



Article Computer Vision Based Deep Learning Approach for the Detection and Classification of Algae Species Using Microscopic Images

Abdullah ¹, Sikandar Ali ¹, Ziaullah Khan ², Ali Hussain ¹, Ali Athar ¹ and Hee-Cheol Kim ^{3,*}

- ¹ Institute of Digital Anti-Aging Healthcare, Inje University, Gimhae 50834, Korea; abdullahbaltistani@gmail.com (A.); sikandarshigri77@gmail.com (S.A.); alihussainnrana@gmail.com (A.H.); ali.athar14@ce.ceme.edu.pk (A.A.)
- ² School of Electrical Engineering and Computer Science, National University of Science and Technology Islamabad, Islamabad 44000, Pakistan; zkhan.msee19seecs@seecs.edu.pk
- ³ College of AI Convergence, Institute of Digital Anti-Aging Healthcare, u-AHRC, Inje University, Gimhae 50834, Korea
- * Correspondence: heeki@inje.ac.kr; Tel.: +82-55-320-3720

Abstract: The natural phenomenon of harmful algae bloom (HAB) has a bad impact on the quality of pure and freshwater. It increases the risk to human health, water bodies and overall aquatic ecosystem. It is necessary to continuously monitor and perform proper action against HAB. The inspection of algae blooms by using conventional methods, like algae detection under microscopes, is a difficult, expensive, and time-consuming task, however, computer vision-based deep learning models play a vital role in identifying and detecting harmful algae growth in aquatic ecosystems and water reservoirs. Many studies have been conducted to address harmful algae growth by using a CNN based model, however, the YOLO model is considered more accurate in identifying the algae. This advanced deep learning method is extensively used to detect algae and classify them according to their corresponding category. In this study, we used various versions of the convolution neural network (CNN) based on the You Only Look Once (YOLO) model. Recently YOLOv5 has been getting more attention due to its performance in real-time object detection. We performed a series of experiments on our custom microscopic images dataset by using YOLOv3, YOLOv4, and YOLOv5 to detect and classify the harmful algae bloom (HAB) of four classes. We used pre-processing techniques to enhance the quantity of data. The mean average precision (mAP) of YOLOv3, YOLOv4, and YOLO v5 is 75.3%, 83.0%, and 91.0% respectively. For the monitoring of algae bloom in freshwater, computer-aided based systems are very helpful and effective. To the best of our knowledge, this work is pioneering in the AI community for applying the YOLO models to detect algae and classify from microscopic images.

Keywords: HAB; algae detection; object detection; YOLO model; microscopic image

1. Introduction

The quality of water is indispensable to public health, industry, and agriculture. Many factors are responsible for the poor quality of water. Among them, algae are a type of organism that degrades water quality and poses risks to ecosystems and people. Algae species are a kind of plant organism that is mostly found in freshwater, ponds, water channels, and on the sides of rivers. Algae bloom is the process of overgrowth of algae, and it is a global problem for the freshwater management system and water bodies. It increases exponentially and covers a large area, which is called harmful algae bloom (HAB). The algae bloom is very dangerous for human health as well as aquatic life. Various factors such as temperature, nutrients, sunlight, and climate changes are responsible for algae bloom [1]. It produces a high toxin compound, which affects the quality of water



Citation: Abdullah; Ali, S.; Khan, Z.; Hussain, A.; Athar, A.; Kim, H.-C. Computer Vision Based Deep Learning Approach for the Detection and Classification of Algae Species Using Microscopic Images. *Water* 2022, *14*, 2219. https://doi.org/ 10.3390/w14142219

Academic Editor: Guangyi Wang

Received: 19 May 2022 Accepted: 11 July 2022 Published: 14 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). and produces undesirable odour and taste. Similarly, it produces neurotoxic, paralytic, amnesia, and many other chemicals, causing the death of marine mammals and even water-dependent creatures [2,3]. In addition, the continuous release of toxins and harmful chemicals covers all the surface of the water, which reduces the sunlight entrance to the water, thus badly influencing the photosynthesis process of the plants under the water. Furthermore, it poorly affects the cycle of the food chain. The HAB consumes oxygen in a huge amount and causes a deficiency of oxygen in the water bodies [4]. Hence it is necessary to save the life of water bodies and the health of humans by installing a proper monitoring and algae bloom identification system. To control and mitigate the HAB in freshwater, the latest artificial intelligence-based techniques, for example neural networks, have been used by researchers to ensure the supply of quality drinking water [5].

There are many techniques to monitor algae bloom. Aircrafts, satellites, or drones are being used to obtain hyperspectral or multi-spectral images to monitor and identify the algae bloom event in a large area [6–8]. It is imperative to monitor the undesirable algae bloom continuously in ponds, water channels, or freshwater reservoirs and to take essential and immediate action in any area to maintain drinking water quality. The traditional method of algae bloom by visual analysis using a microscope is a time-consuming, economically not feasible, and a cumbersome task. An automated system based on the latest object detection algorithms is an effective way for real-time monitoring algae bloom in waterbodies.

