University of Padova

Department of Mathematics

Master Thesis in Erasmus Mundus Joint-Master's Degree in Big Data Management & Analytics

Eco-Friendly Wind Energy: Improving Radar Tracking Systems for Bird Monitoring

Supervisor
Professor Michele Rossi
University of Padova

Co-supervisor
Professor Nuno Simões
Universidad Nacional Autónoma de México

Master Candidate
Diogo Seca Repas Gonçalves

Academic Year

2022-2023

To my dearest family and friends, who have provided the bedrock of emotional support, motivation, and understanding throughout this endeavor. Your faith in my goals has been a constant source of encouragement.

Abstract

The escalating global shift towards sustainable energy sources, particularly wind energy, has accentuated the need for effective measures to mitigate avian fatalities due to the operation of wind turbines. This thesis addresses this pressing ecological issue by enhancing bird tracking in radar systems - a case study on the Birdtrack® radar system, used in wind farms for avian monitoring and collision prevention. The research presents a novel intersection of technology and ecology, significantly contributing to sustainable energy and wildlife conservation efforts.

Central to this research is identifying the existing limitations of the currently adopted Birdtrack® system in differentiating birds from environmental clutter and accurately tracking multiple birds under dynamic conditions. The methodology employed integrates raw radar data preprocessing, visualization techniques, and the development of robust bird detection and tracking algorithms. Notably, incorporating the Observation-Centric SORT (OC-SORT) algorithm for multi-bird tracking and enhanced radar data processing techniques markedly improves the system detection and tracking precision and efficiency.

The effectiveness of the proposed technique is quantified evaluated through case studies in various wind farm environments. Performance metrics such as track velocity, angular trajectory changes, track duration, and quality assessment scores affirm the system superiority over Birdtrack®. The results demonstrate a significant advancement in tracking accuracy, reduced track fragmentation, and increased detection reliability.

The work in this thesis contributes to the improved operation of wind farm systems, while benefiting avian conservation. It addresses the challenges of radar-based bird monitoring and sets a foundation for future research. The successful enhancement of the Birdtrack® system represents a crucial step in mitigating the environmental impact of wind energy operations, thus aligning sustainable energy production with ecological preservation objectives.

Contents

Listing of figures

Listing of tables

Listing of acronyms

BDMA .... Big Data Management & Analytics

BIoU .... Bounding Box Intersection-over-Union

BSJ .... Barão de São João

BT .... Birdtrack®

CIoU .... Complete Intersection-over-Union

CNN .... Convolutional Neural Network

COCO .... Common Objects in Context

CTMC .... Cell Tracking with Mitosis Detection Challenge

DDPM .... Denoising Diffusion Probabilistic Models

DETR .... Detection Transformer

DIoU .... Distance Intersection-over-Union

DPM .... Deformable Parts Model

EMA .... Exponential Moving Average

EMJMD .... Erasmus Mundus Joint Master's Degree

GAN .... Generative Adversarial Networks

GIoU .... Generalized Intersection-over-Union

GPS .... Global Positioning System

GT .... Ground Truth

HDBSCAN .... Hierarchical Density-Based Spatial Clustering

HOG .... Histogram of Oriented Gradients

HOTA .... Higher Order Tracking Accuracy

HSV .... Hue-Saturation-Value

ID .... Identifier

ILSVRC .... ImageNet Large Scale Visual Recognition Challenge

IoU .... Intersection-over-Union

KITTI .... Karlsruhe Institute of Technology and Toyota Technological Institute

LAP .... Linear Assignment Problem

MAE .... Masked Auto-Encoder

MOTA .... Multiple Object Tracking Accuracy

MOTP .... Multiple Object Tracking Precision

MOT .... Multiple Object Tracking

OC-SORT ..... Observation-Centric Simple Online Real-time Tracking

OCM .... Observation-Centric Momentum

OCR .... Observation-Centric Recovery

ORU .... Observation-centric Re-Update

PCA .... Principal Component Analysis

PNG .... Portable Network Graphics

PNTI .... Tejo Internacional Nature Park

R-CNN .... Region-based Convolutional Neural Network

RADAR .... Radio Detection And Ranging

RF .... Radio Frequency

RGB .... Red-Green-Blue

RPN .... Region Proposal Network

SCADA .... Supervisory Control And Data Acquisition

SIFT .... Scale-Invariant Feature Transform

SORT .... Simple Online Real-time Tracking

SOT .... Single Object Tracking

SSD .... Single-Shot Detector

STEP .... Segmenting and Tracking Every Pixel

TSS .... Time Since Saturation

UI .... User Interface

UMAP .... Uniform Manifold Approximation and Projection

VOC .... Visual Object Classes

WAMI .... Wide Area Motion Imagery

XML .... Extensible Markup Language

YOLO .... You Only Look Once

1

Introduction

1.1 Background

The rapid expansion of wind energy as a sustainable power source has brought about unforeseen ecological challenges, particularly in relation to avian wildlife [1]. Among these, the collision of birds with wind turbines stands out as a significant concern. This interaction often results in fatalities, raising environmental and conservation issues. In this context, the implementation of Birdtrack®, a radar-based system designed to mitigate such fatalities, represents a critical intersection of technology and ecology.

Bird migration, a natural and critical process for many bird species, involves regular movements between breeding and wintering grounds. Wind farms, particularly those placed along migratory routes, pose collision risks for these birds, often leading to increased mortality. Understanding these migration patterns and their intersection with wind turbine placements underlines the importance of systems like Birdtrack® in reducing such risks [2].

The application of radar technology in bird conservation within wind energy projects is primarily centered on the detection and monitoring of avian movements. Advanced radar systems, such as Birdtrack®, are equipped to identify and track the flight paths of birds in real time. By continuously scanning the airspace around wind farms, these systems provide critical data on bird presence and movement patterns. This information is crucial for assessing collision risks, particularly during peak migration periods. The radar's ability to detect birds from a considerable distance allows for early warning signals to be generated, enabling timely intervention, such as turbine shutdowns or speed adjustments, to prevent collisions.

Furthermore, the integration of radar technology with ecological data analysis enhances our understanding of bird behavior in relation to wind turbines. By correlating radar-detected movements with specific bird species and environmental conditions, it becomes possible to identify high-risk scenarios and adjust turbine operations accordingly. This not only mitigates the immediate threat to avian life, but also contributes to long-term conservation strategies. It is essential, however, to consider the limitations of radar technology, such as its sensitivity to weather conditions and the challenge of distinguishing between bird species. Research and development are ongoing to refine the accuracy and effectiveness of these systems in diverse environmental settings.

Birdtrack® works by detecting avian entities in the vicinity of wind turbines and subsequently triggering shutdown protocols to prevent collisions. The system employs advanced radar technology, which utilizes electromagnetic waves to sense and track birds. This technology, however, is not without its limitations. The primary challenge lies in its ability to distinguish between actual bird targets and environmental clutter, i.e., radar echoes resulting from non-target objects such as rain, fog, or stationary structures within the radar range.

Radar systems, pivotal in Birdtrack®, use radio waves for the detection and location of objects. These systems transmit electromagnetic waves and analyze the echoes returned from the targets. In the context of bird detection, the effectiveness of radar hinges on the transmitted signal properties, including frequency, wavelength, and pulse duration, which directly influence the system resolution, range, and accuracy.

The sophistication of radar data preprocessing and multiple object tracking algorithms is central to the performance of Birdtrack®. Effective preprocessing is required to refine raw radar data, enhancing the visibility and the separation of real targets from background noise and clutter. In addition, robust algorithms for multiple object tracking are key to maintaining consistent identification and tracking of individual birds over time.

The existing Birdtrack® system, although functional, has its limitations. While the Birdtrack® system attempts to refine raw radar data, the preprocessing methods currently in place fail to adequately enhance the visibility of avian targets amidst background noise and clutter. Additionally, the employed bird detection algorithm is relatively rudimentary and it is heavily dependent on manual thresholds and heuristic approaches. This reliance results in significant challenges, particularly in accurately distinguishing birds from environmental clutter such as atmospheric disturbances or other dynamic clutter sources. Consequently, this leads to inaccuracies in identifying avian targets. Furthermore, the system struggles with maintaining consistent tracking of birds, especially in environments characterized by dynamic and cluttered conditions or intermittent signals. This issue often leads to fragmented tracking, wherein the paths of birds are not followed continuously and coherently. Such fragmentation in tracking undermines the system effectiveness, making it challenging to reliably monitor avian movements over time.

Improving the accuracy and efficiency of Birdtrack® in bird detection and tracking is pivotal. This improvement not only contributes to reducing bird fatalities but also aids in maintaining the operational efficiency of wind turbines by minimizing unnecessary shutdowns. The balance between sustainable energy production and wildlife conservation is thus a critical consideration in developing and optimizing this technology.

This thesis aims to address these challenges by exploring innovative approaches in radar data processing and state-of-the-art tracking algorithms. The expected outcomes include an improvement in the reliability of bird detection and tracking, thereby facilitating environmental studies, and the definition of effective mitigation strategies for avian protection around wind turbines.

1.2 Challenges in Radar Technology for Bird Monitoring

The integration of radar technology into bird monitoring at wind farms, exemplified by Birdtrack®, presents unique challenges that are pivotal to the efficacy of the system. These challenges encompass the complexity of radar signal processing, the intricacies of bird detection algorithms, and the need of tracking multiple targets within a dynamic environment. This section elaborates on these challenges, underlining their significance in the broader context of avian monitoring and protection.

1.2.1 Complexity of Radar Signal Processing

Radar signal processing is inherently complex due to the variable nature of environmental factors and the characteristics of radar signals. The primary difficulty lies in the accurate interpretation of radar echoes [3], which are reflections of electromagnetic waves from objects in the radar's field of view. In the context of Birdtrack®, distinguishing bird echoes from clutter, which includes reflections from stationary objects and atmospheric phenomena, is a significant challenge.

Clutter, which is inherently present in radar data, varies in form. From static objects such as buildings or terrain features to dynamic elements like rain, fog, waves, or even insect clouds. The key challenge is to develop preprocessing algorithms that can effectively identify and filter out such clutter, thereby enhancing the detection of avian targets. This requires sophisticated signal processing techniques that can adapt to changing environmental conditions.

The processing of radar data also involves dealing with issues of signal attenuation, especially in adverse weather conditions, and the varying range resolution, which affects the system ability to detect birds at different distances. These factors prompt the need for the development of algorithms that are capable of maintaining high detection and tracking performance across a variety of operational conditions.

1.2.2 Limitations of Current Bird Detection Algorithms

The detection of birds within radar data is a critical step in the functioning of Birdtrack®. Current detection algorithms, while operational, are limited in their ability to accurately identify birds amidst clutter. These algorithms often rely on predefined thresholds and heuristics [4], which can lead to either missed detections (false negatives) or incorrect identifications (false positives). The challenge is to develop a detection algorithm that can reliably differentiate birds from clutter, by at the same time being robust to non-avian objects or environmental noise.

Furthermore, the variability in bird sizes, flight patterns, and formations adds an additional layer of complexity to the detection process. Birds flying at different altitudes, and speeds, and in varying group sizes pose distinct challenges for radar detection. The algorithm needs to be versatile enough to detect solitary birds as well as large flocks, and capable of adapting to different flight behaviors.

1.2.3 Challenges in Tracking Multiple Birds Simultaneously

Multiple object tracking (MOT) is an essential aspect of Birdtrack®, where the system must accurately track the movements of multiple birds simultaneously. This is particularly challenging in cluttered environments where birds may be flying amidst other moving objects or in varying weather conditions that affect radar visibility. The system needs to distinguish between individual birds, track their paths over time, and handle instances where birds may cross paths or fly close to one another.

Additionally, the radar system must be able to maintain continuity in tracking even when birds move in and out of the radar's range or when the radar signal is momentarily lost due to environmental factors. This requires sophisticated tracking algorithms that can predict/estimate the trajectory of birds during occlusion events and readjust the estimated trajectory as the signal becomes reliable again.

Addressing these challenges is essential for the effective functioning of Birdtrack®. Improved radar signal processing techniques, enhanced bird detection algorithms, and robust multiple bird tracking capabilities are key to accurately monitor and protect avian species in the vicinity of wind farms. This thesis work aims at contributing towards these improvements, offering solutions for an effective monitoring of birds for the operation of wind farms. Such algorithms should strike a balance between the competing objectives of maximizing the operational demands of wind energy production (i.e., maximizing the active time for the wind turbines) and providing bird protection (i.e., momentarily shutting off the wind turbines to prevent collisions).

1.3 Research Objectives and Methodology

The overarching aim of this research is to enhance the efficacy of Birdtrack® by refining its radar data preprocessing and bird detection and tracking algorithms. These goals are decomposed into several specific objectives and corresponding methodologies, as outlined here below.

1.3.1 Objectives

1. **Enhancement of Radar Data Preprocessing:** To develop techniques for processing radar data, resulting in improved clarity of the corresponding images and in a better distinction between avian targets and clutter. This includes reducing noise and environmental interference, thereby facilitating more accurate bird detection and tracking.

2. **Advanced Visualization Techniques:** To implement visualization methods that transform complex radar data into more interpretable forms. This involves converting data from polar to Cartesian coordinates and employing temporal aggregation techniques, enhancing both real-time and the offline analysis of avian activities.

3. **Refinement of Bird Detection Algorithms:** To improve the accuracy and efficiency of bird detection by using sophisticated algorithms capable of distinguishing birds from clutter. This includes addressing the limitations of current heuristics-based detection methods and exploring the use of machine learning approaches.

4. **Enhancement of Multiple Bird Tracking Algorithms:** To develop and integrate advanced algorithms for the simultaneous tracking of multiple birds, with a focus on accuracy, effectiveness, and robustness, especially in cluttered and dynamic environments.

5. **Evaluation and Validation:** To rigorously test and validate the developed techniques and algorithms using real-world data sets, assessing their effectiveness and reliability in practical scenarios.

1.3.2 Methodology

Radar Data Preprocessing

The preprocessing of radar data constitutes a fundamental step in enhancing the overall system performance. This process begins with the transformation, refinement, and filtering of raw radar signals to extract meaningful information while minimizing the impact of noise and clutter. Advanced signal processing techniques will be applied to address the challenges posed by varying environmental conditions and the inherent complexities of radar data. This includes the implementation of adaptive filtering methods, clutter rejection algorithms, and techniques for signal enhancement. The goal is to achieve a clear and reliable representation of radar data, laying the groundwork for effective bird detection and tracking.

Visualization Enhancements

The transformation of complex radar data into an interpretable and user-friendly format is crucial for effective analysis and decision-making. This involves converting radar data from its native polar coordinate system into Cartesian coordinates, which are more intuitive for visual interpretation. Additionally, advanced visualization techniques such as temporal aggregation will be employed to highlight key features and patterns within the data. These enhancements aim at facilitating a more comprehensive real-time understanding of bird movements and behaviors, thereby aiding in more accurate and timely decision-making in bird protection measures.

Bird Detection Algorithm Development

In the development of the Birdtrack® system, a key focus has been placed on enhancing the bird detection algorithm. The methodology adopted for this purpose involves a high-level approach that simplifies radar imagery to facilitate more accurate bird detection. Initially, the radar frames are processed to enhance the clarity and connectivity of the image components, a crucial step in preparing the data for further analysis. This processing is designed to enhance potential bird targets, while reducing background noise and irrelevant data, thereby improving the overall detection accuracy. The visualization enhancements introduced above form a solid foundation for future training of image segmentation or object detection models, which is left for future work.

Multiple Object Tracking (MOT)

The implementation of a robust Multiple Object Tracking (MOT) algorithm is key to monitoring multiple birds simultaneously and accurately. This involves adapting and optimizing state-of-the-art tracking algorithms to meet the specific requirements of bird tracking in a wind farm environment. The proposed MOT system is designed to handle the complexities of tracking multiple moving targets, including occlusion management, varying speeds, and erratic flight patterns. The algorithm should also strive for reduced track fragmentation, ensuring that bird tracks are accurately followed over time, even in the presence of signal interruptions or environmental disturbances.

Evaluation and Validation

The effectiveness of the developed techniques and algorithms is quantitatively evaluated using a comprehensive set of metrics. This includes comparing the performance of the proposed system with the existing Birdtrack® system, focusing on parameters such as track velocity, angular changes in trajectory, and overall track quality. The evaluation process involves analyzing data from various case studies, each presenting unique environmental and operational conditions.

1. *Quantitative Analysis:* Metrics like mean and standard deviation of track velocity, angular change in trajectories, and track duration are used to quantitatively assess the tracking performance. Advanced metrics, such as the Markov chain-based track quality assessment, are employed to provide a deeper understanding of the tracking performance. These metrics consider factors like detection probability, measurement-to-track association probability, and clutter probability to attain a more comprehensive quality score for each track. This provides insights into the precision and stability of the proposed tracking algorithm.

2. *Qualitative Assessment:* Visual inspections of tracking results is as well conducted to qualitatively evaluate the attained performance. This includes analyzing the fragmenta-

tion of tracks, the consistency of track quality, and the ability of the system to maintain continuous tracking of birds over extended time periods. Comparisons of track visualizations between the proposed system and the existing Birdtrack® system is used to highlight improvements in tracking robustness and accuracy.

Through this multi-faceted evaluation approach, the work in this thesis aims to demonstrate the advancements made in radar-based avian monitoring systems, offering tangible improvements in bird detection and tracking, and contributing to the broader goals of wildlife protection in the sector of renewable energy production via wind farms.

Our evaluation serves not only to validate the enhancements made, but also to identify areas for future research and development, paving the way for continued advancements in the field.

The proposed methodological framework is designed to address the specific challenges identified in Birdtrack® and aims to significantly advance the state of avian monitoring technology in wind farm environments. The successful completion of the identified objectives will contribute to more effective and environmentally friendly wind energy production, balancing the need for sustainable energy with the protection of avian species.

1.4 Thesis Structure

This thesis is structured to systematically address the challenges and potential solutions in using radar technology for bird monitoring in the context of wind energy projects. Each chapter builds upon the previous ones, in a progressive manner, thus creating a cohesive and comprehensive exploration of the subject.

Chapter 2 delves into the specifics of the Birdtrack® system, explaining its framework and functionalities. This chapter highlights the source data used by the system and discusses its limitations, underlining the need for improved algorithms.

Chapter 3 provides a comprehensive review of the existing literature relevant to radar technology in bird detection and tracking. This chapter aims to position the current research within the broader academic and practical context, highlighting key developments and identifying gaps in existing methodologies.

Chapter 4, presents the novel approaches and solutions proposed in this thesis. It details the advancements in radar data preprocessing and the development of innovative algorithms for multiple bird tracking. This chapter describes the core contributions of the thesis to the field.

Chapter 5, discusses the datasets used and evaluates the performance of the proposed solutions. It provides a critical analysis of the results, demonstrating the effectiveness of the new methodologies in practical scenarios.

Finally, Chapter 6, summarizes the findings of the thesis and discusses their implications. It offers final remarks and outlines directions for future research. This chapter aims to provide a conclusive summary of the work done, and of its contributions to bird conservation in wind energy contexts.

In summary, this thesis aims to contribute to the field of bird monitoring and protection within wind energy projects, aiding the potential of advanced radar processing techniques and algorithms in mitigating the ecological impacts of wind energy operations.

2

The Birdtrack® System

Birdtrack® is a radar technology designed to mitigate bird fatalities at wind farms by selectively shutting down turbines in response to bird migration or the presence of endangered species. This technology was developed by STRIX, an environmental, social, and sustainability consultancy company focused on providing services globally to leading organizations and assisting them in addressing and managing sustainability challenges. To date, Birdtrack® is the system used in production, being currently deployed in several installations around the world: 4 in Portugal, 2 in Egypt, 3 in the United Kingdom, 3 in Israel, and 3 in Chile. This chapter describes the system's architecture, functionalities, and limitations, and motivates the need for improved algorithms for the detection and tracking of multiple flying objects.

