IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 24, 2019, accepted June 20, 2019, date of publication June 24, 2019, date of current version July 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924660

Single Image Based Algal Bloom Detection Using Water Body Extraction and Probabilistic Algae Indices

CHEOL WOO PARK[1], JONG JU JEON[1], YONG HO MOON[2], AND IL KYU EOM[1]
[1]Department of Electronics Engineering, Pusan National University, Pusan 46241, South Korea
[2]Department of Aerospace and Software Engineering/ReCAPT, Gyeongsang National University, Jinju 52828, South Korea

Corresponding author: Il Kyu Eom (ikeom@ pusan.ac.kr)

ABSTRACT Algal blooms are collections of algae that exist on the surface of the water. Because of their negative effects on aquatic organisms and humans, extensive studies have been performed to detect harmful algal blooms (HABs). However, most of the detection methods are based on remote-sensing imaging and have limitations with regard to resolution, time, and cost. In this paper, we present a new cyanobacterial algal bloom detection algorithm in inland water from a single image. The proposed method can be used as a first step in automatic early detection, warning, and rapid response systems that can be employed to mitigate the detrimental effects of HAB contamination in inland water bodies. We first divide an image into homogeneous regions via a density-based spatial clustering (DBSCAN) algorithm. From the segmented regions, we extract water bodies using wavelet leader-based texture analysis. The entropy and the number of zero wavelet coefficients are used as measures for the water body extraction. For images with a sky region, we introduce a simple sky-region removal method using the average brightness of segmented regions. We propose three probabilistic indices bases on an RGB-based vegetation index, a hue-based index, and a saturation-based index for estimating the degree of green algae in the extracted water body. The final index is obtained via multiplication of these three indices. In experiments on various types of images, our proposed algorithm achieves 94% accuracy for water body extraction. The proposed approach achieves better green algae estimation performance than the conventional vegetation index-based methods.

INDEX TERMS Algal bloom detection, DBSCAN algorithm, entropy, probabilistic algae index, water body extraction, wavelet leader.

I. INTRODUCTION

Algal blooms are collections of algae that grow in eutrophic lakes, slow-flowing rivers, or stagnant oceans, and are accumulated on the surface of water. Algal blooms consume a large amount of oxygen, reducing the amount of dissolved oxygen in water, and can be a major threat to aquatic life. When harmful algal blooms (HABs) occur, the cost of removing unfamiliar tastes and odors in the production of tap water increases, and HABs hinder water activities such as swimming, fishing, and water skiing. In addition, overgrown algal blooms have negative effects on aquatic ecosystems, causing the death of many aquatic animals and plants. To counteract the negative effects of HABs, an automatic HAB monitoring and detection system is needed.

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Asikuzzaman.

Studies for detecting algal blooms commonly utilize remotely sensed images from satellites or aircraft. These approaches are based on the reflectance of chlorophyll-*a*, which is found in all phytoplankton and green algae [1]. Several methods exploiting this spectral property have been proposed for detecting algal blooms using remotely sensed data. The classical normalized difference vegetation index [2] has been used to identify algal blooms in images. In addition, many algorithms based on other indices, such as the enhanced vegetation index [3], the maximum chlorophyll index [4], the floating algae index [5], the index of floating green algae for geostationary ocean color imager [6], and the ocean surface algal blooms index [7], have been proposed.

However, the index-based methods use data from the Landsat TM/ETM+, a moderate resolution imaging spectroradiometer, and a medium resolution imaging spectrometer. These data have a spatial resolution ranging from tens of

meters to over a thousand meter. Therefore, satellite image-based algorithms have difficulty in accurately measuring the proportion of green algae in a single pixel. This problem was denoted as the 'subpixel problem' in [8]. For estimating the floating algae area on a subpixel scale, an algae pixel-growing algorithm [8] using the Rayleigh-corrected reflectance and the floating algae index was proposed. Recently, a spectral-unmixing based green algae area estimation method [9] for the subpixel level was been reported. The temporal resolution of satellite image data is more than one day. Therefore, it is difficult to develop an early warning system for the sudden occurrence of HABs.

In recent years, unmanned aerial vehicles (UAVs) have emerged as useful tools for algal bloom detection [10]–[13]. They can be more efficient than satellite-based imaging systems with regard to time and cost, and can reduce *in situ* sampling costs. In addition, the performance of UAV-based algal bloom detection can be further improved by combining thermal imaging and spectral information from remotely sensed data [14]–[16]. Although the spatial range and spectral capacity of the UAV-based green algae detection method are not comparable to those of the satellite-based method, the former method has the advantage of higher temporal revisit time and spatial resolutions and can perform specific tasks. Current applications of algal bloom detection research based on UAVs are presented in detail in [17].

Other image sources for algal bloom detection are smartphones and online media. With the increasing popularity of online media, there has been a growing trend toward adopting citizen science-based approaches for environmental monitoring. In addition to environment monitoring experts, citizen scientists, such as local residents, fishermen, tourists, and civil environmental monitoring agents, can produce, collect, and transmit image data for algal bloom detection. Thus, citizen scientist-based approaches are intended to complement the traditional *in situ* and remote sensing-based approaches for environmental monitoring [18].

Image data acquired by smartphones and online media have a wide variety of forms. They have different resolutions, scales, and viewpoints, and may be contaminated with noise. In addition, only RGB image data can be used. Fig. 1 shows various image data that were obtained from online sources such as *Google Images*, or recorded by UAVs. As shown in Fig. 1, the images include forests, the sky, and water bodies, as well as artificial structures such as roads, bridges, and buildings.

All natural and artificial objects other than water bodies in image data are obstacles to the automatic detection of algal blooms. Therefore, image processing or computer vision techniques for extracting water bodies are essential in the field of algal bloom detection. Recently, of algal bloom detection algorithms based on various image processing and computer vision techniques, such as clustering [18], [19], the local binary pattern (LBP) [20], [21], and machine learning [18], [21], [22], have been proposed. These methods are promising for HAB detection, as various image processing techniques can be applied and the image data can be acquired easily and quickly.

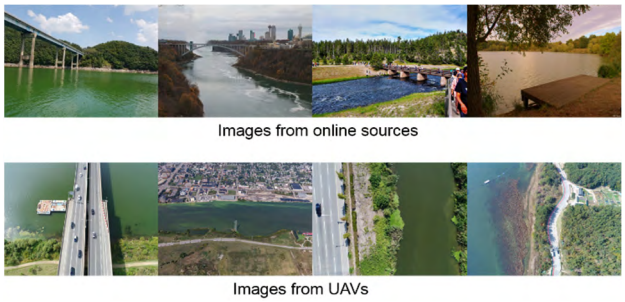

FIGURE 1. Various image data obtained from online sources (top), and recorded by UAVs (bottom).

The aim of this paper is to introduce a new algal bloom detection scheme based on an easily obtainable single image. The test image can be downloaded from Internet, captured by UAVs, or taken from smartphones. We try to develop an image processing-based algorithm that can detect and locate algae in water bodies, irrespective of the variations in camera parameters (resolution, viewpoint and scale) and various environmental conditions. Our algorithm can be used as input to an early algal bloom warning system at a very low cost. From this warning, various post-works, such as, measuring, detecting, and processing for algal blooms, can be performed.

Herein, we propose a single image-based cyanobacterial algal bloom detection scheme in inland water. To detect algal blooms, it is important to identify water bodies in the image. For this purpose, it is helpful to divide the image into similar regions. We use the superpixel and density-based clustering algorithms to divide an image into similar regions. The important task for detecting green algae is to determine which part of the classified image region is the water body. We present a wavelet domain texture analysis to extract water bodies in the divided image regions.