Recently, various approaches like computer vision and deep learning-based techniques have been frequently used for object detection. Especially convolutional neural network (CNN) is getting more focus and has shown promising performance in image analysis, object detection [9], and image segmentation [10]. Very few studies were conducted for the monitoring and identification of algae bloom based on CNN [11]. A study was conducted by Baek et al. [12] to simulate HAB by identifying the bloom initiation and density using regression and classification CNN model. The multi-targeted based Fast Region-Based Convolutional Neural Network (Fast R-CNN)model [13] is studied for the identification and classification of algae species. The latest object detection algorithms, which have high performance and accuracy, can also be used for the detection of algae bloom. Park et.al [14] performed an experiment to detect microalgae using the deep learning based object detection technique YOLOv3. The microscopic images were used to train the model and darknet-53 was used as the backbone. This study also revealed that the training performance of a model with color images is better as compared to grayscale images. Another study was conducted by Park et.al [15] to analyze automatic algae species detection using deep learning based YOLOv3 and YOLOv4. The YOLO is a much faster and powerful algorithm to detect an object in real time from images and video data.

In this study, we investigated the usefulness of three state-of-art models for the detection of algae and compare their performance. YOLOv3, YOLOv4, and YOLOv5 models were used to detect and classify our custom dataset. These models identified the four types of algae species, namely Cosmarium, Closterium, Spirogyra, and Scenedesmus. We used algae microscopic images to train the YOLO models. The latest version of YOLO, i.e., YOLOv5, was used in this study. It is a powerful real-time objection detection model in the AI family. The main contribution of our study in this work is as follows.

- The collection of microscopic algae images;
- Implementation of Dc-GAN to enhance the number of images in the dataset;
- Labelling of images based on their specific classes;
- Trained Yolov3, Yolov4, and Yolov5 models on a custom dataset;
- Comparative analysis of all the models' accuracy and performance.

This paper is organized as follows. The related work and literature review are described in Section 2. Section 3 represents the materials and method. The results and discussion of this study are described in Section 4. The conclusion and future work are explained in Section 5.

2. Related Work

Algae bloom is a natural process that normally occurs in freshwater reservoirs. Whenever blooms occur, it forms colonies and spread on a large scale in the area. The rapid spread of algae creates a dangerous zone for aquatic life. Various factors, like warm temperature, climate change, sufficient light, and increasing nutrition, are responsible for algae bloom in freshwater storage tanks, ponds, and lakes. The high concentration of algae blooms produces different chemical compounds and stops the sunlight reaching inside the water. In addition, it affects the process of photosynthesis and destroys the food chain of the marine ecosystem. This is a worldwide issue and the whole world is suffering from these problems. It is necessary to monitor and identify the algae bloom in an area to avoid a big loss. In the past decade, a lot of research and experiments have been conducted to address this problem.

Paul R. Hill et al. [16] developed a machine learning-based application for the prediction and detection of harmful algal bloom. They used remote sensing data and different machine learning architectures for this purpose. The model showed a detection accuracy of 91%. Derot et al. [17] used a random forest algorithm for the prediction of harmful algae blooming in Lake Geneva due to the cyanobacterium Planktothrix rubescens. The purpose of developing a machine learning-based model was to assist the locals to manage the lake environment. They used 34 years data of P. rubescens concentration and clustered the data into 4 groups using the k-means clustering method. Sönmez et al. [18] applied several CNN models and support vector machines for the classification of algae. They classified cyanobacteria and chlorophyta microalga groups. Seven different CNN models and transfer learning were used in this study. Alex-SVM achieved the highest accuracy. SVM was used to improve the classification accuracy, which improved the accuracy from 98% to 99.66%.

Arabinda Samantaray et al. [19] proposed computer vision and a deep learning-based system for the detection of algae. The authors used state-of-the-art transfer learning techniques to develop their proposed model. Transfer learning techniques are used in machine learning and enable us to take benefit of the pretrain models, which are trained on a huge amount of data. By customizing the pre-trained model and training them on our custom dataset, we can use these models for our customized purposes. They use Faster R-CNN, R-FCN, and Single Shot Detector for the detection of algae. They compare the results of these three transfer learning models, and region-based fully convolutional networks (R-FCN) showed the highest accuracy of 82%, followed by faster R-CNN 72% and SDD at 50%, respectively. Edgar Medina et al. [20] presented a vision inspection system for the detection of algae in underwater pipelines using multilayer perceptron (MLP) and CNN algorithms. The authors used 41,992 samples of data and it was annotated for algae and non-algae manually. They applied data augmentation after splitting the data into training, testing, and validation. The model gave an accuracy rate of 99.39%. Jungsu Park et al. [5] designed an automatic system based on neural architecture search (NAS), which finds the best CNN model for algal genera classification. This system could classify eight classes of algae in watersheds for drinking water supply with an F1-score of 0.95. The experimental results showed that the CNN models developed using neural architecture search could present better performance compared to the conventional ways of developing CNN models.