2.1 Birdtrack® System Framework

Birdtrack® is the world's first shutdown-on-demand technology for wind turbines, significantly reducing bird fatalities. It operates both onshore and offshore, with a sophisticated radar system capable of early collision detection, automatic operation, tracking, and classification of bird targets. However, some challenges remain in accurately identifying static and dynamic clutter sources.

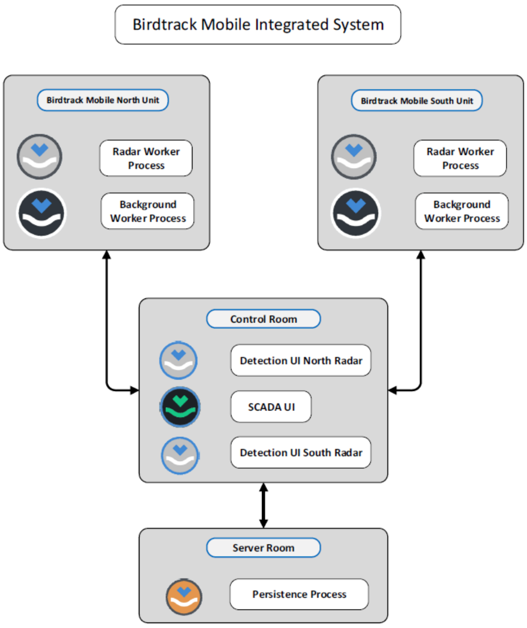


Figure 2.1: Diagram of the distributed framework of Birdtrack®.

The Birdtrack® system is based on a distributed architecture, accommodating diverse operational scenarios, as shown in Fig. 2.1. The following sections present its key software components.

- **Front-end Processes: Detection UI & Player UI** - The Detection UI is responsible for presenting real-time radar data, whereas the Player UI allows for the replay of recorded data for offline analysis. The data is presented to the user as shown in Fig. 2.2.
- **Backend Processes: Radar Worker Process** - This process is tasked with acquiring and processing radar data through specific algorithms, turning raw scans into data that can be easily interpreted.
- **Persistence Process** - Acting as an interface for the storage database, this process organizes and stores data including bird tracks, alarms, and user actions, thus ensuring systematic logging and retrieval.
- **Background Worker Process** - Activated after session closure, this process is responsible for finalizing the associations between bird tags and tracks and filtering the stored data to retain only pertinent information.
- **SCADA UI** - This user interface centralizes alarms from various radars and provides functionalities for actions such as simulating wind turbine power toggles.

- **Operational Configurations** - Birdtrack® can function in two primary configurations: the Standalone Version and the Integrated Configuration.

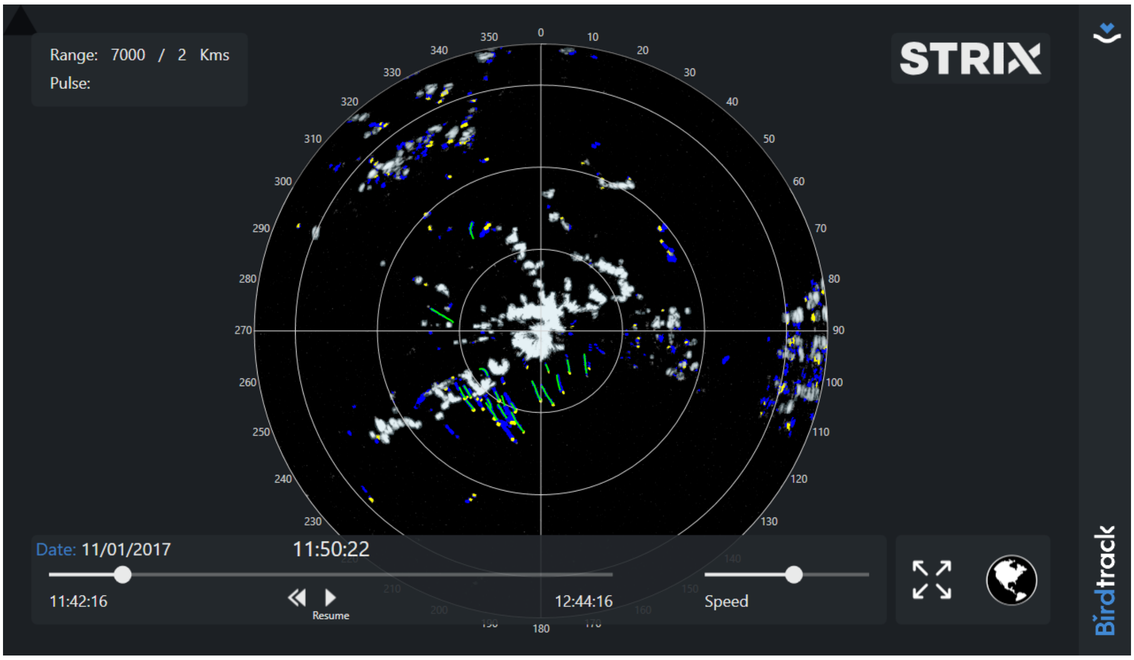


Figure 2.2: Player UI - Graphical User Interface

2.2 Understanding Birdtrack®'s Primary Data

Birdtrack® radar system yields a steady influx of data, providing insights into the dynamics of avian activities. In this section, the integral components of Birdtrack® radar data are described, focusing on the raw horizontal radar images and on the supplementary metadata.

Raw Radar Data: A Closer Look

The Birdtrack® system primarily generates raw radar data, represented as grayscale images. These images contain everything the radar detects: from large bird formations to individual birds, including non-target entities (noise), which will be referred to as clutter. This clutter can be either static or dynamic, making its detection more difficult. The grayscale format is indicative of the intensity of radio frequency (RF) reflections [3], as illustrated in Fig. 2.3. Each radar scanning revolution is digitally stored as a PNG file, and these files are typically large due to the vast area scanned. Over time, the accumulated data amounts to hundreds of gigabytes to terabytes of raw imagery per session. Birdtrack® radars operate at 2.5 seconds per revolution, maintaining a consistent frequency.

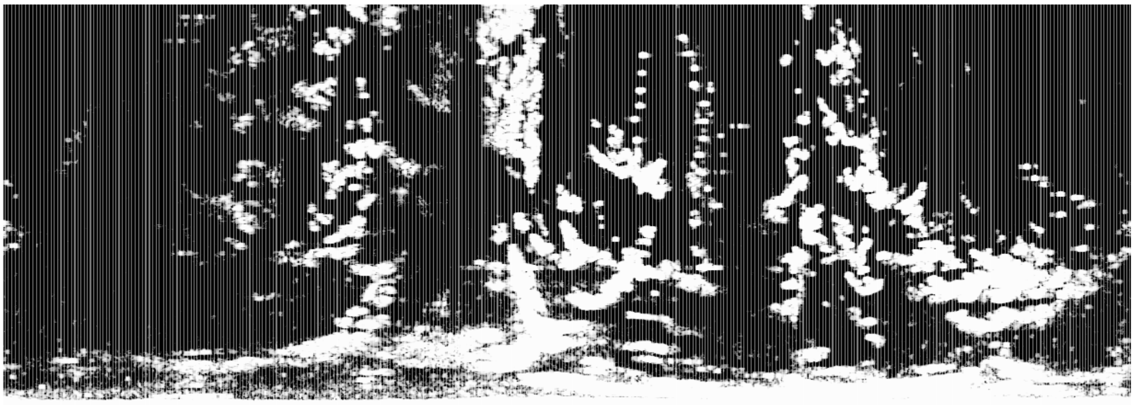


Figure 2.3: Raw radar line data from the BSJ project acquired on October 16th 2022 at 10:30:18. The image was rotated to improve its presentation layout, distance from the radar increases from bottom to top, angle sweeps $360°$ from left to right in the direction specified in the metadata field "ClockwiseImageRotation", and an offset defined by "WestImageHeading". The white vertical lines correspond to missing data.

Metadata: Additional Information

In tandem with the raw radar data, metadata serves as an informative companion, enriching the context of the radar images. For each radar image captured, there exists a corresponding XML file, methodically storing essential details like operational parameters, timestamps, and geographical locations. Some pivotal fields within this metadata include LocationLat, LocationLon, TimeStamp, RadarMLines, and AcquisitionParameters. A detailed description of each field in the metadata can be found in Tab. 2.1.

For the purposes of this thesis, the main metadata used relates to timestamp extraction, missing line preprocessing, and radar range, used in converting pixel distances to real-world distances for analysis.

Field	Description
LocationLat	Latitude of the radar's location in decimal degrees.
LocationLon	Longitude of the radar's location in decimal degrees.
OrientationMode	Mode of orientation of the radar. 0 for horizontal, 1 for vertical.
ClockwiseImageRotation	Indicates whether the image rotation is clockwise. 1 indicates clockwise rotation.
WestImageHeading	Heading of the image towards the west, measured in decimal degrees, ranging from 0 to 359.
ProjectId	ID of the project under which the radar data is being collected.
RadarName	Name of the radar.
TimeStamp	Timestamp of the data collection in milliseconds since the Unix epoch.
Time	Local time of data collection in the format "YYYYMMDD_HHhMMmSSsMSms".
RangeMeters	Range of the radar in meters.
itsDataSizeWidth	Width of the radar data image in pixels.
itsDataSizeHeight	Height of the radar data image in pixels.
DataType	Type of data collected.
RadarLinesFileName	Filename of the PNG image file that stores the radar line data.
GPSEnable	Indicates whether GPS is enabled. 1 indicates that GPS is enabled.
RadarMLines	Indices of all rows in the raw frame that do not have valid data and should be inpainted or ignored.
AcquisitionParameters	Various parameters related to the acquisition of the radar data.

Table 2.1: Description of XML Metadata Fields

2.3 Limitations and Need for Improved Algorithms

Despite its innovative capabilities, the Birdtrack® system presents some limitations. These not only hinder the system performance but also underline the critical need for advancements in detection and tracking algorithms. This section delineates these limitations and the consequent necessity for their rectification.

Challenges Presented by Clutter

The raw radar data is significantly cluttered. This clutter consists of both static and dynamic elements. Static clutter, such as geomorphological features and immobile objects, and dynamic clutter, including meteorological phenomena like rainfall and fog, pose substantial difficulties in the accurate detection and tracking of avian species.

This data, intrinsic to the operations of Birdtrack®, should be partially de-cluttered before any further processing and visualization. Static clutter, due to its static nature, is easier to deal with. The main challenge arises with dynamic clutter, which is usually amorphous and dynamically changes over time.

Data Preprocessing Complications

The system confronts considerable challenges in data preprocessing. The vast volume and variety of raw radar data necessitates preprocessing techniques for refining the data to ensure its clarity and reliability. This preprocessing is vital for both real-time interpretation and subsequent analytical processes.

Visualization and Interpretation Obstacles

The radar's raw data output is not readily interpretable, requiring substantial preprocessing and visualization enhancements for more intuitive real-time analysis. Additionally, developing methods for aggregating temporal information is imperative to distinguish between moving targets and static clutter in individual radar frames.

Limitations of the Bird Detection Algorithm

The current bird detection algorithm, albeit functional, is elementary and largely dependent on manual thresholds and heuristics. This algorithm should be revised [5] to improve its efficacy in accurately detecting avian targets amidst clutter.

Complexities in Multiple Bird Tracking

Tracking multiple birds simultaneously, particularly in a dynamic and cluttered environment, presents significant challenges [6]. The system must reliably differentiate between birds and other entities, maintain as much as possible tracking continuity during signal losses or occlusions, and managing the complexities inherent in tracking multiple birds with varying motion patterns.

Advancing Tracking Algorithms

The incorporation of advanced tracking algorithms, such as the Observation-Centric SORT (OC-SORT), is pivotal for enhancing the system performance. These sophisticated algorithms are crucial for overcoming the existing limitations in tracking accuracy, particularly in scenarios characterized by occlusions and non-linear movements of targets.

2.4 Final Remarks

At the early stages of this thesis project, the company identified some of the existing limitations of the Birdtrack® software, which motivated the approach of this thesis. This initial understanding has been significantly developed throughout this work into the complete assessment that is presented in this chapter. It can thus be claimed that one of the contributions of this thesis is the assessment of the Birdtrack® capabilities and the identification of its limitations.

In summary, the Birdtrack® system, while being innovative, has several limitations. Such limitations include addressing radar data clutter, data preprocessing complexities, challenges in visualization and interpretation, rudimentary bird detection and tracking algorithms, and the difficulty of tracking multiple birds. Addressing them is key to enhance the system detection and tracking capabilities, thereby fulfilling its primary design objectives. This necessitates a concerted effort towards the development and integration of more sophisticated algorithms and processing techniques, promising significant advancements in the field of avian monitoring and protection.

3
Literature Review

The advent and evolution of radar technology have significantly influenced various fields, particularly in environmental monitoring and wind energy. This literature review delves into the development of radar technology, its application in avian monitoring, and the challenges and advancements it presents, especially in the context of wind energy.

3.1 Fundamentals of Radar Technology and Applications in Bird Conservation

Radar technology plays a significant role in various applications within the field of wind energy. It provides valuable insights into wind flow patterns, assesses the impact of wind turbines on weather radar services, and monitors bird movements to mitigate collision risks. These applications are supported by the fundamental principles and techniques of radar technology.

Basics of radar technology

Radar, which stands for "**RA**dio **D**etection **A**nd **R**anging," is based on the emission and detection of radio waves to gather information about objects in the environment. Radar systems consist of antennas and transmitters that emit electromagnetic waves in the form of radio frequency (RF) signals. These signals propagate through the environment and interact with objects, leading to the reflection of some energy back to the radar system. The time it takes for the reflected waves to return can be used to determine the distance to the objects, while the Doppler effect can be used to measure their velocity [3].

The emitted radio waves are received by the radar's antennas, and the reflected waves are processed to extract valuable information about the objects. This process involves advanced signal processing techniques such as pulse compression and Doppler processing, which enhance range resolution and enable velocity measurements [7]. Additionally, radar signals are processed to mitigate noise, clutter, and interference, improving the overall detection and tracking capabilities of the radar system [8]. The principles and techniques of radar technology enable its applications across various fields, including wind energy.

Radar Systems in Avian Monitoring

Radar technology has proven valuable in the early detection, tracking, and analysis of bird behavior in proximity to wind turbines. By continuous monitoring of bird movements, radar systems provide real-time data on bird density, flight altitude, and behavior.

One study conducted in the Netherlands utilized C-band Doppler radar to investigate high-altitude bird migration during spring migration [1]. The researchers found that migration layers of birds occurred predominantly during the presence of a high-pressure system. This information can be crucial for wind farm operators to understand when and where bird migration is likely to occur, allowing for the implementation of measures to minimize the risk of bird collisions with wind turbines [1].

Radar systems can also track the flight paths of individual birds, providing insights into their behavior and response to wind turbines. By analyzing radar data, researchers can determine whether birds actively avoid wind farms or alter their flight paths to navigate around them. This information can inform the design and placement of wind turbines to minimize potential impacts on bird populations.

Furthermore, radar technology can be used to estimate bird density and assess the effectiveness of mitigation measures. By comparing bird activity levels before and after the implementation of mitigation strategies such as the use of deterrent devices or changes in turbine operation, researchers can evaluate the effectiveness of these measures in reducing bird collisions.

The analysis of radar data also contributes to the understanding of bird behavior and migration patterns. By studying the flight altitudes and speeds of birds, researchers can gain insights into their preferred flight corridors and stopover locations. This information can be used to identify important bird habitats and inform conservation efforts.

In wind energy applications, radar has been used to analyze wind flow patterns and assess the impact of wind turbines on weather radar services. Dual-Doppler synthesis with two Doppler radars has been employed to analyze three-dimensional wind flow within and surrounding wind farm wakes [9]. Radar data have also been validated with operational turbine data to project composite power output for multiple turbines, demonstrating the potential of radar technologies in wake modeling, wind farm design, resource assessment, and real-time wind mapping [9].

Radar technology has also been utilized to mitigate avian fatalities due to wind turbines. Selective turbine-stopping programs based on radar observations have been effective in reducing bird mortality rates [10]. Machine learning algorithms and automated target classification in radar ornithology show promise for real-time collision prevention [2]. Continuous monitoring of bird movements using radar technology provides valuable data for estimating collision rates and informing mitigation efforts [11].

Birdtrack® Radar System

The Birdtrack® Radar System is a specialized radar system designed specifically for bird detection near wind turbines. It aims to mitigate the potential negative impacts of wind farms on bird populations. The system utilizes radar technology to detect and track birds in the vicinity of wind turbines, providing valuable information for monitoring and research purposes.

One key advantage of the Birdtrack® Radar System is its ability to provide real-time data on bird movements and behavior near wind turbines. This information can be used to assess the risk of bird collisions with turbine blades and inform mitigation strategies [4]. The radar system can detect and track birds in close proximity to the turbines, allowing for the estimation of flock size, flight altitude, and avoidance behavior [4]. This data can help understand how birds interact with wind turbines and inform the design and placement of turbines to minimize potential impacts on bird populations.

In addition to radar technology, other remote sensing techniques can be employed in conjunction with the Birdtrack® Radar System to gather more comprehensive data on bird flight and behavior. These techniques include the use of video surveillance equipment, microphone systems for detecting bird sounds, laser range finders, ceilometers, and pressure sensors [4]. By combining multiple remote sensing techniques, researchers can obtain a more detailed understanding of bird movements and responses to wind turbines.

The Birdtrack® Radar System and associated remote sensing techniques have been used in various studies to assess the effects of wind farms on bird populations. Research conducted in the Strait of Gibraltar found that wind farms did not appear to be more detrimental to birds than other man-made structures [12]. This suggests that wind farms, when properly designed and located, may not pose significant risks to bird populations.

3.2 Object Detection

Object detection, a crucial component in the field of computer vision, has undergone significant advancements due to the integration of deep learning techniques. This section reviews the major developments in object detection models, focusing on deep learning-based approaches, and provides an insight into their applications, challenges, and future trends.

Problem Statement and Challenges in Object Detection

Object detection, extending beyond mere object classification, involves identifying and localizing objects within images or videos. This task presents several challenges, including intra-class variation, category diversity, and computational efficiency. Intra-class variations arise from differences in appearance, occlusion, and environmental conditions, making it difficult for models to consistently recognize objects. The large number of object categories further complicates this task, requiring extensive, high-quality annotated data for training. Additionally, the computational demands of accurate object detection models necessitate efficient architectures, especially for applications in mobile and edge devices [5].

Datasets and Evaluation Metrics

In the realm of computer vision, object detection datasets are instrumental in training and evaluating algorithms [13], with several key datasets playing pivotal roles. Among them, Pascal VOC [14], introduced in 2005, initially featured four object classes, expanded to 20 over time. It set the foundation for standard evaluation metrics in object recognition, remaining useful for benchmarking despite being eclipsed by larger datasets. ImageNet [15], launched in 2010, marked a significant expansion with 14 million images categorized into fine-grained classes based on WordNet's structure. Its annual challenge, ILSVRC [16], included diverse tasks like image classification and object detection, making it a crucial resource for pre-training complex models. However, challenges such as partial annotations and the need for richer contextual information were noted.

Addressing some of ImageNet's limitations, the Microsoft COCO dataset [17], with 328,000 images, provides detailed annotations and a diverse context for objects, enhancing the generalizability of object detection models. The most recent major dataset, Open Images [18], includes over nine million labeled images with bounding boxes for 600 object classes and segmentation masks added in 2019. Unique for its crowdsourced extension and visual relationship annotations, Open Images offers comprehensive scene understanding. It also employs positive and negative labels, a method shown to significantly improve classification accuracy.

