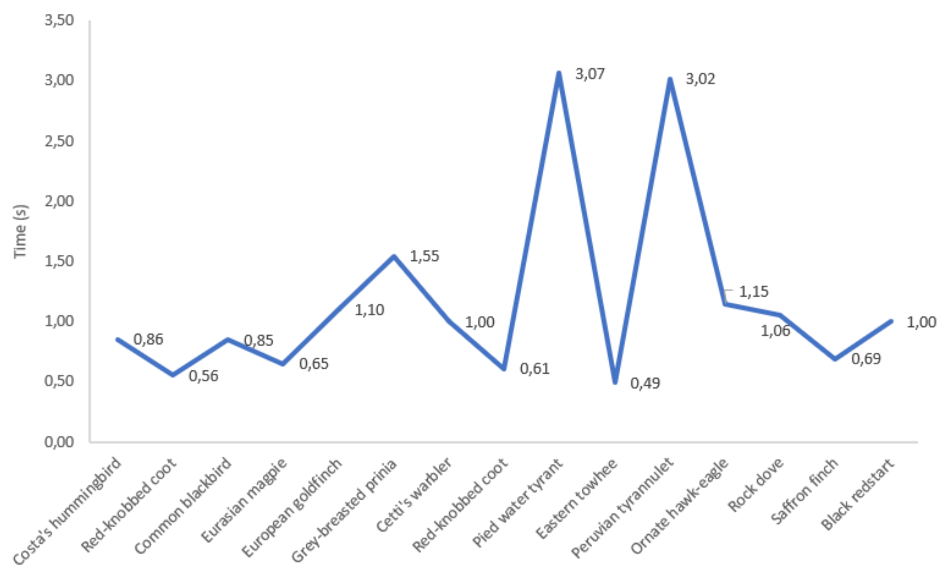

Figure 22. Chart of the average transmission times per bird species.

