2011 International Conference on Environmental and Computer Science
IPCBEE vol.19(2011) © (2011) IACSIT Press, Singapore

Automatic Recognition System for some cyanobacteria using image processing Techniques and ANN approach

Hayat Mansoor [1], Sorayya M [1],Aishah S[1], and Mogeeb A.A Mosleh[2]

[1] Institute of Science Biology, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia
[2] Faculty of Computer Science & Information Technology, University of Malaya Kuala Lumpur, Malaysia

Abstract. The aim of this study is to detect and classify four common cyanobacteria genera automatically that are found in freshwater tropical Putrajaya Lake (Malaysia). Common genera of cyanobacteria such *Microcystis*, *Oscillatoria*, *Anabaena* and *Chroococcus* are given special attention in this study. They are associated with Lake Eutrophication and hence their presence is useful in determining the status of a freshwater lake. Combination of image processing techniques and artificial neural network (ANN) algorithms are used to automate the process of detection and recognition of the four common cyanobacteria genera Water samples were collected and images of the specimens were captured using camera-attached light microscope. Images of the cyanobacteria specimens were subsequently analyzed and matched with cyanobacteria in authoritative references and literatures on cyanobacteria taxonomy. The microscope picture of cyanobacteria were captured and transferred to the computer storage device. Image filtering was then applied to improve image quality and enhanced image features. Then edge detection algorithm was used to analyze the shape and texture of image objects. Image segmentation was applied to obtain the cyanobacteria inside the image based on shape and feature analysis. Finally, ANN was used to identify the input images. ANN was trained using 20 images for each type of cyanobacteria genera. Lastly, system evolution was performed to verify from system accuracy. The system identified 71 images from 80 image samples correctly and detection accuracy rate was more than 95%.

Keywords: Automatic Detection, Cyanobacteria, Image processing, ANN, Algae Recognition, Algae.

1. Introduction

Algae play an important role in both the long and short term determination of water quality in freshwater lakes. Algae are affecting water properties such as water colour, odour, taste, and chemical compounds which may cause potential hazards for human and/or animal health. Many studies have reported that algae can affect drinking water process treatments and some traditional process are used traditionally to enhance algae removal in coagulation process such as pre-oxidation by ozone, chloride dioxide, and chlorine or permanganate [1],[2].

Cyanobacteria have become as one of the critical problem worldwide and their blooms are typical phenomena in eutrophic lakes. Surveys in various countries have demonstrated that about 75% of samples containing cyanobacteria are toxic [6], [7]. World Health Organization (WHO) and several national authorities worldwide have recommended risk assessment plans and safety levels to include cyanobacteria as a parameter for water quality control [8], [9], [10].

Many studies have been done on cyanobacteria to explore more about its features, category, and effect in the environments, but there are few or limited studies on various aspects of tropical freshwater algae. A study to asses eutrophication status on 90 lakes in Malaysia showed that 56 lakes or 62% were eutrophic or in bad situation and requires immediate rehabilitation and restoration, and the other 34 lakes which represent

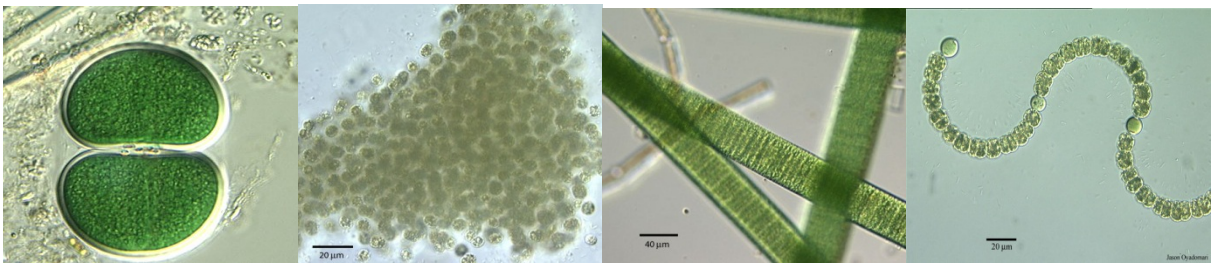
+ Corresponding author. Tel.: + 60132773344; fax: +603-79579249.
E-mail address: MogeebMosleh@yahoo.com.