Bi Xiaolin et al. [21] applied a Support vector machine (SVM) for the detection of microalgae species using hyperspectral microscopic images. They performed several image processing steps to optimize the detection results. The experimental results reported high sensitivity and specificity reaching up to 100%. This study also performed survival competition analysis of microalgae using microscopic imaging technology under pH effect. SS Baek et al. [22] performed classification and quantification of cyanobacteria species using deep learning. Fast regional convolutional neural network (R-CNN) and Convolutional neural network (CNN) were used for the classification of five cyanobacteria species. Microscopic images were used, and post-processing of classified images was conducted, which

helped increase the accuracy of the model. The average precision values range of the model was reported between 0.890 and 0.929.

Jesús Balado et al. [23] used deep learning for the semantic segmentation of macroalgae in coastal environments. Images of five different macroalgal species with high resolutions were used and three CNN models, namely Resnet18, MobilenetV2, and Xception, were applied in this study. Residual Network (ResNet) presented the highest accuracy of 91.9% and all the five classes of macroalgae were segmented correctly. Jesus Salido et al. [24] employed YOLO for the detection and classification of diatoms, i.e., microalgae. They analysed the performance of the model by training and testing the model for the classification of 5, 10, 20, and 30 target species. They also validated the model using the colour images and grayscale images. The model could identify 80 different diatom species with a specificity of 96.2%, sensitivity of 84.6%, and precision of 72.7%.

3. Materials and Methods

In this section, we explain the materials and methods used in this research work. We included data source information, data processing, YOLO background, and other related details in this section. We collected 400 algal images, pre-processed the data, and applied different object detection techniques to detect and classify the different classes of algae.

3.1. Data Source

The dataset used in this research was collected by the main laboratory of Quaid-Azam University Islamabad, Pakistan. It has 400 microscopic images belonging to each of 4 classes, i.e., Cosmarium, Scenedesmus, Closterium, and Spirogyra. These microscopic images were pre-processed and used for the training of the model. The data samples are shown in Figure 1.



Figure 1. Samples of all the four algae classes. (a) Closterium, (b) Cosmarium, (c) Scenedesmus, (d) Spirogyra.

3.2. Data Preprocessing

Data pre-processing is one of the important parts of artificial intelligence model development. Raw data is pre-processed by applying different techniques like data labelling, data augmentation, data sampling, data normalization, etc. To train a deep learning-based model, the size of the dataset is small, which might affect the output performance of the models. So, to overcome this problem, the advanced deep learning-based data augmentation technique, for example, generative adversarial neural (GAN), was used. Figure 2 shows the images generated by applying DC-GAN from the original images. The GAN model considers original images as reference images and generates new images with the same statistic as the original image. The GAN framework is gaining more attention in computer vision due to high capability to generate more useful data based on reference image data.



Figure 2. Example of algae samples used for multiclassification. (**a**) Original samples. (**b**) Preprocessed samples generated using Dc-GAN.

We applied traditional and advanced data augmentation techniques, and 800 images were obtained for each class. This data is sufficient to train the YOLO models more accurately. The DC-GAN is an advanced version of the GAN model and has two main parts generators and discriminators, as shown in Figure 3. The generator synthesizes images based on the original reference image and the discriminator differentiates between real images and synthesis images.



Figure 3. The architecture of DC-GAN. The generator generates fake image samples and the discriminator differentiates between a real image and a fake image.

Deep Convolutional GAN (DC-GAN) works exactly the same as the simple GAN model and is used for unsupervised learning. The DC-GAN uses convolutional stride instead of pooling layer. It does not need a fully connected layer as the GAN model. The batch normalization is used in both the generator and the discriminator. Moreover, the

generator part uses ReLU activation function for all layers except for the output, which uses tanh. Similarly, the discriminator part uses LeakyReLU for all layers except output, which uses the sigmoid activation function, as shown in Figure 3.

The generator and discriminator are completely based on a convolutional neural network. In this experiment, we have 4 classes with 400 real images. The generator part generates fake images and these fake images are fed with reference real images to the discriminator part. Moreover, the discriminator calculates the loss between real and fake images. The loss propagates back to the generator and updates fake images. This process continues, unless the loss becomes minimum. Eventually, the discriminator cannot differentiate anymore between the real and fake images, and then it classifies both images as real. This method was used to generate enough number of images from 400 images.

In YOLO, the images are annotated according to the classes. In our dataset, we had four classes, so all the images were annotated accordingly. We annotated the images by making bounding boxes of the class objects for the classes, namely Cosmarium, Scenedesmus, Closterium, and Spirogyra, as shown in Figure 4.