In the domain of object detection and tracking, bounding box regression is an integral aspect. Traditional methods have predominantly used ln-norm loss for bounding box regression, which is not optimally aligned with the main evaluation metric, Intersection over Union (IoU). IoU calculates the ratio of the intersection area to the union area of the predicted and ground truth bounding boxes. However, it falls short in cases where bounding boxes do not overlap, prompting the development of several improved metrics.

Generalized IoU (GIoU) enhances the standard IoU by considering the smallest enclosing area covering both bounding boxes. This modification allows GIoU to reflect the proximity of shapes even when they do not overlap, although it still faces challenges with slow convergence and inaccurate regression. [19]

Distance-IoU (DIoU) incorporates the normalized distance between the centers of the predicted and target boxes. This metric facilitates faster convergence in training, addressing the slow convergence issue evident in both IoU and GIoU. [20]

Complete IoU (CIoU) builds on the concept of DIoU by integrating three geometric components: overlap area, central point distance, and aspect ratio. CIoU thus offers improved performance by factoring in the shape and orientation of bounding boxes in addition to their overlap and distance. [20]

Bounding Box Intersection over Union (BIoU) is a recent advancement in the evolution of IoU-based loss functions. BIoU addresses some of the limitations of previous metrics by considering additional factors that influence the bounding box regression accuracy. [21]

These metrics represent a progressive refinement of the original IoU, each introducing new parameters to enhance the accuracy and effectiveness of object detection models

Backbone Architectures and Object Detectors

The progress of object detection over the past two decades can be categorized into two significant eras: the traditional object detection period (before 2014) and the deep learning-based detection period (after 2014). In the traditional era, most detection algorithms were built on handcrafted features, with notable examples like the Viola-Jones detectors [22], the HOG (Histogram of Oriented Gradients) feature descriptor [23], and the Deformable Part-based Model (DPM) [24]. These methods laid the groundwork for future developments in object detection.

With the advent of deep learning, there was a paradigm shift in object detection techniques, leading to the development of convolutional neural network (CNN) based detectors. This era is characterized by two main groups of detectors: two-stage and one-stage detectors.

Two-Stage Detectors: Two-stage detectors such as R-CNN [25], Fast R-CNN [26], and Faster R-CNN [27], which follow a coarse-to-fine approach, have been pivotal in the evolution of object detection. These methods initially generate region proposals and then refine them using deep learning techniques. Faster R-CNN, in particular, introduced the Region Proposal Network (RPN), enabling cost-efficient region proposals and accelerating the detection process. Feature Pyramid Networks (FPN) [28] advanced this approach by efficiently detecting objects at various scales. A notable addition to two-stage detectors is ClusterNet [29], designed for detecting small objects in large scenes by exploiting spatio-temporal information, a crucial aspect in Wide Area Motion Imagery (WAMI). ClusterNet addresses the challenges of detecting extremely small, sparse, or densely-packed objects within large search spaces, outperforming state-of-the-art methods in WAMI object detection. Its novel approach effectively combines both appearance and motion information within a deep CNN architecture, proposing regions of objects of interest (ROOBI) that can contain from a single object to clusters of several hundred objects. This method significantly surpasses existing techniques in detecting both moving and stationary objects in WAMI and reduces computational burdens such as the need for computing background-subtracted images

One-Stage Detectors: One-stage detectors, on the other hand, such as YOLO (You Only Look Once) [30, 31, 32, 33, 34, 35, 36, 37] and SSD (Single Shot MultiBox Detector) [38], aim to complete the detection process in one step. These models are generally faster and simpler, but initially struggled with accuracy, particularly in detecting small objects. The development of RetinaNet [39] addressed some of these accuracy issues by introducing the focal loss function, which puts more emphasis on hard, misclassified examples during training. Subsequent improvements and variations in one-stage detectors, such as CornerNet [40] and CenterNet [**?**], shifted focus to keypoint-based detection paradigms, offering more efficiency and simplicity.

Lightweight networks such as SqueezeNet [41], MobileNets [42], ShuffleNet [43], and, more recently, EfficientNets [44] have been developed to facilitate object detection on devices with limited computational resources. These networks strive to maintain detection performance while significantly reducing model size and computational complexity. Comparative studies of these architectures have shown varying degrees of trade-offs between accuracy, speed, and computational requirements [45].

The most recent advancements include the use of Transformers in object detection, as exemplified by DETR (DEtection TRansformer) [46]. DETR views object detection as a set prediction problem and employs an end-to-end detection network with Transformers, marking a significant shift away from traditional CNN-based approaches and anchor box methodologies.

This evolution of object detection, from handcrafted feature-based methods to sophisticated deep learning models, highlights the field's dynamic nature and its continuous pursuit of more accurate, efficient, and robust detection techniques.

Future Trends in Object Detection

The future of object detection lies in addressing current challenges and exploring new methodologies. The integration of deep learning with traditional computer vision techniques, improvements in real-time detection, and advancements in network architectures are areas of ongoing research. The development of more efficient and scalable models, capable of handling diverse and challenging real-world scenarios, is a key focus. Additionally, addressing data bias and skew in training datasets is crucial for the development of robust and accurate object detection systems [45, 5].

3.3 Multiple Object Tracking (MOT)

Multiple Object Tracking (MOT) is a critical and challenging field in computer vision, focusing on locating multiple objects, maintaining their identities, and yielding their trajectories from an input video. These objects can range from pedestrians to vehicles, each presenting unique tracking complexities.

Problem Statement and Challenges on MOT

The core objective of MOT is to ascertain the number of objects, which often varies over time, and to sustain their identities. MOT encompasses common challenges found in both Single Object Tracking (SOT) and MOT, including frequent occlusions, initialization and termination of tracks, objects with similar appearance, and interactions among multiple objects. MOT has been formulated from various perspectives, broadly seen as a multi-variable estimation problem.

Online Tracking exclusively relies on past information up to the current frame, contrasting with offline methods that use both past and future observations. This and the high frame-rate requirements for real-time operation make this setting especially challenging.

MOT Components

The main components of MOT involve discovering multiple objects in individual frames and recovering their identity information across continuous frames. This entails measuring the similarity between objects in frames and recovering identity information based on this similarity across frames. Key aspects include modeling appearance, motion, interaction, exclusion, and occlusion, alongside the inference problem [6]. The appearance model, while not the central focus of MOT as in SOT, remains important for affinity computation. It involves handling issues like occlusion, out-of-plane rotation, illumination changes, and deformation [6]. The motion model is pivotal for capturing the dynamic behavior of objects, estimating potential positions in future frames, thereby reducing the search space. This includes considerations for velocity smoothness, position smoothness, and acceleration smoothness, with linear and non-linear motion models addressing different aspects of object dynamics [6].

MOT Datasets and Evaluation

In the field of Multiple Object Tracking (MOT), the selection of datasets and the criteria for evaluation are critical for the development and benchmarking of algorithms. Over the years, a variety of datasets have been introduced, each with distinct characteristics and challenges. These datasets serve as a standard for evaluating the performance of MOT algorithms under various conditions and scenarios.

MOT Datasets

PETS2009 [47]: This dataset focuses on crowd analysis and person tracking, offering multiple camera views and calibration data. With its diverse set of scenarios ranging from sparse to highly crowded scenes, PETS2009 presents unique challenges for tracking algorithms, particularly in handling occlusions and maintaining consistent track identities in dense crowds.

KITTI Vision Benchmark Suite (KITTI) [48]: A fundamental dataset in autonomous driving research, KITTI encompasses various components like stereo vision, optical flow, and

3D object tracking. Its sequences, captured in real-world urban environments, challenge MOT algorithms with dynamic scenes involving vehicles and pedestrians. The dataset's diversity in scenarios, from crowded urban streets to highways, provides a comprehensive testbed for evaluating the tracking performance of moving objects in varying conditions.

MOT15 [49], **MOT16** [50], and **MOT20** [51]: The MOTChallenge benchmarks, including MOT15, MOT16, and MOT20, have been instrumental in advancing MOT research. MOT15 offers 22 sequences with a variety of scenes, focusing on pedestrian tracking with bounding box annotations. MOT16 improves upon this with higher-quality data and introduces the Region Proposal Network (RPN) for efficient region proposals. MOT20, targeting extremely crowded scenes, provides dense annotations in high-resolution images, focusing on pedestrian tracking in challenging environments. These datasets collectively cover a wide range of real-world scenarios, making them ideal for testing the robustness and versatility of MOT algorithms.

TAO Dataset [52]: A comprehensive dataset designed for tracking any object in diverse environments. It includes over 2,907 high-resolution videos and 833 object classes, making it one of the largest and most diverse datasets in the field. TAO's focus on long-term tracking, occlusions, scale changes, and diverse object interactions offers a comprehensive testbed for evaluating the generalizability of MOT algorithms.

KITTI-STEP and MOTChallenge-STEP [53]: These datasets, introduced in the STEP framework, provide spatially and temporally dense annotations for video panoptic segmentation. KITTI-STEP contains sequences recorded from a car-mounted camera, emphasizing long-term tracking in dynamic urban environments. MOTChallenge-STEP, although smaller in scale, presents its own set of challenges with visually similar objects and crowded scenes.

Crowd of Heads Dataset (CroHD) [54]: Specifically designed for head detection in dense crowds, CroHD offers high-definition images of crowded scenes. Its focus on densely populated environments and varying lighting conditions presents unique challenges for MOT systems, particularly in handling occlusions and maintaining consistent track identities amidst the crowd. CroHD's detailed annotations of over 2.2 million heads in diverse scenes provide valuable data for refining object detection and tracking algorithms in densely crowded contexts.

CTMC Dataset [55]: The Cell Tracking with Mitosis Detection Challenge (CTMC) dataset provides a different perspective, focusing on microscopic cell tracking in time-lapse sequences. It offers challenges such as cell division, varying shapes, and movements. With detailed annotations of cell boundaries, cell IDs, and mitosis events, CTMC pushes the boundaries of MOT into the realm of biological imaging, presenting unique challenges in tracking small, dynamically changing objects.

MOTSynth Dataset [56]: A synthetic dataset tailored for pedestrian detection, tracking, and segmentation, MOTSynth leverages the Grand Theft Auto V (GTA-V) video game to simulate realistic urban environments and pedestrian behaviors. Its diversity in weather conditions, daytime, and pedestrian flows offers a comprehensive platform for training and evaluating MOT algorithms. The dataset's synthetic nature allows for a controlled environment to test and refine algorithms before deploying them in real-world scenarios.

MOT Evaluation Metrics

In the evaluation of Multiple Object Tracking (MOT) systems against an established Ground Truth (GT), a comprehensive set of metrics is utilized, each offering unique insights into the performance aspects of tracking algorithms. Precision and Recall serve as foundational metrics, reflecting the accuracy and reliability of the tracking system. Precision is indicative of the tracker's ability to accurately identify objects, while Recall assesses the tracker's effectiveness in detecting all relevant objects within the scene.

Multiple Object Tracking Precision (MOTP) and Multiple Object Tracking Accuracy (MOTA) are central to MOT evaluation [57]. MOTP measures the tracker's precision in estimating object positions, whereas MOTA provides an overarching assessment of tracking performance by combining the number of misses, false positives, and identity mismatches. These metrics collectively evaluate the tracker's performance in object detection accuracy and identity maintenance.

Additionally, Higher Order Tracking Accuracy (HOTA) [58] offers a balanced assessment by equally emphasizing both detection and association accuracy. HOTA operates by averaging the alignment of trajectories of matched detections across all detections and penalizing mismatches. Unique in its approach, HOTA decomposes into sub-metrics that separately evaluate aspects such as detection recall and precision, and association recall and precision, providing a more detailed analysis of a tracker's capabilities. A notable feature of HOTA is its 'double Jaccard' formulation, which considers both detection and association accuracy, and uniquely integrates localization accuracy into the evaluation by averaging scores across various localization thresholds. This comprehensive approach allows HOTA to provide a more nuanced understanding of tracking performance, addressing limitations of previous metrics like MOTA and IDF1.

Evaluation could also consider the total number of frames, matches, track switches, false positives, misses, detections, unique object appearances, and prediction appearances for a complete contextualization. The total number of frames sets the temporal context of the tracking evaluation. The number of matches indicates the tracker's accuracy in identifying objects, while track switches highlight the consistency in maintaining object identities. False positives and misses shed light on the tracker's precision and recall. The total number of detections, along with unique object and prediction appearances, provides a comprehensive view of the tracker's detection capabilities and the diversity of the tracking environment.

The lifespan of object tracking is scrutinized through metrics such as the number of objects mostly tracked, partially tracked, and mostly lost. These metrics categorize objects based on the percentage of their lifespan during which they are successfully tracked, offering insights into the duration and consistency of tracking. The total number of track fragmentations further complements these metrics by quantifying the instances where tracking continuity is lost and regained.

Finally, there are Identity Management Metrics [59], including ID False Positives (IDFP), ID False Negatives (IDFN), ID True Positives (IDTP), ID Precision (IDP), ID Recall (IDR), and ID F1 Score (IDF1), along with Track Ratios and Frequencies, and the global assignment of identities, are crucial in assessing the tracker's performance in managing object identities. These metrics evaluate the accuracy of identity matching, the consistency of identity maintenance across frames, and the effectiveness of the tracker in complex scenarios where maintaining consistent object identities is challenging.

MOT Algorithms

The cornerstone of multiple object tracking (MOT) is laid in the development of foundational tracking and detection techniques. "Simple Online and Realtime Tracking" [60] is a seminal work that highlights the importance of detection quality and combines the Kalman Filter and Hungarian algorithm for efficient tracking. Similarly, "Joint Detection and Embedding for Real-Time Multiple Object Tracking" [61] furthers this domain by introducing a system where detection and embedding are unified, significantly reducing computational overhead and enhancing real-time tracking capabilities.

Occlusion management is a pivotal aspect of MOT. "Efficient Online Multi-Object Tracking Framework with GMPHD Filter and Occlusion Group Management" [62] introduces an innovative framework that incorporates the GMPHD filter and a novel occlusion group management scheme. This approach is adept at reducing false negatives due to missed detections and improves the handling of false positives through the SIOA metric. Furthermore, "Online

Multi-Object Tracking with Unsupervised Re-Identification Learning and Occlusion Estimation" [63] enhances MOT by focusing on unsupervised re-identification learning and occlusion estimation, adding a new dimension to tracking performance in occluded scenarios.

The evolution of MOT is marked by advanced architectural innovations. "Tracking Objects as Points" [64] introduces CenterTrack, which simplifies the detection and tracking process by focusing on the center points of objects. The integration of Transformers in MOT, as seen in "TrackFormer: Multi-Object Tracking with Transformers" [65], and "TransCenter: Transformer with Dense Queries for Multiple-Object Tracking" [66], marks a significant shift in methodologies, offering enhanced data association and global inference capabilities. "TransMOT: Spatial-Temporal Graph Transformer for Multiple Object Tracking" [67] further advances this trend by employing graph transformers to model complex spatial-temporal interactions.

Integrating appearance-based methods and global tracking strategies has been crucial in advancing MOT. "FairMOT: On the Fairness of Detection and Re-Identification in Multiple Object Tracking" [68] addresses the balance between detection and re-identification, ensuring equal emphasis on both tasks. "Global Multi-Object Tracking with Transformers" [69] presents a transformative approach to MOT, processing entire sequences for global trajectory prediction. Additionally, "Multiple Object Tracking from Appearance by Hierarchically Clustering Tracklets" [70] leverages appearance features for clustering tracklets, showcasing the effectiveness of appearance-based tracking.

Real-time tracking remains a vital goal in MOT. "Observation-Centric SORT: High-Speed Simple Online and Real-Time Tracking under Occlusion" [71] and "Sparse Graph Tracker: Rethinking Recovery in Online Multi-Object Tracking" [72] contribute significantly to this area. The former enhances tracking robustness under occlusion and non-linear motion, while the latter introduces higher-order relational features by converting video data into a graph. "ByteTrack: Multi-Object Tracking by Associating Every Detection Box" [73] and "Deep-OC-SORT: Deep Motion Model for Online and Offline Multi-Object Tracking" [74] further enrich the field by offering methods to associate detection boxes effectively and integrating appearance matching into motion-based tracking methods, respectively.

Enhancing existing models and developing unified frameworks for diverse tracking tasks mark recent advancements in MOT. "StrongSORT: Make DeepSORT Great Again" [75] revisits the classic DeepSORT algorithm, introducing novel algorithms for improved tracking. "Unicorn: Unified Tracking Framework for Single, Multiple Object and Video Object Segmentation" [76] provides a unified solution for various tracking tasks. Specialized domains such as sports tracking are addressed in "SportsMOT: A Large Multi-Object Tracking Dataset in Multiple Sports Scenes" [77], while "Video-kMaX: Unified Online and Near-Online Video Panoptic Segmentation" [78] demonstrates the potential of unified approaches in video segmentation. Additionally, "SMILEtrack: Similarity Learning for Multiple Object Tracking" [79] and "MOTRv2: Bootstrapping End-to-End Multi-Object Tracking with A Pretrained Detector" [80] contribute novel perspectives on similarity learning and the use of pre-trained detectors, respectively.

Together, these papers present a comprehensive overview of the state-of-the-art MOT, showcasing a wide array of innovations and methodologies that are shaping the evolution of this dynamic field. The future of MOT research is poised to explore several directions. These include enhancing algorithms to better handle occlusions, improving models for maintaining object identities and integrating MOT with other computer vision tasks like object detection and scene understanding.

3.4 Concluding Remarks

Summary of Key Findings

This literature review has explored the significant advancements and challenges in radar technology, specifically focusing on its applications in bird conservation and environmental monitoring within the wind energy sector. The fundamentals of radar technology, including the basics of its operation and the key principles behind its functionality, have been discussed. These principles provide the foundation for the wide-ranging applications of radar, particularly in avian monitoring, where it is crucial for understanding bird behavior, migration patterns, and the interactions of birds with wind turbines.

The review also covered the development and effectiveness of the Birdtrack® Radar System, designed to mitigate the impacts of wind turbines on bird populations. This system highlights the integration of radar technology with conservation efforts, offering real-time data crucial for informed decision-making in the placement and operation of wind turbines.

In the realm of computer vision, advancements in object detection and multiple object tracking (MOT) have been extensively reviewed. The evolution from traditional methods to deep learning-based techniques marks a significant shift in the field. The review detailed various datasets and evaluation metrics crucial for the development and benchmarking of algorithms in object detection and MOT, illustrating the ongoing progress in these areas.

The review also highlighted the challenges in MOT, including managing occlusions, maintaining object identities, and integrating MOT with other computer vision tasks. The discussion of various algorithms and their contributions to the field underscored the dynamic nature of MOT research and its importance in real-world applications.

Future Research Directions

Looking forward, several gaps and areas for future investigation have emerged from this review. In the context of radar technology and its applications in wind energy and bird conservation, future research could focus on improving the accuracy and range of radar systems. This includes developing algorithms that can more precisely distinguish between different species of birds and their flight patterns, which is crucial for reducing bird fatalities due to wind turbines. Additionally, exploring the integration of radar data with other environmental monitoring tools could provide a more comprehensive understanding of wildlife and its interaction with renewable energy infrastructure.

In computer vision, specifically in object detection and MOT, future research should continue to refine algorithms for enhanced accuracy, speed, and robustness. This includes developing methods that can effectively deal with diverse and challenging scenarios, such as crowded scenes or rapidly changing environments. Moreover, addressing issues like data bias and skew in training datasets is critical for ensuring that these systems are fair and equitable.