For the extracted water bodies, we define three values for estimating the amount of algal bloom s in a pixel. The first quantified value is obtained from RGB-based vegetation indices, such as the normalized green-red difference index (NGRDI), normalized green-blue difference index (NGBDI), and green leaf index (GLI) [16]. The second value indicates how much green is in the pixels of the extracted water body. This value is defined by applying the error function to the hue value of the pixel. Because green algae reduces the transparency of water bodies, the saturation of the pixel is defined as the third quantified value, representing this phenomenon. In this paper, we use these three quantified values to estimate the degree of algal bloom in the range from 0 to 1 for the extracted water bodies. Through experiments, we demonstrate that our proposed algorithm can effectively detect HABs in a single image.

The remainder of this paper is organized as follows. Section II describes the basic framework for the image-based algal bloom detection scheme. The proposed water body extraction method is presented in Section III. In Section IV,

the green algae estimation algorithm based on RGB images is presented. Section V presents the experimental results obtained using the proposed approach, and the conclusions are drawn in Section VI.

II. BASIC FRAMEWORK FOR IMAGE-BASED ALAGAL BLOOM DETECTION

We investigate image-based approaches from the viewpoint of the basic framework for algal bloom detection. In [18], an agglomerative clustering algorithm, that groups superpixels to maximize the likelihood of a heterogeneous or cluttered surface category, was presented. This method combines multiple cues, such as color, textural, and contextual, while minimizing the pixel-level entropy function via grouping of superpixels. For classified image segments, the algorithm identifies five categories (lake clear, tree, grass, sky, and lake HAB) using Fisher vector pooling of a convolutional neural network filter bank [23]. Because this method labels each image segment, the water body is extracted during the segmentation. To evaluate the HAB detection performance, this method extracts the pixels that correspond to lake regions, and deploys an instance-level binary support vector machine classifier trained to further classify the lake regions as HAB vs. clear lake surfaces. However, the algorithm can only determine the presence or absence of the HABs in the lake.

A hybrid image segmentation scheme that uses visual features from a camera and inertial features measured by camera inertial sensors was reported for aquatic environment monitoring [19]. In this method, several lightweight and robust computer vision algorithms are used to detect harmful aquatic processes in a dynamic environment. The segmented patches are passed to patch identification for probability thresholding, noise removal, and patch recognition.

The LBP was used as a texture feature to segment objects in images [20], [21]. According to pre-defined object labels (algal, grass, water, and ground), an unsupervised texture segmentation method with the LBP was utilized. Each segmented class was identified by the Chi-square distance. However, the performance of the proposed system, was not evaluated, and the effects of comparable objects, such as trees, plants, and seaweeds, in the image's background were not examined. A fully machine learning-based HAB detection approach was proposed in [22]. This method indicates the percentage of algae in the boxed area of an image. However, image labeling using an annotation software is required before training.

Many image-based algal bloom detection approaches involve labeling image segments [18], [20]–[22]. However, for detecting green algae, it is sufficient to extract the water bodies from the image. If only water bodies can be extracted properly, the labeling process is not necessary. According to the aforementioned studies, the basic framework for image-based algal bloom detection is composed of three phases, as shown in Fig. 2: clustering (phase I), water body extraction (phase II), and algal bloom detection (phase III). Depending on the algal bloom detection method, phase I and II can be combined into one phase. Additionally, the three phases can be combined into one by using machine learning.

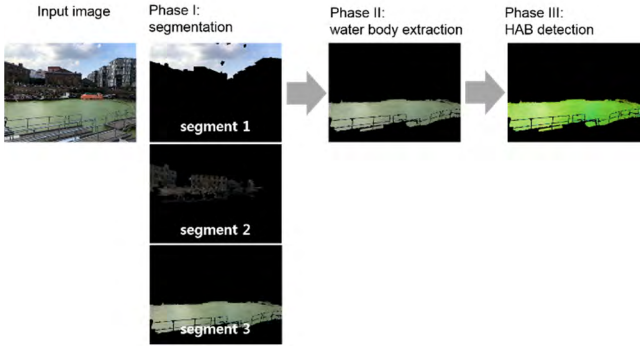

FIGURE 2. Basic framework for image-based algal bloom detection.

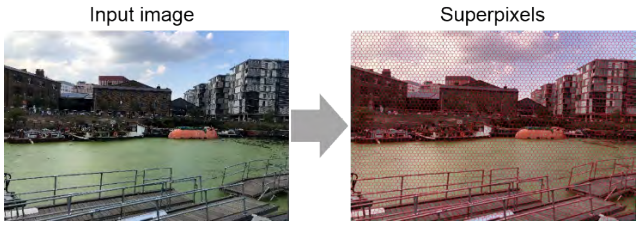

FIGURE 3. Example of an over-segmented image using superpixels.

In this paper, we present an efficient cyanobacterial algal bloom detection algorithm based on RGB images obtained from smartphones, the Internet, and UAVs. Our method involves three phases (segmentation, water body extraction, and algal bloom estimation), and dose not employ a machine learning scheme.

III. PROPOSED WATER BODY EXTRACTION METHOD

A. IMAGE SEGMENTATION

A test image for algal bloom detection can have various contents, such as a water body, a forest, the sky, a bridge, or a road. Therefore, it is important to classify the image into homogeneous regions to extract water bodies from the test image. A superpixel is a group of connected pixels with similar colors or gray levels. Therefore, the superpixel segmentation algorithm can be usefully applied before region segmentation. In this paper, we use the simple linear iterative clustering (SLIC) method [24] to over-segment the input image. Fig. 3 shows an example of the over-segmented image obtained using the SLIC algorithm. In this case, the number of superpixels is 3,000.

The purpose of the segmentation step (phase I) in the proposed method is to group the superpixels into several regions with similar properties. For this, we exploit the density-based spatial clustering of applications with noise (DBSCAN) algorithm [25]. DBSCAN is a density-based clustering algorithm that finds a number of clusters starting from the estimated density distribution of corresponding nodes. The advantage of DBSCAN is that it does not need to determine the number of clusters. In addition, the DBSCAN algorithm classifies noise

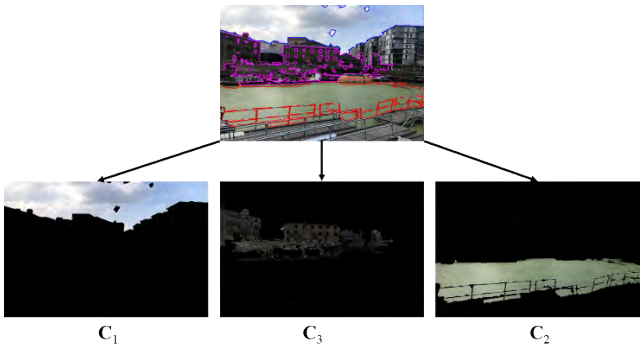

FIGURE 4. Example of three segments with artificial structures.

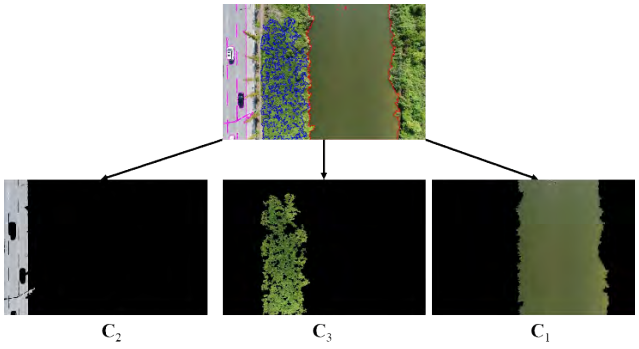

FIGURE 5. Example of three segments with trees and grasses.

pixels separately. These characteristics are considered to be suitable for our desired image clustering.

