Analysis of Aerial Images for Marine Ecosystem Protection

Report prepared by: **NOEL Shanti**

M1 Computer Science Student

Université Côte d'Azur

Supervisor: Martinet Jean

2024/2025

Contents

Abstract

This study aims to develop a generative model based on computer vision methods to predict the distribution of small boats around large vessels in the region of the Lerins Islands, a site where Posidonia meadows face significant anthropogenic pressures. Using drones and aerial imagery, a Gaussian Mixture Model (GMM) was designed to predict the arrangement of small boats around larger ones. This report presents the methodology, results, and perspectives of this project, contributing to a better understanding and monitoring of fragile marine ecosystems.

Figure 1: The Lerins Abbey

1 Introduction

This multidisciplinary project typically involves students from diverse academic backgrounds, especially those from the MARRES program. However, this year, no students from the MARRES program participated in the project. The team consisted of two Master's students in Computer Science: **Alexis Dubarry** and myself. We worked on two separate directions, resulting in two separate reports.

1.1 Context

Posidonia play a crucial role in Mediterranean marine ecosystems. These underwater grasslands, vital for biodiversity, serve as refuges and breeding grounds for numerous marine species, while contributing to seabed stabilization and carbon sequestration. However, these fragile ecosystems are under threat from human activities such as boat anchoring, pollution, and climate change, leading to a gradual decline in their coverage.

The Lérins Islands, located off the coast of Cannes, represent an area where Posidonia is particularly affected. Anthropogenic pressure, especially from ship anchoring in the Frioul Channel, poses a direct threat. It has become imperative to monitor these ecosystems to assess the impact of human activities and develop appropriate conservation strategies.

Drone technology, combined with advanced computer vision methods, offers a unique opportunity to monitor these ecosystems effectively. Aerial images captured by drones provide detailed data on the extent of the meadows, facilitating their analysis and mapping.

1.2 State of the Art

Marine ecosystems, particularly Posidonia oceanica meadows in the Mediterranean, face significant threats from human activities like anchoring and pollution. Recent advancements in remote sensing, including drones and multispectral imaging, provide high-resolution data for habitat monitoring. Computer vision techniques, such as object detection (e.g., YOLO) and segmentation (e.g., U-Net), automate the analysis of aerial and underwater imagery. Predictive models like Gaussian Mixture Models (GMM) and spatial-temporal methods assess human impact and support decision-making. Despite challenges in data availability and generalization, integrating AI with conservation strategies offers a powerful approach to protect sensitive marine environments, as demonstrated in this project.

1.3 Objectives and Research Question

The main objective of this study is to develop a generative model to predict the distribution of small boats around large vessels in sensitive areas like the Lerins Islands. This approach will:

- Improve the understanding of spatial interactions between large and small boats.
- Identify areas with a high potential impact on Posidonia meadows.
- Assist marine managers in decision-making to protect these fragile ecosystems.

Research Question: Can the distribution of small boats around large vessels be accurately predicted using a generative model based on position data and aerial imagery?

2 Materials and Methods

2.1 Data and Tools Used

Data:

- **Aerial Images:** Drone captures showing Posidonia meadows and boat distributions.
- **Manual Annotations:** Positions of large and small boats annotated on the images for model training.

Tools and Libraries:

- **Python:** Main programming language for data processing.
- **Scikit-learn:** Implementation of Gaussian Mixture Models (GMM).
- **Matplotlib and OpenCV:** Visualization and image processing.
- **Numpy:** Used for numerical computations, including the calculation of relative positions between boats, distance computations, and efficient array manipulations essential for processing spatial data.

2.2 Approach and Implementation

The project was carried out in three main phases:

1. **Training Phase:**
    - Users clicked on the images to record the positions of large and small boats.
    - Relative positions of small boats concerning large ones were calculated to feed the GMM.
    - The GMM was trained to learn the spatial distribution of small boats around larger ones.
    - Influence circles (400-pixel radius) were used to constrain predictions to a realistic area.