38% of the study is classified as mesotrophic. Moreover, it is reported that the most common phytoplankton in Malaysian lakes is cyanobacteria and *anabaena* was responsible for most of cyanobacteria bloom [11].

Recently, image processing and computer vision has been growing at a fast pace, and computers and workstations have become powerful enough to process image data. Image processing has to be a standard scientific tool and image processing techniques are now applied to virtually all natural sciences and technical disciplines [12].

Most research applied an image processing to detect, count, identify, and classify the algae groups; some of this approach was efficient with 92% in accuracy [13], [15]. Some other was used to determine the abundance, size, shape, volume and surface area for different microorganism with a great accuracy [14]. At the end of 90's image processing with some other technique such as fuzzy logic, genetic algorithm, and ANN began to be used for better accuracy in classification and recognition of microorganisms. Some developed tools is used effectively for online monitoring, some tools is used for measurements density of microorganism in water, other tools were developed to assist in recognition process such as enhancing images, noise elimination, edges extracted segmentation [19], [20]. Others techniques used image processing with genetic algorithms or ANN for recognition purpose [16], [17].

Unfortunately, most of existing research developed tools for one or two certain type of algae thus because the difficulties in detecting all types of algae in one tool, and also due to the differences in algae shapes [18], [21]. Therefore the aim of this study is planned to develop automatic system using image processing with ANN for detection and recognition four types of cyanobacteria genera (*Microcystis, Oscillatoria, Chroococcus, Anabaena*) of freshwater tropical Putrajaya Lake. Figure 1 illustrates the four common genera of cyanobacteria studied.



*(a)*Microcystis *(b)* Oscillatoria *(c)* Chroococcus *(d)* Anabaena

Fig. 1: Cyanobacteria target research in this study.

2. Methodology

Water samples were collected from different sampling sites at Putrajaya Lake. The water samples were analyzed and examined using electronic microscope (MTC#B1-220ASA). A microscope Eye – Piece camera (AM432X) is attached to the microscope lens and connected with pc via USB port for image acquisition. It was used to capture, load, and store the images into computer directly. Our data set contained four genera of cyanobacteria; 100 image samples were collected for each genus. All sample images collected and divided into two groups, one group which were used for ANN training purpose and another group for system testing to avoid biasness in results.

2.1. System Development

Matlab ver.14 used for system development process because it has an integrated technical computing environment that is suitable for algorithm design and development, and it is a high-level programming language that includes hundreds of functions in several areas.

The image processing consists of the following process which must be performed sequentially, Image acquisition, Image enhancement, Image segmentation, Feature extraction, Classification and neural training, and Detection/recognition process.

Each task must execute in sequence because output from each process is the input for next step. System flowchart is shown in Figure 2, and described in more details below:

2.1.1. Image pre-processing:

Images uploaded manually to the system, might suffer from low contrast quality and it may contain some noise and/or unwanted area. A pre-processing for captured images is necessary to enhance the images in order to produce clearer details which can improve the overall appearances. A filter model was used to make the uploaded image clearer.