The potential impact of these advancements on sustainable energy and wildlife conservation is substantial. Improved radar technology could lead to more wildlife-friendly renewable energy solutions, while advancements in computer vision could provide tools for better monitoring and understanding of wildlife behavior. These developments not only contribute to the conservation of biodiversity but also support the sustainable expansion of renewable energy, a critical component in combating climate change.

In conclusion, the fields of radar technology and computer vision hold great promise for furthering our capabilities in environmental monitoring and conservation. By continuing to explore these technologies and addressing the identified research gaps, significant strides can be made towards a more sustainable and environmentally conscious future.

4

Proposed Solution

In response to the identified limitations of the Birdtrack® system, this chapter is dedicated to the description of an improved approach using state-of-the-art multiple object tracking algorithms. The enhancements in detection and tracking capabilities are critical to effectively achieve Birdtrack®'s primary design goals. The approach proposed in this chapter aims to significantly improve the accuracy and efficiency of the system. It involves the integration of radar data preprocessing techniques, for clutter reduction and image interpretability enhancement, coupled with the adoption of better MOT algorithms. The primary goal is the development of a system proficient in the accurate and efficient identification and tracking of birds within complex and dynamic environments, significantly improving the current technology. This chapter methodically illustrates the transition from processing raw radar data to achieving a sophisticated tracking system, capable of fulfilling the requirements of avian activity monitoring.

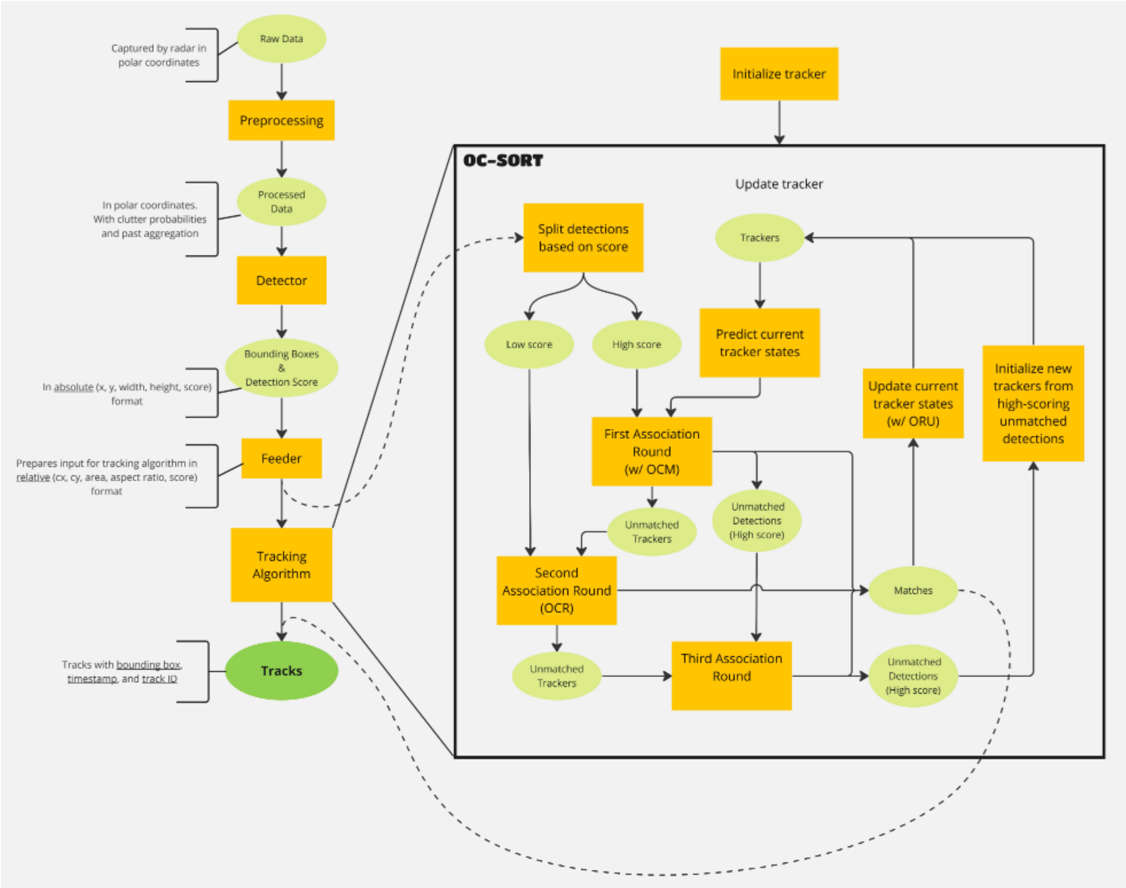


Figure 4.1: Flowchart illustrating the radar data preprocessing and multi-stage tracking algorithm for avian activity detection using the Birdtrack® radar system.

4.1 Radar Line Data Preprocessing

In this section, the essential step of radar data preprocessing is addressed, focusing on data from the Birdtrack® radar system. The raw radar data, while insightful, is cluttered with various disturbances, such as noise and clutter. Clutter can be classified as static, as in the geomorphological characteristics of the region, trees, or other non-moving objects, or it can be dynamic, such as in rainfall, clouds, fog, waves, or insect clouds), with incidence rates depending on the region of the world where the radar is situated. Hence, refining this data to ensure its clarity and reliability is paramount to facilitate subsequent steps in the pipeline.

This chapter focuses on visualizing Raw Radar Data. Visualization aids in comprehending intricate data. This segment discusses the transformation from polar coordinates into Cartesian coordinates to facilitate a more intuitive visualization and make pixel distances interpretable as real-world distances. It finished with a temporal aggregation method that allows the identification of targets from individual frames, opening the door to one-shot machine-learnable online detectors such as SSD [38], YOLO variants [30, 31, 32, 33, 34, 35, 36, 37], or ClusterNet [29]. These results elucidate the processes involved in radar data preprocessing, specifically for the Birdtrack® system's objectives.

Radar Imagery Visualization

To aid visualization, detection, and consequently tracking, several steps need to be performed, which include: (1) interpolation of missing lines; (2) conversion to the Cartesian coordinate system in a lower resolution for size reduction, and easier, more interpretable processing; (3) and finally, online temporal aggregation for live visualization and clutter rejection. These steps are explained in the following.

Conversion to cartesian coordinates

After the raw data and associated metadata have been appropriately read, missing lines identified in the metadata are interpolated using an image in-painting technique based on the Fast Marching Method developed by Alexandru Telea [81], this provides a cleaner image with no evident missing lines, while still being fast for online processing of large images.

To align all images to the same frame of reference, they are rotated so the up direction is aligned with the north. This can be achieved in polar coordinates by moving the final N lines to the top of the image buffer before the coordinate conversion. Here N is defined by dividing 360 degrees by the number of lines per revolution and multiplying this value with the desired offset in decimal degrees.

The image is then converted from the original polar coordinates to Cartesian coordinates more amenable for processing and interpretation. In this step, a resolution of 1024x1024 with 8-bit grayscale depth is fixed for all frames. This leads to an overall reduction in the size of the data while keeping the resolution above the minimum required for target detection, this will be elaborated in subsequent sections. The result of this processing step can be seen in Fig. 4.2.

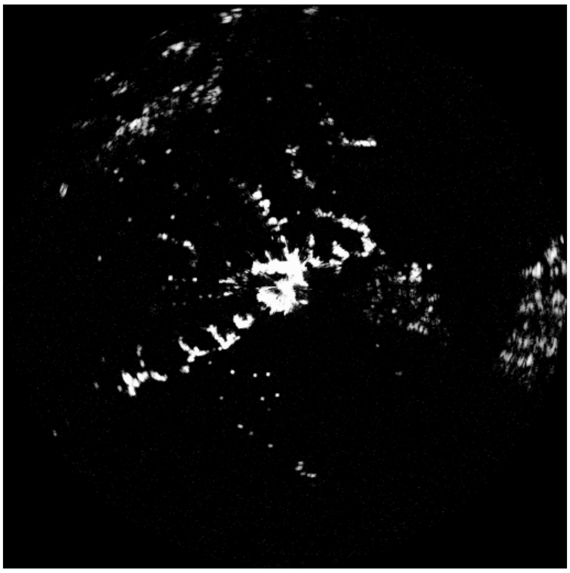


Figure 4.2: Raw frame after conversion to cartesian coordinates. The radar is now located in the center of the frame.

Attempts have been made to further clean this data. One such attempt was the use of Non-Local Means Denoising [82] followed by a power-law transformation (also known as gamma correction), to improve the signal-to-noise ratio in the data. These methods proved very successful in some sessions while providing unusable images in others, either by removing small targets in the denoising step or by raising the noise floor levels to unusable values in the gamma correction step. Due to these inconsistencies and taking into account the strong filtering capabilities of the tracking algorithms used further in the pipeline, these techniques were not used throughout the remainder of this analysis.

Data aggregation

For aggregating temporal information, two different methods were simultaneously employed: Exponential Moving Average (EMA), and Time Since Saturation (TSS).

Both intend to aggregate single-pixel detection statistics so the first step here is to define a pixel as being saturated at timestep t, when it is above the dynamic threshold of the image at that timestamp. Otsu's [83] method for dynamic thresholding involves building a histogram of pixel intensities and finding the threshold that minimizes intra-class variance (which is equivalent to maximizing inter-class variance) through exhaustive search. After this threshold is calculated we can classify pixels as saturated or non-saturated and attribute a binary mask to each frame, where the threshold is dynamic and frame-dependent. The main advantage of using this method is that it is more robust to background noise: if the noise floor of the data is increased, the threshold will be increased as well to reflect this change in distribution.

The first method described is an Exponential Moving Average (EMA) of the saturated pixels. When used with a high alpha it serves as a good proxy metric for the detection of static clutter, as this kind of clutter tends to saturate the same or neighboring pixels consistently throughout the session. One way to interpret this aggregation is as a value proportional to the parameters of an inhomogeneous Poisson process that can be used to model the static clutter's distribution. This will be used for the calculation of track quality metrics as well as the development of fast clutter rejection filters. The result of this processing step can be seen in Fig. 4.3.

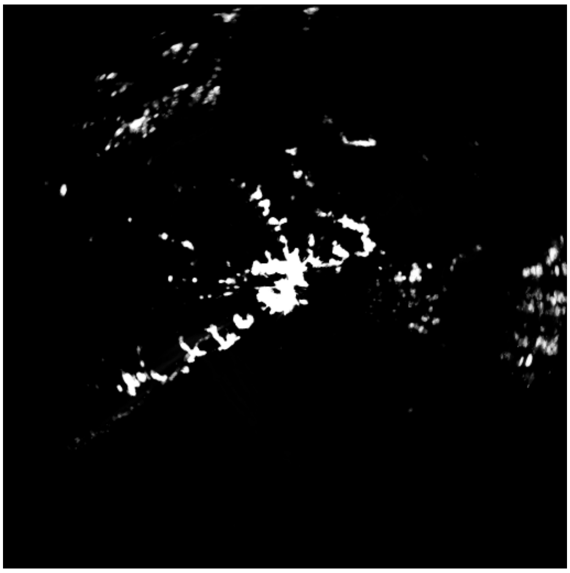


Figure 4.3: Result of EMA aggregation.

The second method described is the Time Since Saturation (TSS). This method provides information on when was the last time a single pixel was saturated. For its calculation, a window size is defined (in the experiments a value of 24, corresponding to 2 minutes, was used), the intensity of each pixel is maxed out using the last frame's saturation mask, and the values are linearly decayed to 0 over the defined time window. Alg. 4.1 showcases how to perform this aggregation. The major advantage of this aggregation is that it creates very distinct trails behind moving objects, making movement patterns very evident when looking at a single frame of the session. This in turn allows a simplification of human annotation tasks in future projects, or single-shot online detectors to be trained on these cheap-to-acquire aggregations. The result of this processing step can be seen in Fig. 4.4.

Algorithm 4.1 Time Since Saturation

Require: mask of saturated pixels per frame

past $\leftarrow \mathbf{0} \in \mathbb{R}^{1024\times1024}$ {Initialization matrix}

window_size $\leftarrow 24$ {Sliding window size}

for each mask in masks

  past $\leftarrow$ clip(past $- \frac{1}{\text{window\_size}}, 0, 1)$ {Linear decay}

  past[mask $\neq 0$] $\leftarrow 1$ {Recently saturated pixels}

end for

=0

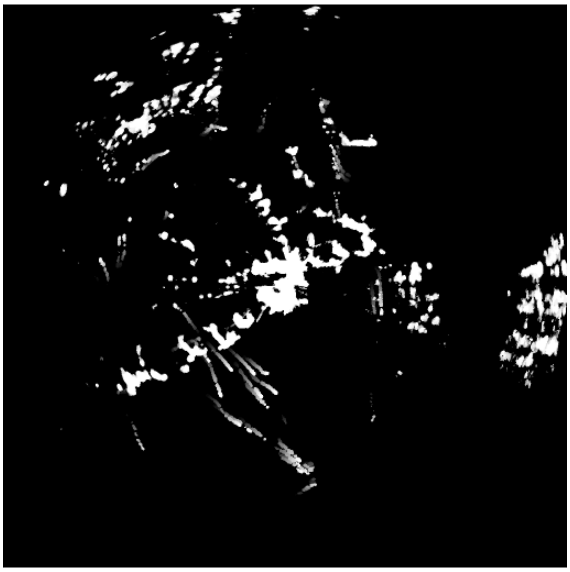


Figure 4.4: Result of TSS aggregation.

The main outcome of these aggregations is their very interpretable visualizations which maintain the full fidelity of the original raw frame after conversion to cartesian coordinates. These allow for a single frame, now with 3 channels (which were chosen to be: red - TSS, green - original, and blue - EMA) instead of the original grayscale version, to contain enough information to allow for the identification of moving targets, rejection of static clutter, as well as discern movement patterns of distinct objects, thus providing workable data for target-clutter binary classification. The results can be seen in Fig. 4.5.

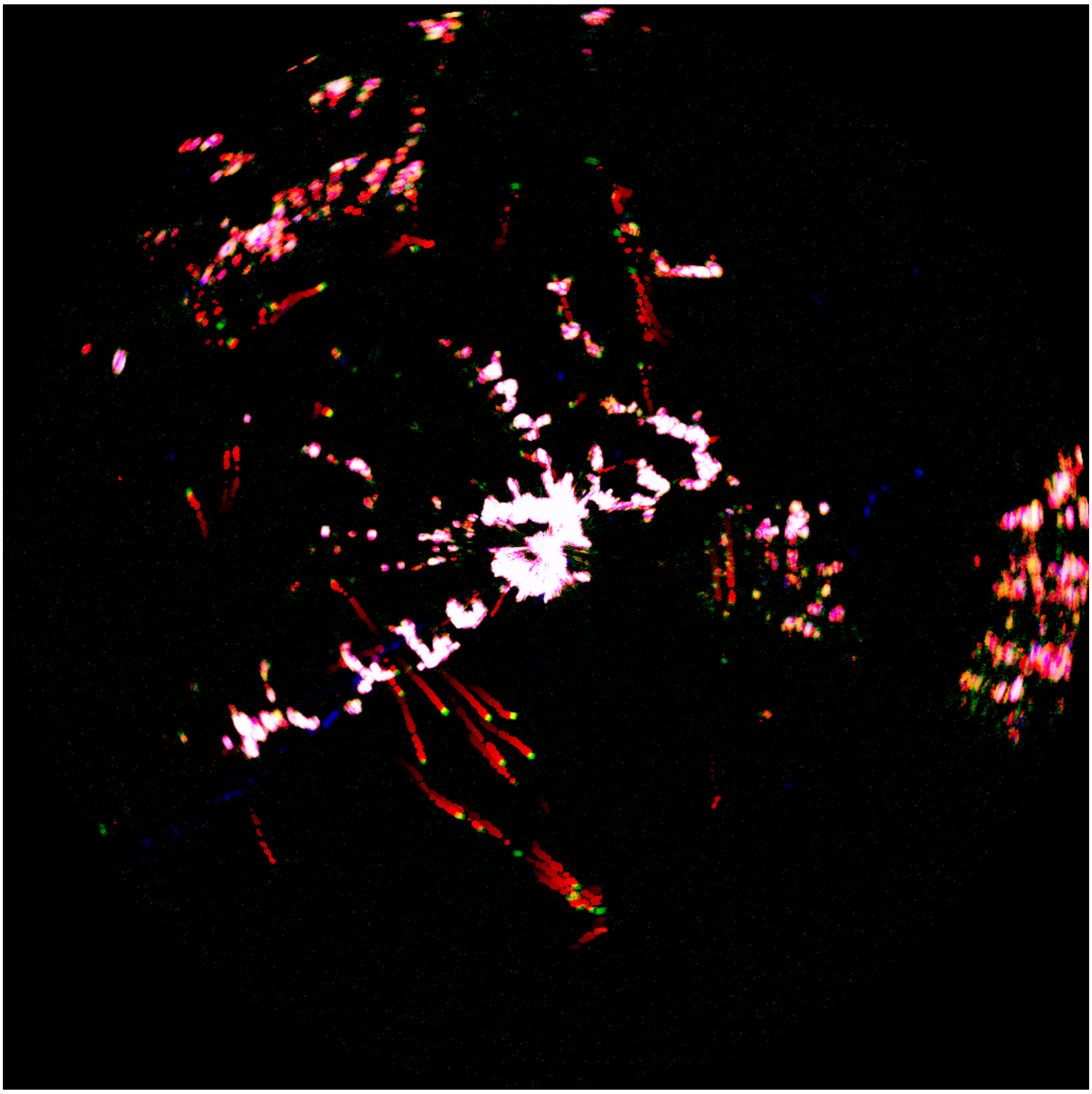


Figure 4.5: Visualization with temporal aggregations.

4.2 Multiple Bird Tracking using Radar

In this section, the focus shifts to the application of radar data for tracking multiple birds simultaneously. The overarching goal is to develop a system capable of accurately and efficiently tracking the movements of birds within the radar's range in real-time. This involves a series of steps, beginning with the detection of birds from the preprocessed radar imagery, followed by the implementation of a tracking algorithm that can handle multiple targets. The challenge lies in accurately identifying and following individual birds amidst a dynamic and often cluttered environment. The system's effectiveness hinges on its ability to distinguish between birds and other objects, maintain continuity of tracking even with intermittent loss of signal, and handle the complexities of tracking multiple birds simultaneously. The subsequent subsections delve into the specifics of bird detection, the possibilities for tracking a single bird, and the intricacies involved in extending this to multiple bird tracking.

Bird Detection

The bird detection mechanism is an important algorithmic component of the radar-based avian tracking system, designed with an emphasis on high recall (low false-negative rate) to ensure comprehensive detection of birds within the radar's range. several false-positive detections are filtered out by the subsequent tracking algorithm so the objective at this stage is to avoid the potential loss of tracking targets.