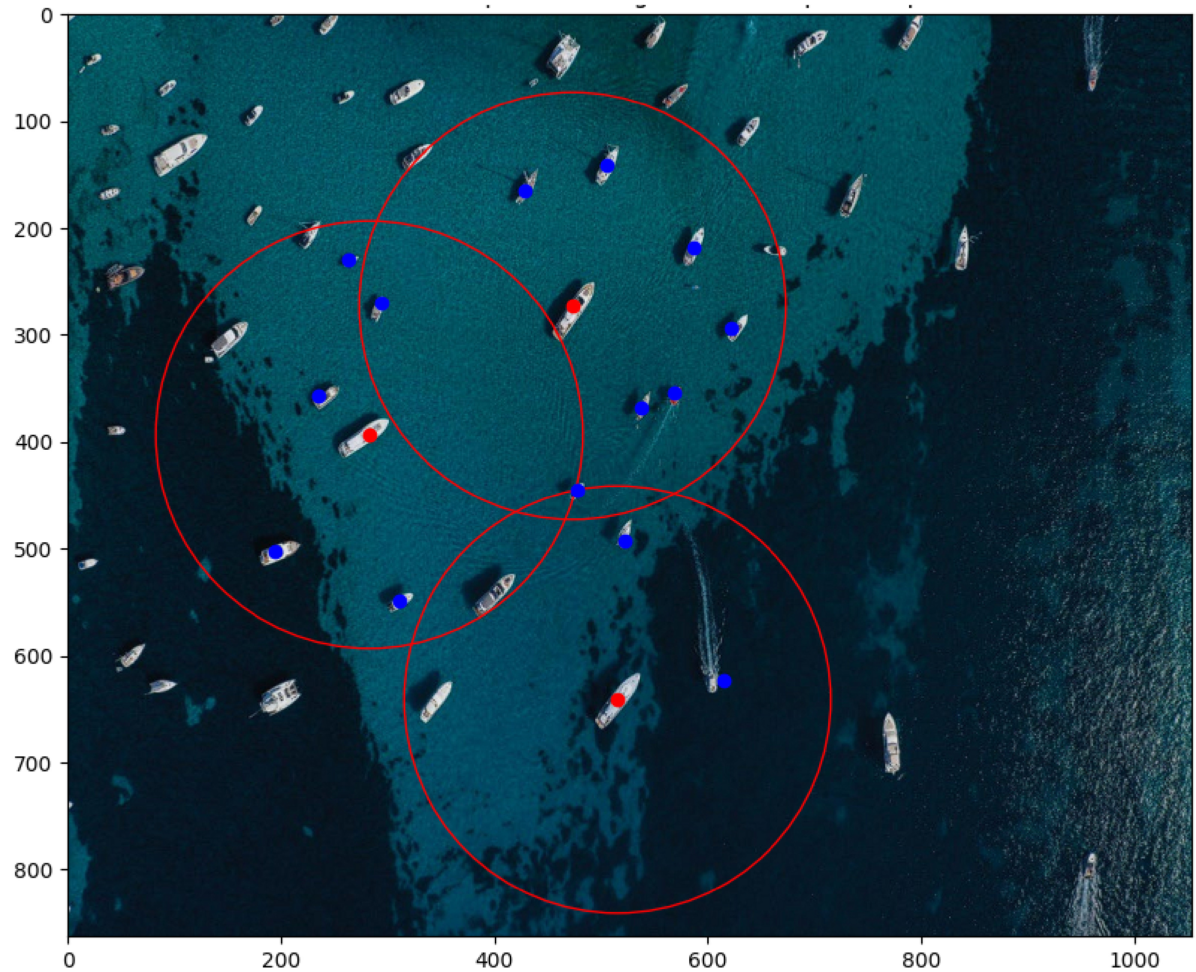
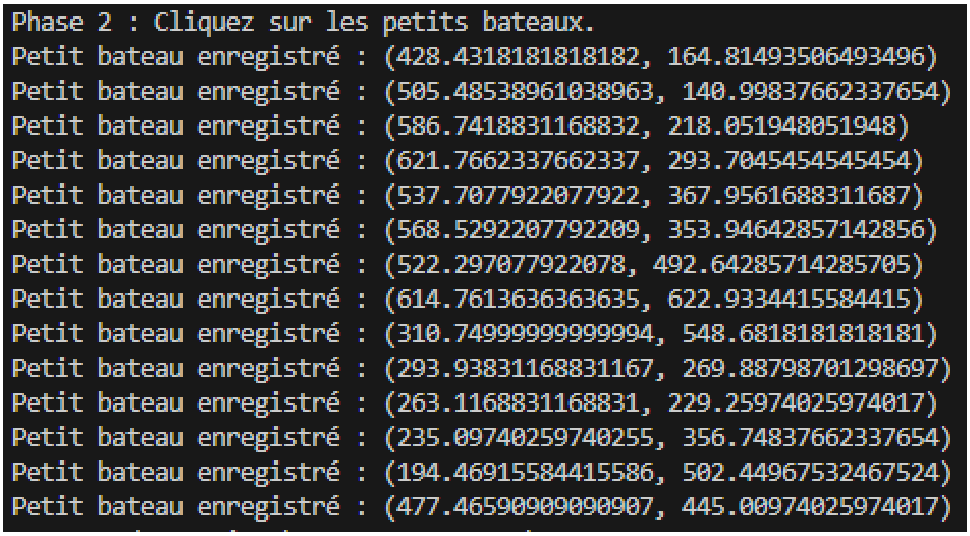