Figure 4. Annotation of images using the annotating tool.

As a result of the annotation, text files were generated corresponding to each image, which contains information about the objects in the images. The text file has the class Id, coordinate of the bounding box, and height and width of the bounding box.

3.3. YOLO Background

There are several artificial intelligence-based techniques for the classification and detection of objects such as CNN [25], Fast Regions with Convolutional Neural Network (Fast R-CNN) [26], and faster R-CNN [27], however, among them, You Only Look Once (YOLO) [28,29] presents good object detection results with high accuracy and precision. Therefore, we used YOLO for the detection of algae. This approach has been used in many areas and applications. It has been used in industries [30], agriculture [31], medical [32], and many other areas.

Generally, there are two categories of object detection algorithms, namely one-stage detectors and two-stage detectors. One-stage detectors [33–35] form a very simple architecture of fully convolutional networks and give classification probabilities of each class as an output. However, two-stage detectors [9,36] have complicated architecture and the objects are regressed twice. First, the high probability region having an object is filtered out, and then it is fed into the region convolutional network. Finally, the classification score will be taken as an output. The one-stage detectors are considered to be more robust.

YOLO falls in the category of a one-stage detector. The main architecture behind YOLO is convolutional neural networks that help with computer vision tasks. YOLO [37] is faster and more accurate compared to the rest of the object detection techniques like R-CNN, etc. YOLO was developed to overcome the computational complexity problems associated with other object detection techniques. It performs detection and classification of the object in one step, therefore it is called You Only Look Once (YOLO) [38]. It can compute the bounding boxes and probabilities of the classes simultaneously. YOLO splits the image into grid blocks and detects the objects of interest while two-stage detectors use the proposal approach for recognizing the objects. It shows the bounding box, class score, and object score as the result of the output of the model. YOLO can detect various objects with single inference. There are different versions of the YOLO model, i.e., YOLOv1, YOLOv2, YOLOv3, YOLOv4, and YOLOv5. YOLOv2 [37] was proposed to enhance detection efficiency by introducing a batch normalization process. It uses Darknet-19. Although YOLOv2 showed good results compared to YOLOv1, however for small object detection its accuracy was low. Thus, YOLOv3 [39] came into the picture, which uses a variant of Darknet architecture, residual block, skip connections, and upsampling. These new features enable it to detect objects of different sizes and scales, empowering it to track small objects as well. Furthermore, YOLOv4 [40] was proposed, aiming to get high accuracy and speed. It uses Cross Stage Partial Network (CSPDarknet-53) and other novel methods like spatial pyramid pooling (SPP) and path aggregation (PAN).

In May 2020, YOLOv5 [41] was proposed by Ultralytics LLC whic is situated in Clarksburg, MD 20871, USA, which is the latest version of the YOLO family with 140 FPS and a size of 27 MB. It is 90% smaller and 180% faster than YOLOv4. YOLOv5 is lighter and faster compared to the previous versions. For the real-time detection of objects, YOLOv5 is very suitable with a high inference speed. It possesses several properties of previous versions. For example, it uses Spatial pyramid Pooling Network (SPP-NET), which was used in YOLOv4. The Common Object in Context (COCO) dataset was used to train YOLOv5. It was developed by using the Pytorch framework, however, the previous versions were developed in the Darknet framework. Furthermore, YOLOv5 has four versions, namely YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. These four versions have different architectures. Moreover, the convolutional kernel and feature extraction modules are also different. The architecture of YOLOv5 consists of three modules, namely the backbone network, neck network, and detect the network, as shown in Figure 5. The backbone module extracts the feature information from the input images while the neck network combines all those feature information extracted by the backbone network, and it creates three scales of feature maps. Lastly, the objects in the images are detected by the detect part of the architecture. Three important factors make YOLOv5 conspicuous among other object detectors. First, YOLOv5 fuses cross-stage partial network (CSPNet) [42] into Darknet and resolves the repeated gradient information issue.



Figure 5. YOLOv5 architecture. (a) Input image, (b) backbone, (c) neck and (d) detection.

Consequently, it reduces the FLOPS (floating-point operations per second) and parameters, thus increasing the accuracy and inference speed of the model. Furthermore, it also reduces the size of the model. Secondly, a path aggregation network (PANet) [43] has been applied in YOLOv5, which helps improve the propagation of low-level features. Thirdly, the YOLOv5 head has a multi-scale ability that produces three different sizes of feature maps, and it can detect objects of any size, i.e., small, medium, and large.

