International Journal of Grid and Distributed Computing
Vol. 11, No. 9 (2018), pp.93-102
http://dx.doi.org/10.14257/ijgdc.2018.11.9.08

A Study on Red Tide Detection Technique by using Multi-Layer Perceptron

Su-Ho Bak[1], Do-Hyun Hwang[1], Heung-Min Kim[1], Bum-Kyu Kim[1], Unuzaya Enkgjargal[1], Seong-Yeol Oh[2] and Hong-Joo Yoon[1*]

[1]Division of Earth Environmental System Science, Pukyong National University, Yongso-ro 45, Nam-gu, Busan, 48513, South Korea
[2]Oceanic Climate & Ecology Research Division, National Institute of Fisheries Science, 216, Gijanghaean-ro, Gijang-eup, Busan, 46083, South Korea
**yoonhj@pknu.ac.kr, Romeo98x@nate.com*

Abstract

This study proposed a red tide detection technique for Cochlodinium polykrikoides(C. polykrikoides) by using Normalized water-leaving radiance from GOCI (: Geostationary Ocean Color Imager) data and 2 stage filtering algorithm based on multi-layer perceptron. We designed the algorithm to classify the type of seawater into 3 classes (Red tide, Clear water, Turbid water). The proposed algorithm has been developed to remove the clear water pixels first and then the turbid water pixels in the satellite images. As a result of the evaluation of the detection accuracy using the verification data, the total accuracy was confirmed to be about 97%. Multi-layer perceptron based algorithm can extract features from input data using their internal structures that is called hidden layer. Nevertheless, it showed similar performance to the logistic regression model using statistical variable selection method and showed higher accuracy than the decision tree model. The results of this study can contribute to reduction of red tide damage through early detection and monitoring of red tide.

Keywords: *Harmful Algal Bloom, Red Tide, Machine Learning, Artificial Neural Network, Multi-Layer Perceptron*

1. Introduction

Red tide is a phenomenon in which seawater is discolored or physically damaged by the massive growth of phytoplankton in special environment condition. In Korea it had suffered 7.2 million dollars a loss due to *Cochlodinium polykrikoides(C. polykrikoides)* red tide since 1995. In Korea, red tide occurred due to diatoms until 1980s but since 1990s increasing frequency of red tide due to dinoflagellates. Especially, *C. polykrikoides*, which is a kind of dinoflagellates, is the most frequent occurrence, and the amount of damage caused by it is also considerable.

Detection to red tide occurrence at the early stage of occurrence and preparing countermeasures is the most effective in mitigating the damage (Kim *et al.*, 2007). It is difficult for monitoring red tide due to dinoflagellates using ship because dinoflagellates are occurred wide areas. Therefore, we need for remote sensing based on satellite (Oh and Yoon, 2012).

The initial satellite image-based red tide detection technique used mainly Chlorophyll-*a* concentration from satellite images (Stumpf *et al.*, 2003; Tomlinson *et al.*, 2004). However, the Chlorophyll-*a* concentration from satellite images has a disadvantage that the accuracy

Received (June 6, 2018), Review Result (August 24, 2018), Accepted (August 31, 2018)
* Corresponding Author

ISSN: 2005-4262 IJGDC

is low in areas with high concentrations of dissolved organic matter and suspended matter (Bak *et al.*, 2016). In order to solve these problem, detection techniques using optical properties of seawater have been attempted recently (Kim *et al.*, 2007; Kim *et al.*, 2009; Son *et al.*, 2011; Son *et al.*, 2012; Bak *et al.*, 2016). The algorithms proposed by previous studies are based on the gradient change in the specific band interval observed after red tide species has blooming. The method using the optical properties of seawater is advantageous in that it is less distorted by the atmosphere.

However, this method requires a high understanding of the optical properties of the phytoplankton to be detected. Detection is possible using a simple mathematical model for red tides from Case-1 seawater with low concentrations of dissolved organics and suspended matter. When the red tide is blooming, the concentrations of chlorophyll-*a*, a photosynthetic pigment present in phytoplankton cells, and other carotenoid pigments are also increased.

