Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 2, No. 2: 2023
ISSN : **1906-9685**

DETECTION OF ALGAE USING CNN AND SVM ALGORITHM AND ITS UTILIZATION

Mrs. Shobha D, Assistant Professor, DoS in Computer Science, SBRR Mahajana First Grade College (Autonomous) PG Wing, Pooja Bhagavat Memorial Mahajana Education Centre, Metagalli,K.R.S Road, Mysuru, Karnataka: shobhadgps@gmail.com

ABSTRACT

Now day's microalgae are using more and more in biotechnology, foodstuff and pharmacology. Microalgae are found in water resources. To identify a large group of organisms which have photosynthetic properties and are very important to life. It's necessary to identify its behavior of harmful and to represent the health of the ecosystem and water quality, safety risks. In order to determine the unclear algae boundary and noisy background, KNN and SVM algorithm are used to segmenting algae bodies from an image background and this method is used for classification based on classifiers fusion and a new method for computing texture descriptors from a blurry texture object are considered. Feature combination approach is applied to handle a variation of algae shapes of the same genus. In segmentation process, the Binary conversion, and edge detection are used.

Keywords: Microalgae, Machine learning, classification.

1. Introduction

Algae have a common feature with vascular plants where they are eukaryotes capable of photosynthesis with chlorophyll as their primary pigment but other morph anatomical features among vascular plants such as true roots, stems, and leaves [1]. To identify algae concentration in the field, researchers have developed a handheld, low-cost in-situ device employing spectrophotometry and optical filtering. Most algae are aquatic while others are terrestrial that can be found on moist soil, trees, and rocks. Some of them are unicellular and others are multicellular, they can live in colonies or have a leafy appearance such as seaweeds, also the size of species varies from microscopic to giant kelp with millions. Algae are important microscopic aquatic life forms as they are primary producers in an aquatic food chain and oxygen producers in an aquatic ecosystem.

2. Machine learning algorithms to detect the algae

The machine learning process is the analysis and learning of data with algorithms, Machine learning contains supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning, depending on the training model. Many different models can be used for machine learning training, and a comprehensive description of centralized representative models are in the following and subsequent judgment and prediction about the actual situations are made automatically (Wei et al., 2019) . A framework with many parameters is first built, and then the prepared data is fed into the model. The parameters are continuously adjusted until they match or close to the correct result (Bishop, 2013).

3. Microalgae detection and classification with machine learning

Method

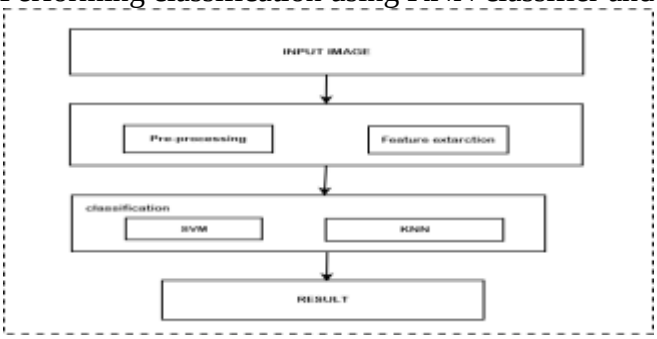
As unicellular organisms, microalgae are microscopic and they do not differ much from one species to another. The thousands of species of microalgae may be present in a tiny sample, microalgae Classification is a very challenging work. Traditional manual classification under a

microscope is not only laborious, but also requires a high level of skill and experience for the operators and its time consuming. The manual classification method of Microalgae accuracy is usually inefficient and unsatisfied. Machine learning algorithms are used in detecting microalgae and classification . This algorithm allows computers to automatically learn the characteristics of different algae based on existing data and give classification results for new data. The data microalgae images obtained through microscopy processed by machine learning algorithms.

The dataset consist of microalgae images and labels. Segmentation of microscopic algae manually from Image is a difficult task. It depends on how well structural perceive the image. Similarities between different types of algae tissues are very high resemblance. The contribution presenting classification of algae based automatic detection of features of algae based on their area, centroid, maximum angle, minimum angle and shape. This method helps in detecting algae features based on their area of the images. Microalgae counting are used to measure biomass quantity. Usually, it is performed in a manual way using a Neubauer chamber and expert criterion, with the risk of a high error rate. To improve this process, classification is performed. Once calculated, these moments provide measurements for colour similarity between images. These values of similarity can be compared to the value of images indexed in a database for tasks like image retrieval.