3.4. Experimental Environment

The data analysis, data pre-processing, and all the experiments were conducted using a 64-bit Windows operating system with Intel(R) Core (TM) i7-7700 CPU @ 2.60 GHz, 3.60 GHz processor, and 32 GB installed Random Access Memory (RAM),manufactured by intel and sourced from Gimhae, Korea. For the experiment and training of the models, Google Colab was used. Python language (Version 3.8), Open-Source Computer Vision Library OpenCV, Tensorflow (Version 2.8), Pytorch, Keras (Version 2.8), and Scikit-learn libraries were also used in this experiment.

3.5. Performance Measurement

To evaluate the model, we used different performance measure metrics, i.e., precision, recall, F1 score, and mean average precision (mAP). The object detection models are commonly evaluated by using these performance measure metrics. The mean average precision (mAP) gives the score by comparing the ground-truth bounding box to the detected box of the object in the image.

$$Precision = \frac{TP}{TP + FP},$$
(1)

$$\operatorname{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}},\tag{2}$$

$$F1 \text{ score} = \frac{2 \times (Precision \times Recall)}{Precision + Recall},$$
(3)

$$mAP = \frac{1}{N} \sum_{N}^{i=1} AP_i, \qquad (4)$$

where TP represents true positive, TN denotes true negative, FP is false positive, and FN denotes false negative. Precision gives the true positive values out of all positive predictions. In object detection tasks it is calculated using the threshold value of intersection over Union (IOU). Recall calculates how well the model can recognize the true positive out of all the predictions, i.e., true positive values and false negative values. F1 score shows the composition means of precision and recall, and its values lie between 0 and 1.

4. Results and Discussion

In this section, we elaborated the experimental details of this research. We applied three one-stage object detection models, namely, YOLOv3, YOLOv4, and YOLOv5. The models could detect all four algae classes with good precision.

4.1. Training the Models

The dataset consists of 3200 microscopic images after performing the data augmentation technique of 4 algae classes and it was split into 80% training and 20% testing sets. All the three object detectors were trained on the training dataset and the training specifications have been shown in Table 1. The YOLOV3 and YOLOV4 were trained on 80% of the total dataset with 100 epochs and a batch size of 32. The Adam optimizer was used, and the learning rate is set to be 0.01

Model	Train	Test	Epoch	Batch Size	LR	Optimizer
YOLOv3	80%	20%	100	32	0.01	Adam
YOLOv4	80%	20%	100	32	0.01	Adam
YOLOv5	80%	20%	100	16	0.01	SGD

Table 1. Specification of object detection models.

Since YOLOv5 is the latest and most robust object detection model, our focus was thus YOLOv5. Moreover, our experimental results also revealed the robustness and high performance of this object detector. The YOLOv5 was trained on 80% of the total dataset with 100 epochs. Stochastic gradient descent (SGD) optimizer was used, and the batch size was set to 16. Furthermore, the learning rate was set to 0.01. After the completion of model training, the weights of the model were saved, and the model was evaluated on the test dataset. The model could recognize the four types of algae species with correct labels for each class and their probability belonging to a particular class.

4.2. Models Evaluation

The models were evaluated by analysing different performance measure metrics for example precision, recall, mAP, confusion matrix, etc. The confusion matrix is the tabular representation and the overall summary of the YOLOv5 model performance has been shown in Figure 6. It is a two-dimensional matrix that shows actual class and predicted class, respectively. It provides a better and deeper instinct regarding the overall performance of the model. The diagonal values represent the right prediction of true values, as shown in Figure 6. The recall, precision, F1 score, and mAP are calculated based on confusion metrics. The experimental results revealed that the YOLOv5 outperformed the other models. The evaluation indicators of each model are shown in Table 2. The overall performance of YOLOv3, YOLOv4, and YOLOv5 is good in terms of precision, recall, F1 score, and mAP. The mAP of YOLOv3, YOLOv4, and YOLOv5 is 75.3%, 83.0%, and 90.1%, respectively. This shows that the performance of YOLOv5 is high and more accurate.



Figure 6. Confusion matrix of YOLOv5 model on test dataset.

Table 2. Evaluation indicators of the models.

Model	Precision	Recall	F1 score	mAP
YOLOv3	77.0%	84.40%	80.47%	75.3%
YOLOv4	88.0%	69.0%	77.35%	83.0%
YOLOv5	88.0%	85.0%	86.5%	90.1%

Each model is tested with a separate dataset and the test result of YOLOv5 is shown in Figure 7. It shows the result of YOLOv5 and the performance of the trained model with mAP against each image.



Figure 7. The detection and classification performance of YOLOv5.

4.3. Discussion for the Algae Detection

In this study, three object detection models have been employed for the detection of four different species of algae, namely Cosmarium, Closterium, Scenedesmus, and Spirogyra. The harmful algae blooms in freshwater may pollute the fresh water and badly affect marine life. Furthermore, they make the water contaminated, toxic, and unfit for drinking. In order to monitor the water quality, an automated monitoring system is indispensable. Thanks to the latest technology, like artificial intelligence and computer vision techniques, we can monitor the quality of water by detecting the algae blooms in water pools and other water storage. Researchers have used deep learning-based techniques, for example, CNN, R-CNN, Fast R-CNN, Faster R-CNN, and YOLO family-based object detectors for detecting different kinds of objects. Since these object detectors have high performance for detecting the object, we therefore used YOLO for the detection of four types of algae. For this purpose, we first collected the algae data, and to preprocess the data we used three YOLO family object detectors, namely YOLOv3, YOLOv4, and YOLOv5.