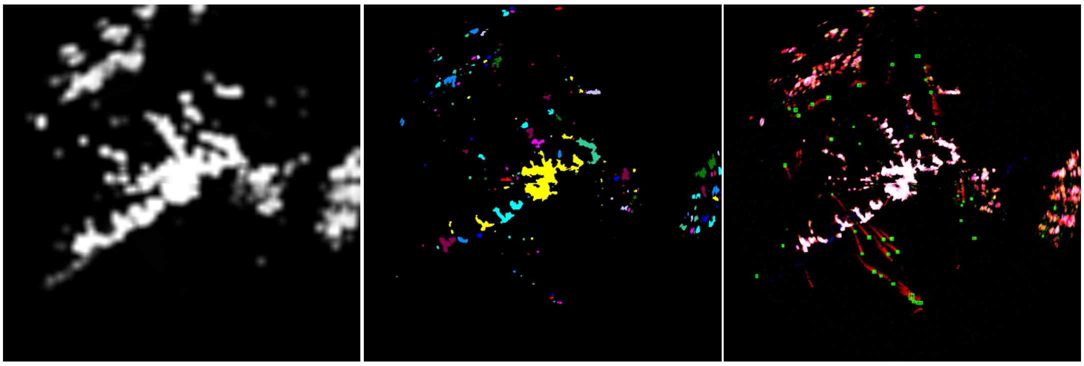
The developed detection algorithm operates on the processed radar imagery, specifically leveraging the Exponential Moving Average (EMA) aggregation. This aggregation aids in discerning potential clutter by emphasizing areas of persistent saturation indicative of static objects. The EMA is then subjected to a dilation process with a kernel size of 17x17, followed by a blurring step with a kernel size of 21x21. This sequence generates a probability mask that signifies the likelihood of clutter presence, based on the proximity of saturation points to the long-duration static clutter detected by the EMA.

Subsequently, the radar frame is blurred, and Otsu's method [83] is applied to binarize the image. This preparatory step enhances the connectivity of components, facilitating the subsequent application of a connected components algorithm. The algorithm's robustness is bolstered by the blurring, which simplifies the image features and attenuates noise-induced disconnected components.

The detection process entails the exclusion of components whose area falls below a 6-pixel threshold or exceeds 256 pixels, thereby filtering out noise and objects not within the expected size range for birds. These values were suggested by the developers of Birdtrack® after extensive validation in production. Each remaining component undergoes analysis to ascertain the clutter probability, derived from the intersection of the component's mask with the clutter probability mask, taking the maximum within this intersection.

The detector outputs bounding boxes in absolute top-left (x,y)-coordinates, width, and height. The confidence in detection (detection score) for this algorithm is the complement of the calculated clutter probability, $P(detection) = 1 - P(clutter)$.

Figure 4.6 presents a visual progression of the bird detection process in the radar-based avian tracking system. The first subfigure (a) displays the Dilated and Blurred Exponential Moving Average (EMA) heatmap, highlighting areas of persistent saturation that may indicate static clutter. The second subfigure (b) illustrates the Connected Components, showcasing the impact of blurring and Otsu's binarization method on the radar imagery, which enhances component connectivity. The final subfigure (c) depicts the Detections, where bounding boxes identify the birds, filtered based on size and clutter probability. This figure encapsulates the critical steps of the detection process, from clutter identification to the final detection of birds, demonstrating the algorithm's ability to discern and isolate avian targets in radar data. Further filtering is achieved by the tracker so the focus is on high recall (low false negative rate).



(a) Dilated and blurred EMA heatmap **(b)** Connected Components **(c)** Detections

Figure 4.6: Detector pipeline steps visualized.

While this detector is rudimentary, relying on simple heuristics, it establishes a baseline for bird detection. Future work (see section 6.1) may refine this process, potentially employing learned detectors that could provide a measurable enhancement in detection efficacy over the methodology presented in this work.

Single Bird Tracker

The foundational step towards understanding Multiple Object Tracking (MOT) lies in the domain of Single Object Tracking (SOT), where the focus is on the continuous observation of an individual bird's trajectory. Within this context, the Kalman Filter emerges as a quintessential tool, providing a probabilistic approach to estimate the state of a system over time, amidst noise and uncertainty in the measurements.

A Kalman Filter is particularly useful in this scenario for its efficiency and its recursive nature, allowing for real-time updating of the bird's position and velocity. While standard Kalman Filters are well-suited for linear systems, variants such as the Extended Kalman Filter and the Unscented Kalman Filter are designed to handle non-linearities in the system dynamics or the observation models. Particle filters and their variants stand as another alternative, offering a non-parametric solution to the tracking problem by representing the posterior distribution of the state with a set of random samples.

In the case of tracking a single bird, a standard Kalman Filter is employed to model the bird's motion under the assumption of constant velocity linear dynamics. The state space of the tracker is formulated as a seven-dimensional vector, encapsulating the bird's position, area, aspect ratio, and their respective temporal derivatives:

$$\mathbf{x} = [c_x, c_y, area, aspect\_ratio, \dot{c_x}, \dot{c_y}, ar\dot{e}a]^T \tag{4.1}$$

Here, c_x and c_y denote the center coordinates of the bounding box, with $area$ calculated as the product of width and height, and $aspect\ ratio$ as the quotient of width by height. These values are expected to be in relative format, which is achieved by rescaling the absolute pixel coordinates and sizes to relative values in the $[0 - 1]$ range.

The transition model of the system is dictated by the following system of linear differential equations:

$$\begin{cases} c_x(t + \Delta t) = c_x(t) + \Delta t \times \dot{c_x}(t) \\ c_y(t + \Delta t) = c_y(t) + \Delta t \times \dot{c_y}(t) \\ area(t + \Delta t) = area(t) + \Delta t \times ar\dot{e}a(t) \\ aspect\_ratio, \dot{c_x}, \dot{c_y}, ar\dot{e}a \text{ are kept constant in this model.} \end{cases} \tag{4.2}$$

Which can be represented by the following matrix for a discrete linear transition model:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.3}$$

where Δt is the time elapsed between frames. This transition matrix facilitates the prediction of the bird's next state by incorporating inter-frame linear motion and consistent changes in the area.

The Kalman Filter operates in two fundamental steps: predict and update. In the prediction step, the current state estimate is advanced to the next time step. In the update step, the predicted state is corrected using the new measurement. This cycle proceeds iteratively, with each iteration refining the estimate of the bird's trajectory.

The measurements fed into the Kalman Filter, denoted by $\mathbf{z}$, are in the form:

$$\mathbf{z} = [c_x, c_y, area, aspect\,ratio] \tag{4.4}$$

The measurement model, which relates the true state to the measurements, is assumed to be direct observation with added noise.

This simple linear dynamics model suffices for approximating the inter-frame dynamics of the bird's motion, ensuring that the tracker is responsive to the bird's apparent motion while remaining robust against minor fluctuations in the observed measurements. The efficacy of this model forms the cornerstone for further exploration into MOT, as delineated in the next section.

Multiple Bird Tracker

The challenge of tracking multiple birds simultaneously from radar data presents unique complexities. Unlike Single Object Tracking (SOT), which focuses on the trajectory of an individual bird, Multiple Object Tracking (MOT) must efficiently and accurately track numerous birds, each with distinct motion patterns, amidst a dynamically changing environment. This subsection delves into the development and tuning of a multiple bird tracker tailored for radar-based avian activity detection, as illustrated in Figure 4.1.

Algorithm Selection and Tuning

For this purpose, the Observation-Centric SORT (OC-SORT) algorithm, a more advanced variant of the Simple Online and Real-time Tracking (SORT) algorithm, was adapted to better achieve the goals of this system. OC-SORT is particularly suitable for high frame rate scenarios and exhibits superior capability in handling occlusions and non-linear movements. The algorithm's primary advantage lies in its observation-centric approach, which heavily relies on actual observations rather than purely on estimations. This feature is particularly beneficial in radar-based tracking, where direct measurements are often more reliable due to the lack of movement in the sensing device, leading to a fixed point of view for all measurements.

The Kalman Filter, a core component of the SORT family algorithms, was tuned specifically for the tracking of birds. The state space model of the Kalman Filter includes the position, size, and velocity of each bird. The process noise covariance matrix $\mathbf{Q}$ and measurement noise covariance matrix $\mathbf{R}$ were adjusted based on empirical observations of bird movements, ensuring that the tracker remains sensitive to actual movements while being robust against measurement noise.

Speed Gating and Tracker Association

To manage the computational complexity and improve tracking accuracy, a speed-gating mechanism was implemented. This approach involves setting a maximum speed threshold, beyond which a detected object is not considered for association with an existing track. The threshold is based on a reasonable estimation of the maximum speed at which birds are likely to move, in this analysis it was set at 150 km/h.

The most common algorithm used for minimum cost association is the Hungarian method. The tracker association process was improved using the Jonker-Volgenant algorithm for solving the Linear Assignment Problem (LAP) [84]. This algorithm is more efficient and accurate compared to the heuristic-based methods used in the existing Birdtrack® system, particularly in scenarios with a large number of detections or overlapping tracks. This is due to the inherent sparsity of the association matrix and the ability of the Jonker-Volgenant algorithm to reduce the problem size considerably before associating tracks to detections while keeping a near-optimal global association nature. Although the theoretical worst-case complexity of this algorithm is the same as the Hungarian method's, $O(N^3)$, in practice it has a much shorter execution time, especially for sparse inputs.

Observation-Centric Methodology

The OC-SORT algorithm employs multiple association rounds to efficiently match detections to tracks:

- **First Association Round (OCM):** This round incorporates the Observation-Centric Momentum (OCM) method, which combines motion direction and distance-intersection-over-union (DIoU) metrics to provide a robust association measure.

- **Second Association Round (ByteTrack):** This round is inspired by ByteTrack [73] and is used to assist in re-associating lost tracks, leveraging low-scoring detections as candidates for each unmatched track.
- **Third Association Round (OCR):** The Observation-Centric Recovery (OCR) handles any remaining detections and tracks using a more permissive association threshold, in an attempt to minimize the number of unmatched trackers while maintaining high track quality.

New trackers are initialized based on high-scoring unmatched detections. Matched trackers are updated using the associated detection and Observation-centric Re-Update (ORU) in the case of previously missed detections, minimizing the accumulation of error in the Kalman filter's estimations. Trackers with consecutive hits above a configurable threshold are returned as the currently locked tracks. Trackers with consecutive missed detections above a configurable threshold are deleted.

Efficiency and Efficacy

The adapted OC-SORT algorithm, as depicted in Figure 4.1, provides a robust solution for real-time multiple bird tracking. The system's efficacy is rooted in the tuning of its components and its efficiency is due to the employment of carefully chosen algorithms designed to handle the specifics of radar-based avian tracking. The main advantage of this tracker lies in its ability to maintain continuity of tracking even in the presence of intermittent signal loss or occlusions, a common challenge in radar imagery that the existing BirdTrack® algorithm struggles with. This specific adaptation of OC-SORT can consistently achieve high frame rates from 60 to 80 frames per second in the experiments detailed in the next chapter.

5
Experimental Results

In this chapter, the experimental methodology and results are presented. The main challenges that we faced lies in the *lack of reliable ground truth detections and trajectories*, as their acquisition would require multiple human experts in the field simultaneously monitoring the radar imagery and identifying birds within its range and this would be a prohibitively costly process.

Nevertheless, to compare the proposed system against the currently adopted bird tracking technology, performance proxy metrics, along with a visual inspection of the results are utilized for performance assessment. Our results indicate that our new system, developed as part of this thesis work, is more effective than the currently adopted algorihm and it is capable of: decreasing the number of bird tracks that go undetected and detecting longer bird trajectories, even in the presence of missing data or disturbances that hamper their reliable detection in some portions of the data.

5.1 Datasets used

The data sets used for the development, validation, and evaluation of this work resulted from radar sessions of the following case studies, where radar was used to monitor bird activity on operational wind power projects.

Case study 1 - BSJ wind power project

Located in the Southwest mainland of Portugal Barão de São João Wind Power Project is the largest wind farm in the region. It spans areas near the village of Barão de São João and the Charrascosa woods. The BSJ is a 50MW Wind Power Project consisting of 25 wind turbines aligned along the ridges of Barão de São João and Charrascosa. To the southeast, the wind farm borders the Barão de São João National Forest, between the geodesic markers of Relvas and Pedra Branca. To the northwest, it is bounded by the Mata da Charrascosa. The farm encompasses the Vinha Velha valley.

The project area is within the Natura 2000 site Costa Sudoeste (PTCON0012) covering a mosaic of habitats, including pine forests, eucalyptus groves, acacia, shrubs, cork oak forests (with and without shrubs), riparian galleries, and small ponds. This region is also home to many migratory soaring birds as part of an important migratory flyway, crossed by 5000 individuals of 30 soaring bird species every autumn. North and northwest of the wind park area lie elevations of minor altitude (up to 297 m), known as the Espinhaço do Cão mountain range. This range is interspersed with relatively deep valleys traversed by watercourses like the Paraíso stream and the Sinceira stream. To the south and southeast of the ridges, there are low-altitude flat areas marked by agricultural lands, orchards, and Mediterranean scrubland. These lands are dominated by limestone soils, commonly known as Algarve barrocal. The west and southwest feature low-altitude plateaus (around 100 m) marked by temporary ponds and gently sloping valleys.

The data set consists of a radar session on October 16th, 2022, with the radar operating in horizontal mode. The radar range was set to cover a distance of 7 kilometers. The session started at 09:07 and concluded at 12:00. Throughout this period, the radar system was actively scanning the designated area for birds under heavy weather with rainfall.

Radar Session Information	BSJ
Date	16.10.2022
Session Start	09:07
Session End	12:00
Radar model	Birdtrack SS-X25 Series
Radar sensor	Furuno DRS25A-NXT
Orientation Mode	Horizontal Mode
Radar Range	7 km
Coordinates	Lat 37.183872, Lon -8.777694

Table 5.1: BSJ Radar Session Details

Case study 2 - Genesis wind power project

Located in the Golan Heights, the Genesis Wind Power Project is the largest wind farm in Israel consisting of 39 wind turbines with an installed capacity of 207 MW. It spans areas near Yonatan and Keshet with topographical features and vegetation zones with significant ecological and geological interest. Its volcanic past and present-day climatic conditions have resulted in a landscape that is both rugged and fertile, supporting a diverse range of plant life and ecosystems. The project is also positioned along one of the world's major bird migration routes. This unique placement accentuates the ecological richness of the region, teeming with diverse bird and bat species, including the endangered griffon vulture.

The data set consists of a radar session on May 10th, 2022, with the radar operating in horizontal mode. The radar range was set to cover a distance of 7 kilometers. The session started at 08:12 and concluded at 15:37. Throughout this period, the radar system was actively scanning the designated area for birds while being actively cluttered by large amorphous insect clouds.

Radar Session Information	Genesis Wind
Date	10.05.2022
Session Start	08:12
Session End	15:37
Radar model	Birdtrack SS-X25 Series
Radar sensor	Furuno DRS25A-NXT
Orientation Mode	Horizontal Mode
Radar Range	7 km
Coordinates	Lat 32.934099, Lon 35.841742

Table 5.2: Genesis Wind Radar Session Details

Case study 3 - PNTI national park test site

Located in the heart of mainland Portugal, the Tejo Internacional Nature Park (PNTI) represents a stunning example of natural preservation. This test site, nestled within the park boundaries, is particularly notable for its location along the border section of the International Tagus River and its encompassing valleys. The landscape here is characterized by a "wild" charm, with its scenic valleys and rugged canyons, notably in Segura where steep, scarped outcrops form dramatic rocky formations.

The geology of the area is predominantly shale, supporting a diverse range of flora. This includes characteristic species from southern landscapes, such as cork and holm oak woods, dense olive groves—sometimes terraced—and steppe-like areas with cereals. In the less fertile soils, the gum rockrose dominates, adding to the unique Mediterranean character of the park's vegetation.

Spanning over 26,490.43 hectares, the park is not just a haven for plant life but also an essential habitat for a variety of bird species. It plays a crucial role in providing nesting sites on rocky cliffs and nearby areas for top endangered soaring birds. These include the Spanish imperial eagle (Aquila adalberti), golden eagle (A. chrysaetos), Bonelli's eagle (A. fasciata), cinereous vulture (Aegypius monachus), Egyptian vulture (Neophron percnopterus), black stork (Ciconia nigra, the symbol of the park), black wheatear (Oenanthe leucura), and the red kite (Milvus milvus). The presence of these species, especially in such significant numbers, underscores the ecological importance of the Tejo Internacional Nature Park in conserving these majestic birds.

The data set consists of a radar session on Jan 11th, 2017, with the radar operating in a horizontal mode. The radar range was set to cover a distance of 7 kilometers. The session started at 11:42 and concluded at 12:44. Throughout this period, extremely numerous flocks of birds can be seen within the radar's range.

Radar Session Information	PNTI
Date	11.01.2017
Session Start	11:42
Session End	12:44
Radar model	Birdtrack SS-X25 Series
Radar sensor	Furuno DRS25A-NXT
Orientation Mode	Horizontal Mode
Radar Range	7 km
Coordinates	Lat 39.703600, Lon -7.149430

Table 5.3: PNTI Radar Session Details

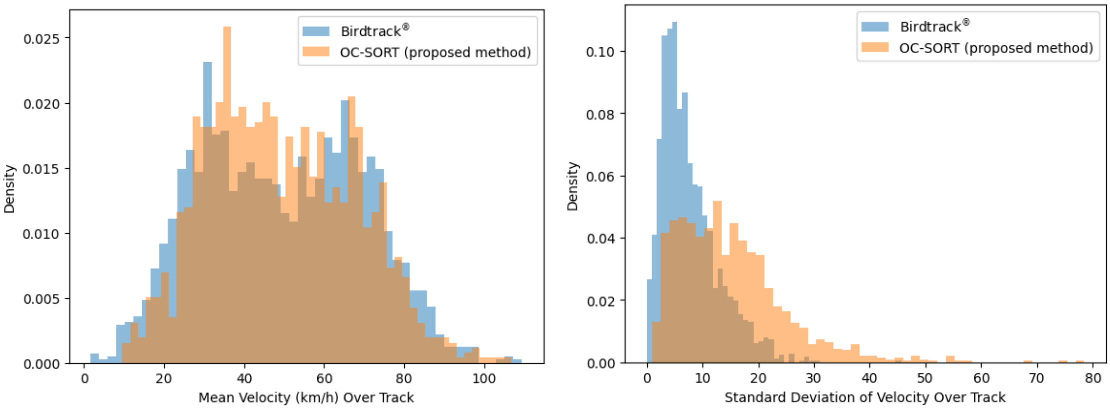
Case studies 1 and 2 were used for the development and validation of the algorithm due to their challenging, high static- and dynamic-clutter environments. These feature rain and insect clouds that make the problem of accurately detecting and tracking birds much more challenging.

Case study 3 was used for the final evaluation of the developed algorithms. This session was selected by the company as a scenario commonly used for demonstration purposes to their clients, where their algorithm is expected to perform robustly. The Birdtrack® parameters were carefully tuned by experts for this specific session to get a comparison of the developed algorithm against a known-to-be-good baseline of the existing algorithm.

5.2 Evaluation results

The Birdtrack® system does not make available intermediate detections, making direct comparison of the detectors infeasible. For this reason, the evaluation presented here is focused on the tracker.

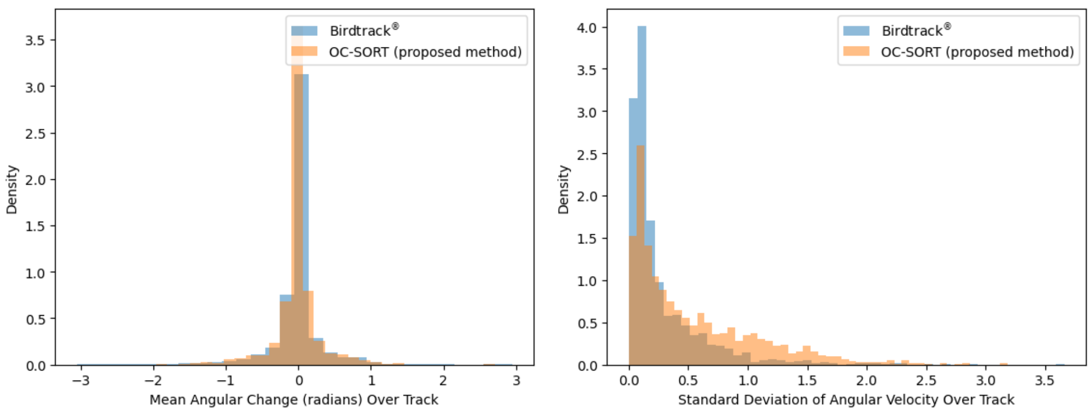
The proxy metrics proposed by the company experts consist of simple metrics such as the mean and standard deviations of both track velocity and angular changes in the trajectory. These metrics were used to validate whether the acquired trajectories are within plausible ranges for flying birds. The results are shown in Fig. 5.1, and Fig. 5.2. These histograms indeed show plausible distributions and acceptable variances for realistic bird flight trajectories.