4. Methodologies:

- ❖ Collecting standard algae image in database
- ❖ Converting RGB image into grayscale image
- ❖ Converting grayscale image into binary image
- ❖ Extracting features using standard Area, centroid, maximum angle, minimum angle and shape from matlab code
- ❖ Performing classification using KNN classifier and SVM



4.1 Modules:

- Image acquisition
- Preprocessing
- Feature extraction
- Classification

4.1.1 Image Acquisition:

The image acquisition is first step and is essential for the rest of the system. In order to capture high quality images the iphone 5s camera is used,equipped with 8 megapixels and 1.5 pixels. Each pixel represents the measurement of some property of a scene measured over a finite area. The bright nesses of the image filtered through red green and blue filters (three values).

The Image viewer app provides all the image display capabilities; it provides access to several other tools for navigating and exploring images.

4.1.2 Preprocessing:

Image Resize: when scaling a vector graphic image, the graphic primitives that make up the image can be scaled using geometric transformations, with no loss of image quality.

Image edge: Then the image smoothing is further processed with sobel edge detection. Then the featured image is used as test image in the classification and then the sobel logic output is determined.

4.1.3 Feature extraction:

Textural features are extracted for both query image and images in the database. The distance (i.e., similarities) between the features vectors of the query image and database are then computed and ranked. The database images that have highest similarity to the query image are retrieved. Then the performance analysis is carried out using precision and recall. Content based image retrieval is working with different types of image database.

All databases are having two types of images like labeled and unlabelled. A preprocessing is a process for preparing an input image to be suitable for processing (i.e, segmentation and feature extraction). The first preprocessing process is to resize an input image. Since sizes of our input images are largely varied, we need to resize them into the same scale in order to correctly compute algae shape features, particularly, shape measurement features. The images whose longest side is larger than 256 pixels are resized to 256 pixels while the other side remains in the same aspect ratio.

4.1.4 Classification:

K- Nearest Neighbour

KNN is a non-parametric lazy learning algorithm it means that it does not make any assumptions on the underlying data distribution. It is also a lazy algorithm. More exactly, all the training data is needed during the testing phase. KNN – makes decision based on the entire training data set.

Assumptions in KNN

KNN assumes that the data is in a feature space. More exactly, the data points are in a metric space. The data can be scalars or possibly even multidimensional vectors. Since the points are in feature space, they have a notion of distance .Each of the training data consists of a set of vectors and class label associated with each vector. In the simplest case, it will be either + or – (for positive or negative classes). But KNN, can work equally well with arbitrary number of classes.

It is also given a single number "k". This number decides how many neighbours (where neighbours is defined based on the distance metric) influence the classification. This is usually an odd number if the number of classes is 2. If k=1, then the algorithm is simply called the nearest neighbour algorithm. KNN (K Nearest Neighbours) finds the k nearest elements of testing pattern from training data without considering the class labels. KNN assigns the testing pattern to that class which has higher number of elements in k elements.

Algorithm: K -Nearest Neighbour

Input: Training data and testing pattern.
Output: classified tested pattern.
Step 1: Determine the value of k.
Step 2: Find the k nearest elements of testing pattern from training data.
Step 3: Calculate the frequency of elements of all classes present in the k elements.
Step 4: Assign the testing pattern to the class, which has highest frequency.

- **Euclidean Distance**

$$D_A(x, y) = \sum\sqrt{x_i}\,2 - y_i 2$$

The feature vectors must be normalized before classification algorithm run. The overall

KNN algorithm is running in the following steps:

1. KNN algorithm uses training set that consist features and labelled classes store in database.
2. Calculate distance between test features and all training feature.
3. Sort the distance and determine k nearest neighbour.
4. Use simple majority of the category of nearest neighbour assign to the test.

The KNN classification algorithm is performed by using a training set which contains both the input feature and the labelled classes and then by comparing test feature with training feature a set of

distance of the unknown K nearest neighbours determines. Finally test class assignment is done by either averaging the class numbers of the K nearest reference points.