We conducted comparative analysis of our study with other state-of-art models, as shown in Table 3.

Method	Model	Precision	Recall	F1 score	mAP
Park et.al [14]	YOLOv3	80%	-	-	81.0
	YOLOv3	-	-	-	40.9
Deals at al [1]	YOLOv3 Tiny	-	-	-	88.8
Park et.al [15]	YOLOv4	-	-	-	84.4
	YOLOv4 Tiny	-	-	-	89.8
	YOLOv3	77.0%	84.40%	80.47%	75.3
Our approach	YOLOv4	88.0%	69.0%	77.35%	83.0
	YOLOv5	88.0%	85.0%	86.5%	90.1

Table 3. Comparative analysis of our approach with other research works based on the state-of-the art.

Table 3 shows that among all the models, the performance of YOLOv5 is high and the mean average precision (mAp) score is 90.1. We have used DC-GAN based generated data to train our model, so we can say that its performance is better than the mentioned state-of-art model.

The main objective of this study is to develop a deep learning-based automatic detection and classification model to detect and classify the microscopic image of HAB. The YOLO-based models were used in this paper to classify the four various classes of algae species: (1) Cosmarium, (2) Closterium, (3) Scenedesmus, (4) Spirogyra. The performance of the YOLO family shows that the YOLOv5 is performing very well on the test dataset, as shown in Figure 7.

5. Conclusions

In this paper, we presented three state-of-the-art object detection models, namely YOLOv3, YOLOv4, and YOLOv5 for the detection of four classes of algae, i.e., Cosmarium, Scenedesmus, Closterium, and Spirogyra. These models are popular and mostly used for real-time object detection purposes because these object detectors are robust and have high accuracy. Among the three object detectors, YOLOv5 outperformed other models with high inference and accuracy. Our research work can be summarized as follows: first, we collected 400 microscopic images of each class belonging to 4 Algae species. Second, we applied preprocessing techniques to our custom dataset so that object detection models can be trained using the preprocessed data. We used DC-GAN, an advanced version of Generative adversarial networks (GAN) for the generation of new images from the original images. As a result of DC-GAN, we generated 3200 images. Third, we trained the YOLOv3, YOLOv4, and YOLOv5. We performed hyperparameter tuning of the models to get the optimal performance of the model. Lastly, we evaluated the models while testing the models with the testing dataset. The evaluation results revealed that YOLOv5 outperformed the other two object detection models and showed good performance.

We have proposed a novel model for the detection and classification of algae species by merging the two approaches, i.e., DC-GAN and real time object detection algorithm (YOLOv5). This model is robust and efficient for detecting and classifying the algae species in real time environment. The architecture of this model is not too complex, which helps to detect the objects within seconds.

YOLOv5 has a wide range of advantages over the conventional object detectors and is suitable for real-time object detection. Furthermore, our experimental results also proved the potential of this model in terms of high accuracy and inference time for the detection of algae, therefore this model can be deployed for accurate and rapid detection and classification of algal species in a real-time environment. The presented model could classify the four algal species with 88.0 precision and 85.0 recall. Thus, this model can be used for the early warning and real-time monitoring of HAB in water.

Based on accuracy, it is stated that YOLOv5 is suitable for automatic detection and identification of micro algal plants. This technique may be put into tiny drones or air crafts to detect algal blooms (algae colonies) in real time environment. Because algae are microorganisms, drones can only identify colonies for detection and classification of each micro algae. We may utilize microcontrollers and attach a camera to it, then integrate the camera with a microscope to detect and identify micro algae in real time. Even though this system was trained on an algae dataset with only four classes, it could only classify and identify these classes. Furthermore, the accuracy of the model can be increased by adding a greater number of real-world pictures. In the future, we may be able to apply new models, such as RetinaNet, by modifying the model architecture for more precision.

Author Contributions: Conceptualization, Abdullah and S.A.; methodology, Abdullah, S.A. and A.H.; validation, Abdullah, S.A., Z.K. and H.-C.K.; formal analysis, A.H., A.A. and H.-C.K.; data curation, Abdullah, S.A., Z.K. and H.-C.K.; writing—original draft preparation, Abdullah and S.A.; writing—review and editing, Abdullah and S.A.; supervision H.-C.K.; project administration, H.-C.K.; funding acquisition, H.-C.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was supported by the 2021 Inje University Research grants (No.1711139492).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: All the participants gave their consent to participate in this study.