(a) Mean velocity (in km/h)

(b) Standard deviation of velocities (in km/h)

Figure 5.1: Evaluation of track velocities. Both methods present values within the plausible range for flying birds. Due to its fragmented shorter tracks, Birdtrack® shows an overall smaller deviation. Nonetheless, OC-SORT's deviation still presents itself within an acceptable range for realistic flight trajectories.



(a) Mean angular change (in radians)

(b) Standard deviation of angular change (in radians)

Figure 5.2: Evaluation of angular changes. Both present values within the plausible range for flying birds (normal distribution with 0 mean and small standard deviation), which tend to move in approximately rectilinear trajectories with small angular adjustments. Again, Birdtrack® shows lower deviation due to its fragmented tracks having shorter lengths.

A more complex track quality metric used in [85] models a track evolution as a two-state Markov chain, where the target being tracked can either exist (state 1) or not exist (state 2). Track quality can then be interpreted as the complement of the probability of transition from state 1 to state 2, or equivalently the probability that a target will keep existing in the following time step.

This approach relates track quality to the number and quality of hits (expressed in terms of detection probability, measurement-to-track association probability, and clutter probability) and the number of misses. This can be expressed in the following recursive formulation:

$$Q(k+1) = \begin{cases} \frac{1-\pi_d(1-\Lambda(k+1))}{1-\pi_d(1-\Lambda(k+1))Q(k)} Q(k), & \text{if hit} \\ \frac{1-\pi_d}{1-\pi_d Q(k)} Q(k), & \text{if miss} \end{cases}$$

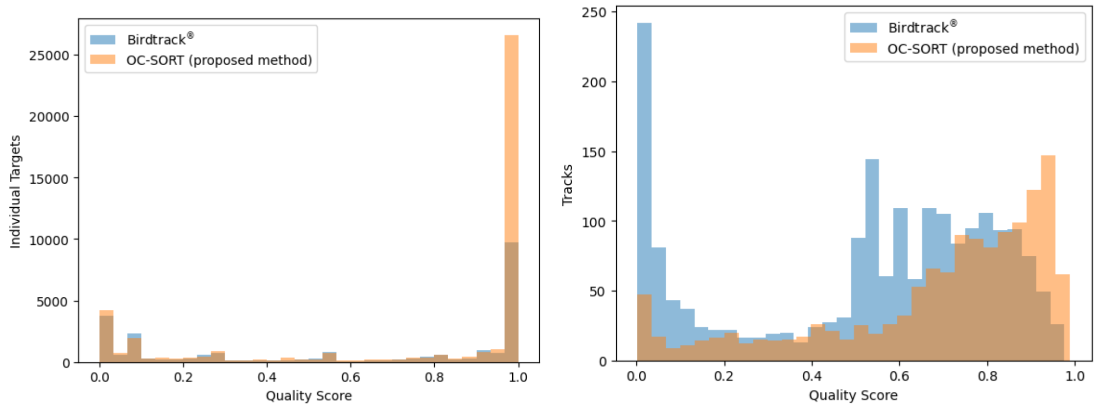
Where π_d is the probability of detection of a target, if it exists, and,

$$\Lambda(k+1) = \frac{f(z_{k+1} \mid x_k)}{f(z_{k+1} \mid \varnothing)}$$

is the likelihood ratio of two probability density functions, the first corresponds to the probability of a measurement z_{k+1} given the current track state x_k, which can be interpreted as the measurement-to-track association probability, and the second corresponds to the probability density of a measurement given no tracks, which can be interpreted as the probability that a measurement is clutter-induced.

The EMA aggregation produced in the preprocessing step is used to estimate the denominator, and the Kalman filter's predicted likelihood is used to estimate the numerator. This provides a robust metric for track quality that decreases when no measurements are associated in a given time step or when a measurement whose likelihood ratio Λ is smaller than 1 is associated. The quality increases when a measurement with a likelihood ratio larger than 1 is associated.

An initial quality value of 0.1 is used for all tracks at initialization. The probability of detection was set at 0.8, estimated from the number of missed hits in the obtained trajectories. These parameters resulted in a per-target mean quality of 60.2% for Birdtrack®, and 75.3% for OC-SORT. When aggregated by track, the mean track qualities achieved were 52.3% for Birdtrack®, and 68.8% for OC-SORT. These results can be further visualized in Fig. 5.3.



(a) Quality scores of individual targets

(b) Mean quality scores per track

Figure 5.3: Evaluation of quality scores. It should be highlighted the favorable skew of the proposed algorithm towards higher quality.

Another important metric is the total duration of tracks, which can be an indicator of fragmentation when numerous shorter tracks are identified. Track fragmentation happens when intermittent signals lead the tracker to lose, and subsequently re-acquire, tracks, assigning them different IDs. Reducing track fragmentation is extremely important when using these systems for impact assessment and environmental monitoring, as it helps count the number of birds that are identified in a specific period more accurately and reliably. Fig. 5.4 shows the comparison of track durations for Birdtrack® and OC-SORT, where OC-SORT shows a very favorable skew towards longer track durations.

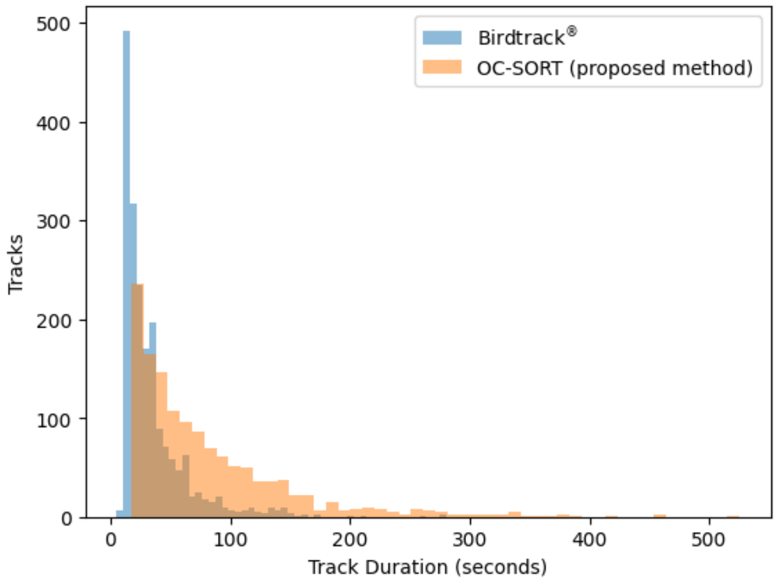
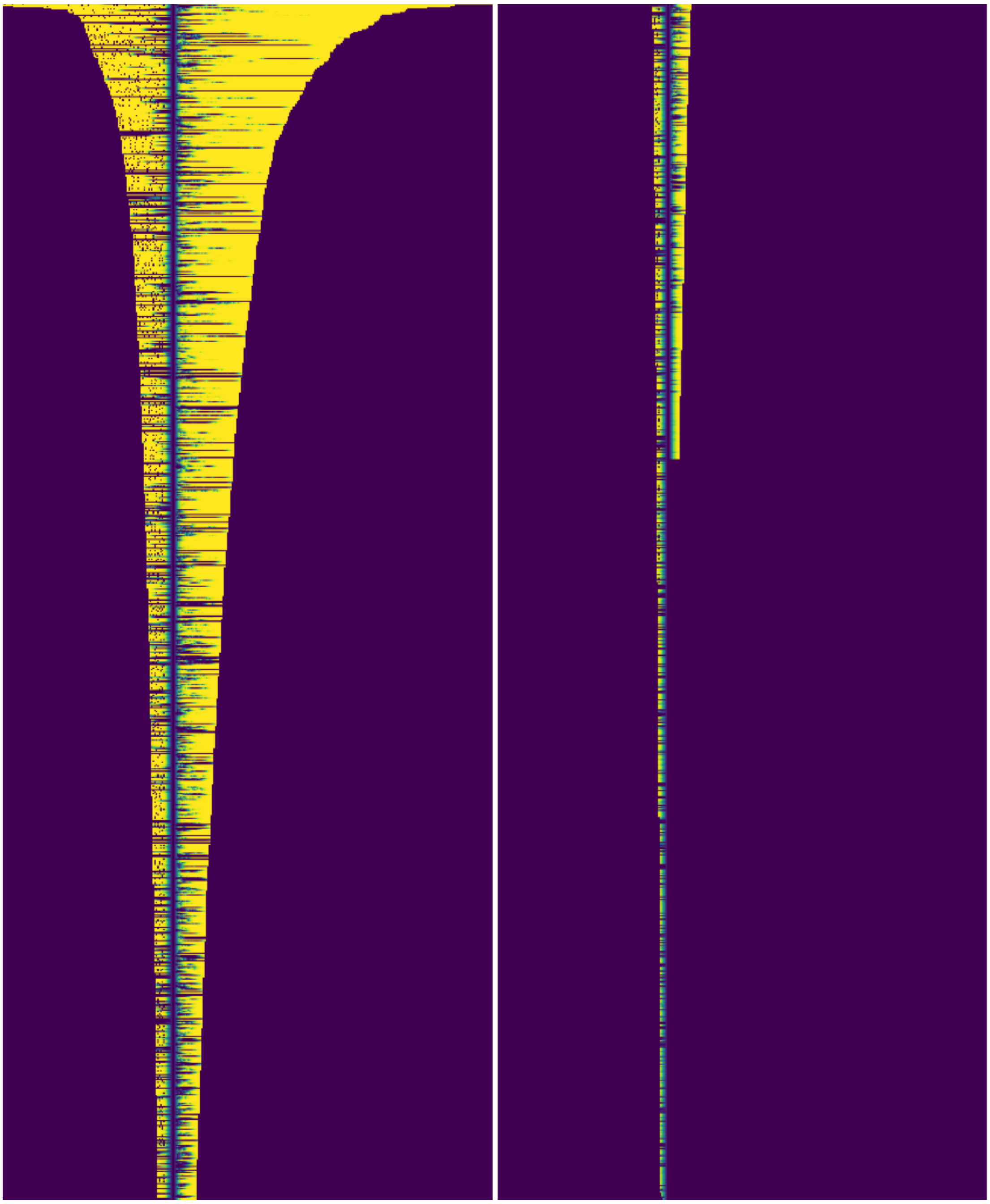


Figure 5.4: Histogram comparing track durations for both systems. OC-SORT shows a very favorable skew towards longer-duration tracks. The area under the histogram for OC-SORT is 1.64 times larger than for Birdtrack®, indicating that even though Birdtrack® has more tracks in total, the OC-SORT tracks are of longer duration achieving higher $track \times seconds$.

Finally, Fig. 5.5 shows a different kind of visualization to evaluate the robustness of each algorithm in terms of occlusions or missed detections, track quality, track quantity, and track length. The dividing line in the middle represents the start of each track and the length from this starting point to the left (for Birdtrack® tracks), or to the right (OC-SORT tracks) corresponds to the number of frames elapsed since the first detection. The intensity of each pixel corresponds to the track quality at that specific instant. Here we see Birdtrack® detects more, although much shorter, and consequently lower quality, tracks. This is attributed to track fragmentation, easily identified when visually inspecting the tracks each method outputs (see Fig. 5.6). It should also be noted the consistency of the OC-SORT tracks, which, even if not perfect, has a perceptually more uniform quality.



(a) Comparison of largest duration tracks

(b) Comparison of smaller duration tracks

Figure 5.5: Comparison of quality (intensity), length (width of each horizontal bar), and quantity (number of horizontal bars) of Birdtrack® (left of the middle line), and OC-SORT (right of the middle line). In this visualization, tracks were sorted by the total length in descending order. The image was split in half for a better layout, on the left is the top half of the lengthier tracks, and on the right is the bottom half showing the shorter tracks. Even though Birdtrack® has more tracks in total, OC-SORT presents itself as a more robust algorithm for multiple bird tracking.

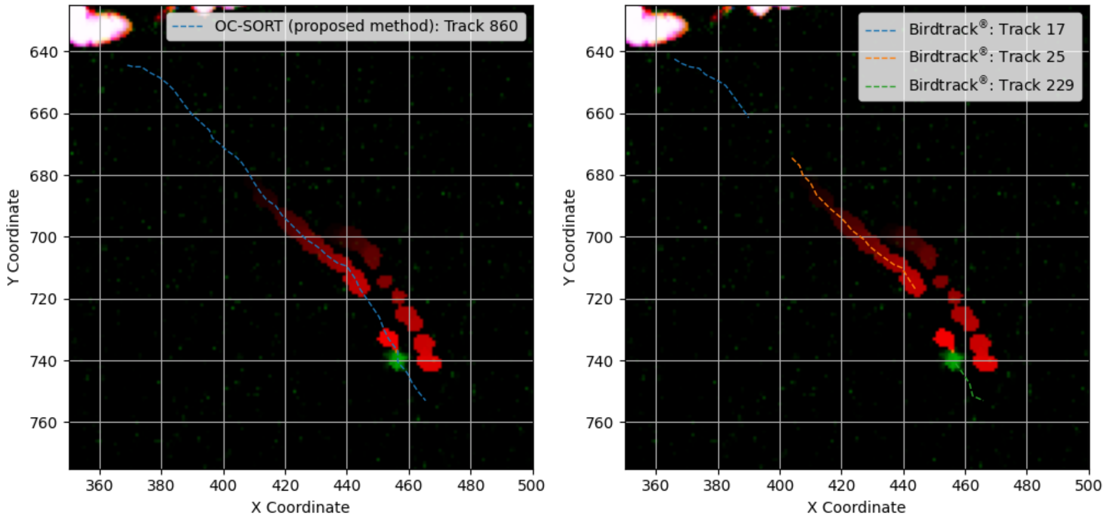


Figure 5.6: Example of track fragmentation in Birdtrack® (right) that is correctly tracked with the proposed OC-SORT algorithm (left).

6
Discussion

6.1 Conclusions and Future Work

This master's thesis has endeavored to advance the capabilities of existing radar-based bird monitoring systems. The Birdtrack® radar system, a critical tool in mitigating bird fatalities at wind farms, was identified as a relevant case study. The research focused on refining the system radar data preprocessing, bird detection, and tracking algorithms. These enhancements are crucial in the broader context of sustainable energy and avian conservation, addressing the delicate balance between energy production and ecological preservation.

The thesis presented a thorough examination of the existing Birdtrack® system, identifying its limitations in differentiating birds from environmental clutter and in accurately tracking multiple birds under dynamic conditions. The proposed methodology integrates both classical and innovative radar data preprocessing methods, intuitive visualization techniques, and the development of robust bird detection and tracking techniques. A significant contribution was the employment of the Observation-Centric SORT (OC-SORT) algorithm, demonstrating major improvements in the accuracy and efficiency of bird detection and tracking.

Experimental results, derived from diverse real-world wind farm case studies, validated the effectiveness of the proposed system. Metrics such as track velocity, angular trajectory changes, track duration, and quality assessment scores highlighted the system superiority over the existing Birdtrack® technology. The findings revealed substantial advancements in tracking accuracy, reduced fragmentation, and increased the bird detection reliability, by significantly contributing to the field of radar-based avian monitoring.

Looking forward, there are several potential directions to further refine and enhance real-time bird tracking systems, as we discuss next.

Improvement of Radar Data Preprocessing

Future research could focus on enhancing the missing line in-painting and denoising process in radar data preprocessing. The current method, while effective, could benefit from the application of advanced techniques like Masked Auto-Encoder (MAE-based) [86], Generative Adversarial Networks (GAN-based) [87], or Denoising Diffusion Probabilistic Models (DDPM-based) [88] interpolators. These approaches could offer a more accurate and realistic reconstruction of missing radar lines, as well as noise reduction, potentially improving the overall quality of the radar imagery, given their current success as generative models.

Advanced Bird Detection Algorithms

The adoption of learned models for bird detection represents a promising avenue. Techniques such as Single-Shot Detector (SSD) [38], You Only Look Once (YOLO) [30, 31, 32, 33, 34, 35, 36, 37], and innovative approaches like ClusterNet [29] or Video-kMaX [78] could be explored. These models, particularly those aimed at handling spatio-temporal information, could provide more precise detection capabilities, especially for small objects in large scenes.

Enhanced Multiple Bird Tracking

Incorporating latent features into the tracker state or as an association metric could significantly improve the tracking algorithm. For instance, NVIDIA visual tracker based on the discriminative correlation filter (NvDCF), or advanced versions of Deep SORT [74] like StrongSORT [75], could be explored. These enhancements could provide more accurate and reliable tracking of birds by leveraging the latest advancements in machine learning and computer vision.

Integration with Environmental Data

Further integration of prior and side information such as environmental data (e.g., weather conditions), bird migration patterns, and ecological studies, could augment the system performance. This integration could lead to more informed and context-aware detection and tracking, adjusting system parameters dynamically based on environmental factors.

Developing a framework for real-time adaptive system configuration, where the system parameters are adjusted dynamically based on the current operational conditions, could further enhance the system effectiveness.

Extensive Validation with Broader Data Sets

Conducting extensive validation and testing using a broader range of data sets from various geographical locations and under different environmental conditions would be beneficial. This would help in understanding the system performance in diverse scenarios and further refining the algorithms.

Collaboration with Ecological Experts

Collaborating closely with ecological experts and ornithologists could provide valuable insights into bird behaviors, which could be used to fine-tune the detection and tracking algorithms, as well as to come up with better metrics for measuring the efficacy of the algorithms being developed. This collaboration could also aid in the validation of the system with more reliable ground truth data. In particular, the data that we used for the current evaluation is unlabeled and this poses significant challenges in the verification of the performance as no ground truth is available. Further study would be needed to either define reliable performance proxies that can be computed in an unsupervised manner from unlabeled data or automatic data labelling tehcniques for the problem at hand.

In summary, this thesis has made significant contributions to radar-based avian monitoring, addressing key challenges in bird detection and tracking in wind farm environments. The present work paves the way to future research and development in this domain, contributing to the harmonious coexistence of sustainable energy production and wildlife conservation.

6.2 Final Remarks

This thesis started with an ambitious goal: the semantic segmentation of dynamic clutter sources in radar data. However, as the research progressed, several significant challenges became evident, necessitating a pragmatic shift in the original focus.

The main challenge was the absence of reliable ground truth data, a critical element for the training and validation of semantic segmentation models. The complexity of training unsupervised models at high resolutions, coupled with the lack of a comprehensive data catalog for archived radar sessions, further compounded the difficulty of the original objective. Moreover, the limited monitoring performance of the existing Birdtrack® system hindered the ability to effectively benchmark our new techniques and enhance their functionalities.

These difficulties led to refocusing the research scope towards a more achievable, yet challenging, target: enhancing bird detection and tracking capabilities. This new goal, while necessitated by the encountered constraints, allowed for significant advancements within the available resources and timeframe.

The success achieved in enhancing the detection and tracking algorithms, as evidenced by the improved performance of the proposed tracking algorithm, is remarkable. Notably, this work was undertaken by a single developer over a few months, contrasting with the efforts of a team of experts over several years. This context highlights the notable achievements made under challenging circumstances.