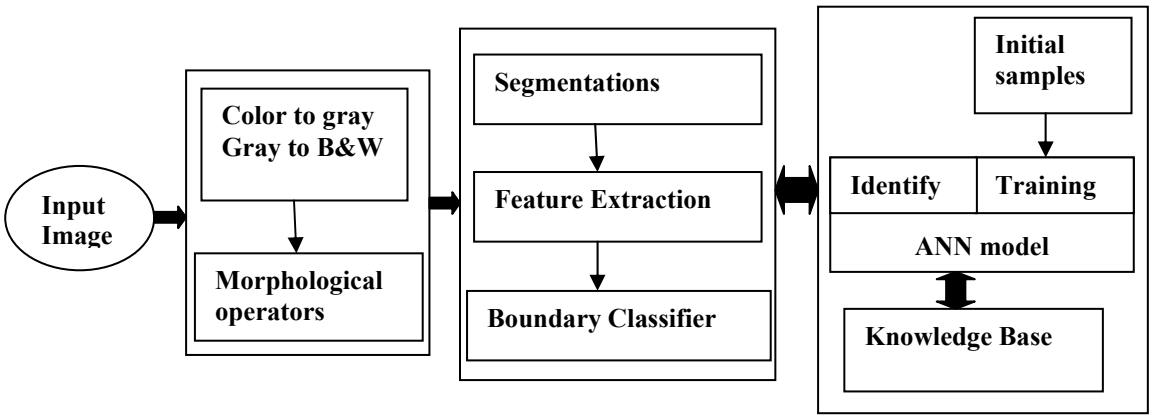


Fig. 2: Steps of automatic algae recognition system.

Neighbourhood averaging was used to reduce image noise. Then, morphological operations were performed for noise elimination and edge extraction of selected cyanobacteria. Morphological model was used to smooth the original image, suppress noise, keep the cyanobacteria body clearer, and make its branches thinner. Morphological operator was applied to reconstruction image via extract some lost fragments of cyanobacteria contour and accompanied with edge detection simultaneously.

Then, Image segmentation was used to cluster pixels into significant image regions. Particularly, it is used to identify the location of feasible objects and their boundaries. Most microorganism images contain many other objects inside the images. Therefore image segmentation was used to consolidate the objects included inside the images for sub images to process separately. The segmentation process is applied to the images using threshold techniques based on Laplace of the Gaussian (LOG) operator. Threshold is used to convert gray image to a binary images, and LOG operator is used to be as mask to the binary image which gives three different values to the pixel value for output image, where object pixels have a value 1, background pixels have a value 2, and image border pixels have a value of 0.

2.1.2 Feature Extraction:

Feature extraction was used to transform the input data into a set of values that describe the object features under consideration. In our study there are four genera of cyanobacteria needed to be identified, and each genera has it is own specific feature which vary with the others. Feature extraction techniques used was independent features such as color, texture, and shape for extraction of object features. A principal component analysis (PCA) method was used for extracted feature; the image of N x M pixels containing P points is represented as a superposition of the eigenvectors as given below:

$$R_i(\omega_j)=\sum_{k=1}^{P} a_{ik} w_k(\omega_j),$$

Where aik ≡ ak (x,y) are position-dependent expansion coefficients, and Ri(xj) ≡ R(x,y,xj) are responses at different spatial locations. PCA allows simplifying and decorrelating the multivariate statistical data, and its components are used as an input for the recognition using ANN.

After, feature extraction, a classifier based on object boundary is applied for classifying cyanobacteria to three different categorizations selected based on their shape boundary, the selected categorization is circularity, spiral, and irregular. The classifier is used to optimize ANN recognition time process and to satisfy better accurate results. Then, the output of this classifier is used as input to the ANN, the output equal '0' if circular, '1' for spiral, and '2' if irregular. This classifier can be extended in further study to assist for recognition all types of exist algae.

2.1.3. Artificial Neural Network (ANN):

Multilayer perceptron, feed forward ANN was used to perform identification process for selected cyanobacteria genera. The ANN architecture consists of 6 inputs, 3 outputs, 3 neuron in hidden layer, 0.78 for learning rate, and 0.5 for momentum. The database included 4 genera of cyanobacteria, 80 variation images for each type. The classifier is used to index the database content during training mode for categorizing purposes. The target output of ANN is (0,0,1) for Microcystis , (0,1,0) for Oscillatoria , (1,0,1) for Chroococcus , and (1,1,1) for Anabaena. More importantly, the characteristic response is close to 1 within the each cyanobacteria type, and is essentially zero outside.