Let us assume that the given image $\mathbf{I}$ is divided into Nnon-overlapped segments using the DBSCAN algorithm as follows.

$$\mathbf{I} = \mathbf{C}_1 \cup \mathbf{C}_2 \cup \cdots \cup \mathbf{C}_N, \tag{1}$$

where $\mathbf{C}_i$ (i =1, 2, …, N) is the i-th segment of the image, and $\mathbf{C}_i \cap \mathbf{C}_j = \phi$ for $i \neq j$. In (1), the segments are ordered by the number of pixels that they contain as follows.

$$|\mathbf{C}_1| \geq |\mathbf{C}_2| \geq \cdots \geq |\mathbf{C}_N|, \tag{2}$$

where $|\mathbf{C}_i|$ is the number of pixels in segment $\mathbf{C}_i$. Because a segment that may contain a water body will occupy a certain area in the image, we remove clusters that do not satisfy the condition of $|\mathbf{C}_i|/|\mathbf{I}| > \delta$, where δ is a positive small value. Through this process, small segments with a very low probability of containing a water body are eliminated. Finally, we can obtain $K(K \leq N)$ segments to extract water bodies.

Fig. 4 shows an example of three segments ($K = 3$) obtained for a given image. As shown in Fig. 4, the buildings on the right side, small ships in the middle, and floors at the front of the image are removed, because they are all composed of small homogeneous regions. Fig. 5 shows another example with trees and grasses. Trees and grasses can have similar colors to water bodies with green algae. Therefore, they can be an obstacle in the extraction of water bodies having green algae. As shown in Fig. 5, the trees and grasses on the right side of the image were removed. Only homogeneous grasses were present in $\mathbf{C}_2$, and the small cars on the left side of the image were eliminated. This elimination process facilitates the extraction of water bodies.

B. WATER BODY EXTRACTION

We assume that we can only estimate the location where the image is captured and that no other information is available beyond the RGB values of the image. The test image does not have spectral information because it has RGB form. The image includes forests, the sky, and water bodies, as well as artificial structures. To differentiate the tree, grass, sky, and water body, machine learning-based methods [18], [22] and LBP-based texture analysis [20], [21] are used. Since the waters are homogeneous and have similar patterns, texture-based analysis will be suitable for water body extraction. In addition, in order to detect algal blooms, we only need to distinguish the water body and other areas.

In particular, wavelet-based representations have been proposed by many researchers for texture analysis and classification [26]–[28]. These methods have the advantage of using information in both the frequency and spatial domains. In this paper, we introduce a simple wavelet domain texture analysis algorithm for water body extraction. We exploit the wavelet leader, and propose a water body extraction measure using the ratio of zero coefficients and entropy.

For a given image $\mathbf{I}$, the discrete wavelet transform (DWT) decomposes $\mathbf{I}$ into four subbands as follows.

$$\{W_o(\mathbf{I};\mathbf{x})\} = \mathrm{DWT}(\mathbf{I};\mathbf{x}), \quad o \in \{A, H, V, D\}, \tag{3}$$

where $\{W_o(\mathbf{I};\mathbf{x})\}$ is the set of four wavelet subbands at spatial location $\mathbf{x}$, DWT($\mathbf{Z}$) is the DWT on $\mathbf{Z}$, and o indicates the direction of the wavelet subband (A: low-frequency subband, H: horizontal direction, V: vertical direction, D: diagonal direction). For an image segment $\mathbf{C}_i$, we can obtain the corresponding four wavelet subbands $W_A(\mathbf{C}_i;\mathbf{y})$, $W_H(\mathbf{C}_i;\mathbf{y})$, $W_V(\mathbf{C}_i;\mathbf{y})$, and $W_D(\mathbf{C}_i;\mathbf{y})$ from (3), where $\mathbf{y}$ is the coordinate indicating the spatial location of segment $\mathbf{C}_i$.

In this paper, we use the wavelet leader [29], which is proposed for multi-fractal analysis of images. The wavelet leader can improve the robustness of certain statistical measurements of conventional wavelet coefficients [28]. The wavelet leader for $\mathbf{C}_i$, $W_L(\mathbf{C}_i)$ is defined as

$$W_L(\mathbf{C}_i;\mathbf{y}) = \max(|W_H(\mathbf{C}_i;\mathbf{y})|, |W_V(\mathbf{C}_i;\mathbf{y})|, |W_D(\mathbf{C}_i;\mathbf{y})|). \tag{4}$$

It is mathematically justified that wavelet leaders allow accurate measurement of the multi-fractal properties of two-dimensional measuring fields [28]. However, the conversion of wavelet coefficients into wavelet leaders does not remove a large amount of information from texture images, because this conversion is based on the maximum operation. Therefore, we remove small wavelet leaders using a simple thresholding operation as follows.

$$W_{Lf}(\mathbf{C}_i;\mathbf{y}) = \begin{cases} W_L(\mathbf{C}_i;\mathbf{y}), & W_L(\mathbf{C}_i;\mathbf{y}) > t \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

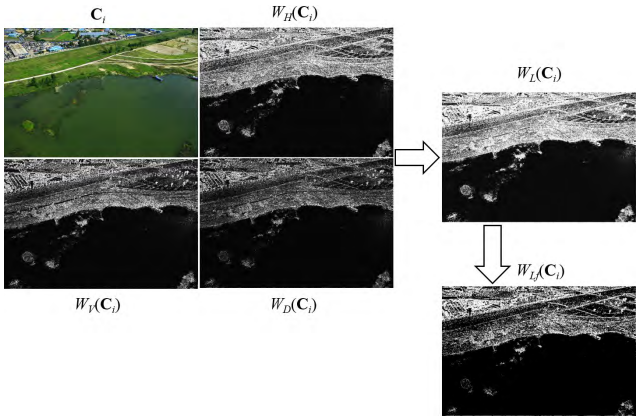

FIGURE 6. Example of wavelet leaders and their thresholding version.

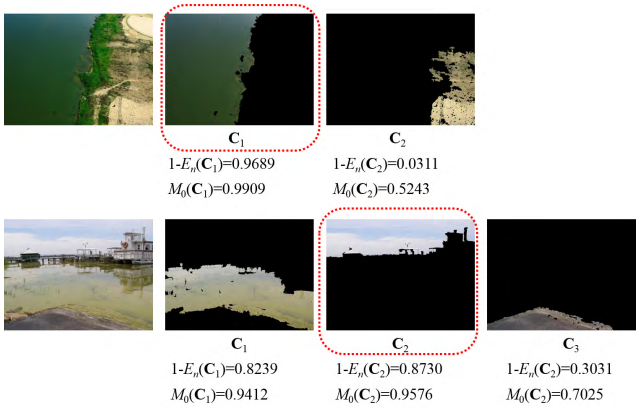

FIGURE 7. Example of water body extraction. Top: Extraction for a UAV-based image, Bottom: Extraction result for an online-based image.

where $W_{Lf}(\mathbf{C}_i)$ represents the wavelet leaders after the thresholding operation, and t is the threshold. In this paper, for t, we use the standard deviation of $W_L(\mathbf{C}_i)$. Fig. 6 depicts the high-frequency wavelet coefficients and wavelet leaders of $\mathbf{C}_i$. $W_{Lf}(\mathbf{C}_i)$ is illustrated in the bottom right of Fig. 6. As shown in Fig. 6, a large amount of small wavelet coefficients are removed in the thresholding operation for wavelet leaders.

