**UNIVERSITY OF SCIENCE AND TECHNOLOGY OF HANOI**

Information and Communication Technology Department



Data Mining and Machine Learning II

**FLOATING WATER OBJECT DETECTION**

Final group project

*Student:*

Đặng Tuấn Hưng - BA12-082

Nguyễn Quốc Huy - BA12-089

Phùng Nam Khánh - BA12-092

Vũ Thành Long - BA12-110

Trương Quang Nam - BA12-134

Ngô Quang Trường - BA12-180

1. **Introduction**

This report presents a machine learning pipeline for detecting and classifying floating objects in aerial images using the AFO dataset, a publicly available dataset from Kaggle. The dataset consists of 3,641 high-resolution drone images, with 39,991 hand-annotated objects, designed to support maritime search and rescue operations. The detection system is implemented in Python and combines traditional computer vision techniques with machine learning algorithms to handle challenges such as small object sizes and class imbalance. The goal is to identify floating objects like humans, boats, and buoys, enabling faster and more accurate responses in emergency situations. This report details the pipeline, methodology, and key components of the system, highlighting its effectiveness in real-world scenarios.

1. **Data analysis**
2. **General Information**

* Dataset name: AFO - Aerial dataset of floating objects.
* Data source: [Kaggle](https://www.kaggle.com/datasets/jangsienicajzkowy/afo-aerial-dataset-of-floating-objects).
* Description: The AFO dataset is designed to train machine learning and deep learning models for maritime search and rescue. It contains approximately 40,000 manually annotated objects in images extracted from drone videos over water. Many objects are small, posing a detection challenge.
* Problem type: Object detection and Classification.
* Annotation Format:Darknet YOLO format (normalized bounding box coordinates).

1. **Images**

* Total Images: 3,641 high-resolution images
* Image Size:
  + Minimum: 237.83 KB
  + Maximum: 3,399.41 KB
  + Average: 1,257.17 KB
* Image Resolution:
  + Minimum: 1920x1080 pixels
  + Maximum: 3840x2160 pixels
* Image Format: .jpg
* Color Mode: RGB

1. **Annotations**

* Total Annotation Files: 3,641 (one per image)
* Total Objects: 39,991 manually annotated objects
* Object Distribution Across Files:
  + Minimum: 0 objects (729 files with no objects)
  + Maximum: 94 objects (ev\_1030.txt)
  + Average: ~11 objects per file
* Classes: 6 categories
  + 0: Human (33,174 instances)
  + 1: Wind/SUP board (3,922 instances)
  + 2: Boat (702 instances)
  + 3: Buoy (587 instances)
  + 4: Sailboat (160 instances)
  + 5: Kayak (1,446 instances)
* Annotation quality:
  + No images lack annotations.
  + 28 incorrect annotations (bounding boxes outside [0,1] range) were identified and removed during preprocessing, as they represent a negligible fraction (0.07%) of the total objects.
* Features**:**

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | Discrete or Continuous | Quantitative or  Qualitative | Numerical or  Categorical |
| Class label | Discrete | Qualitative | Categorical |
| X Center | Continuous | Quantitative | Numerical |
| Y Center | Continuous | Quantitative | Numerical |
| Width | Continuous | Quantitative | Numerical |
| Height | Continuous | Quantitative | Numerical |

* Class label is discrete (finite values) and categorical because it represents different object types.
* Bounding Box Features (x, y, width, height) are continuous because they can take any real value within a range.

1. **Data Preprocessing**

* Dataset Split:
  + Training: 76.6% (2,787 images)
  + Testing: 14.1% (514 images)
  + Validation: 9.3% (339 images)
* Handling Incorrect Annotations: 28 annotations with out-of-range bounding boxes were removed.
* File Organization: Images and annotations are organized into ***dataset/working/images*** and ***dataset/working/labels*** directories, subdivided into train, test, and val folders using the split\_data function.

1. **Pipeline**