2.2. System evaluation and experimental results

System interface is shown in figure 4 below. The system functions have been tested separately for evaluation purpose.

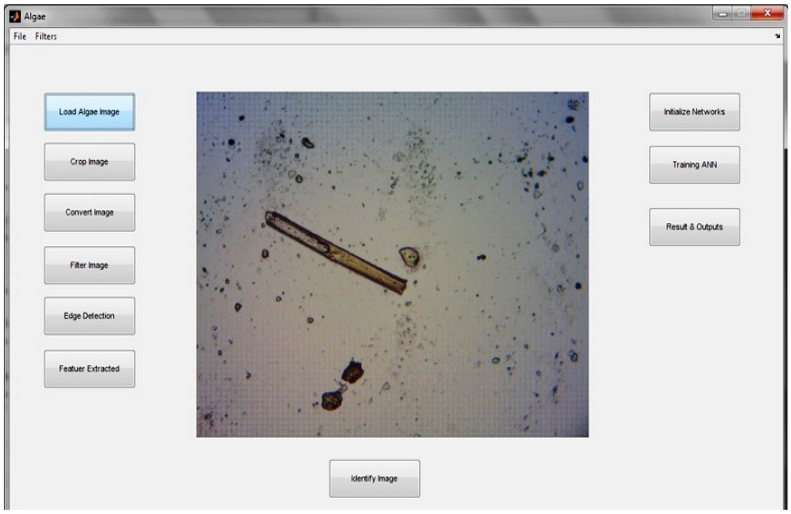


Fig. 4: System Interface.

In these study 400 samples of 4 cyanobacteria genera is considered. Each genus has 100 samples that were used for training and testing the ANN. In the training mode 320 samples was used to train the ANN, 80 samples for each type. For system testing the remaining 80 samples, 20 samples for each genus were used. The classification results are shown in table 1 for each cyanobacteria genus. The results illustrate as more than 95% success in identifying and classification the input samples of 4 genera of cyanobacteria.

Table 1. System testing for cyanobacteria image recognition

Cyanobacteria genera	No. of Tested Case	Predicated results				
		1	2	3	4	others
Microcystis	20	19	0	0	0	1
Oscillatoria	20	18	0	0	1	1
Chroococcus	20	17	0	0	0	3
Anabaena	20	17	2	0	0	1

3. Conclusion

This study has shown that the automated recognition of freshwater tropical algae is certainly important and feasible. Using image processing techniques with combination of ANN can achieve accurate results over 95% for identifying and classifying cyanobacteria and algae without any user interaction. The pre-processing image technique are enhanced the appearance of cyanobacteria significantly. Extracting texture feature for algae and/or spices is given more accurate information about their type, and it is very suitable for classifying them using ANN.

4. Future work

This research can be considered as a preliminary study towards the development of computer software that can identify and recognize all different types of species and/or algae. The new technique that developed as classifier can be extended to categorize all exist type of algae based on their boundary for optimization ANN process time.

5. Acknowledgements

The authors would like to thank University of Malaya, for supporting and providing all the equipments and facility to perform this study.

6. References

[1] S. Gao, J. Yang, J. Tian, F. Ma, G. Tu, M. Du. Electro-coagulation–flotation process for algae removal. *Journal of Hazardous Materials*. 2009, 177: 336–343.

[2] Gilbert, J. J. Effect of food availability on the response of planktonic rotifers to a toxic strain of the cyanobacterium Anabaena flos-aquae. Limnol. Oceanog. 1996, 41:1565–72.

[3] Komarek, J. Coccoid and colonial cyanobacteria.In Freshwater algae of north America. J.D Wehr and R.G Sheath, eds.Maryland Heights,Mo. 2003.