It is assumed that water body regions are homogenous compared with other image segments, such as grasses and buildings, and can have uniform or regular texture patterns, such as small waves as shown in Fig. 4 and Fig. 5. Therefore, we assume that water bodies have a large amount of zero wavelet coefficients. In this paper, we introduce the first measure $M_0(\mathbf{C}_i)$ for extracting water bodies as the ratio of zero coefficients in a given segment as follows.

$$M_0(\mathbf{C}_i) = \frac{Z\left(W_{Lf}(\mathbf{C}_i)\right)}{\left|W_{Lf}(\mathbf{C}_i)\right|}, \tag{6}$$

where $Z(W_{Lf}(\mathbf{C}_i))$ is the number of zero coefficients in $W_{Lf}(\mathbf{C}_i)$, and $|W_{Lf}(\mathbf{C}_i)|$ is the number of wavelet coefficients in the segment $W_{Lf}(\mathbf{C}_i)$. We expect that water bodies have a large amount of zero coefficients, as shown in Fig. 6.

The distribution of wavelet coefficients in the segment $W_{Lf}(\mathbf{C}_i)$ can be a clue for extracting water bodies. The probability distribution of a water body region will have a high peak at zero, and its entropy will be relatively low compared with other regions. The entropy of segment $\mathbf{C}_i$, $E(\mathbf{C}_i)$ is calculated as follows.

$$E(\mathbf{C}_i) = -\sum_{\mathbf{y}} p\left[W_{Lf}(\mathbf{C}_i;\mathbf{y})\right] \ln p\left[W_{Lf}(\mathbf{C}_i;\mathbf{y})\right], \tag{7}$$

where $p[W_{Lf}(\mathbf{C}_i;\mathbf{y})]$ is the probability of the wavelet coefficients of $W_{Lf}(\mathbf{C}_i;\mathbf{y})$ at position $\mathbf{y}$. In this paper, we develop a water body extraction measure by combining $M_0(\mathbf{C}_i)$ and $E(\mathbf{C}_i)$ as follows.

$$\mathbf{C}_W = \arg\max_{\mathbf{C}_i} \left(M_0(\mathbf{C}_i)\left(1 - E_n(\mathbf{C}_i)\right)\right), \tag{8}$$

where $\mathbf{C}_W$ is the extracted water body, and $E_n(\mathbf{C}_i)$ is the normalized entropy such that

$$E_n(\mathbf{C}_i) = \frac{E(\mathbf{C}_i)}{\sum_{i=1}^{K} E(\mathbf{C}_i)}. \tag{9}$$

We assume that the test image has at least one water body region, because the image is captured for algal bloom detection. However, the measurement shown in (8) can extract only one water body, because the maximum operation is used to extract a water body. To extract two or more water bodies in the test image, the segment having a value more than 95% of the maximum $M_0(\mathbf{C}_i)(1\text{-}E_n(\mathbf{C}_i))$ value is regarded as a water body.

Fig. 7 illustrates the water bodies extracted using (8). As shown in Fig. 7(a), because both $1\text{-}E_n(\mathbf{C}_1)$ and $M_0(\mathbf{C}_1)$ for segment $\mathbf{C}_1$ have the maximum value, $\mathbf{C}_1$ is extracted as the water body. On the other hand, the $\mathbf{C}_2$ segment having a sky region has the maximum $1\text{-}E_n(\mathbf{C}_2)$ and $M_0(\mathbf{C}_2)$ as shown in Fig. 7(b). This situation can occur frequently in the process of extracting the water body from online-based images. In this paper, we introduce a simple sky region block algorithm using the *value* component of the hue, saturation, and value (HSV) color space.

C. SKY REGION REMOVAL

As shown in Fig. 7, the sky region has similar colors and a homogeneous pattern, akin to a water body. This can make it difficult to distinguish between a water body and the sky region. We present a sky region block method to be applied before water body extraction. Because we have already segmented regions, it is possible to block the sky region in a simple manner. Generally, the sky region has a high brightness value. Therefore, we introduce a simple sky region block method using the average brightness (V) of the HSV color space. Let $\mu(V(\mathbf{C}_i))$ be the average brightness value for each segment $\mathbf{C}_i$. $V(\mathbf{C}_i;\mathbf{y})$ is given as

$$V\left(\mathbf{C}_i;\mathbf{y}\right) = \max\left(R(\mathbf{C}_W;\mathbf{y}), G(\mathbf{C}_W;\mathbf{y}), B(\mathbf{C}_W;\mathbf{y})\right), \tag{10}$$

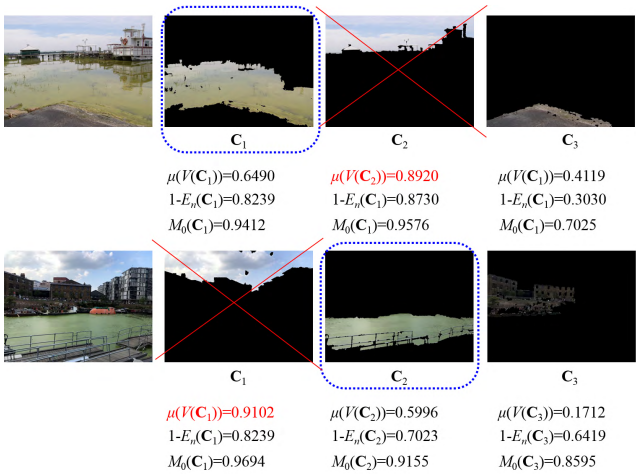

FIGURE 8. Examples of water body detection after using the sky region removal method.

where $R(\mathbf{C}_i;\mathbf{y})$, $G(\mathbf{C}_i;\mathbf{y})$ and $B(\mathbf{C}_i;\mathbf{y})$ are the red, green, and blue pixel values, respectively, at a location $\mathbf{y}$ in $\mathbf{C}_i$. If $\mu(V(\mathbf{C}_i))>\gamma$,we can remove the segment $\mathbf{C}_i$ before the water body extraction process. Fig. 8 shows sky region block examples. We observe that the sky segments that were erroneously detected as the water body were properly removed.

IV. ALGAL BLOOM ESTIMATION

We introduce a quantified value between 0 and 1 for estimating the degree of HAB in a pixel. This value is given by the multiplication of the modified vegetation index, hue-based index, and saturation-based index in the extracted water body region.

A. RGB-BASED VEGETATION INDEX

From remotely sensed data, several vegetation index-based algal bloom estimation approaches using the spectral property of algal blooms have been presented. However, these indices are not directly applicable to RGB images. In recent years, various RGB-based vegetation indices have been introduced. In [16], four types of indices including the NGRDI, NGBDI, GL) and excess green index (ExG) were used to identify green algae. In 2018, the relationship between commonly used vegetation indices extracted from UAV-based RGB and multispectral images was investigated to estimate the number of oilseed rape flowers [30]. These studies demonstrated the capabilities of various RGB-based vegetation indices for different applications. Table 1 presents examples of RGB-based vegetation indices along with their formulas.