Figure 2: Phase 1 and 2.

```
Gros bateau enregistré : (473.262987012987, 272.68993506493496)
Gros bateau enregistré : (282.7305194805195, 393.1737012987012)
Gros bateau enregistré : (515.2922077922078, 641.1461038961038)
```

Figure 3: Recording of large boats



```
Phase 2 : Cliquez sur les petits bateaux.
Petit bateau enregistré : (428.4318181818182, 164.81493506493496)
Petit bateau enregistré : (505.48538961038963, 140.99837662337654)
Petit bateau enregistré : (586.7418831168832, 218.051948051948)
Petit bateau enregistré : (621.7662337662337, 293.7045454545454)
Petit bateau enregistré : (537.7077922077922, 367.9561688311687)
Petit bateau enregistré : (568.5292207792209, 353.94642857142856)
Petit bateau enregistré : (522.297077922078, 492.64285714285705)
Petit bateau enregistré : (614.7613636363635, 622.9334415584415)
Petit bateau enregistré : (310.74999999999994, 548.6818181818181)
Petit bateau enregistré : (293.93831168831167, 269.88798701298697)
Petit bateau enregistré : (263.1168831168831, 229.25974025974017)
Petit bateau enregistré : (235.09740259740255, 356.74837662337654)
Petit bateau enregistré : (194.46915584415586, 502.44967532467524)
Petit bateau enregistré : (477.46590909090907, 445.00974025974017)
```

Figure 4: Recording of small boats

2. **Prediction Phase:**
   - Users provided hypothetical positions of large boats.
   - The GMM generated probable positions of small boats around these large vessels.