Moving forward, several future directions are proposed to continue the advancement of radar-based avian monitoring systems. These include the creation of a comprehensive data catalog, the acquisition of reliable ground truth data, the integration of robust performance monitoring tools, the exploration of supervised learning techniques, and securing adequate resources for high-resolution model training. Additionally, an iterative approach to model development is suggested, starting with simpler models and gradually increasing its complexity.

In conclusion, this thesis journey, from its ambitious inception to its strategic adaptation and significant achievements, reflects the dynamic and evolving nature of interdisciplinary research at the intersection of technology and ecology. The insights gained and the future directions proposed not only set the stage for continued progress in radar-based avian monitoring, but may also serve as valuable guidelines for similar research endeavors in wildlife conservation.

A

Dynamic Clutter Identification

During the initial stages of this master's thesis, various exploratory techniques were investigated to address the challenge of dynamic clutter rejection in radar data. The focus was on unsupervised methods that could potentially enhance the system's ability to distinguish between avian targets and environmental clutter. This appendix details some of these preliminary explorations, which even though weren't used in the final approach, served as an in-depth exploration of potential creative approaches to the problem of dynamic clutter. The techniques discussed here include interpretations of radar data in a (2+1)-dimensional space, feature extraction methods such as Scale-Invariant Feature Transform (SIFT), Uniform Manifold Approximation and Projection (UMAP), and Gunnar Farnebäck's two-frame Dense Optical Flow Estimation.

A.1 Interpreting Radar Data in (2+1)-dimensional Space

An initial exploration involved interpreting radar data within a (2+1)-dimensional framework. This approach aimed to incorporate both spatial dimensions and the temporal aspect of radar scans, providing a more comprehensive understanding of the dynamic environment. The idea was to leverage this multi-dimensional interpretation to better identify and track the movement of avian species, while effectively filtering out static and dynamic clutter.

By layering radar frames on top of each other, a 3D volume representing the horizontal radar data over time can be created. Two axes show the spatial data, and the third axis shows the time of capture. This 3D approach can reveal how bird activities change over periods (Refer to Fig. A.1 for a 3D plot).

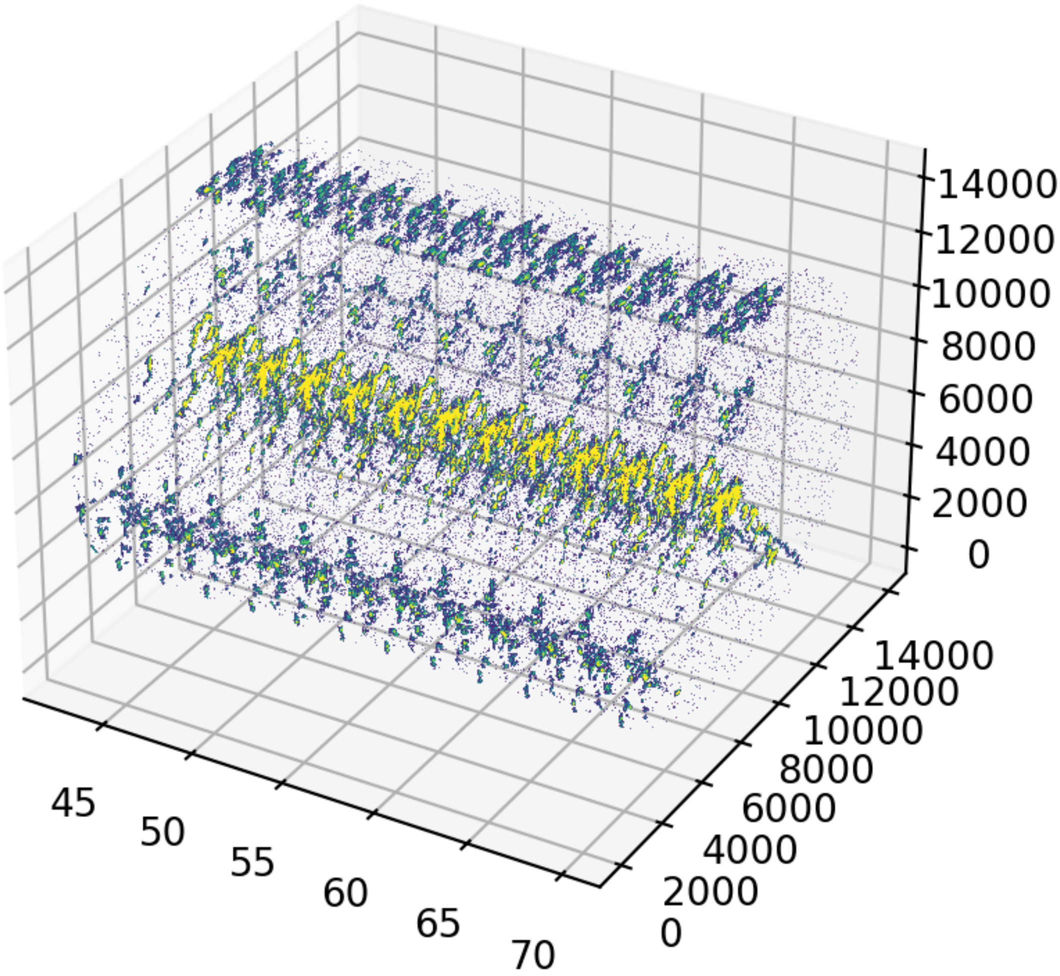


Figure A.1: Radar data in 3D Cartesian coordinates. In this image time, in seconds goes from left to right, with the spatial coordinates being on the remaining two axis. This provides a new way to visualize the radar data sessions from a volumetric point of view, allowing different analysis techniques that incorporate the temporal dimension.

A.2 Scale-Invariant Feature Transform (SIFT)

The application of the Scale-Invariant Feature Transform (SIFT) in radar data analysis was an exploratory step taken in this research. SIFT, a robust feature detection algorithm widely used in computer vision, was employed to identify consistent patterns or features across multiple radar frames. The intention was to utilize SIFT for tracking the movement trajectory of bird flocks, leveraging its resilience to changes in image scaling, rotation, and illumination. This section details the implementation of SIFT in radar data, its strengths, challenges encountered, and the rationale behind its eventual exclusion from the final system.

SIFT's Application and Challenges in Radar Data

SIFT was applied to radar data to detect key points that could be matched across frames, thereby facilitating the tracking of avian movements. Its capability to consistently identify features despite variances in scaling and rotation was deemed advantageous for radar imagery analysis. However, the application of SIFT in this context presented notable challenges:

1. **Detection of Positive and Negative Space:** SIFT demonstrated a tendency to detect features in both positive space (indicating the presence of an object) and negative space (indicating the absence of an object). This characteristic, while useful in some contexts, posed a challenge in accurately isolating bird movements from the radar background.

2. **Sensitivity to Object Edges:** A significant issue encountered was SIFT's heightened sensitivity to object edges. This led to the over-detection of features, particularly around the peripheries of larger objects, resulting in an excess of smaller, often irrelevant features.

Observations and Findings

Upon executing the SIFT algorithm on radar data, a collection of key points along with their corresponding feature descriptors was generated. These key points were then cross-referenced across different frames to identify consistent objects. Figure Fig. A.2 illustrates the efficacy and limitations of SIFT in radar data:

- **BSJ Project – Clear Conditions:** The left frame from the BSJ project shows SIFT's performance in a scenario devoid of dynamic clutter, highlighting its potential in cleaner environments.

- **BSJ Project – Rainfall Conditions:** The central frame, also from the BSJ project, captured during heavy rainfall, exhibits the challenges of applying SIFT in conditions with dynamic clutter.

- **Genesis Project – Insect Clouds:** The rightmost frame from the Genesis project demonstrates the impact of insect clouds on SIFT key points, further emphasizing the difficulty in distinguishing relevant features.

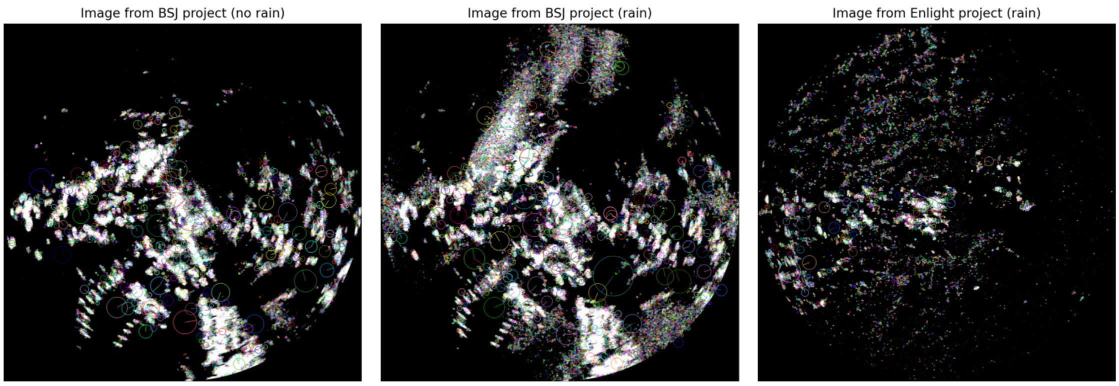


Figure A.2: This figure showcases the results of running the SIFT detector on three distinct frames. (left) A clean frame from the BSJ project without any dynamic clutter. (center) Shows the results of SIFT detection on radar from the BSJ project operating under heavy rain. (right) SIFT key points of the Genesis project under the influence of rainfall and insect clouds.

Conclusion and Decision to Discard SIFT

Despite its initial promise, the SIFT algorithm was eventually discarded in favor of a simpler Connected Components algorithm. The latter demonstrated a higher reliability in consistently identifying birds across various test sessions. The abundance of irrelevant keypoints generated by SIFT, and the associated challenge in filtering these out, rendered it less suitable for the specific requirements of radar-based avian monitoring. The decision to adopt the Connected Components algorithm was driven by its effectiveness and simplicity, aligning better with the practical needs of the Birdtrack® system

A.3 Hierarchical Density-Based Spatial Clustering (HDBSCAN)

Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) was explored as a potential clustering algorithm for analyzing radar data, particularly in classifying bird movements. HDBSCAN's unique approach to clustering, based on varying data point densities, offered an intriguing solution for segmenting radar data into distinct clusters.

Application and Challenges of HDBSCAN

HDBSCAN's main advantage in radar data analysis is its ability to adaptively form clusters without predefining the number of clusters. This feature is especially beneficial for handling unpredictable or varied data patterns typically observed in radar scans.

Key challenges encountered with HDBSCAN include:

- **Difficulty with Overlapping Objects:** In instances where objects overlap or intersect, HDBSCAN tends to merge them into a single cluster. This is due to its density-based clustering approach, which does not effectively distinguish between closely situated but distinct objects.

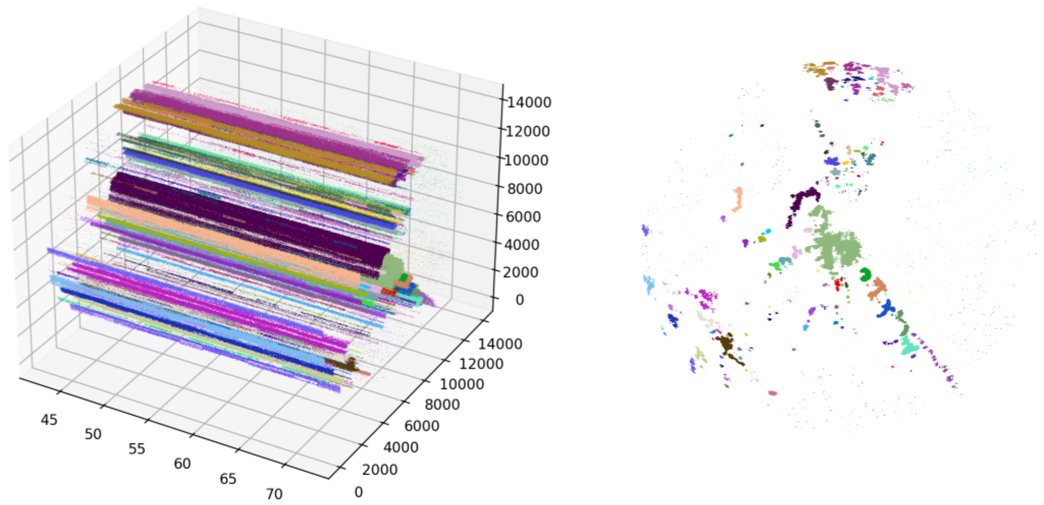
- **Misinterpretation of Low-Density Clusters:** The algorithm struggled with clustering bird flocks exhibiting low density, often merging them with clutter. This resulted in difficulties in differentiating between actual bird movements and environmental noise.

Visual Representation and Analysis

For a more intuitive analysis, HDBSCAN clusters were visually represented, with each cluster being marked by unique colors or labels. This approach provided clear insights, particularly in radar data with high-density patterns.

- **Volumetric Representation:** The left panel in Fig. A.3 illustrates a volumetric representation of radar data, where the temporal dimension extends from left to right, and the spatial dimensions are shown on the other axes. To ensure density-connection across time and maintain consistent object categorization, noise was added to the temporal dimension.

- **Slice Visualization:** The right panel in Fig. A.3 displays a slice of the volumetric data, emphasizing HDBSCAN's capacity to cluster related components within a single frame, even if they are disjoint.



(a) 3D plot **(b)** 2D slice of 3D volume

Figure A.3: (left) A volumetric representation of the radar data is presented with time from left to right and space represented in the remaining two axes. Noise was added to the temporal dimension of the session represented as a 3D volume in an attempt to keep regions of interest in each frame density-connected across time. This leads to a uniform categorization of objects across time (as shown by the unique colors assigned to them). (right) A slice of the volume on the left demonstrates the clustering capabilities within a single frame.

Decision Against Using HDBSCAN

Despite its potential, HDBSCAN was ultimately not employed in the final system due to its limitations in distinguishing bird flocks from dynamic clutter. The consistent merging of low-density bird clusters into single entities, and the subsequent confusion with environmental clutter, posed significant challenges. As a result, alternative methods were sought that could offer more precise and discernible clustering of avian movements in radar imagery.

A.4 Uniform Manifold Approximation and Projection (UMAP)

Uniform Manifold Approximation and Projection (UMAP) is an advanced nonlinear dimensionality reduction technique explored in this research to address the complexities of rainfall clutter in radar imagery. The application of UMAP involved its use for dimensionality reduction and data visualization, aiming to represent the neighborhood around each radar pixel in a lower-dimensional space while preserving essential data structures.

UMAP's Principle and Application in Radar Data

UMAP operates on the principle of preserving both local and global structures within data. It constructs a topological representation of the data's underlying manifold and projects this onto a lower-dimensional space, maintaining the integrity of neighborhood relationships. The algorithm's balance of repulsive and attractive forces ensures uniform data point spacing and the preservation of local structures.

UMAP's strengths in radar data analysis include:

- **Preservation of Data Structures:** Its ability to capture both local and global data patterns, crucial for identifying complex relationships within radar imagery.
- **Nonlinear Embedding:** Unlike linear techniques like PCA, UMAP effectively captures nonlinear patterns, crucial for analyzing intricate radar data.
- **Data Visualization:** UMAP's utility in visualizing high-dimensional data in a lower-dimensional space helps reveal hidden clusters and patterns.

However, UMAP's performance is subject to certain limitations:

- **Hyperparameter Sensitivity:** Careful selection of hyperparameters like the number of neighbors and minimum distance is crucial for optimal performance.
- **Dependence on Distance Function:** The algorithm's effectiveness relies heavily on the selected distance function. Ideal results require a distance function that leverages learned embeddings of different clutter sources.
- **Random Initialization Variability:** UMAP's random initialization may lead to slight variations in results across runs.
- **Computational Complexity:** Larger datasets and dimensions increase UMAP's computational demands.

UMAP Diagnostic and Hyperbolic Projections

UMAP's application to radar data, particularly after temporal context aggregation, provided insights into embedded patterns and structures. Diagnostic plots, as seen in Fig. A.4, display the data distribution in reduced space, indicating the retention of intrinsic data features during dimensionality reduction.

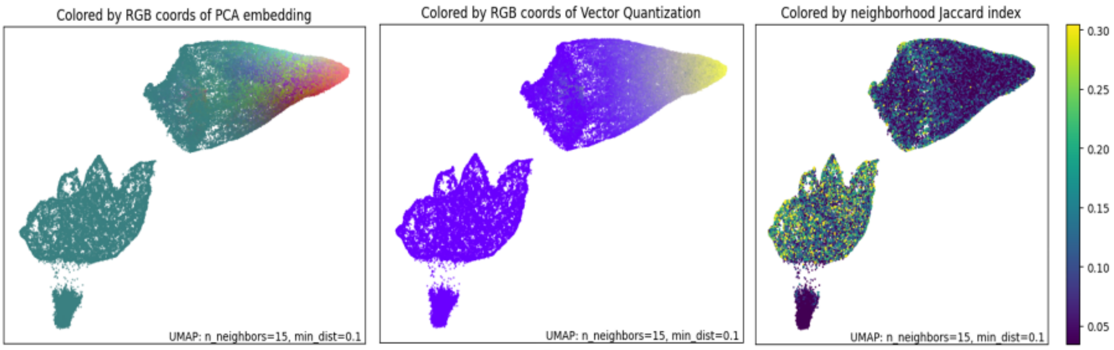
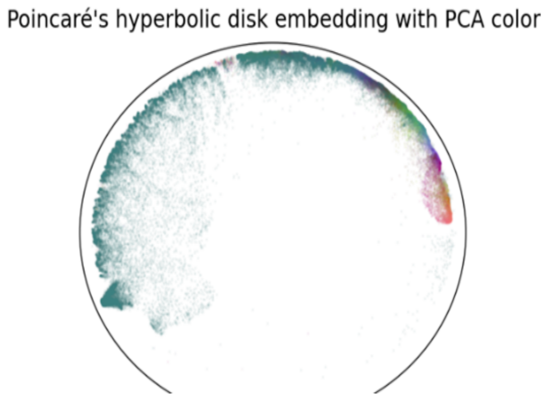


Figure A.4: Diagnostic plots of UMAP on 32x32 RGB crops of the radar images after temporal context aggregation. Smooth gradients are good indicators of high-quality embeddings.

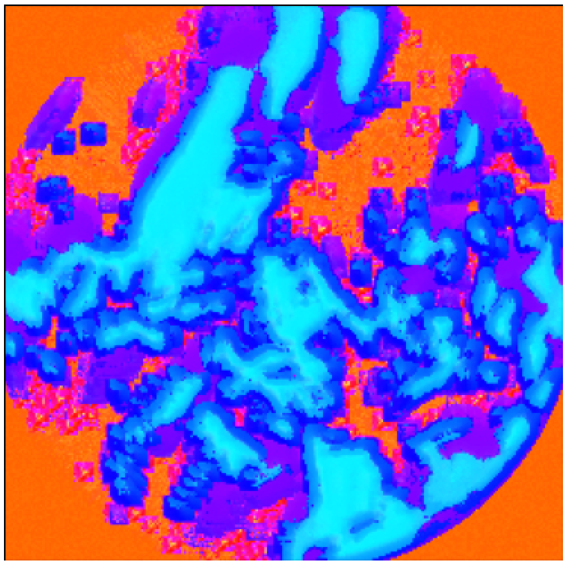
Further visualization in Fig. A.5 offers additional insights:

- **Poincaré Hyperbolic Disk Projection:** The left subfigure visualizes UMAP projections onto the Poincaré Disk, a tool for mapping infinite planes into a bounded disk while maintaining meaningful distance representation.

- **HSV Color Space Visualization:** The right subfigure utilizes HSV color space for radar frames during heavy rainfall (Fig. A.6). Hue and Saturation are derived from the embedding's angle and magnitude, respectively, in the Poincaré Disk, enhancing visual interpretation.