Original Algae images

5. Advantages:

- This algorithm can correctly separate the regions that have the same properties which are define.
- These methods can provide the original images which have clear edges the good segmentation results.

5.1Algae utilization:

In water resource management, algae are used as a biological index to indicate a quality of water because they are sensitive to environmental changes. Therefore, recognition of microalgae is one of the most important issues in water resource management.

Microalgae counting are used to measure biomass quantity. Usually, it is performed in a manual way using a Neubauer chamber and expert criterion, with the risk of a high error rate. To improve this process, classification is performed.

Microalgae in the ocean are usually single-celled organisms that play a crucial part in marine ecology (Chew et al., 2017).

Algae are important microscopic aquatic life forms as they are primary producers in an aquatic food chain and oxygen producers in an aquatic ecosystem. In water resource management, algae are used as a biological index to indicate a quality of water because they are sensitive to environmental changes.

6. Conclusion:

In this work classification of algae based on classifiers fusion and the classification is based on the extract shape features like area, centroid, angle at centroid, maximum angle, and minimum angle of the datasets. It is concluded that it is very challenging task build up an automatic algae classification for all type of algae

Future Enhancement :

Classification is executed sequentially in proposed method but for huge training data and testing data classification takes more time. Parallel computing can be applied on classification process to get result faster.

References

[1] S.Z. Boujelbene, D. Ben AyedMezghani, and N. Ellouze. Vowel phoneme classification using SMO algorithm for training support vector machines. In Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008. 3rd International Conference on, pages 1–5, April 2008.

[2] John Canny. A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, PAMI-8(6):679–698, November 1986.

[3] S. Conseil, S. Bourennane, and L. Martin. Comparison of Fourier descriptors and Hu moments for hand posture recognition. In European Signal Processing Conference (EUSIPCO), 2007.

[4] RiaRodette G. de la Cruz, Trizia Roby-Ann C. Roque, John Daryl G. Rosas, Charles Vincent M. Vera Cruz

[5] Oana Astrid Vătămanu, MirceaJivulescu, “Image Classification Using Local Binary Pattern Operators for Static Images”,8th IEEE International Symposium on Applied Computational Intelligence and Informatics, Timisoara, Romania, pp. 173-178, May 23- 25, 2013.
[6] V. Karunakaran, “A Survey on Image Classification Methods”, International journal of advance research in Electronics and Communication Engineering (IJAREC) , Vol. 1 , No. 4, pp. 69-72, 2012.
[7] PoojaKamavisdar , SonamSaluja, Sonu Agrawal, “A Survey on Image Classification Approaches and Techniques”, International Journal of Advanced Research in Computer andCommunication Engineering, Vol. 2, No. 1, pp. 1005- 1009, January 2013.
[8] S. Arunadevi, Dr. S. DanialMadan Raja, “A Survey on ImageClassification Algorithm Based on Per-pixel”, International Journal of Engineering Research and General Science, Vol. 2, No. 6, pp. 387-392, November 2014.
[9] S.Z. Boujelbene, D. Ben AyedMezghani, and N. Ellouze. Vowel phoneme classification using SMO algorithmfor training support vector machines. In Information and Communication Technologies: From Theory toApplications, 2008. ICTTA 2008. 3rd International Conference on, pages 1–5, April 2008.
[10] John Canny. A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEETransactions on, PAMI-8(6):679–698, November 1986.
[11] S. Conseil, S. Bourennane, and L. Martin. Comparison of Fourier descriptors and Hu moments for handposture recognition. In European Signal Processing Conference (EUSIPCO), 2007.
[12] RiaRodette G. de la Cruz, Trizia Roby-Ann C. Roque, John Daryl G. Rosas, Charles Vincent M. Vera Cruz,MacarioO. Cordel, Joel P. Ilao, Adrian Paul J. Rabe, and J.ParungaoPetronilo. SMO-based system for identifying common lung conditions using histogram. In Medical Information and Communication Technology(ISMICT), 2013 7th International Symposium on, pages 112–116, March 2013.
[13] S. N. Deepa and B.A. Devi. Neural networks and SMO based classification for brain tumor. In Informationand Communication Technologies (WICT), 2011 World Congress on, pages 1032–1037, December 2011.