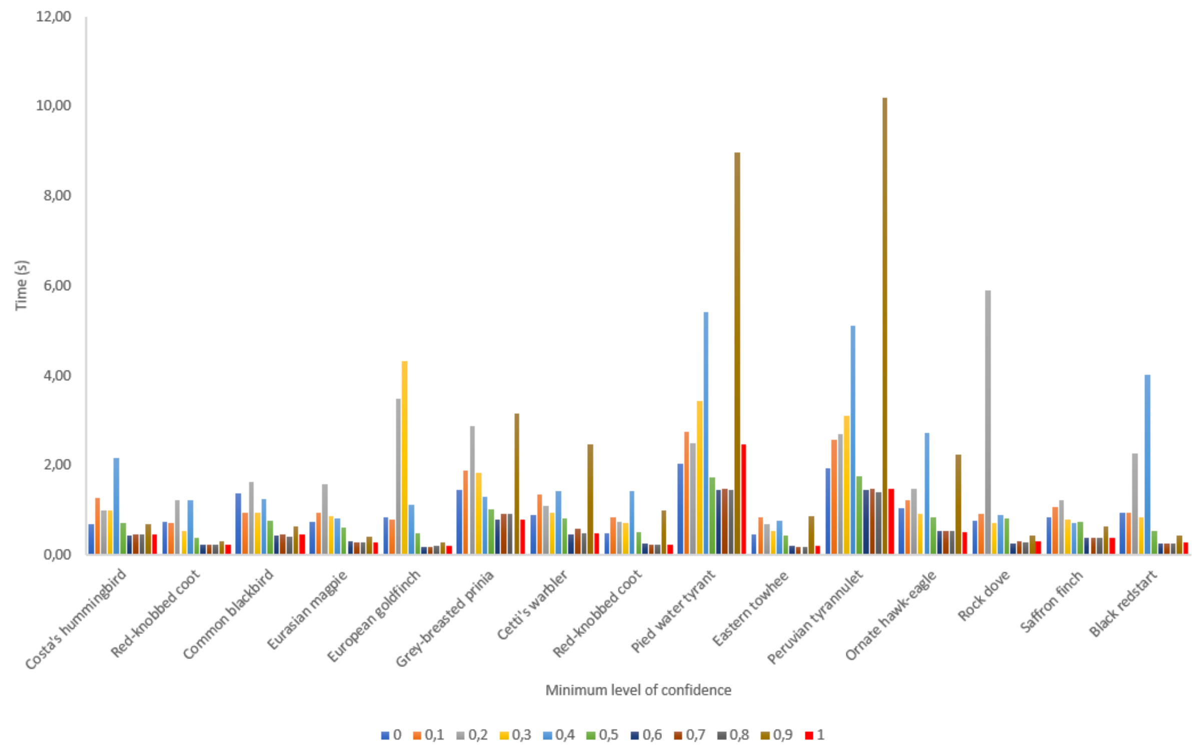

Figure 23. Chart comparing transmission times per bird species for different minimum levels of confidence.

Nonetheless, an advantage worth highlighting is the fact that a more substantial investment at the outset provides for easier and more efficient expansion. Endowing the system with enhanced hardware resources and processing power capability at the central node, will allow to handle an increase in the scale of the system without the immediate need for large additional investments in each distributed node. This can result in greater flexibility and scalability, allowing for smoother adaptation to the growing demands of the agricultural environment. This approach also ensures that any necessary improvements to the processing can be implemented efficiently and cheaply, without requiring the individual exchange of all nodes present in the solution.

After analyzing the results of the tests carried out on the two approaches, we concluded that the second approach is the most viable. This is because the first approach presented considerably high processing times. Therefore, the following tests will be conducted using the second approach.

4.2.4 Impact of Compression Type

Due to the limited resources of the intended solution, audio compression plays a crucial role. There are several types of audio compression, each with specific characteristics that can affect the quality of the sound, which can be reflected in the accuracy of the bird species analysis and processing time. Choosing the right type of compression is therefore essential to ensure the most efficient and reliable system performance.

To assess the impact of the type of audio compression on the analysis, we used two compression formats: WAV and MP3. The two were chosen because WAV is usually the format required by audio analysis tools, while MP3 is a very compact format that still retains some audio quality.

The chart displayed in Figure 24 shows that compressing files in WAV format is generally associated with a shorter analysis time. In some cases, this time difference is practically insignificant, while in others it is obvious. These discrepancies are possibly due to the MP3 format's tendency to selectively compress parts of the audio that are considered less perceptible to the human ear [19]. This process implies the need to decode these parts, adding time to the analysis procedure. In contrast, the WAV format, characterized by the preservation of the original audio quality, allows analysis to be carried out without prior procedures.

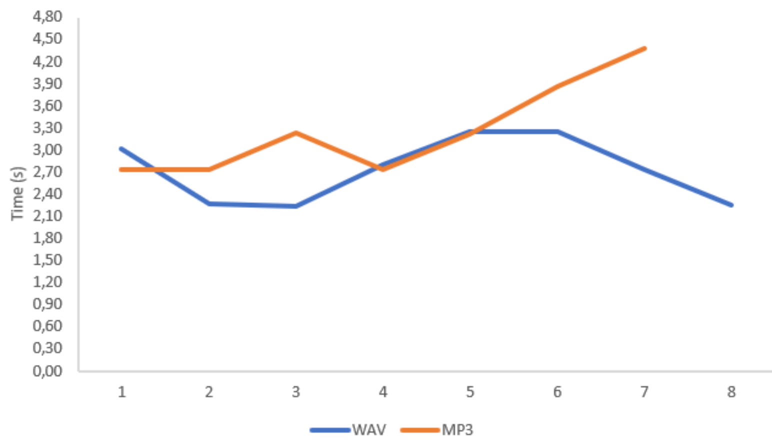

Figure 24. Chart comparing analysis times for WAV and MP3.