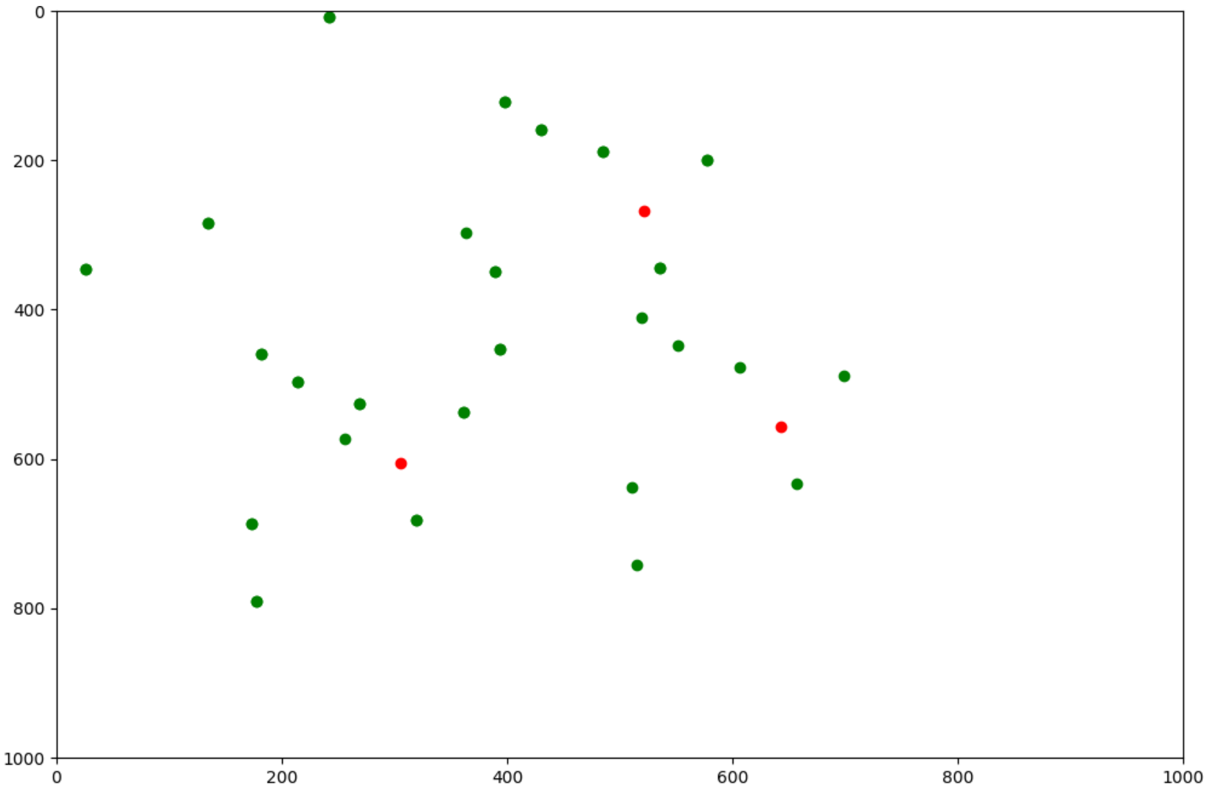
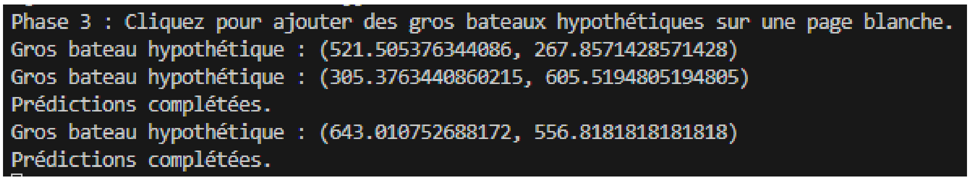
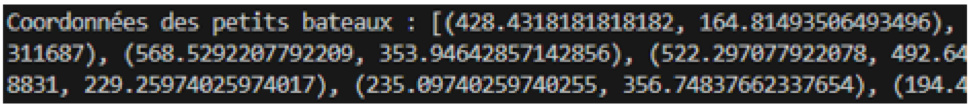


Figure 5: Phase 3.



```
Phase 3 : Cliquez pour ajouter des gros bateaux hypothétiques sur une page blanche.
Gros bateau hypothétique : (521.505376344086, 267.8571428571428)
Gros bateau hypothétique : (305.3763440860215, 605.5194805194805)
Prédictions complétées.
Gros bateau hypothétique : (643.010752688172, 556.8181818181818)
Prédictions complétées.
```

Figure 6: Placement of hypothetical large boats



```
Coordonnées des petits bateaux : [(428.4318181818182, 164.81493506493496),
311687), (568.5292207792209, 353.94642857142856), (522.297077922078, 492.64
8831, 229.25974025974017), (235.09740259740255, 356.74837662337654), (194.4
```

Figure 7: Prediction of hypothetical small boat

2.3 Generative Model and GMM

2.3.1 Generative Model

A generative model is a class of machine learning models used to learn a joint distribution $P(X, Y)$ from data. These models can generate new data similar to the training data by simulating the characteristics of the underlying distribution. In our context, a generative model allows us to simulate the spatial distribution of small boats around large vessels.

Generative models, such as Gaussian Mixture Models (GMM), are particularly useful for:

- Capturing complex patterns in data.
- Generating realistic synthetic data for predictive scenarios.
- Representing probabilistic relationships between different components in a multidimensional space.

2.3.2 GMM Implementation in the Code

In this project, the GMM is used to model the spatial distribution of small boats around large vessels. The process is structured as follows:

1. **Training the GMM:** The relative positions of small boats with respect to large boats are calculated and used as input data for the GMM. For example:

$$(x_{\text{rel}}, y_{\text{rel}}) = (x_{\text{small}} - x_{\text{large}}, y_{\text{small}} - y_{\text{large}}) \tag{1}$$

2. **Prediction with the GMM:** The GMM generates relative positions of small boats around each hypothetical large boat. These positions are transformed into absolute positions:

$$(x_{\text{abs}}, y_{\text{abs}}) = (x_{\text{large}}, y_{\text{large}}) + (x_{\text{rel}}, y_{\text{rel}}) \tag{2}$$

3. **Spatial Filtering:** A filter excludes predictions outside the influence radius (e.g., 400 pixels).