The green algae detection results for the extracted water bodies obtained using the RGB-based vegetation indices are shown in Fig. 9. When there is green algae in the image, it is detected very strongly, even when there is a small amount of green algae, the algae is detected. As shown in Fig. 9, for most of the test images, the RGB-based vegetation indices do not have the discriminating power for detecting the green algae. Therefore, these indices are insufficient for detecting green algae. For this reason, we introduce three indices, including an RGB-based vegetation index, for algal bloom detection.

TABLE 1. Various RGB-based vegetation indices and their formulas.

Vegetation Index	Formula	Reference
Normalized Green-Red Difference Index (NGRDI)	(G-R)/(G+R)	[31]
Normalized Green-Blue Difference Index (NGBDI)	(G-B)/(G+B)	[31]
Green Leaf Index (GLI)	(2G-R-B)/(2G+R+B)	[32]
Excess Green Index (ExG)	2G-R-B	[33]

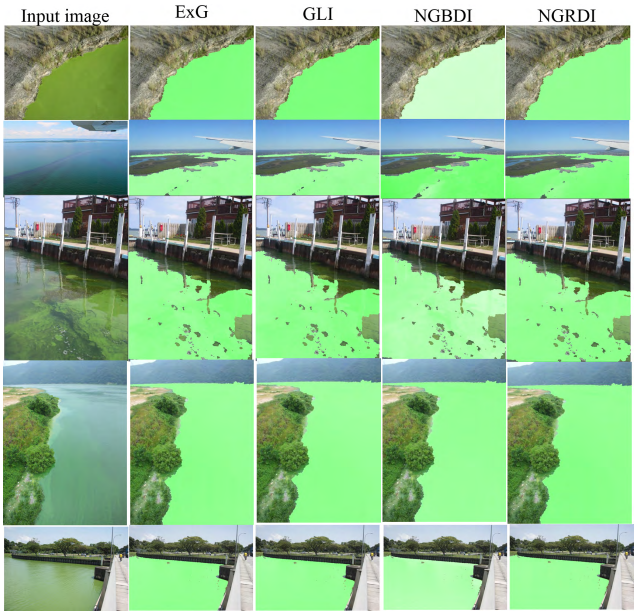

FIGURE 9. Green algae detection results for extracted water body using various RGB-based vegetation indices.

In this paper, we first adjust the NGRDI between 0 and 1 to define the first index for green algae detection. For an extracted water body $\mathbf{C}_W$, the first index based on the NGRDI, $P_V[\mathbf{C}_i;\mathbf{y}]$, is defined as

$$P_V[\mathbf{C}_W;\mathbf{y}] = \frac{G(\mathbf{C}_W;\mathbf{y})}{G(\mathbf{C}_W;\mathbf{y}) + R(\mathbf{C}_W;\mathbf{y})}, \tag{11}$$

where $G(\mathbf{C}_W;\mathbf{y})$ and $R(\mathbf{C}_W;\mathbf{y})$ are the green and red pixel values, respectively, at a location $\mathbf{y}$ in $\mathbf{C}_W$.

B. HUE INDEX

Green algae have a green color. It is useful to exploit the hue value in the HSV color space to define the hue-based index for algal bloom detection. Hue is the attribute of color and is discernible as red, green, blue, and so on. It is calculated

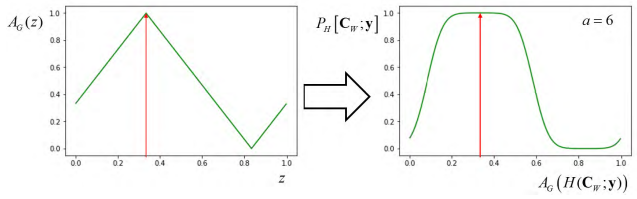

FIGURE 10. Affinity function for the green color using the hue values and the detection function based on the error function.

using the RGB value as follows.

$$H(\mathbf{C}_i;\mathbf{y}) = \frac{1}{360} \times \begin{cases} 0, & V(\mathbf{y}) = m(\mathbf{y}) \\ 60^\circ \times \left(\dfrac{G(\mathbf{y}) - B(\mathbf{y})}{V(\mathbf{y}) - m(\mathbf{y})} \bmod 6 \right), & V(\mathbf{y}) = R(\mathbf{y}) \\ 60^\circ \times \left(\dfrac{B(\mathbf{y}) - R(\mathbf{y})}{V(\mathbf{y}) - m(\mathbf{y})} + 2 \right), & V(\mathbf{y}) = G(\mathbf{y}) \\ 60^\circ \times \left(\dfrac{R(\mathbf{y}) - G(\mathbf{y})}{V(\mathbf{y}) - m(\mathbf{y})} + 4 \right), & V(\mathbf{y}) = B(\mathbf{y}), \end{cases} \tag{12}$$

where $m(\mathbf{C}_i;\mathbf{y})$=min($R(\mathbf{C}_i;\mathbf{y})$,$G(\mathbf{C}_i;\mathbf{y})$,$B(\mathbf{C}_i;\mathbf{y})$), and $H(\mathbf{C}_i;\mathbf{y})$ is the hue value at the location $\mathbf{y}$ in the segmented region $\mathbf{C}_i$. In (12), $\mathbf{C}_i$ is omitted for simplicity. The hue is represented by an angle ranging from 0° to 359°. For example, green is represented by 120°, and blue is represented by 240°. The hue value starts from 0° (red) and the color changes every 120°, first to green and then to blue. In this paper, we normalize the hue value between 0 and 1.

We define the affinity function for the green color, and modify it using the error function. The green affinity function $A_G(z)$ is defined as

$$A_G(z) = \begin{cases} 2z + \dfrac{1}{3}, & 0 \le z < \dfrac{1}{3} \\ -2z + \dfrac{5}{3}, & \dfrac{1}{3} \le z < \dfrac{5}{6} \\ 2z - \dfrac{5}{3}, & \dfrac{5}{6} \le z \le 1. \end{cases} \tag{13}$$

For the hue value of the extracted water region $\mathbf{C}_W$, the second metric for algal bloom detection, $P_H[\mathbf{C}_i;\mathbf{y}]$ is give as

$$P_H[\mathbf{C}_W;\mathbf{y}] = \frac{\text{erf}\{a\,(A_G\,(H(\mathbf{C}_W;\mathbf{y})) - 0.5)\} + 1}{2}. \tag{14}$$

$P_H[\mathbf{C}_i;\mathbf{y}]$ is used on a pixel-by-pixel basis. Fig. 10 presents the affinity function defined for the green color and the probability based on the error function.

C. SATURATION INDEX

While the hue refers to the color in an image, the saturation describes the intensity or purity of the hue. Therefore, the saturation index should be used together with the hue-based index for algal bloom detection. We introduce the saturation index for green algae detection as the ratio of the saturation values of $\mathbf{C}_W$ to their maximum value. Let $S(\mathbf{C}_W;\mathbf{y})$ be the saturation value. It is defined as

$$S(\mathbf{C}_W;\mathbf{y}) = 1 - \frac{m(\mathbf{C}_W;\mathbf{y})}{V(\mathbf{C}_W;\mathbf{y})}. \tag{15}$$

Using this value, we can obtain the probabilistic index based on the saturation, $P_S[\mathbf{C}_W;\mathbf{y}]$ as follows.

$$P_S[\mathbf{C}_W;\mathbf{y}] = \frac{S(\mathbf{C}_W;\mathbf{y})}{S_{\max}(\mathbf{C}_W)}, \tag{16}$$

where $S_{\max}(\mathbf{C}_W)$ is the maximum value of $S(\mathbf{C}_W;\mathbf{y})$.

