Assignment 2

Project Proposals

Group of two (up to three) persons.

For the second assignment take into account the following considerations:

1. *Computer vision applications should be developed in **Pytorch**;*
2. *The **code file** and **report** will be the **colaboratory** to be submitted in Moodle;*
3. *The colaboratory needs to be ready to work (the import of dataset must be automatically performed. Save the colab after running all cells of the program);*
4. *The report in the colaboratory should include:*
   - ***Introduction** summarizing the objectives of the work;*
   - ***Methodology** describing the algorithm or the NN model as well as, the description of the dataset used for training/evaluation;*
   - ***Results** section that demonstrate the improvement/work that was made. Try to focus on benchmarking your algorithm/model or showing quantitative metrics. Special attention must be made in presenting the results.*
5. *Each group must do a 5 minute presentation of the work (focused on the method and the results).*

Proposal 1

Obstacle Semantic Segmentation

Complexity: ★★★ (High)

Motivation:

Semantic segmentation for Autonomous Surface Vehicles (ASVs) is critical in enabling real-time obstacle detection and navigation in complex maritime environments. By accurately segmenting objects like buoys, boats, debris, and natural obstacles such as rocks and marine vegetation at the pixel level, ASVs can safely chart paths, avoid hazards, and make real-time course adjustments, even in dynamic waters. Given the challenges of varied lighting, water reflections, and wave interference, robust segmentation models enhance ASV autonomy, allowing them to reliably interpret and respond to their surroundings. This capability not only improves the safety and efficiency of ASV operations in industries like shipping, environmental monitoring, and offshore maintenance but also broadens their application to more challenging and congested waters where precise obstacle segmentation is essential.

Dataset:

- *Dataset Size:* 2916 LWIR images.
- *Annotations:* Each image is labeled with pixel-level segmentation across seven classes (sky, water, bridge, obstacle, living obstacle, background, self)
- *Location:* Captured in and around Boston Harbor. Collected over two years, the images reflect a wide range of marine environments and conditions, including busy harbors and open waters, across different seasons and times of day.
- *Format:* Each data entry includes an image filename and a corresponding semantic segmentation mask, structured as: *data/filename.png mask/filename.png*

Data: `https://drive.google.com/file/d/1T572f0oqy5JmuTvVEwkSUeXLWOSHl4hL/view`
Label: `https://drive.google.com/file/d/1pHp480_Q-s72RoDf1nD7ERzsv9yZTDE1/view`
(Download Links)

Challenge:

1. Propose a custom AI-model that performs semantic segmentation from the LWIR image.

2. Train and test the AI-model without and with data augmentation.
3. Use the following metrics to assess the quality of your implementation:
   - IoU of training and testing;
   - Precision and recall;
   - Model complexity (# parameters).
4. Compare the results of your AI-model with at least 1 existing models (e.g., Unet with VGG backbone).
5. Discuss the obtained results taking into consideration the following paper: `https://journals.sagepub.com/doi/10.1177/02783649231153020`

References:

- `https://github.com/uml-marine-robotics/MassMIND`
- `https://journals.sagepub.com/doi/10.1177/02783649231153020`
- `https://ieeexplore.ieee.org/document/9659477`

Proposal 2

Buoy Detection and Classification

Complexity: ★ (Low)

Motivation:

Buoy detection is a critical component of autonomous maritime navigation, as buoys serve as essential markers for safe passage, obstacle avoidance, and navigational guidance. However, reliably detecting and classifying buoys in diverse marine environments is challenging due to variable lighting, weather conditions, and reflections on water surfaces. A robust, real-time buoy detection model can transform marine navigation by enabling Autonomous Surface Vehicles (ASVs) and other autonomous vessels to operate safely and efficiently, even in busy harbors or open ocean settings. With advancements in computer vision, especially object detection and classification, the potential to automate this essential navigational task is within reach. Success in this challenge will contribute to enhanced safety, operational autonomy, and environmental monitoring in marine industries, reducing human intervention and setting new standards for autonomous maritime technology.

Dataset:

- *Dataset Size:* 744 Images.
- *Annotations:* Bounding box and 4 classes (east, green, red, west).