Data Availability Statement: Private dataset (Microscopic image dataset for multi-class classification) is not available online. The data used in this study are available on request from the corresponding author.

Acknowledgments: This work was supported by the Commercialization Promotion Agency for R&D Outcomes (COMPA) grant funded by the Korean Government (Ministry of Science and ICT)" (R&D project No.1711139492).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Paerl, H.W.; Otten, T.G. Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. *Microb. Ecol.* 2013, 65, 995–1010. [CrossRef] [PubMed]
- 2. Bhat, S.; Matondkar, S.P. Algal blooms in the seas around India–networking for research and outreach. *Curr. Sci.* 2004, *87*, 1079–1083.
- 3. Okaichi, T.; Yanagi, T. *Sustainable Development in the Seto Inland Sea, Japan*; Terra Scientific Publishing Company: Tokyo, Japan, 1997; pp. 251–304.
- 4. Anderson, D.M. Approaches to monitoring, control and management of harmful algal blooms (HABs). *Ocean Coast. Manag.* 2009, 52, 342–347. [CrossRef] [PubMed]
- Park, J.; Lee, H.; Park, C.Y.; Hasan, S.; Heo, T.-Y.; Lee, W.H. Algal Morphological Identification in Watersheds for Drinking Water Supply Using Neural Architecture Search for Convolutional Neural Network. *Water* 2019, *11*, 1338. [CrossRef]
- Goldberg, S.J.; Kirby, J.T.; Licht, S.C. Applications of Aerial Multi-Spectral Imagery for Algal Bloom Monitoring in Rhode Island; SURFO Technical Report No. 16-01; University of Rhode Island: Kingston, RI, USA, 2016; p. 28.
- Kudela, R.M.; Palacios, S.L.; Austerberry, D.C.; Accorsi, E.K.; Guild, L.S.; Torres-Perez, J. Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sens. Environ.* 2015, 167, 196–205. [CrossRef]

