Deep-Learning Models Benchmarking with Multi-Class Plankton Image

Benchmarking CNN and GCN for Plankton Classification

Group: G1

C4IM2508 Jaronchai Dilokkalayakul

C4IM2501 Daffa Akbar Aprilio

C5IM2015 Ganchimeg Namuunbayar

01 May, 2025

Update 1

- Selected and prepared plankton dataset from IFCB source
- Chose 4 plankton classes with balanced data samples
- Preprocessed data: resized, normalized, augmented
- Performed EDA to visualize class distribution

Introduction

- **Objectives**: To benchmark and compare the performance of various deep learning models, including Convolutional Neural Networks (CNNs) and Graph Convolutional Networks (GCNs), on plankton classification tasks.

- **Goals**: Identify the most effective architecture for accurately classifying plankton species.

- **Dataset**: Images from the Imaging FlowCytobot (IFCB)

- **Variables**: Class labels (species), image files

- **Data Cleaning**: Choosing 4 classes from the raw data folder:

- Represents a genus (no mixing, no non-plankton image, etc)

- Every class contains similar amount on files

Data Science Methods

- Image preprocessing: resizing, normalization, augmentation
- Modeling methods: CNNs (ResNet50, EfficientNet, DenseNet, MobileNet) and GCNs
- Tools: TensorFlow, seaborn, matplotlib

Exploratory Data Analysis

Explanation of Dataset

Plankton Class	Count
Dinobryon	588
Pseudonitzschia	578
Dactyliosolen	532
Corethron	447

- Data classes: Image files (PNG)
- Labels: Plankton genus

Data Preparation

- Resized images to 224x224 pixels
- Normalized pixel values to [0, 1]
- Performed data augmentation (flipping, rotation)
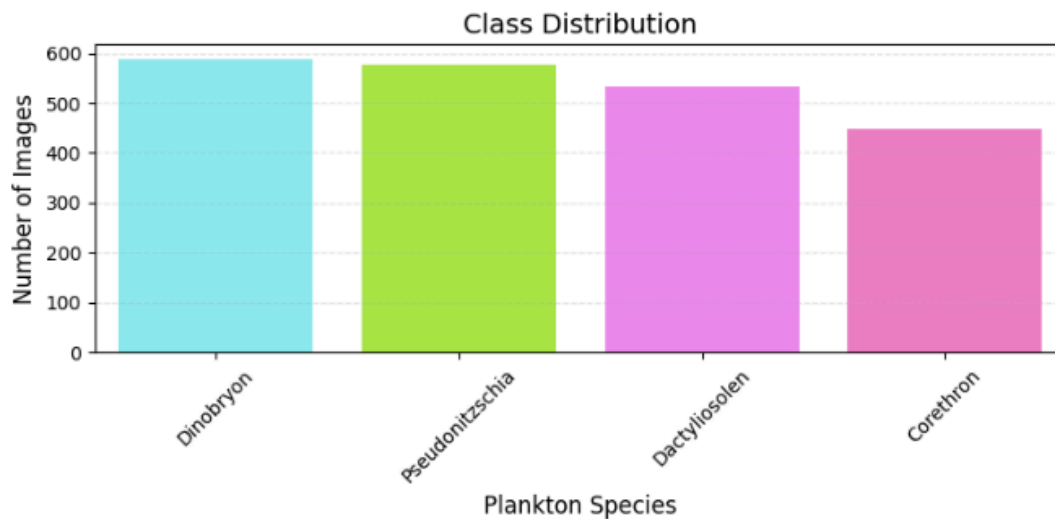- Split into training and test datasets

Data Visualizations

Figure 1: Figure 1. Distribution of image count across selected plankton classes.
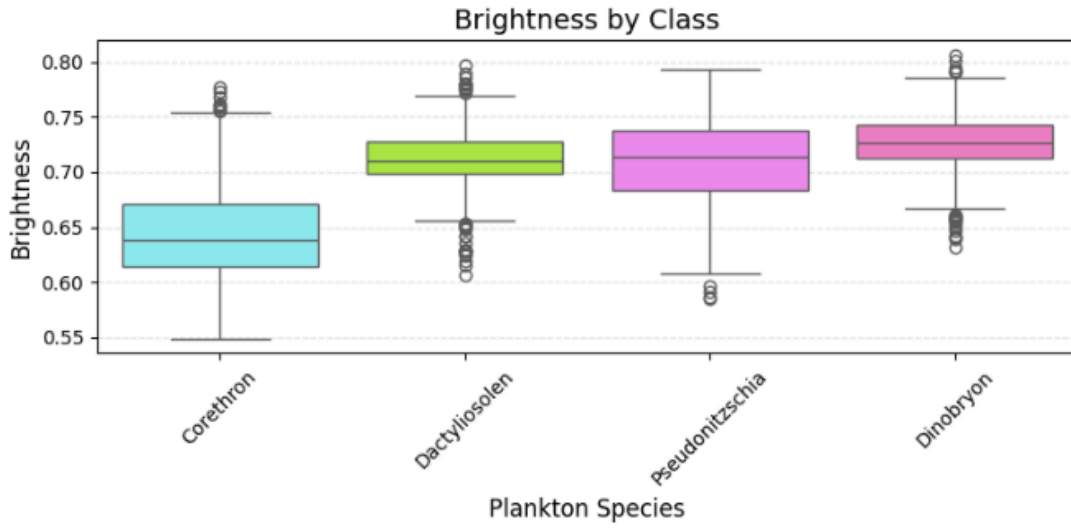
Variable Correlations