The photosynthetic pigments and auxiliary pigments of phytoplankton absorb mainly blue light of 430~500nm wavelengths, resulting in changes in the water-leaving radiance gradient between blue and green wavelengths. In the case of seawater without red tide, the water-leaving radiance of blue wavelength band is generally higher than that of green wavelength band, but in case of red tide sea water, this gradient between blue and green band is reversed and the water-leaving radiance of green wavelength band is higher. However, in the case of Case-2 seawater with a high concentration of dissolved organic matter and suspended matter, it is impossible to detect red tide using this property. This is because dissolved organic matter also shows a high absorbance at the blue wavelength range, and therefore the water-leaving radiance of the green wavelength is observed higher than the blue wavelength in the sea water in which no red tide is generated. Therefore, to detect the red tide in Case-2 seawater, we need to add more terms to explain the complexity of Case-2 in the mathematical model for existing Case-1 seawater.

However, since Case-2 seawater produces a wide variety of background fields depending on the ratio of dissolved organic matter and suspended matter, and the respective concentration, it is very difficult to produce a mathematical model corresponding to each background field. In addition, it is also a difficult task to interpret the complex differences in the blue band where dissolved organic matter and phytoplankton commonly cause absorption. To solve this problem, we propose a red tide detection algorithm using multi-layer perceptron.

2. Data and Method

2.1. Dataset

In order to study the machine learning model, we used the spectral profiles of visible light and near infrared ray exported light spectra in the red tide, the clear water (Case-1), and the turbid water (Case-2). The spectral profiles were produced from the GOCI Level 1B data from COMS (: Communication, Ocean and Meteorological Satellite) provided by Korean Ocean Satellite Center. In order to obtain the spectral profile in the red tide area, red tide breaking data from the NIFS (: National Institute of Fisheries Science) was used (Figure 1).

GOCI (: Geostationary Ocean Color Imager) is a satellite payload of COMS with a spatial resolution of 500m and has 6 channels (412, 443, 490, 555, 660, 680nm) in the visible light region and 2 channels (745, 865nm) in the near infrared region (Table 1). A Level 1B image is a radiometric and geometric corrected image. We calibrated this image using the GDPS (: GOCI Data Processing System). The water-leaving radiance of Level 2 data produced after atmospheric correction was used as a dataset for model training.

Table 1. Band Composition of GOCI

Band	Centroid Wavelength (nm)	Bandwidth (nm)
1	412	20
2	443	20
3	490	20
4	555	20
5	660	20
6	680	10
7	745	20
8	865	40

The red tide breaking data provided by NIFS is provided once a day when red tide occurs, and includes the location of the red tide area and the information about the causative organism, biological density, and water temperature at the corresponding location. In this study, we used the map of red tide occurrence included in the red tide breaking data (Figure 1).

The map of red tide occurrence is images, it is not included location information in the image data. Therefore, we extracted latitude and longitude data in red tide occurred area from images (map of red tide occurrence) after assigning coordinate system to images by using georeferencing.

Georeferencing is performed by inputting actual coordinates to a reference point on a photograph or a map by assigning real-world coordinates to a map of an aerial photograph or an image having no coordinate value, and all other coordinates are relative to the reference point is input as a coordinate value.