(a) UMAP projection to Poincaré's Hyperbolic Disk.



(b) Radar frame under heavy rain colored by HSV representation of the projection on the left. Hue and saturation are defined as the angle and magnitude of the crop's embedding on the left image. Value is set as 1 in this image, but it can also be modulated by the reflection intensities at each point painting a more informative picture of the current frame.

Figure A.5: Using UMAP for feature visualization.

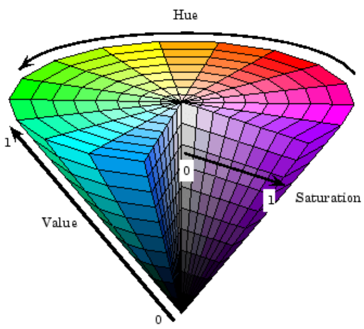


Figure A.6: HSV colorspace representation

Future Directions

While UMAP successfully differentiated clutter from targets, its primary signal appeared reliant on overall neighborhood intensity. Key areas for future exploration include:

- **Robust Feature Extraction for Distance Metrics:** Implementing deep learning methods for feature extraction to improve the distance metric selection.

- **Aligned UMAP for Consistent Initialization:** Utilizing Aligned UMAP to align new embeddings with previous frames, enhancing consistency.
- **Parametric UMAP with Convolutional Encoding:** Adapting UMAP to expedite the inference process on full images and reduce computational complexity.

Due to time constraints, these optimizations were not fully implemented and are recommended for future research.

A.5 Gunnar Farnebäck's two-frame Dense Optical Flow Estimation

This section delves into the application of Gunnar Farnebäck's two-frame Dense Optical Flow Estimation algorithm for motion analysis in radar imagery. Focused on addressing the challenge of rainfall clutter, this method is pivotal in enhancing object-tracking capabilities within the radar data context.

Application and Strengths of the Algorithm

Gunnar Farnebäck's algorithm estimates the dense optical flow between successive radar frames, elucidating the movement of objects within the scene. It computes a flow field for each frame, capturing the magnitude and direction of object motion, essential for tracking birds amid rainfall clutter.

Key strengths of this algorithm include:

- **Comprehensive Motion Capture:** The dense optical flow estimation provides a detailed understanding of movement patterns, offering advantages over sparse methods that only track discrete points.
- **Enhanced Object Tracking:** By accurately capturing motion, the algorithm aids in distinguishing moving objects like birds from static background elements.

Challenges and Limitations

Despite its effectiveness, Gunnar Farnebäck's method encounters certain limitations:

- **Difficulty with Repetitive Textures and Fast-Moving Objects:** The algorithm may struggle in areas with uniform textures or rapidly moving objects, leading to inaccurate flow estimations.
- **Susceptibility to Noise and Occlusions:** Radar image noise or blockages can adversely affect the motion estimation, potentially resulting in errors.

Optimizing Time Intervals and Visualization

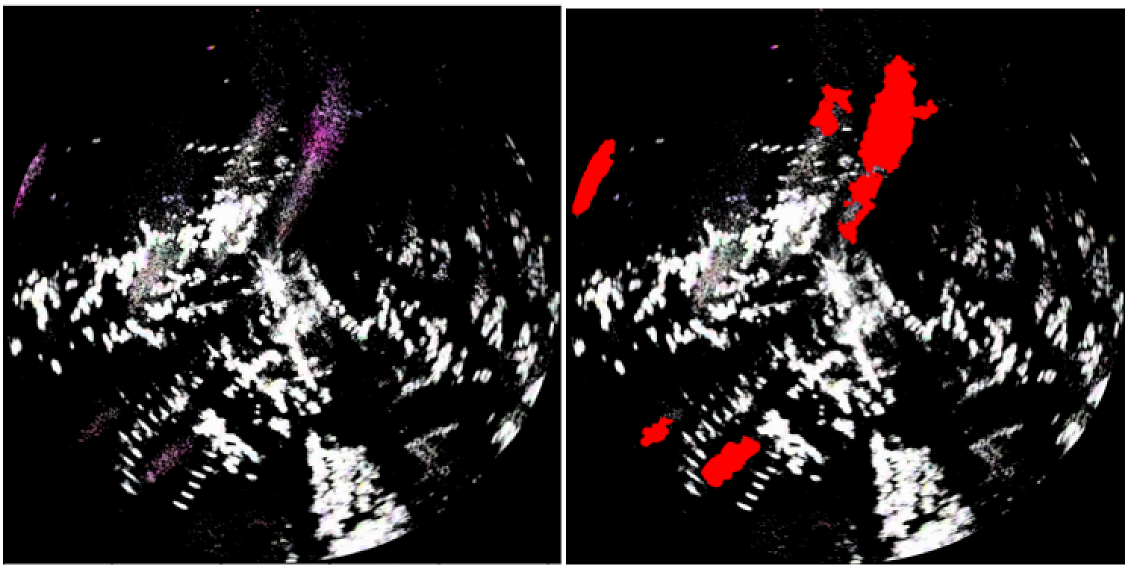
To optimize motion estimation, various time intervals between frames were tested. The most effective results were observed with a 10-frame (25-second) interval. Shorter intervals failed to consistently detect movement, while longer ones exaggerated motion, leading to inaccuracies. Fig. A.7 showcases results for different frame intervals, highlighting the algorithm's capacity to track object movement within radar data.

(a) Raw frame for reference **(b)** Optical Flow with frame deltas of (left) 1, (center) 10, and (right) 20.

Figure A.7: Optical Flow parameter optimization

Dense Optical Flow for Clutter Rejection

The dense optical flow results were utilized to filter the base image, focusing on extracting the largest connected flow components. This approach serves as a baseline for cloud filtering, targeting moving objects that cover extensive radar range areas. Fig. A.8 presents the outcomes of this filtering process, demonstrating the algorithm's efficacy in enhancing object detection in challenging rainfall clutter conditions.



(a) Raw frame overlaid with optical flow estimation **(b)** Masks obtained using this procedure.

Figure A.8: Optical Flow for Dynamic Clutter Rejection

Reflections

The application of Gunnar Farnebäck's Dense Optical Flow Estimation algorithm significantly enriches the analysis of object movement in radar imagery. This enhancement is instrumental in advancing radar-based bird detection systems, offering a potential solution to the rainfall clutter challenge.

Looking forward, exploring a combination of various feature extraction methods might further mitigate the effects of rainfall clutter. Such an approach would necessitate a careful consideration of computational demands, ensuring an optimal balance between accuracy improvement and resource allocation.

References

[1] A. M. Dokter, J. Shamoun-Baranes, M. U. Kemp, S. Tijm, and I. Holleman, "High altitude bird migration at temperate latitudes: A synoptic perspective on wind assistance," *Plos One*, 2013.

[2] I. M. Rosa, G. Palminha, H. Costa, M. Mascarenhas, C. Fonseca, and J. Bernardino, "Classification success of six machine learning algorithms in radar ornithology," *Ibis*, 2015.

[3] M. A. Richards, J. A. Scheer, and W. Holm, "Principles of modern radar: Basic principles," 2010.

[4] M. Desholm, A. D. Fox, P. D. L. Beasley, and J. Kahlert, "Remote techniques for counting and estimating the number of bird-wind turbine collisions at sea: A review," *Ibis*, 2006.

[5] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye. (2019) Object detection in 20 years: A survey.

[6] W. Luo, J. Xing, A. Milan, X. Zhang, W. Liu, and T.-K. Kim, "Multiple object tracking: A literature review," 2014.

[7] J. Ou, J. Zhang, and R. Zhan, "Processing technology based on radar signal design and classification," *International Journal of Aerospace Engineering*, 2020.

[8] J. Lien, N. Gillian, M. E. Karagozler, A. P. M, C. Schwesig, E. R. Olson, H. K. Bhai Raja, and I. Poupyrev, "Soli," *Acm Transactions on Graphics*, 2016.

[9] B. D. Hirth, J. L. Schroeder, W. S. Gunter, and J. Guynes, "Coupling doppler radar-derived wind maps with operational turbine data to document wind farm complex flows," *Wind Energy*, 2014.

[10] M. d. Lucas, M. Ferrer, M. J. Bechard, and A. Muñoz, "Griffon vulture mortality at wind farms in southern spain: Distribution of fatalities and active mitigation measures," *Biological Conservation*, 2012.

[11] R. Fijn, K. L. Krijgsveld, M. Poot, and S. Dirksen, "Bird movements at rotor heights measured continuously with vertical radar at a <scp>d</scp> utch offshore wind farm," *Ibis*, 2015.

[12] M. d. Lucas, G. F. E. Janss, and M. Ferrer, "The effects of a wind farm on birds in a migration point: The strait of gibraltar," *Biodiversity and Conservation*, 2004.

[13] A. Salari, A. Djavadifar, X. Liu, and H. Najjaran, "Object recognition datasets and challenges: A review," *Neurocomputing*, vol. 495, pp. 129–152, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092523122200039X

[14] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, p. 303–338, Sep. 2009. [Online]. Available: http://dx.doi.org/10.1007/s11263-009-0275-4

[15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[16] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.

[17] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár. (2014) Microsoft coco: Common objects in context.

[18] A. Kuznetsova, H. Rom, N. Alldrin, J. Uijlings, I. Krasin, J. Pont-Tuset, S. Kamali, S. Popov, M. Malloci, A. Kolesnikov, T. Duerig, and V. Ferrari, "The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale," 2018.

[19] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese. (2019) Generalized intersection over union: A metric and a loss for bounding box regression.

[20] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren. (2019) Distance-iou loss: Faster and better learning for bounding box regression.

[21] N. Ravi, S. Naqvi, and M. El-Sharkawy, "Biou: An improved bounding box regression for object detection," *Journal of Low Power Electronics and Applications*, vol. 12, no. 4, p. 51, Sep. 2022. [Online]. Available: http://dx.doi.org/10.3390/jlpea12040051

[22] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. I–I.

[23] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, 2005, pp. 886–893 vol. 1.

[24] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[25] R. Girshick, J. Donahue, T. Darrell, and J. Malik. (2013) Rich feature hierarchies for accurate object detection and semantic segmentation.

[26] R. Girshick. (2015) Fast r-cnn.

[27] S. Ren, K. He, R. Girshick, and J. Sun. (2015) Faster r-cnn: Towards real-time object detection with region proposal networks.

[28] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. (2016) Feature pyramid networks for object detection.

[29] R. LaLonde, D. Zhang, and M. Shah. (2017) Clusternet: Detecting small objects in large scenes by exploiting spatio-temporal information.

[30] J. Terven and D. Cordova-Esparza. (2023) A comprehensive review of yolo: From yolov1 and beyond.

[31] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. (2015) You only look once: Unified, real-time object detection.

[32] J. Redmon and A. Farhadi. (2016) Yolo9000: Better, faster, stronger.

[33] ——. (2018) Yolov3: An incremental improvement.

[34] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao. (2020) Yolov4: Optimal speed and accuracy of object detection.

[35] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei. (2022) Yolov6: A single-stage object detection framework for industrial applications.

[36] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao. (2022) Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors.

[37] D. Reis, J. Kupec, J. Hong, and A. Daoudi. (2023) Real-time flying object detection with yolov8.

[38] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," 2015.

[39] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. (2017) Focal loss for dense object detection.

[40] H. Law and J. Deng. (2018) Cornernet: Detecting objects as paired keypoints.

[41] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. (2016) Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size.

[42] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications.

[43] X. Zhang, X. Zhou, M. Lin, and J. Sun. (2017) Shufflenet: An extremely efficient convolutional neural network for mobile devices.

[44] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," 2019.

[45] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee. (2021) A survey of modern deep learning based object detection models.

[46] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko. (2020) End-to-end object detection with transformers.

[47] J. Ferryman and A. Shahrokni, "Pets2009: Dataset and challenge," in *2009 Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, 2009, pp. 1–6.

[48] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.

[49] L. Leal-Taixé, A. Milan, I. Reid, S. Roth, and K. Schindler. (2015) Motchallenge 2015: Towards a benchmark for multi-target tracking.

[50] A. Milan, L. Leal-Taixe, I. Reid, S. Roth, and K. Schindler. (2016) Mot16: A benchmark for multi-object tracking.

[51] P. Dendorfer, H. Rezatofighi, A. Milan, J. Shi, D. Cremers, I. Reid, S. Roth, K. Schindler, and L. Leal-Taixé. (2020) Mot20: A benchmark for multi object tracking in crowded scenes.

[52] A. Dave, T. Khurana, P. Tokmakov, C. Schmid, and D. Ramanan. (2020) Tao: A large-scale benchmark for tracking any object.

[53] M. Weber, J. Xie, M. Collins, Y. Zhu, P. Voigtlaender, H. Adam, B. Green, A. Geiger, B. Leibe, D. Cremers, A. Ošep, L. Leal-Taixé, and L.-C. Chen. (2021) Step: Segmenting and tracking every pixel.

[54] R. Sundararaman, C. D. A. Braga, E. Marchand, and J. Pettre. (2021) Tracking pedestrian heads in dense crowd.

[55] S. Anjum and D. Gurari, "Ctmc: Cell tracking with mitosis detection dataset challenge," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 4228–4237.

[56] M. Fabbri, G. Braso, G. Maugeri, O. Cetintas, R. Gasparini, A. Osep, S. Calderara, L. Leal-Taixe, and R. Cucchiara. (2021) Motsynth: How can synthetic data help pedestrian detection and tracking?

[57] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: The clear mot metrics," *EURASIP Journal on Image and Video Processing*, vol. 2008, p. 1–10, 2008. [Online]. Available: http://dx.doi.org/10.1155/2008/246309

[58] J. Luiten, A. Osep, P. Dendorfer, P. Torr, A. Geiger, L. Leal-Taixe, and B. Leibe, "Hota: A higher order metric for evaluating multi-object tracking," 2020.

[59] E. Ristani, F. Solera, R. S. Zou, R. Cucchiara, and C. Tomasi. (2016) Performance measures and a data set for multi-target, multi-camera tracking.

[60] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," 2016.

[61] Z. Wang, L. Zheng, Y. Liu, Y. Li, and S. Wang. (2019) Towards real-time multi-object tracking.

[62] Y. min Song, K. Yoon, Y.-C. Yoon, K.-C. Yow, and M. Jeon, "Online multi-object tracking framework with the gmphd filter and occlusion group management," 2019.

[63] Q. Liu, D. Chen, Q. Chu, L. Yuan, B. Liu, L. Zhang, and N. Yu. (2022) Online multi-object tracking with unsupervised re-identification learning and occlusion estimation.

[64] X. Zhou, V. Koltun, and P. Krähenbühl. (2020) Tracking objects as points.

[65] T. Meinhardt, A. Kirillov, L. Leal-Taixe, and C. Feichtenhofer. (2021) Trackformer: Multi-object tracking with transformers.

[66] Y. Xu, Y. Ban, G. Delorme, C. Gan, D. Rus, and X. Alameda-Pineda. (2021) Transcenter: Transformers with dense representations for multiple-object tracking.

[67] P. Chu, J. Wang, Q. You, H. Ling, and Z. Liu. (2021) Transmot: Spatial-temporal graph transformer for multiple object tracking.

[68] Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, "Fairmot: On the fairness of detection and re-identification in multiple object tracking," 2020.

[69] X. Zhou, T. Yin, V. Koltun, and P. Krähenbühl. (2022) Global tracking transformers.

[70] A. Girbau, F. Marqués, and S. Satoh. (2022) Multiple object tracking from appearance by hierarchically clustering tracklets.

[71] J. Cao, J. Pang, X. Weng, R. Khirodkar, and K. Kitani. (2022) Observation-centric sort: Rethinking sort for robust multi-object tracking.

[72] J. Hyun, M. Kang, D. Wee, and D.-Y. Yeung. (2022) Detection recovery in online multi-object tracking with sparse graph tracker.

[73] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang. (2021) Bytetrack: Multi-object tracking by associating every detection box.

[74] G. Maggiolino, A. Ahmad, J. Cao, and K. Kitani. (2023) Deep oc-sort: Multi-pedestrian tracking by adaptive re-identification.

[75] Y. Du, Z. Zhao, Y. Song, Y. Zhao, F. Su, T. Gong, and H. Meng. (2022) Strongsort: Make deepsort great again.

[76] B. Yan, Y. Jiang, P. Sun, D. Wang, Z. Yuan, P. Luo, and H. Lu. (2022) Towards grand unification of object tracking.

[77] Y. Cui, C. Zeng, X. Zhao, Y. Yang, G. Wu, and L. Wang. (2023) Sportsmot: A large multi-object tracking dataset in multiple sports scenes.

[78] I. Shin, D. Kim, Q. Yu, J. Xie, H.-S. Kim, B. Green, I. S. Kweon, K.-J. Yoon, and L.-C. Chen. (2023) Video-kmax: A simple unified approach for online and near-online video panoptic segmentation.

[79] Y.-H. Wang, J.-W. Hsieh, P.-Y. Chen, M.-C. Chang, H. H. So, and X. Li. (2022) Smile-track: Similarity learning for occlusion-aware multiple object tracking.

[80] Y. Zhang, T. Wang, and X. Zhang. (2022) Motrv2: Bootstrapping end-to-end multi-object tracking by pretrained object detectors.

[81] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of Graphics Tools*, vol. 9, 01 2004.

[82] A. Buades, B. Coll, and J.-M. Morel, "Non-Local Means Denoising," *Image Processing On Line*, vol. 1, pp. 208–212, 2011, https://doi.org/10.5201/ipol.2011.bcm_nlm.

[83] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[84] M. Dell'Amico and P. Toth, "Algorithms and codes for dense assignment problems: the state of the art," *Discrete Applied Mathematics*, vol. 100, no. 1, pp. 17–48, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166218X99001729

[85] A. Sinha, Z. Ding, T. Kirubarajan, and M. Farooq, "Track quality based multitarget tracking approach for global nearest-neighbor association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1179–1191, 2012.

[86] C. Zhang, C. Zhang, J. Song, J. S. K. Yi, K. Zhang, and I. S. Kweon. (2022) A survey on masked autoencoder for self-supervised learning in vision and beyond.

[87] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks in computer vision: A survey and taxonomy," *ACM Computing Surveys*, vol. 54, no. 2, p. 1–38, Feb. 2021. [Online]. Available: http://dx.doi.org/10.1145/3439723

[88] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion models: A comprehensive survey of methods and applications," *ACM Comput. Surv.*, vol. 56, no. 4, nov 2023. [Online]. Available: https://doi.org/10.1145/3626235

Acknowledgments

I am immensely grateful to Dr. Mahmoud Sakr and Maxime Schoemans of ULB for their invaluable contributions, insights, and guidance throughout this exploration. Their expertise and dedication have been instrumental in shaping the direction and depth of this work. Thank you for your unwavering support and for sharing your knowledge generously.

I appreciate the opportunity to collaborate with a reputable institution like Kalanbio. I am deeply grateful, especially to Dr. Nuno Simoes, for hosting me during my research period. Their supportive environment allowed for this work to exist.

I would also like to thank Professor Michele Rossi for his mentorship and support during my research. His knowledge in the field and useful insights have been critical for completing this thesis. I appreciate his time and commitment to supporting my academic development.

I wish to extend my gratitude to STRIX for providing the datasets essential for this research. Their contribution has been crucial in ensuring the depth and accuracy of my analysis. My thanks to the team at STRIX for their support, namely Dr. Luis Pina and Ricardo Oliveira, for their input, challenges to my ideas, and enhancements to my research. Their collaborative approach and support have added value to this study and improved my teamwork and communication abilities.

An Erasmus scholarship provided by ULB has funded this thesis.