TABLE 2. Overall algorithm of the proposed method.

Input: Input image $\mathbf{I}$
Output: Probabilistic map, $P_G[\mathbf{C}_W;\mathbf{y}]$ for the extracted water body.
1) Superpixel segmentation for $\mathbf{I}$.
2) Obtain N non-overlapped segments $\mathbf{C}_i$ using the DBSCAN algorithm.
3) Select K ($K \le N$) segments using $|\mathbf{C}_i|/|\mathbf{I}| > \delta$.
4) Apply the sky region removal method using $\mu(V(\mathbf{C}_i)) > \gamma$.
5) Perform the discrete wavelet transform for each $\mathbf{C}_i$ using (3).
6) Obtain the thresholded wavelet leader $W_{Lf}(\mathbf{C}_i)$ using (5).
7) Select $\mathbf{C}_i$ with the maximum value of $M_o(\mathbf{C}_i)(1 - E_n(\mathbf{C}_i))$ as the water body using (8).
8) Calculate the three indices based on (11), (14), and (16).
9) Obtain the final green algae estimation index using (17).

D. PROPOSED INDEX

All three measures presented in this paper have values between 0 and 1. We can express the occurrence of green algae in the extracted water body as a probabilistic value for each pixel. That is,

$$P_G[\mathbf{C}_W;\mathbf{y}] = P_V[\mathbf{C}_W;\mathbf{y}]\,P_H[\mathbf{C}_W;\mathbf{y}]\,P_S[\mathbf{C}_W;\mathbf{y}], \tag{17}$$

where $P_G[\mathbf{C}_W;\mathbf{y}]$ is the final green algae estimation index at the location $\mathbf{y}$ in the extracted water body region $\mathbf{C}_W$. The overall algorithm for detecting green algae in an image is presented in Table 2.

V. SIMULATION RESULTS

A. EXPERIMENTAL SETUP

To verity the effectiveness of the proposed algal bloom detection method, we test it on various types of images. The test images are composed of 161 images captured by UAVs, 135 aerial images downloaded from K-water [34], and 170 images from online sources (*Google Images*). We randomly collected images with water bodies. These images may or may not have algal blooms. The online images may have included sky regions. We have four parameters to perform our algorithm. The parameters are determined as shown in Table 3.

B. WATER BODY EXTRACTION RESULTS

Fig. 11 presents examples of the water body extraction results for various test images. The extraction results for UAV-based images without a sky region are shown in Fig. 11(a). The water bodies are well extracted using our method.

TABLE 3. Parameter values used in our simulations.

Parameter	Usage	Value
δ	To remove small segmented regions	0.1
γ	To block sky region	0.7
t	To remove small wavelet leaders	std($W_L(\mathbf{C}_i)$)
a	Shape parameter used in (14)	6

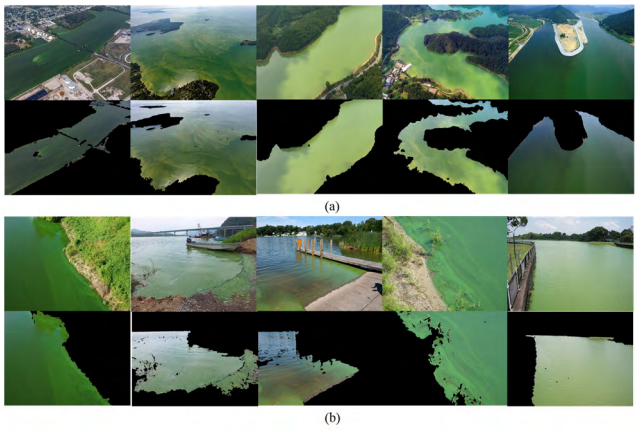

FIGURE 11. Water body extraction results. (a) UAV-based images without a sky region, (b) online images with or without a sky region.

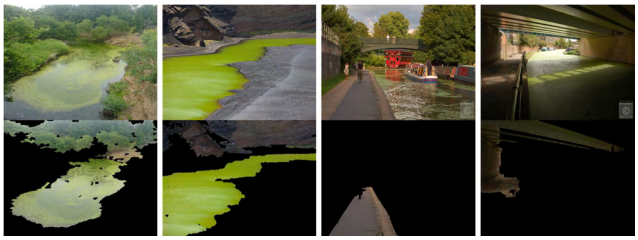

FIGURE 12. Failure cases for water body extraction.

Fig. 11(b) presents the extraction results for the online images, which may have contained sky regions. Our results exhibit good extraction performance regardless of the presence of sky regions as shown in Fig. 11(b).

Fig. 12 depicts some failure cases. When the water body and other regions have a similar pattern, the water body extraction may fail (first and second columns in Fig. 12). Additionally, the water body extraction may fail if the pattern of the water body is distorted by sunlight or shadows (third and fourth columns in Fig. 12).

Table 4 shows the water body extraction performance after application of the sky region block filter for all the test images. In Table 4, the accuracy is defined as the percentage of the image from which the water body was extracted correctly for the entire test image. The accuracy is measured by manually observing the water bodies extracted from the image. As shown in Table 4, erroneous detection results are obtained for 28 of the total 466 images. The proposed water body extraction method has an accuracy of 94.0%. The extraction accuracy is highest for the UAV-based images, because they do not contain sky regions. The extraction

TABLE 4. Water body extraction performance for various test images.

	Image sources			
	UAV	Aircraft	Online	Total
Number of image (A)	161	135	170	466
Correct extraction (B)	158	122	158	438
Incorrect extraction (C)	3	13	12	28
Accuracy (B/A × 100 %)	98.1	90.4	92.9	94.0

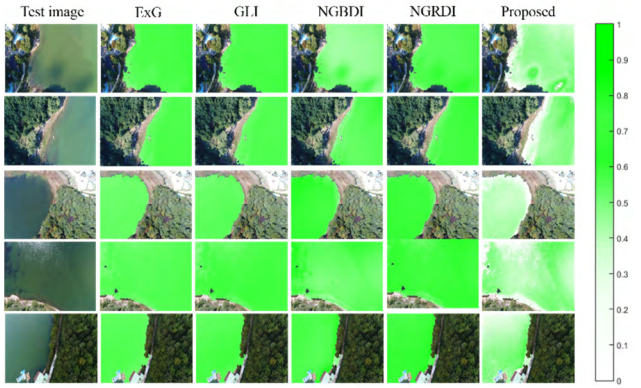

FIGURE 13. Algal bloom detection results for UAV-based test images with green algae.

accuracy is lowest for the aerial images without sky regions, because the resolution of these images are low owing to the aerial photographing. The accuracy of the water body extraction for the online images is 92.9%. Our algorithm achieves high water body extraction performance through the simple method of removing the sky region.

C. ALGAL BLOOM ESTIMATION RESULTS

We perform algal bloom estimation experiments for the following three cases: 1) water bodies with green algae, 2) clean water bodies without green algae, and 3) erroneously extracted water bodies.

Fig. 13 shows the algal bloom estimation results for UAV-based images with green algae. Here, all the vegetation indices are normalized between 0 and 1. The conventional vegetation indices do not reflect the degree of green algae, whereas the proposed index reflects the green algae accurately depending on the characteristics of the water body.

Fig. 14 shows the detection results for UAV-based images assumed to contain no green algae. The existing indices express green algae even when there is no algae. The existing vegetation indices are composed only of ratios based on the combination of RGB values. In contrast, the proposed method does not indicate that there are algal blooms in such cases, as shown in Fig. 14. This is because the proposed method uses an index based on the hue and saturation.