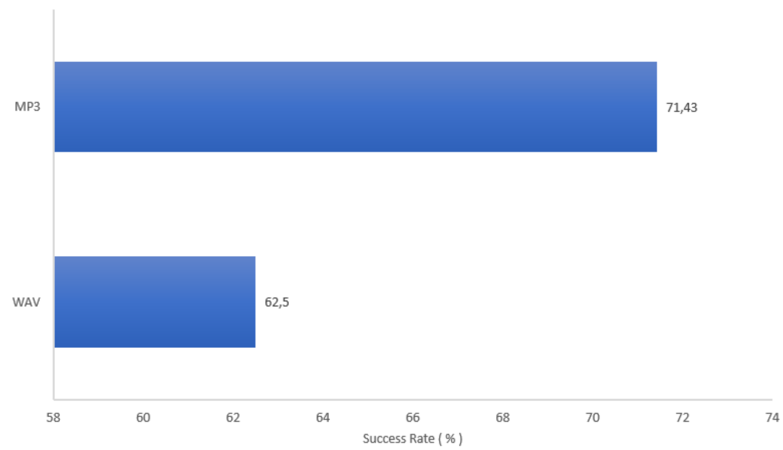

Figure 25. Chart comparing the overall success rate for WAV and MP3.

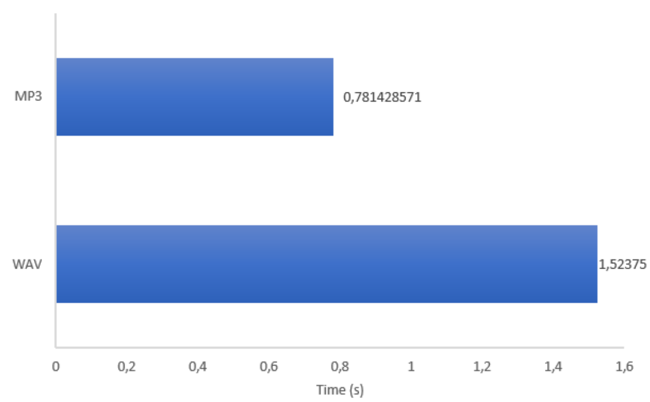

Figure 26. Chart comparing average transmission times per compression format.

In the tests carried out, the audio files with MP3 compression had a success rate of 71.43%, while the files in WAV format had success rate of 62.5%, as can be seen in Figure 25.

Although the analysis time was slightly longer for the MP3 files, this format showed greater accuracy in correctly identifying the species. This suggests that although MP3 adds a decompression step to the process, the quality of the resulting audio is still high enough to allow accurate bird species identification. In contrast, WAV files, despite being processed faster and having better quality, did not achieve the same success rate.

We can see in Figure 26 that the time required to transmit MP3 files is significantly lower, almost half that of WAV files. This result is because MP3 compressed files are much smaller in size than WAV files.

Analysis of the tests carried out shows that although WAV files offer slightly shorter analysis times, MP3 files offer a higher success rate and significantly faster sending times. These factors suggest that MP3 compression may be better choice for a solution that has bandwidth and storage restrictions, without compromising the accuracy of bird species identification.

4.2.5 Impact of Audio Quality

Implementing such a system in an orchard presents challenges regarding the quality of the audio captured due to atmospheric and environmental conditions that can vary significantly. Factors such as wind, rain, background noise from other agricultural activities and distance from sound sources can deteriorate the clarity and accuracy of recordings.

Given that our solution will be implemented in such a scenario, it is crucial to assess the impact of the audio quality on the BirdNET algorithm's effectiveness. Understanding how the system handles different quality levels will allow us to determine the minimum hardware requirements that will provide a reasonable level of performance performance, even under adverse conditions.

The chart displayed in Figure 27 allows a comparison both within the audio quality levels themselves and between the different audio quality levels (A being the highest and E the lowest). For level A, despite its high quality, there were two cases in which the bird species were not successfully detected. Nevertheless, when a successful detection occurs, the BirdNET algorithm's confidence is over 80%.