[4] Tisdale, E. S. Epidemic of intestinal disorders in Charleston, occurring simultaneously with unprecedented water supply conditions. *American Journal of Public Health.* 1931, 21: 198-200.

[5] W. Y., Chen, J. P., Wang, X. M., and Xu, L. H. Altered expression of p53, Bcl-2 and Bax induced by microcrystal-LR in vivo and in vitro. *Toxicon.* 2005*,* 46: 171-177.

[6] Chorus, I., Falconer, I.R., Salas, H.J., Bartram, J. Health risk caused by freshwater cyanobacteria in recreational waters*. J.Toxicol. Environ. Health B*. 2000, 3: 323–347.

[7] Azevedo, S. New Brazilian regulation for cyanobacteria and cyanotoxins in drinking water. In: 5th International Conference on Toxic Cyanobacteria, Noosa, Australia, 2001.

[8] Falconer, I.R., Toxic cyanobacterial bloom problems in Australian waters: risk and impacts on human health*. J.Phycologia.* 2001, 40: 228–233.

[9] Codd, G.A., Morrison, L.F., Metcalf, J.S. Cyanobacterial toxins: risk management for health protection*. Toxicol. Appl. Pharmacol.* 2005, 203: 264–272.

[10] Walsby, A.E., Avery, A. Measurement of filamentous cyanobacteria by image analysis. *J. Microbiol. Methods*. 1996, 26:11–20.

[11] M.Y. FATIMAH, A.K. MOHSIN and A.S. M. KAMAL. Phytoplankton Composition and Productivity of a Shallow. *Tropical Lake.Pertanika.* 1984, **7**(3): 101-113.

[12] Jähne, B. Digital image processing. 5th rev.*London. Springer*.2002.

[13] Jefferies, H. P., Berman, M. S., Poularikas, A. D. Automated sizing, counting and identification of zooplankton by pattern recognition. *Mar. Biol*. 1984, 78: 329–334.

[14] Estep, K. W., Maclntyre, F., Hjorleifsson, E., and Sieburth, J. MeN. 1986. MacImage: *a user-friendly image-analysis system for the accurate mesuration of marine organisms*. Marine Ecology Progress Series, 33: 243-253.

[15] Katsinis. C., Poularikas, A.D. Image processing and pattern recognitionwith applications to marine biological images. In: *Proc. SPIE 7th meeting on applications of digital image processing*, San Diego.1984, pp. 324-329.

[16] Schultze-Lam, S., Harauz, G., and Beveridge, T.J. Participation of a Cyanobacterial- S Layer in Þ ne-grain mineral formation. *Journal of Bacteriology.* 1992, 174: 7971-7981.

[17] Blackburn N, Hagström Å, Wikner J, Cuadros-Hansson R, Bjørnsen PK. Rapid determination of bacterial abundance, biovolume, morphology, and growth by neural network-based image analysis. *Appl Environ Microbial.* 1998, 64: 3246–3255.

[18] Wilkins MF, Boddy L, Morris CW, Jonker RR. Identification of phytoplankton from flow cytometry data by using radial basis function neural networks. *Applied and Environmental Microbiology*. 1999, 65: 4404–4410.

[19] Sandesh B. K., Shalini C. , B.R. Brinda , M.A. Kumar. Digital image processing—an alternate tool for monitoring of pigment levels in cultured cells with special reference to green alga Haematococcus pluvialis, *Biosensors and Bioelectronics*. 2005, 21: 768–773.

[20] Junna Cheng, Guangrong Ji, Chen Feng, Haiyong Zheng Application of Connected Morphological Operators to Image Smoothing and Edge Detection of Algae. *International Conference on Information Technology and Computer Science*. 2009, pp. 73-76.

[21] Z. Yao, M. Fei, K. Li, H. Kong, B. Zhao.(2007).Recognition of blue-green algae in lakes using distributive genetic algorithm-based neural networks.*Journal of Neuro computing*.Vol. 70, pp 641–647.