`https://universe.roboflow.com/vortexbuoytrainingset/buoy-detection-qzjg1`
(Download Link)

Challenge:

1. Propose a custom AI-model that performs buoy detection and classification from the images. For the output consider the object geometrical center and class. In case of multiple detections consider only the object with the largest bounding box.
2. Train and test the AI-model without and with data augmentation.
3. Use the following metrics to assess the quality of your implementation:
   - RMSE of training and testing;

- Confusion matrix;
- Model complexity (# parameters).

4. Compare the results of your AI-model with at least 1 existing models (e.g., MobileNet, VGG, EfficientNet and ResNet).
5. Discuss the obtained results.

Proposal 3

Classification of defects in photovoltaic modules

Complexity: ★★ (Medium)

Motivation:

Thermal inspection of photovoltaic (PV) modules is a non-invasive technique used to assess the health and performance of solar panels. By capturing and analyzing the heat signatures emitted by PV modules, thermal inspections can detect anomalies such as hotspots, dust accumulation, or damaged cells. Hotspots, in particular, can indicate potential issues like cell degradation or electrical faults. This proactive approach to maintenance allows for the early identification of problems, reducing downtime and maximizing the overall efficiency and lifespan of solar installations. Thermal inspection plays a crucial role in ensuring the reliability and energy yield of PV systems, making it an essential tool in the field of solar energy management and maintenance.

Dataset:

- *Dataset Size:* 20000 images.
- *Annotations:* Each image belongs to one of 12 classes (cell, cell-multi, cracking, hot-spot, hot-spot-multi, shadowing, diode, diode-multi, vegetation, soiling, offline-module and no-anomaly).
- *Format:* image.jpg and JSON file with image path and respective anomaly label.

`https://github.com/RaptorMaps/InfraredSolarModules`
(Download Link)

Challenge:

1. Develop three AI-models to evaluate the status of the PV module using thermal signatures including:
   (a) Model 1: Binary classification (anomaly or no-anomaly);
   (b) Model 2: Classification with 11 anomaly classes (cell, cell-multi, cracking, hot-spot, hot-spot-multi, shadowing, diode, diode-multi, vegetation, soiling and offline-module);

(c) Model 3: Classification with 12 classes (cell, cell-multi, cracking, hot-spot, hot-spot-multi, shadowing, diode, diode-multi, vegetation, soiling, offline-module and no-anomaly).

2. Describe data augmentation techniques that were used.

3. Compare the results of your AI-model with at least 1 existing models (e.g., MobileNet, VGG, EfficientNet and ResNet). Use the following metrics to assess the quality of your implementation:

   - Accuracy (%) and F1-Score (%) of training and testing;
   - Confusion matrix;
   - Model complexity (# parameters).

4. Discuss the results, taking into consideration the following paper:
   `https://www.sciencedirect.com/science/article/pii/S0196890424006599`

References:

- `https://www.sciencedirect.com/science/article/abs/pii/S0263224123006991?via%3Dihub`
- `https://www.sciencedirect.com/science/article/pii/S0196890424006599`

Proposal 4

Protection Anode Detection in Underwater Structures

Complexity: ★★ (Medium)

Motivation:

Sacrificial anodes play a critical role in protecting maritime infrastructure from electrochemical corrosion, which makes them essential components in marine engineering and naval maintenance. By systematically detecting and assessing the condition of these protective elements, the operational life of ships, offshore platforms, marine pipelines, and other critical maritime assets can be substantially extended. Traditional manual inspection methods are time-consuming, potentially hazardous, and prone to human error, making automated detection through advanced computer vision techniques a compelling technological solution.

Dataset:

a) ANODE Dataset:

- *Dataset Size*: 18230 images.
- *Annotations:* Each image has a label file (.txt) that contains 5 values: (i) object class; (ii) x coordinate center point of the bounding box; (iii) y coordinate center point of the bounding box; (iv) bounding box width; (v) bounding box height.
- *Location:* ATLANTIS Test Centre, Viana do Castelo, Portugal; and CRAS (INESC TEC) Interior Pool, Porto, Portugal.
- *Format:* image.png, and label.txt.

`https://rdm.inesctec.pt/dataset/nis-2024-005`
(Download Link)

b) NEREON (Negative class):

- *Dataset Size*: 7000 images.
- *Annotations:* No annotations are provided, as this dataset serves as a negative class for non-anode scenes (i.e., no anode presence or detection).

- *Location:* ATLANTIS Test Centre, Viana do Castelo, Portugal.
- *Format:* image.png.

`https://rdm.inesctec.pt/dataset/nis-2023-002`
(Download Link)

Challenge:

1. Propose a custom AI model that predicts detections of sacrificial anodes from the images.
2. Train and test the AI-model without and with data augmentation.
3. Use the following metrics to assess the quality of your implementation:
   - IoU of training and testing;
   - Precision and recall;
   - Model complexity (# parameters).
4. Compare the results of your AI-model with at least 1 existing model (e.g., MobileNet, VGG, EfficientNet and ResNet).
5. Discuss the obtained results.

Proposal 5

Captcha decoding

Complexity: ★ (Low)

Motivation:

A CAPTCHA (Completely Automated Public Touring test to Tell Computers and Humans Apart) is a commonly used feature in web applications to block non-human access. CAPTCHAs' purpose is to prevent spam on websites, such as promotion spam, registration spam, and data scraping, and bots are less likely to abuse websites with spamming if those websites use CAPTCHA. Many websites use CAPTCHA to prevent bot raiding, and it works effectively. CAPTCHA's design is that humans can complete CAPTCHAs, while most robots can't.

Dataset:

- *Dataset Size:* 14860 images.
- *Format:* Two datasets divided in train and test with the label in the image name (contains 4 or 5 characters):
  - Soft dataset is formed by CAPTCHAs that are more simple. Students must start the project with this dataset;
  - Hard dataset is formed by CAPTCHAs with strange elements added, to make the identification more difficult to predict.

`https://uporto-my.sharepoint.com/:u:/g/personal/up488707_up_pt/EQA2GIkSLUpFrdFDUv4QXcoBrFXHDp-rKySVEMOgZXSftg?e=b2D8xx`
(Download Link)
Login to UP required

Challenge:

1. Propose a custom AI model that decodes CAPTCHA images considering 4 and 5 encoders. The model of the CNN needs to be designed, implemented and trained (no fine tuning approaches should be applied).
2. Use the following metrics to assess the quality of your implementation:

- Train and test accuracy;
- Confusion matrix;
- Others evaluation methodologies (e.g., histograms, model complexity).

3. Discuss the result of your approach, in particular, limitations.

Proposal 6

Semantic Segmentation in Underwater Imagery

Complexity: ★★★ (High)

Motivation:

Underwater object segmentation is crucial for enabling autonomous underwater robotics, providing essential scene understanding capabilities that directly support critical marine exploration and intervention tasks. AUVs (Autonomous Underwater Vehicles) and ROVs (Remotely Operated Vehicles) require robust object segmentation techniques to be able to distinguish between critical elements like marine structures, biological organisms, geological features, and potential hazards, which is fundamental for safe and effective underwater operations. Leveraging Machine Learning models to transform visual data into precise spatial understanding allows robotic systems to navigate, interact with, and analyze complex underwater environments with enhanced perception and decision-making capabilities.

Dataset:

- *Dataset Size:* 1635 images.
- *Annotations:* The objects in each image belong to one of 8 classes (waterbody background, human divers, aquatic plants/flora, wrecks/ruins, robots and instruments, reefs and other invertebrates, fish and other vertebrates, sea-floor and rocks). The pixels of each object are colored according to the corresponding class. The segmentation labels are provided as masks, using a set of 3-bit binary RGB colors to represent the eight categories.
- *Format:* image.jpg and mask.bmp

`https://irvlab.cs.umn.edu/resources/suim-dataset`
(Download Link)

Challenge:

1. Propose a custom AI model to perform object segmentation in underwater images.
2. Train and test the AI-model without and with data augmentation.

3. Use the following metrics to assess the quality of your implementation:
   - IoU of training and testing;
   - Model complexity (# parameters).
4. Compare the results of your AI-model with at least 2 existing models (e.g., ResNet, VGG, or other), one of which should be SUIM-Net (`https://arxiv.org/pdf/2004.01241`).
5. Discuss the obtained results.

References:

- `https://arxiv.org/pdf/2004.01241`

Open Proposal

Self-proposed project

Students can develop a project in CV that is related to their MSc Thesis. The teams should send a project proposal (*maximum 2 pages*) containing the following topics:

- Motivation;
- Objectives;
- Problem statement (eg, classification, regression, etc);
- Dataset.

Minimum Requirement:

At least one dataset fully labeled with an adequate size for the task!