For level B, there is only one case of detection failure, while the successful cases show a high confidence rate, similar to level A. At level C, which is intermediate, all detections were successful. However, on average, the BirdNET algorithm's confidence is lower, although there are still cases with a higher confidence, which shows that even with some noise the algorithm can still detect the bird species with confidence.

At levels D and E, which represent the lowest audio quality, there is only one case of detection failure at level D, but the confidence rates are generally below 50%, indicating less certainty in the detections made.

In Figure 28 we can see that level C, with an average of 63.64%, has on average the highest success rate in detecting species, suggesting a good balance between audio quality and algorithm accuracy. Level A, despite its higher quality, only has a success rate of 52.73%, while level B has 45.45%. Levels D and E, due to their lower audio quality, have average success rates of 18.18% and 36.36% respectively. This indicates that despite the apparent greater difficulty in detection with lower audio quality, detections are still possible.

In summary, audio quality significantly affects the success rate, but despite the various levels, C proved to be the most efficient overall, with no detection failures and acceptable confidence rates. With all the data together, we conclude that the BirdNET algorithm can cope quite well with variations in audio quality and successfully detect the species most of the time, even in noisy conditions.

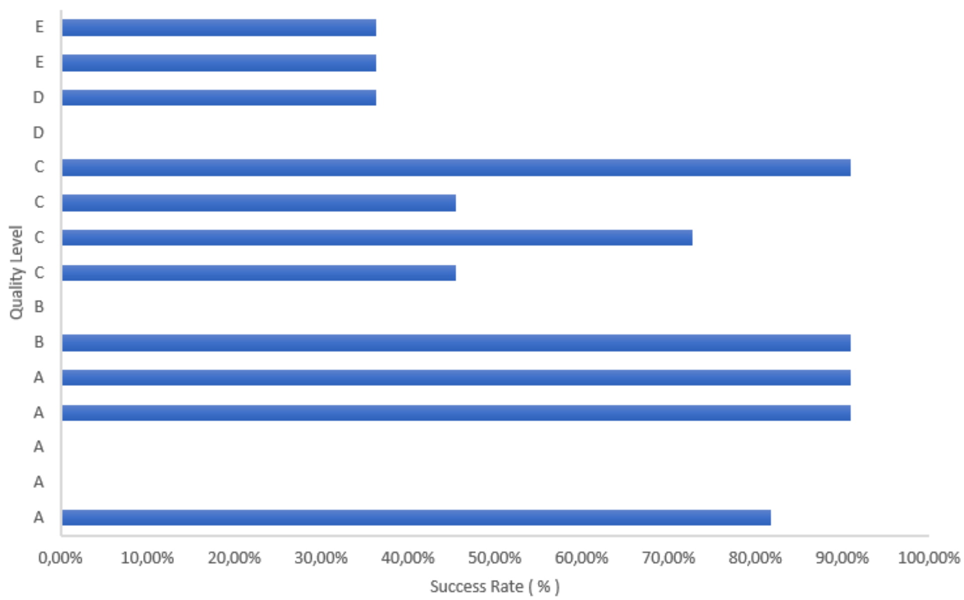

Figure 27. Success rates at different audio quality levels.