2.3.3 Mathematical Formula Used in the Code

The GMM is implemented using the `GaussianMixture` class from Scikit-learn. The model adjusts the input data by learning the parameters of the normal (Gaussian) distributions that compose the clusters.
In the code, training is performed with the command:

```
model = GaussianMixture(n_components=3, random_state=0).fit(data_relative)
```

Figure 8: GaussianMixture

where `data_relative` contains the relative positions $(x_{\text{rel}}, y_{\text{rel}})$.

- **Mean (μ):** Represents the center of each cluster.
- **Covariance Matrix (Σ):** Describes the shape and size of the cluster.
- **Weights (π):** Proportion of each cluster in the overall distribution.

The mathematical formula used to calculate the probability of an observation X is:

$$P(X) = \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(X|\mu_k, \Sigma_k) \tag{3}$$

Where:

- K is the number of clusters (components).
- π_k is the weight of the k-th component.
- $\mathcal{N}(X|\mu_k, \Sigma_k)$ is the probability density function of a normal distribution, given by:

$$\mathcal{N}(X|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2}(X-\mu)^T \Sigma^{-1} (X-\mu)\right) \tag{4}$$

- d is the dimension of the data (here $d = 2$ for spatial positions).

3 Results

3.1 Model Training

The GMM was trained on the relative positions of small boats concerning large vessels. The results show that:

- Small boats follow a concentrated distribution around large vessels, with an average density of 3 to 5 small boats per large vessel.
- The Gaussian components learned reflect realistic clusters aligned with observations.

3.2 Predictions and Visualization of Distributions

- Predictions of small boats exhibit a distribution consistent with the training data.
- The addition of Gaussian noise diversifies patterns to avoid overly regular layouts.
- Visualizations reveal concentrated clusters within the radius of influence.

4 Conclusions and Perspectives

4.1 Achievements

This study demonstrated the successful application of a Gaussian Mixture Model (GMM) to predict the spatial distribution of small boats around large vessels, offering valuable insights for marine ecosystem protection. Key achievements include:

- Development of an interactive workflow to collect data and generate predictions.
- Effective training and validation of a GMM to replicate realistic spatial patterns of small boats around large vessels.
- Application of predictive modeling to highlight areas of potential ecological impact.

4.2 Future Perspectives

The findings of this project open avenues for further research and practical applications:

- **Dynamic Modeling:** Extending the approach to incorporate temporal data for real-time predictions.
- **Geographical Expansion:** Testing the methodology in other sensitive marine ecosystems.
- **Automated Data Collection:** Integrating detection systems, such as YOLO, to streamline image annotation and improve scalability.
- **Conservation Strategies:** Using the predictive insights to inform policies for anchoring regulation and protected area design.

This project highlights the potential of combining AI-driven generative models with drone technology to address pressing environmental challenges. It sets a strong foundation for further innovation in marine conservation and provides a practical tool to mitigate the impacts of human activities on fragile ecosystems.

Acknowledgments

I would like to thank Jean Martinet and Alexis Dubarry for their support throughout this project. I am also grateful to Jean-Francois Maitre for providing access to the data and images, as well as to Christophe Mocquet.

References

- Gaussian Mixture Models - Wikipedia: https://en.wikipedia.org/wiki/Gaussian_mixture_model
- Scikit-learn Documentation: Gaussian Mixture Models https://scikit-learn.org/stable/modules/mixture.html
- Scikit-Learn: Gaussian Mixture Models (GMM) - YouTube - Engineering Educator Academy https://www.youtube.com/watch?v=djBukz4iamY
- Python Matplotlib Library for Visualization: https://matplotlib.org/
- Numpy Documentation https://numpydoc.readthedocs.io/en/latest/format.html,https://numpy.org/doc/
- Studies on Posidonia Meadows and Anthropogenic Impacts: https://hal.science/tel-03092784
- Alireza Foroutan Github : https://github.com/AlirezaForoutanT/Boat-Detection-and-Classification-YOLOv8
- Cover Page Image: https://www.horizon-lerins.com/billetterie/
- Image of Lerins Abbey: https://divinebox.fr/abbaye-lerins/

Github

- https://github.com/Shanto0o/Boat-Distribution-Prediction