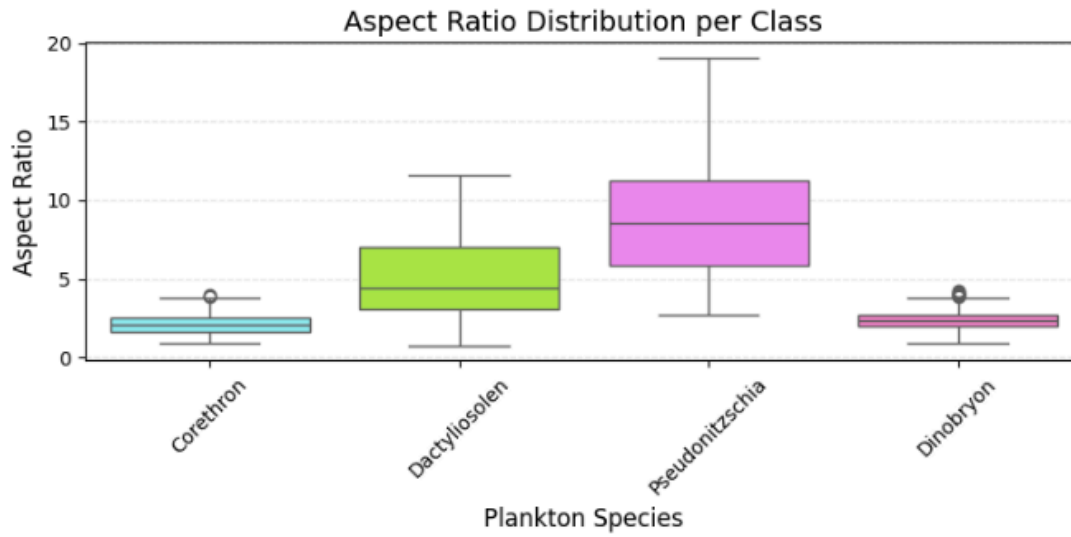Figure 2: Figure 2. Brightness distribution across the selected plankton classes.

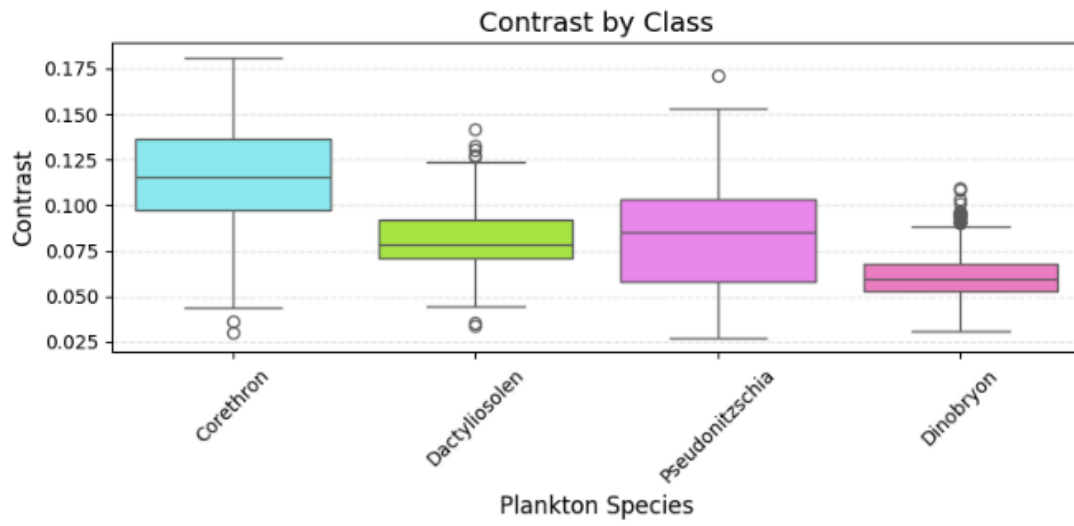Figure 3: Figure 3. Distribution of aspect ratios for images in each plankton class.

Figure 4: Figure 4. Contrast values distribution across the selected plankton classes.