DOI: xxxxxxxxxxxxxxxxxxxxxxxxx

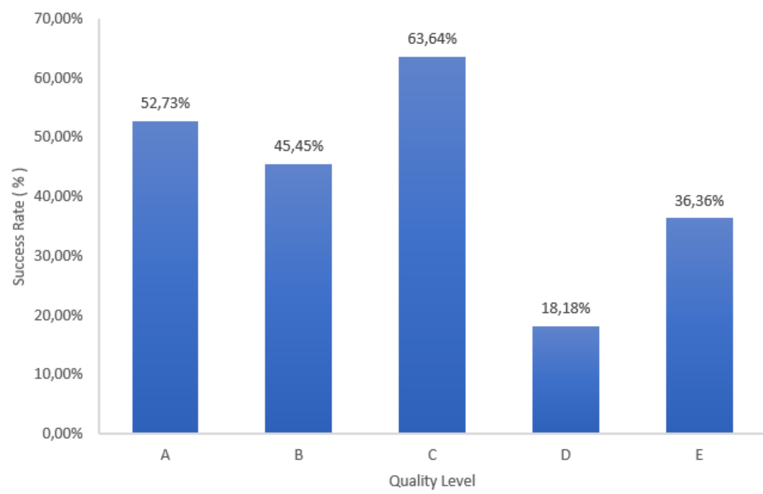

Figure 28. Chart of the average success rate for each quality level.

4.3 Critical Analysis

So far, tests have been conducted to evaluate response times and analyze the effects of various parameters on the bird detection process. The results of these tests have provided a clear view of the advantages and challenges associated with each proposed approach.

In the first approach the network is not overloaded because bird species identification occurs at the distributed nodes and only the result is sent to the central node. However, the analysis times are longer and the high costs in extensive orchards represent a significant disadvantage.

On the other hand, the second approach offers significantly shorter analysis times, as the birds sound data is processed at the central node. But there is the possibility of overloading the network with sound data collected at the distributed nodes or exceeding the processing capacity of the central node.

A hybrid option could be an interesting solution to combine the strengths of both approaches. For example, such an approach could involve using the central node as the main method, resorting to local analysis in critical situations. This would provide an efficient balance between fast response times and effective network management.

Regarding the disparities in the different categories of audio quality, categories lower than C significantly compromise the reliability of accurate bird species detection. However, the BirdNET algorithm proved capable of detecting birds even in conditions considered less than ideal.

Thus, the implementation of a hybrid approach, together with the consideration of audio quality limitations, could optimize the bird detection process, reconciling efficiency and accuracy.

5. CONCLUSION

The work presented in this paper focuses on using smart agriculture to address the problem of preventing bird-related losses in orchards. It represents the first stage of the development of a prototype for the detection and dispersal of birds in orchards. This initial stage focused on proposing and evaluating solutions for detecting the presence and classifying bird species based on sound acquisition.

Two main approaches for this issue were identified. In the first approach, nodes distributed across the orchard are responsible for detecting and identifying the bird species and sending the result to a central node. This minimizes the burden on the network but causes longer analysis times and high costs in extensive plantations.

In the second approach, nodes distributed across the orchard only capture the birds sound and send it to the central node that will process it to detect and identify the bird species. This results in faster analysis times, but with the risk of overloading the network resources or the central node hardware resources, which can compromise effectiveness under certain conditions.

Additional tests were conducted to analyze how different parameters can influence the results, such as the quality of the audio captured, and the type of audio compression used. It was shown that audio quality is a crucial factor in detection accuracy. Audio categories lower than C significantly compromise the reliability of the system.

The results presented in this paper will guide the prototype implementation in the right direction, which will occur in the next phase of this research project.

AUTHOR CONTRIBUTIONS

Conceptualization, D.C. and F.C.; methodology, D.C. and F.C.; validation, P.D.G., J.M.L.P.C. and V.N.G.J.S.; formal analysis, P.D.G., J.M.L.P.C. and V.N.G.J.S.; investigation, D.C. and F.C.; writing—original draft preparation, D.C. and F.C.; writing—review and editing, P.D.G., J.M.L.P.C. and V.N.G.J.S.; supervision, J.M.L.P.C. and V.N.G.J.S.; funding acquisition, P.D.G., J.M.L.P.C. and V.N.G.J.S. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] A. Anderson et al., "Bird damage to select fruit crops: The cost of damage and the benefits of control in five states,"

Crop Protection, vol. 52, 2013, doi: 10.1016/j.cropro.2013.05.019.