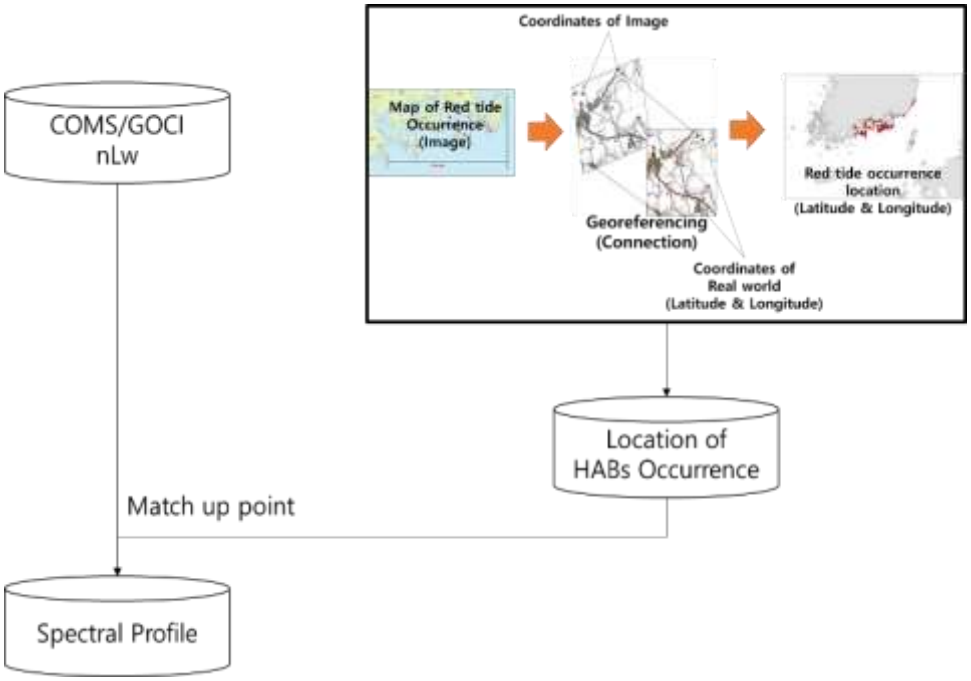

Figure 1. Spectral Profile Production Process

Location information of the red tide area obtained as a result of georeferencing was used to find the match-up point with the GOCI image. A spectral profile of the red tide occurred area was obtained from the match-up point. The spectral profiles of the red tide not occurred area were obtained by random sampling at GOCI images at the time when no red tide occurred.

The final dataset was divided into Training Dataset and Test Dataset for over-fitting. 70% of the total data was used as the training dataset. The remaining 30% was used as a test dataset. At this time, each class of Test Dataset was adjusted to have the same number of data. This is to avoid imbalance data. Each class information is one-hot encoded.

2.2. Red Tide Detection using Spectral Profile

Sea-surface discoloration is caused by phytoplankton, dissolved organic matter, and suspended matter. When red tide occurs, the number of phytoplankton is increased and discoloration of seawater occurs. The discoloration caused by phytoplankton is due to the pigment of phytoplankton (Yoon, 2012; Bak *et al.*, 2016). Depending on the species of phytoplankton, the kinds of pigment present in the cell are also different.

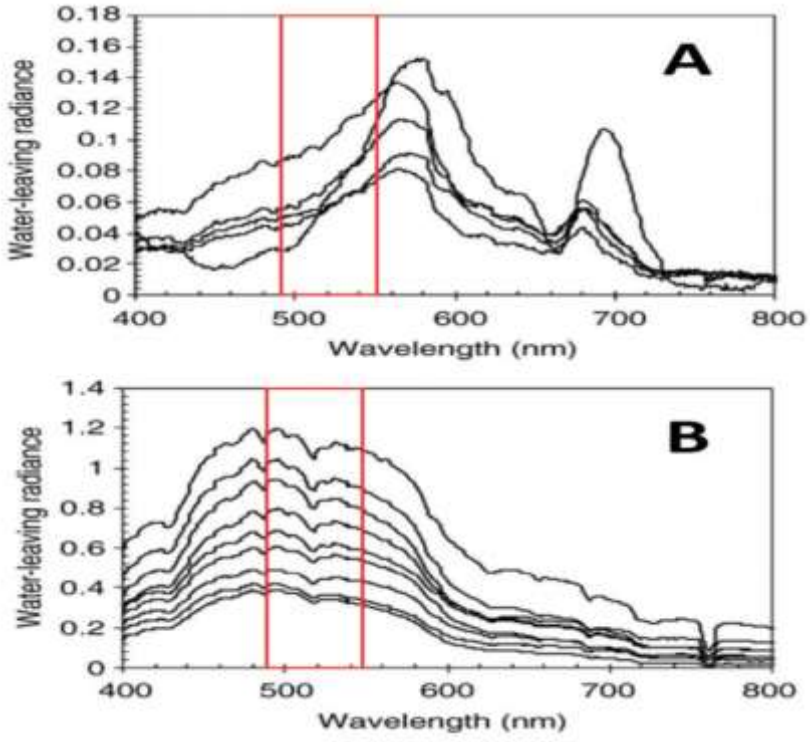

Figure 2. Water-leaving Radiance Spectral Profile in Visible Band. Water-Leaving Radiance Spectral Profile of Red Tide Occurred Area(A). Water-Leaving Radiance Spectral Profile of Clear Water Area(Case-1)(B). The Red Square Shows a Wavelength Range where 490nm(Blue) to 555nm(Green)(Ahn et al., 2006)

C. polykrikoides has pigments such as Chlorophyll-*a*, -c, β-carotein, and Peridinin, and discolors the seawater to brown in the blooming period (Yoon, 2012). Figure 2 is a graph showing the spectral profiles of seawater in which *C. polykrikoides* blooms occurred and clear water in which no red tide occurred. As shown in the graph, red tide and clear water have opposite spectral gradient between blue and green bands. In case of clear water, water-leaving radiance tends to decrease as the wavelength

increases from 490nm to 555nm. Therefore, it is possible to distinguish the red tide area from the clear water area using the difference between the water-leaving radiance of the 490 nm band and the water-leaving radiance of 555 nm (formula 1). The pixels satisfying the formula (1) are the area where the red tide occurs, and if not satisfied, they can be classified as clear water.

$$\mathrm{nLw}(490) - \mathrm{nLw}(555) < 0 \quad (1)$$

However, Turbid water (Case-2), which has a high concentration of dissolved organic matter and suspended matter, cannot be distinguished using equation (1) alone. Figure 3 is a graph showing the spectral profiles of the seawater in which the *C. polykrikoides* red tide occurred and the turbid water in which no red tide occurred. In turbid water, unlike clear water, the spectral slope between the blue and green bands is similar to the sea water in which the red tides occur. Therefore, when red tide is detected using equation (1), it can be detected together with not only red tides but also turbid waters.

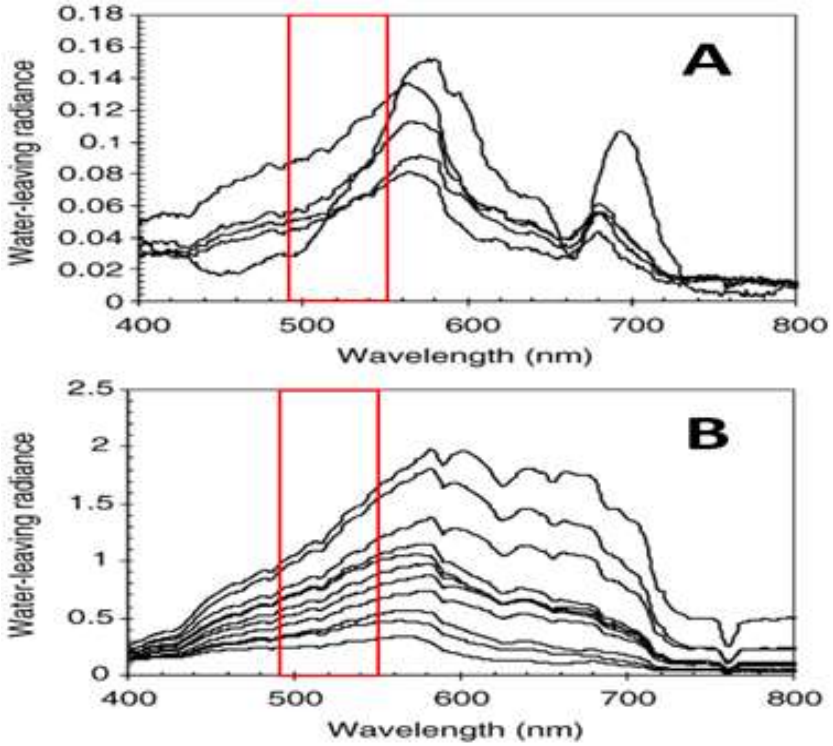

Figure 3. Water-leaving Radiance Spectral Profile in Visible Band. Water-leaving Radiance Spectral Profile of Red Tide Occurred Area(A). Water-Leaving Radiance Spectral Profile of Turbid Water Area(Case-2)(B). The Red Square Shows a Wavelength Range where 490nm(Blue) to 555nm(Green)(Ahn *et al.*, 2006)

To solve these problems, previous studies used the difference in water-leaving radiance of red bands (Kim *et al.*, 2007) or the difference in absorption coefficient in blue bands (Bak *et al.*, 2016). However, since the water-leaving radiance of the red band is the result of absorption and fluorescence by Chlorophyll-*a*, it is possible that the detected result is phytoplankton other than red tide species. In addition, when the difference in the light absorption in the blue band is used, distortion due to dissolved organic matter may occur.

In the case of fresh water close to Case-1, only phytoplankton is a cause of discoloration of seawater. However, in case of turbid water near Case-2, it is more complicated than Case-1 because it causes discoloration of seawater in three kinds of phytoplankton, dissolved organic matter and suspended matter. Therefore, even if the same red tide organisms grow at the same density, they may show different discoloration patterns depending on the concentration of dissolved organic matter and suspended matters. Therefore, in order to detect red tide in the turbid waters near Case-2, modeling that can cope with various background colors is required. In this study, we propose a multi-layer perceptron based red tide detection algorithm to cope with such complexity.

2.3. Red Tide Detection Algorithm based on Multi-Layer Perceptron

The algorithm proposed in this study detects red tides through a two-step filtering process (Figure 4).

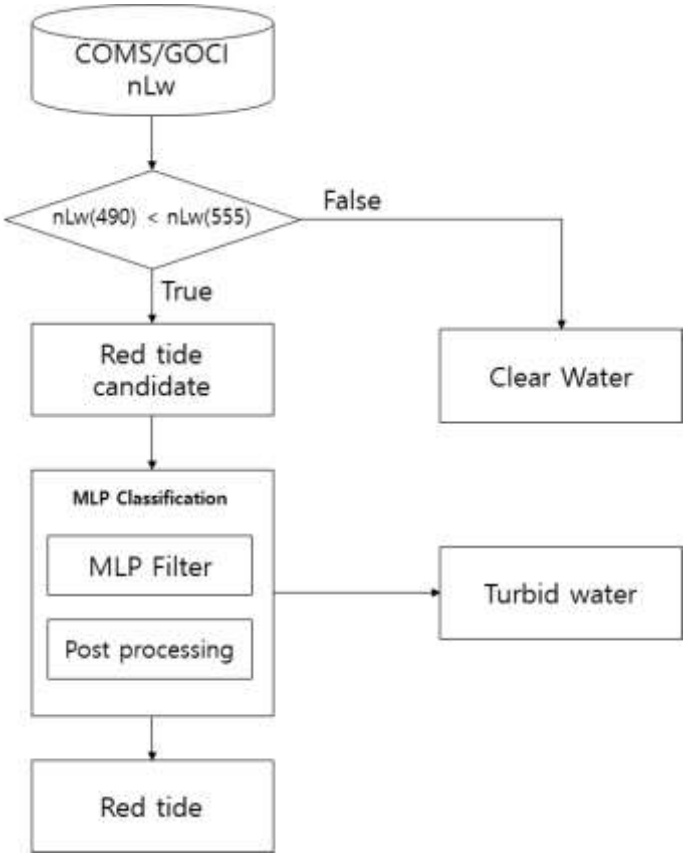

Figure 4. Detection Algorithm for *C. polykrikoides* Red Tide

The first step is to find the clear water pixels in the satellite image and use the slope between the blue and green bands. At this stage, a pixel whose water-leaving radiance (nLw) of the blue band (490 nm) is lower than the water-leaving radiance of the green band (555 nm) is classified as red tide candidate, and the remaining pixels are classified as clear water.

The second step is to separate the turbid water and red tide pixels using the multi-layer perceptron for the pixels classified as red tide candidates in the first step. In the first step, pixels classified as red tide candidates are re-classified into red tide, clear water, and turbid water using a previously learned multi-layer perceptron model, and finally red tide pixels are separated through post-processing. The multi-layer perceptron used in this study has 8

input layers and 3 output layers. The number of nodes in the input layer coincided with the eight channels of GOCI, and the output layer used three nodes to divide seawater into red tide, fresh water, and turbid water. Four hidden layers are placed between the input layer and the output layer, and each hidden layer is designed to have 100 nodes. The activation function uses ReLU (Rectifier Linear Unit) and Sigmoid function is used only in the layer from the last hidden layer to the output layer. This is to normalize the probabilities of belonging to each class from 0 to 1. The post-processing is a process for filtering the case where the multilayer perceptron model outputs similar probability values to two or more classes in classifying the class of the seawater. In this study, when a multilayer perceptron model outputs probability values of 0.5 or more in two or more classes, it is classified as pixels other than red tide.

3. Result and Discussion

In this study, the detection accuracy was evaluated using Confusion Matrix (Table 2). The confusion matrix is a tool that is often used as an evaluation index when evaluating the performance of a machine learning model.

Table 2. Confusion Matrix

		Prediction	
		True	False
Reference	True	TP	FN
	False	FP	TN

※ TP : True Positive / FP : False Positive
FN : False Negative / TN : True Negative

In this case, the total accuracy is defined as in formula (2).

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} \tag{2}$$

The performance of the proposed algorithm using the verification dataset was about 97% (Table 3).

Table 3. Confusion Matrix of Detection Algorithm for *C. polykrikoides*

		Detection		
		R	C	T
Reference	R	152	0	6
	C	0	155	3
	T	3	0	155

※ R: Red tide / T : Turbid water / C : Clear water
※ Total Accuracy≒0.97

It can be confirmed that there is a high accuracy improvement when compared with the detection accuracy (52.76%) of the most recently proposed seawater optical characteristics based red tide detection algorithm (Son *et al.*, 2012). Compared with the results obtained by using other machine learning techniques, the accuracy was higher than the decision tree model (86%) and the accuracy was similar to that of the logistic regression model (98%) (Bak *et al.*, 2018).

The decision tree model and the logistic regression model proposed in previous studies are methods that require additional variable selection. In decision trees, decision tree models have their own variable selection process, but they have limitations in that they do not exceed the predefined set of input variables. Therefore, the accuracy depends on the input variables defined by the researcher. In the case of the logistic regression model, it is possible to use features other than the predefined input variables through the combination of the variables. However, this method also has the maximum performance after the statistical variable selection method. On the other hand, in the case of the multilayer perceptron, the researchers showed high accuracy without performing a separate parameter selection process.

4. Conclusion

In this study, red tide pixels and normal seawater pixels were classified in satellite image using multi-layer perceptron. As a result of evaluating the performance of the detection algorithm, it is confirmed that the accuracy improvement is about 30 ~ 40% compared with the previous research based on the optical characteristics of seawater. Also, it is confirmed that the accuracy is high when compared with the result of previous study (86%, 98%) based on machine learning. The detection technique proposed in this study can detect red tides with high accuracy even in optically complex Case-2 seawater. Based on these results, it can be used for future red tide prevention.

Feature selection is essential for the development of remote sensing products. This process selects the band that is most effective to distinguish the interest class from the other classes among the bands of the satellite image or the aerial photograph. In the case of red tide detection, much of this process has relied on analyst intuition. Spectroscopic profile differences between red tides and general seawater have been interpreted by simply reading the spectroscopic profile graphs obtained from the field or laboratory. Thus, most of the features that could be used for red tide detection were based on the differences obtained visually through the spectral profile graphs.

In Case-1 seawater, it was possible to classify only the characteristics obtained through visual observation of the analyst, but it was impossible in Case-2 seawater. Case-2 The background color of seawater produced through the interaction of constituents of seawater is very diverse. Therefore, it was a very difficult task to select features corresponding to each of these background colors.

Most of the previous studies performed detection by binary rules using some pre-defined features of the analyst. This resulted in the analysts failing to deal with many exceptions that they did not understand. However, machine learning-based algorithms have the ability to cope with more exceptions by using various features not available through observation.

This study confirms that machine learning can be one of the most efficient ways to solve these problems. In particular, the multi-layer perceptron is superior to other machine learning techniques because it has superior feature selection ability and it is less dependent on the researcher 's background knowledge.

Acknowledgments

This research was a part of the project titled “Establishment and demonstration of red tide detection and prediction system for minimizing red tide damage” funded by the Ministry of Oceans and Fisheries, Korea (PM60650).

References

[1] Y. M. Kim, Y. G. Byun, Y. Huh and K. Y. Yu, “Detection of Cochlodinium polykrikoides Red Tide Using MODIS Level 2 Data in Coastal Waters”, KSCE Journal of Civil Engineering, vol. 27, no. 4D, **(2007)**, pp. 535-540.

[2] S. Y. Oh and H. J. Yoon, "Red-Tide Detection by using MODIS/AQUA Data", Journal of the Korea Institute of Electronics Communication Sciences, vol. 6, no. 2, **(2012)**, pp. 309-312.
[3] R. P. Stumpf, M. E. Culver, P. A. Tester, M. Tomlinson, G. J. Kirkpatrick, B. A. Pederson, E. Truby, V. Ransibrahmanakul and M. Soracco, "Monitoring Karenia brevis blooms in the Gulf of Mexico using satellite ocean color imagery and other data", Harmful Algae., vol. 2, **(2003)**, pp. 147-160.
[4] M. Tomlinson, R. P. Stumpf, V. Ransibrahmanakul, E. Truby, G. J. Kirkpatrick, B. A. Pederson, G. Vargo and C. Heil, "Evaluation of the use of SeaWiFS imagery for detecting Karenia brevis harmful algal blooms in the eastern Gulf of Mexico", Remote Sensing of Environment, vol. 91, **(2004)**, pp. 293-303.
[5] S. H. Bak, H. M. Kim, D. H. Hwang, H. J. Yoon and W. C. Seo, "Detection technique of Red Tide Using GOCI Level 2 Data", Korean Journal of Remote Sensing, vol. 32, no. 6, **(2016)**, pp. 673-679.
[6] Y. M. Kim, Y. G. Byun, Y. G. Kim and Y. D. Eo, "Detection of Cochlodinium polykrikoides red tide based on two-stage filtering using MODIS data", Desalination, vol. 249, **(2009)**, pp. 1171-1179.
[7] Y. B. Son, J. Ishizaka, J. C. Jeong, H. C. Kim and T. H. Lee, "Cochlodinium polykrikoides Red Tide Detection in the South Sea of Korea using Spectral Classification of MODIS Data", Ocean Sci. J., vol. 46, no. 4, **(2011)**, pp. 239-263.
[8] Y. B. Son, Y. H. Kang and J. H. Ryu, "Monitoring Red Tide in South Sea of Korea(SSK) Using the Geostationary Ocean Color Imager(GOCI)", Korean Journal of Remote Sensing, vol. 28, no. 5, **(2012)**, pp. 531-548.
[9] M. R. Park, K. H. Park and J. S. Ahn, "A Study on the Regional Classification Method using Dynamic Time Warping", The geographical Journal of Korea, vol. 45, no. 3, **(2011)**, pp. 387-395.
[10] Y. H. Ahn and P. Shanmugam, "Detecting the red tide algal blooms from satellite ocean color observations in optically complex Northeast-Asia Coastal waters", Remote Sensing of Environment, vol. 103, **(2006)**, pp. 419-437.
[11] Y. H. Yoon, "Sea rebellion", Red tide, Jipmoondang, Paju, South Korea (in Korean), **(2012)**.
[12] S. H. Bak, H. M. Kim, B. K. Kim, D. H. Hwang, U. Enkhjargal and H. J. Yoon, "Study on Detection Technique for Cochlodinium polykrikoides Red tide using Logistic Regression Model and Decision Tree Model", J. of the Korea Institute of Electronics Communication Sciences, vol. 13, no. 4, **(2018)**, pp. 777-786.

Authors

Su-Ho Bak, received the M.S. degree from Dept. of Spatial Information Engineering of Pukyong National University, Busan in 2017. He is doing a Ph.D. course in Dept. of Spatial Information Engineering of Pukyong National University.

Bum-Kyu Kim, received the B.S. degree from Dept. of Spatial Information Engineering of Pukyong National University, Busan in 2015. He is doing a graduate school student in Dept. of Spatial Information Engineering of Pukyong National University.

Do-Hyun Hwang, received the M.S. degree from Dept. of Spatial Information Engineering of Pukyong National University, Busan in 2013. She is doing a Ph.D. course in Dept. of Spatial Information Engineering of Pukyong National University.

Heung-Min Kim, received the M.S. degree from Dept. of Spatial Information Engineering of Pukyong National University, Busan in 2017. He is doing a Ph.D. course in Dept. of Spatial Information Engineering of Pukyong National University.

Unuzaya Enkgjargal, received the B.S. degree from Dept. of Wireless communication engineering of Mongolian university of science and technology, Ulaanbaatar in 2014. She is doing a graduate school student in Dept. of Spatial Information Engineering of Pukyong National University.

Seung-Yeol Oh, S.Y. Oh received his M.S. and Ph.D. degree from Pukyong National University, Busan in 2012, 2017. His special fields are Satellite Remote Sensing(Ocean Color Remote Sensing).

Hong-Joo Yoon, Dr. H.J. Yoon is professor and his academic degrees are Master in Ocean Dynamics (Pukyong National University, Korea), and Doctor in Geodynamics (Grenoble I University, France). His special fields are Satellite Remote Sensing (Satellite Oceanography and Meteorology) and Marine & Fisheries GIS. His recently studying the main subjects are change of climate, variations of SSH&SST, ocean circulation and application of Big data in ocean.