Fig. 15 presents the estimation results for online-based images with green algae. The proposed algorithm exhibits reasonable estimation performance and captures the detailed information of the green algae. For the case of no green algae,

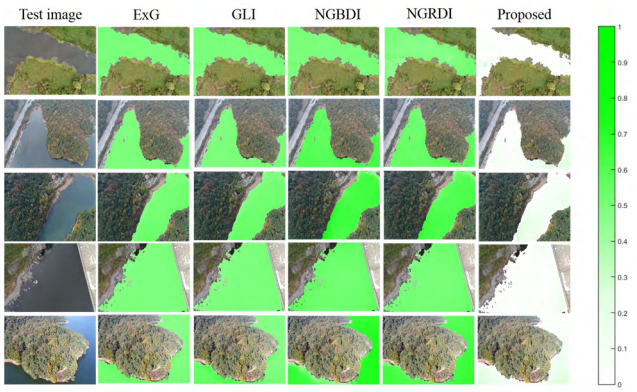

FIGURE 14. Algal bloom detection results for UAV-based test images without green algae.

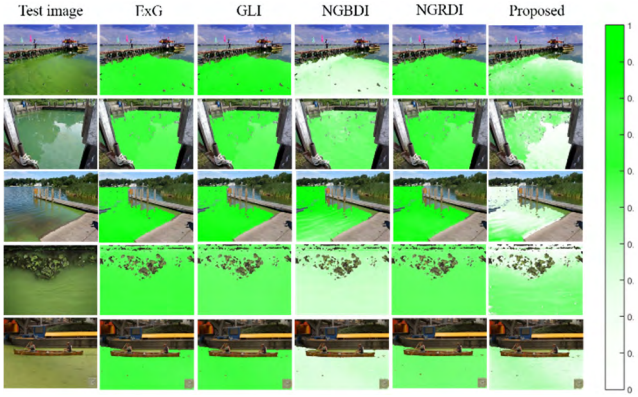

FIGURE 15. Algal bloom detection results for online-based test images with green algae.

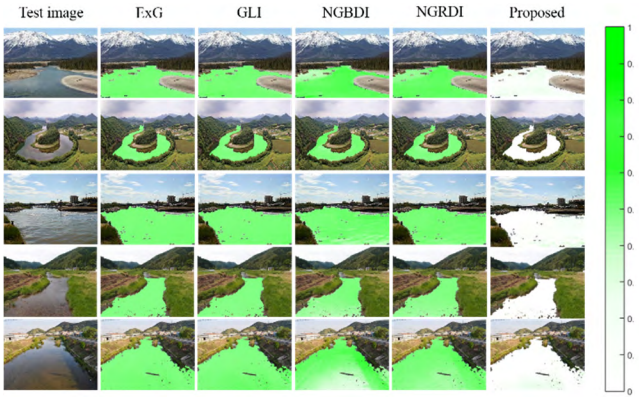

FIGURE 16. Algal bloom detection results for online-based test images without green algae.

our method achieves superior results to the existing algorithms as shown in Fig. 16. Similar to the results shown in Fig. 14, all the existing methods indicate that there is green algae.

The estimation results for the aerial test images are shown in Fig. 17. All the images in this set include green algae in the water bodies. Because these images are captured by aircraft,

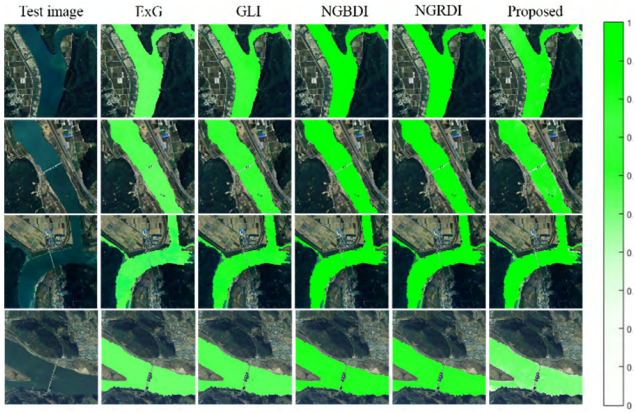

FIGURE 17. Algal bloom detection results for aircraft -based test images with green algae.

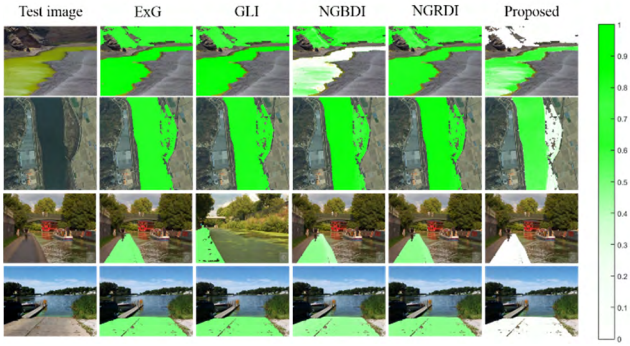

FIGURE 18. Algal bloom detection results for cases where the water body is erroneously extracted.

the image resolution is low, and the color ranges are limited. Therefore, all the methods exhibits similar performance.

Various vegetation indices were originally designed for satellite imagery. Therefore, they show good results for images taken from a long distance. However, the conventional indices often show poor estimation results for high resolution images captured by UAVs or smartphone cameras. Conversely, the proposed method exhibits similar estimation results regardless of the image type. This is because our method uses not only the vegetation index but also the saturation and hue of the water body.

Another advantage of the proposed method is that even when the water body is erroneously extracted, it is not identified as green algae as shown in Fig. 18. Here, the first and second rows show examples where two segments are detected as the water body. One segment is a real water body, and the other is not a water body. In contrast to the existing methods, the proposed method does not detect green algae in areas other than the water body. The third and fourth rows in Fig. 18 show the case where a non-water body is detected as a water body. The proposed algorithm indicates that there is no green algae in this area.

The results of the experiments indicate that our method can obtain better algal bloom estimation performance than the conventional methods for various types of images.

In addition, the proposed algorithm performs well even when the water body is erroneously detected.

D. LIMITATION

The proposed method introduced a new scheme for detecting algal blooms using a single RGB image. For this reason, there exist some difficulties in verifying the performance or usefulness of the proposed approach. Our method is not able to extract the water perfectly under various distortions such as sunlight and shadows. This is a problem that many image processing methods face and can be improved using new methods.

At present, it is very difficult to objectively verify the green algae detection performance, because there is no database or evaluation indicator for algal bloom detection. The performance of the proposed method can only be subjectively evaluated using the experimental results presented. In the future, if the image-based green algae detection method is activated, we expect that a database and criteria for objectively evaluating the performance will be established. A more important problem is how to use the generated probability map to monitor or mitigate harmful algal blooms. Our results can be used as input to an early algal bloom warning system after properly processing the probability map, or used as initial input to an algal bloom removal robotic system [21].

VI. CONCLUSION

In this paper, we presented a new single image-based algal bloom detection scheme. We first extracted the water body using wavelet leader-based texture analysis. We used the entropy and the number of zero wavelet coefficients as measures for the water body extraction. We presented a simple sky region removal method using the average brightness of the segmented regions for an online image dataset. For the extracted water body, we developed three indices bases on the RGB-based vegetation index, hue component, and saturation component. The final index was obtained via multiplication of these three indices. In various experiments, we achieved 94% accuracy for water body extraction, and our method exhibited better estimation results than the existing methods. In addition, we showed that the proposed method performs well even if the water body is extracted incorrectly.