[2] J. Coleman and E. B. Spurr, “Farmer perceptions of bird damage and control in arable crops,” New Zealand Plant Protection, vol. 54, pp. 184–187, Aug. 2001, doi: 10.30843/nzpp.2001.54.3719.

[3] C. Sausse et al., “Contemporary challenges and opportunities for the management of bird damage at field crop establishment,” Crop Protection, vol. 148, 2021, doi: 10.1016/j.cropro.2021.105736.

[4] E. B. Micaelo, L. G. P. S. Lourenço, P. D. Gaspar, J. M. L. P. Caldeira, and V. N. G. J. Soares, “Bird Deterrent Solutions for Crop Protection: Approaches, Challenges, and Opportunities,” Agriculture, vol. 13, no. 4, p. 774, Mar. 2023, doi: 10.3390/agriculture13040774.

[5] R. Shete and S. Agrawal, “IoT based urban climate monitoring using Raspberry Pi,” in 2016 International Conference on Communication and Signal Processing (ICCSP), IEEE, Apr. 2016, pp. 2008–2012. doi: 10.1109/ICCSP.2016.7754526.

[6] A. K. Saha *et al.*, “IOT-based drone for improvement of crop quality in agricultural field,” in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, Jan. 2018, pp. 612–615. doi: 10.1109/CCWC.2018.8301662.

[7] L. G. P. S. Lourenço, E. B. Micaelo, P. D. Gaspar, J. M. L. P. Caldeira, and V. N. G. J. Soares, “Protótipo de solução de deteção e dispersão de aves para proteção de colheitas,” *Revista de Sistemas e Computação*, vol. 13, no. 3, pp. 34–42, 2023, doi: 10.36558/rsc.v13i3.8490.

[8] J. Höchst *et al.*, “Bird@Edge: Bird Species Recognition at the Edge,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2022. doi: 10.1007/978-3-031-17436-0_6.

[9] J. Höchst, “Bird@Edge Mic,” https://github.com/umr-ds/BirdEdge-Mic.

[10] P. McGuire, “BirdNET-PI.” Accessed: Feb. 26, 2024. [Online]. Available: https://github.com/mcguirepr89/BirdNET-Pi

[11] T. P. Hakala, “birdnet-go,” https://github.com/tphakala/birdnet-go.

[12] Cornell Lab of Ornithology, “Merlin Bird ID - Free, instant bird identification help and guide for thousands of birds - Identify the birds you see,” https://merlin.allaboutbirds.org.

[13] B. L. Sullivan, C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling, “eBird: A citizen-based bird observation network in the biological sciences,” *Biol Conserv*, vol. 142, no. 10, 2009, doi: 10.1016/j.biocon.2009.05.006.

[14] C. Pérez-Granados, “BirdNET: applications, performance, pitfalls and future opportunities,” *Ibis*, vol. 165, no. 3, 2023, doi: 10.1111/ibi.13193.

[15] S. Cass, “Nvidia makes it easy to embed AI: The Jetson nano packs a lot of machine-learning power into DIY projects - [Hands on],” *IEEE Spectr*, vol. 57, no. 7, pp. 14–16, Jul. 2020, doi: 10.1109/MSPEC.2020.9126102.

[16] Raspberry Pi (Trading) Ltd, “Raspberry Pi 4 Model B Datasheet,”https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-datasheet.pdf.

[17] Raspberry Pi (Trading) Ltd, “Raspberry Pi Zero 2 W.”

[18] S. Kahl, C. M. Wood, M. Eibl, and H. Klinck, “BirdNET: A deep learning solution for avian diversity monitoring,” Ecol Inform, vol. 61, Mar. 2021, doi: 10.1016/j.ecoinf.2021.101236.

[19] L. C. Gay, “MP3: The Meaning of a Format,” Ethnomusicology, vol. 58, no. 3, 2014, doi: 10.5406/ethnomusicology.58.3.0548.