- Lekki, J.; Anderson, R.; Avouris, D.; Becker, R.; Churnside, J.; Cline, M.; Demers, J.; Leshkevich, G.; Liou, L.; Luvall, J.; et al. Airborne Hyperspectral Sensing of Monitoring Harmful Algal Blooms in the Great Lakes Region: System Calibration and Validation; National Technical Information service: Hampton, VA, USA, 1 February 2017. Available online: https://ntrs.nasa.gov/api/ citations/20170002298/downloads/20170002298.pdf (accessed on 2 May 2022).
- Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Adv. Neural Inf. Processing Syst.* 2015, 28. Available online: https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028 a21ed38046-Paper.pdf (accessed on 2 May 2022).
- 10. Fu, J.; Liu, J.; Jiang, J.; Li, Y.; Bao, Y.; Lu, H. Dual attention network for scene segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
- 11. Li, X.; Liao, R.; Zhou, J.; Leung, P.T.; Yan, M.; Ma, H. Classification of morphologically similar algae and cyanobacteria using Mueller matrix imaging and convolutional neural networks. *Appl. Opt.* **2017**, *56*, 6520–6530. [CrossRef]
- Baek, S.S.; Pyo, J.; Kwon, Y.; Chun, S.J.; Baek, S.; Ahn, C.Y.; Oh, H.M.; Kim, Y.O.; Cho, K. Deep learning for simulating harmful algal blooms using ocean numerical model. *Front. Mar. Sci.* 2021, *8*, 1446. Available online: https://www.frontiersin.org/articles/ 10.3389/fmars.2021.729954/full (accessed on 2 May 2022). [CrossRef]
- Qian, P.; Zhao, Z.; Liu, H.; Wang, Y.; Peng, Y.; Hu, S.; Zhang, J.; Deng, Y.; Zeng, Z. Multi-target deep learning for algal detection and classification. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020.
- 14. Park, J.; Baek, J.; You, K.; Nam, S.W.; Kim, J. Microalgae Detection Using a Deep Learning Object Detection Algorithm, YOLOv3. *J. Korean Soc. Water Environ.* **2021**, *37*, 275–285.
- 15. Park, J.; Baek, J.; Kim, J.; You, K.; Kim, K. Deep Learning-Based Algal Detection Model Development Considering Field Applica-tion. *Water* **2022**, *14*, 1275. [CrossRef]
- Hill, P.R.; Kumar, A.; Temimi, M.; Bull, D.R. HABNet: Machine learning, remote sensing-based detection of harmful algal blooms. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 3229–3239. [CrossRef]
- 17. Derot, J.; Yajima, H.; Jacquet, S. Advances in forecasting harmful algal blooms using machine learning models: A case study with Planktothrix rubescens in Lake Geneva. *Harmful Algae* 2020, *99*, 101906. [CrossRef] [PubMed]
- Sonmez, M.E.; Eczacioglu, N.; Gumuş, N.E.; Aslan, M.F.; Sabanci, K.; Aşikkutlu, B. Convolutional neural network—Support vector machine based approach for classification of cyanobacteria and chlorophyta microalgae groups. *Algal Res.* 2021, *61*, 102568. [CrossRef]
- 19. Samantaray, A.; Yang, B.; Dietz, J.E.; Min, B.C. Algae detection using computer vision and deep learning. *arXiv* 2018, arXiv:1811.10847. Available online: https://arxiv.org/abs/1811.10847 (accessed on 2 May 2022).
- Medina, E.; Petraglia, M.R.; Gomes, J.G.R.; Petraglia, A. Comparison of CNN and MLP classifiers for algae detection in underwater pipelines. In Proceedings of the 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, Canada, 28 November–1 December 2017.
- Bi, X.; Lin, S.; Zhu, S.; Yin, H.; Li, Z.; Chen, Z. Species identification and survival competition analysis of microalgae via hyperspectral microscopic images. *Optik* 2019, 176, 191–197. [CrossRef]
- 22. Baek, S.-S.; Pyo, J.; Pachepsky, Y.; Park, Y.; Ligaray, M.; Ahn, C.-Y.; Kim, Y.-H.; Chun, J.A.; Cho, K.H. Identification and enumeration of cyanobacteria species using a deep neural network. *Ecol. Indic.* 2020, *115*, 106395. [CrossRef]
- Balado, J.; Olabarria, C.; Martínez-Sánchez, J.; Rodríguez-Pérez, J.R.; Pedro, A. Semantic segmentation of major macroalgae in coastal environments using high-resolution ground imagery and deep learning. *Int. J. Remote Sens.* 2021, 42, 1785–1800. [CrossRef]
- 24. Salido, J.; Sánchez, C.; Ruiz-Santaquiteria, J.; Cristóbal, G.; Blanco, S.; Bueno, G. A Low-Cost Automated Digital Microscopy Platform for Automatic Identification of Diatoms. *Appl. Sci.* **2020**, *10*, 6033. [CrossRef]
- Sardogan, M.; Tuncer, A.; Ozen, Y. Plant leaf disease detection and classification based on CNN with LVQ algorithm. In Proceedings of the 2018 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, Bosnia and Herzegovina, 20–23 September 2018.
- Wang, X.; Shrivastava, A.; Gupta, A. A-Fast-Rcnn: Hard positive generation via adversary for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
- Nieuwenhuizen, A.; Hemming, J.; Suh, H. Detection and classification of insects on stick-traps in a tomato crop using Faster R-CNN. In Proceedings of the Netherlands Conference on Computer Vision, Eindhoven, The Netherlands, 26–27 September 2018.
- Shafiee, M.J.; Chywl, B.; Li, F.; Wong, A. Fast YOLO: A fast you only look once system for real-time embedded object detection in video. *arXiv* 2017, arXiv:1709.05943. Available online: https://arxiv.org/abs/1709.05943 (accessed on 2 May 2022).
- 29. Stavelin, H.; Rasheed, A.; San, O.; Hestnes, A.J. Applying object detection to marine data and exploring explainability of a fully convolutional neural network using principal component analysis. *Ecol. Inform.* **2021**, *62*, 101269. [CrossRef]
- Kou, X.; Liu, S.; Cheng, K.; Qian, Y. Development of a YOLO-V3-based model for detecting defects on steel strip surface. *Measurement* 2021, 182, 109454. [CrossRef]
- Morbekar, A.; Parihar, A.; Jadhav, R. Crop disease detection using YOLO. In Proceedings of the 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 5–7 June 2020.
- 32. Loey, M.; Manogaran, G.; Taha, M.H.N.; Khalifa, N.E.M. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain. Cities Soc.* **2021**, *65*, 102600. [CrossRef] [PubMed]

- 33. Tian, Z.; Shen, C.; Chen, H.; He, T. Fcos: Fully convolutional one-stage object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019.
- Zhang, S.; Wen, L.; Bian, X.; Lei, Z.; Li, S.Z. Single-shot refinement neural network for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
- Zhu, C.; He, Y.; Savvides, M. Feature selective anchor-free module for single-shot object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.
- Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
- Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
- Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
- Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* 2018, arXiv:1804.02767. Available online: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html (accessed on 2 May 2022).
- 40. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934. Available online: https://arxiv.org/abs/2004.10934 (accessed on 2 May 2022).
- 41. Ultralytics. Yolov5. 2020. Available online: https://github.com/ultralytics/yolov5 (accessed on 2 May 2022).
- Wang, C.Y.; Liao, H.Y.M.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 28 July 2020.
- 43. Wang, K.; Liew, J.H.; Zou, Y.; Zhou, D.; Feng, J. PANet: Few-shot image semantic segmentation with prototype alignment. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October 2019–2 November 2019.