REFERENCES

[1] L. Li, L. Li, and K. Song, "Remote sensing of freshwater cyanobacteria: An extended IOP Inversion Model of Inland Waters (IIMIW) for partitioning absorption coefficient and estimating phycocyanin," *Remote Sens. Environ.*, vol. 157, pp. 9–23, Feb. 2015.

[2] J. W. Rouse, Jr., R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," in *Proc. 3rd Earth Resour. Technol. Satell. Symp.*, Washington, DC, USA, 1974, pp. 309–317.

[3] A. Huete, C. Justice, and W. Van Leeuwen, "MODIS vegetation index (MOD13) Algorithm. Theoretical basis document," NASA Goddard Space Flight Centre, Greenbelt, MD, USA, 1999. [Online]. Available: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf

[4] J. Gower, S. King, G. Borstad, and L. Brown, "Detection of intense plankton Blooms using the 709 nm band of the MERIS imaging spectrometer," *Int. J. Remote Sens.*, vol. 26, no. 9, pp. 2005–2012, 2005.

[5] C. Hu, "A novel ocean color index to detect floating algae in the global oceans," *Remote Sens. Environ.*, vol. 113, pp. 2118–2129, Oct. 2009.

[6] Y. B. Son, J.-E. Min, and J.-H. Ryu, "Detecting massive green algae (*Ulva prolifera*) Blooms in the Yellow Sea and East China Sea using Geostationary Ocean Color Imager (GOCI) data," *Ocean Sci. J.*, vol. 47, no. 3, pp. 359–375, 2012.

[7] P. Shanmugam, M. Suresh, and B. Sundarabalan, "OSABT: An innovative algorithm to detect and characterize ocean surface algal Blooms," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 6, no. 4, pp. 1879–1892, Aug. 2013.

[8] Y. Zhang, R. Ma, H. Duan, S. A. Loiselle, J. Xu, and M. Ma, "A novel algorithm to estimate algal Bloom coverage to subpixel resolution in lake Taihu," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 3060–3068, Jul. 2014.

[9] B. Pan, Z. Shi, Z. An, Z. Jiang, and Y. Ma, "A novel spectral-unmixing-based green algae area estimation method for GOCI data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 2, pp. 437–449, Feb. 2017.

[10] D. Van der Merwe and K. P. Price, "Harmful algal Bloom characterization at ultra-high spatial and temporal resolution using small unmanned aircraft systems," *Toxins*, vol. 7, no. 4, pp. 1065–1078, 2015.

[11] S. Shang, Z. Lee, G. Lin, C. Hu, L. Shi, Y. Zhang, X. Li, J. Wu, and J. Yan, "Sensing an intense phytoplankton Bloom in the western Taiwan Strait from radiometric measurements on a UAV," *Remote Sens. Environ.*, vol. 198, pp. 85–94, Sep. 2017.

[12] H.-M. Kim, H.-J. Yoon, S. W. Jang, S. N. Kwak, B. Y. Sohn, S. G. Kim, and D. H. Kim, "Application of unmanned aerial vehicle imagery for algal Bloom monitoring in river Basin," *Int. J. Control Automat.*, vol. 9, no. 12, pp. 203–220, 2016.

[13] S. W. Jang, H. J. Yoon, S. N. Kwak, B. Y. Sohn, S. G. Kim, and D. H. Kim, "Algal Bloom monitoring using UAVs imagery," *Adv. Sci. Technol. Lett.*, vol. 138, pp. 30–33, 2016.

[14] J. A. J. Berni, P. J. Zarco-Tejada, L. Suárez, and E. Fereres, "Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 722–738, Mar. 2009.

[15] R. Aguirre-Gómez, O. Salmerón-García, G. Gómez-Rodríguez, and A. Peralta-Higuera, "Use of unmanned aerial vehicles and remote sensors in urban lakes studies in Mexico," *Int. J. Remote Sens.*, vol. 38, no. 8, pp. 2771–2779, 2017.

[16] F. Su, Z. Gao, X. Jiang, W. Shang, J. Ning, D. Song, and J. Ai, "A UAV and S2A data-based estimation of the initial biomass of green algae in the South Yellow Sea," *Mar. Pollut. Bull.*, vol. 128, pp. 408–414, Mar. 2018.

[17] C. Kislik, I. Dronova, and M. Kelly, "UAVs in support of algal Bloom research: A review of current applications and future opportunities," *Drones*, vol. 2, no. 4, p. 35, 2018.

[18] A. C. S. Kumar and S. M. Bhandarkar, "A deep learning paradigm for detection of harmful algal Blooms," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Santa Rosa, CA, USA, Mar. 2017, pp. 743–751.

[19] Y. Wang, R. Tan, G. Xing, J. Wang, X. Tan, and X. Liu, "Energy-efficient aquatic environment monitoring using smartphone-based robots," *Trans. Sensor Netw.*, vol. 12, no. 3, 2016, Art. no. 25.

[20] S. W. Jung, D. H. Kim, K. W. Kim, and H. Myung, "Image-based algal Blooms detection using local binary pattern," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Daejeon, South Korea, 2016, pp. 3539–3540.

[21] S. Jung, H. Cho, D. Kim, K. Kim, J. Han, and H. Myung, "Development of algal Bloom removal system using unmanned aerial vehicle and surface vehicle," *IEEE Access*, vol. 5, pp. 22166–22176, 2017.

[22] A. Samantaray, B. Yang, J. E. Dietz, and B.-C. Min, "Algae detection using computer vision and deep learning," 2018, *arXiv:1811.10847*. [Online]. Available: https://arxiv.org/abs/1811.10847

[23] M. Cimpoi, S. Maji, and A. Vedaldi, "Deep filter banks for texture recognition and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 3828–3836.

[24] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[25] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discovery Data Mining (KDD)*, Portland, OR, USA, 1996, pp. 226–231.

[26] A. Arneodo, N. Decoster, P. Kestener, and S. G. Roux, "A wavelet-based method for multifractal image analysis: From theoretical concepts to experimental applications," *Adv. Imag. Electron Phys.*, vol. 126, pp. 1–92, Jan. 2003.

[27] S. Arivazhagan and L. Ganesan, "Texture classification using wavelet transform," *Pattern Recognit. Lett.*, vol. 24, nos. 9–10, pp. 1513–1521, 2003.
[28] H. Ji, X. Yang, H. Ling, and Y. Xu, "Wavelet domain multifractal analysis for static and dynamic texture classification," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 286–299, Jan. 2013.
[29] H. Wendt, S. G. Roux, S. Jaffard, and P. Abry, "Wavelet leaders and bootstrap for multifractal analysis of images," *Signal Process.*, vol. 89, no. 6, pp. 1100–1114, 2009.
[30] L. Wan, Y. Li, H. Cen, J. Zhu, W. Yin, W. Wu, H. Zhu, D. Sun, W. Zhou, and Y. He, "Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape," *Remote Sens.*, vol. 10, no. 9, 2018, Art. no. 1484.
[31] A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist, "Novel algorithms for remote estimation of vegetation fraction," *Remote Sens. Environ.*, vol. 80, pp. 76–87, Apr. 2002.
[32] X. Wang, M. Wang, S. Wang, and Y. Wu, "Extraction of vegetation information from visible unmanned aerial vehicle images," *Trans. Chin. Soc. Agricult. Eng.*, vol. 31, no. 5, pp. 152–159, 2015.
[33] D. M. