UNIVERSITY OF SCIENCE AND TECHNOLOGY OF HANOI
UNDERGRADUATE UNIVERSITY

Group Project

Plankton Classification Platform for Supporting Research Activities of Environmental Scientists

Report 1

Submitted by

LE Nguyen Khoi	USTHBI7-086
NGUYEN Xuan Bach	USTHBI7-021
LE Minh Anh	USTHBI7-007
DOAN Thanh Huy	USTHBI7-084
PHAM Minh Duc	USTHBI7-041

Under the guidance of:

Dr. NGHIEM Thi Phuong ICT Lab

Hanoi, January 2019

TABLE OF CONTENT

I. The Purpose of the Project

Plankton is a diverse group of organisms that live in the oceans and cannot swim against the water current [1]. And they have an important role to ecosystem and human life that we cannot undeniable. The carbon cycle of the Earth is almost controlled by these tiny organisms. In order to study plankton's population that is imbalance nowadays because of the ecosystem changed. Underwater camera systems are used to capture microscopic, high –resolution images of plankton over large study areas. However, there are a huge amount of plankton images, it's raises a key question that how we can analyze the collected plankton images in order to assess the population and distribution of plankton species. A natural approach to this question is to manually classify collected plankton images for further human analysis This approach, however, is infeasible since it would be extremely time and cost consuming to manually classify millions of images [2]. To overcome this problem, the project has two main goals:

(1) To develop a classification platform to automatically put new plankton images in their correct categories;

(2) To build a plankton website so that the environmental scientists can contribute plankton data for easier use of plankton images in research and education.

II. Assumptions/Contrait

Our main goal of this work is to support the easy use of plankton images in research and education using the proposed plankton classification platform with web user interface.

Our targets are scientists and student who want to research and study about plankton. In order to make it work. We have two tasks:

- Task 1 - Front-end: We use CSS (most is we using Sass), Javascript, HTML to create a web user interface to use plankton classification platform.
- Task 2 - Back-end: We use image classifier using convolutional neural network to train and test images in order to make the computer know how to classified each images of plankton to which category they belong to.

The specific objectives of the project:

- Understand the basic background about plankton and types of plankton.
- Understand the basic background about machine learning and image classification.
- Study the image classification models for the case of big datasets.
- Apply image classification models for the case of big datasets.
- Apply image classification models to build plankton classification platform
- Develop web user interface for environmentalists to use plankton classification platform and let people also can contribute plankton data for easier use of plankton images in research and education.

III. Scope of Management

In scope of management, we have two section:

- Section 1 - Personnel Management: It's including our recruitment, training and manpower planning.
- Section 2 - Development Management: We design and sketch our website. Research and development function of the website include back-end and front-end.

IV. Work Breakdown

- Preparation
    - Obtaining instruction and information from supervisors
    - Research article
    - Discussion
    - Planning
    - Writing Report 1
- Development
    - Front-end
        - Sketch website

        - ➢ Design and Programming
    - Back-end
        - ➢ Collect Dataset
        - ➢ Divide dataset to each domain, type
        - ➢ Divide data to 7-3 (70% data training, 30% data testing)
        - ➢ Train and test
        - ➢ Apply to front-end
    - Test the application
    - Debug
- Conclusions
    - Discussion
    - Writing Report 2
    - Prepare for the presentation

V. Project Management Team/Structure

Our team have 5 people, divide into 2 small groups:

- Front-end team – Group 1:
    - Pham Minh Duc
    - Doan Thanh Huy
- Back-end team – Group 2:
    - Le Nguyen Khoi
    - Nguyen Xuan Bach
    - Le Minh Anh

Also, Le Nguyen Khoi will be change to front-end team if the process the plan is not like what we planned on schedule.

VI. Milestones with Timeframe and Person in Charge

ID	Task Name	Start	Finish
	The Whole Project	1/12/2018	12/3/2019
1	Rally member	10/11/2018	15/11/2018
2	Planning and Divide work	15/11/2018	20/11/2018
3	Research on techniques	20/11/2018	26/11/2018
4	Obtaining project from supervisor	26/12/2018	30/11/2018
5	Research article	01/12/2018	05/12/2018
5	Sketch website (Group 1) Collect Data (Group 2)	05/12/2018	10/12/2018
6	Design, programming, debug (Group 1)	10/12/2018	Middle February
7	Train, test, debug (Group 2)	10/12/2018	Middle February
8	Writing report 1	22/01/2019	28/01/2019
9	Incorporate Back-end and Front-end	Middle February	Middle February
10	Run, test and debug	Middle February	Late February
11	Discussion	Late February	Late February
12	Writing Report 2	Middle February	Late February
13	Preparing to Presentation	Late February	Early March
14	Presentation	12/03/2019	12/03/2019

VII. Cost and Budget Management

- Software:

- Visual Studio Code (Free)
- Local Server - XAMPP (Free)

- Hardware:
  - Laptop (Self-provided)
  - Desktop PC (Self-provided)

VIII. Risk Management

- Because our team are not skillful so the project has to work with a long time, cannot be done it faster, especially with back-end.
  Solution: Work as a group and asking each other if there is any problem.
- The data is huge, and our laptop is not strong enough. So, it will take time to training models. It will maybe have some problems during the training.
  Solution: Make sure there are not any mistakes before training.

IX. Communication Methods and Channels Internally and Externally

- In order to exchange file to each other, we use: GitHub, Google Drive
- Communication between member: E-mails, Facebook, Discord, Direct meeting at class
- Communication with supervisor: Dr. Nghiem Thi Phuong – ICT Lab, 5th floor, USTH Building.

References:

[1] Castro Pond, Stephen, and George L. Pickard. Introductory dynamical oceanography. Elsevier, 2013

[2] Culverhouse, Phil F., et al. “Do experts make mistakes? A comparison of human and machine identification of dinoflagellates.” Marine Ecology Progress Series 247.17-25 (2003):5.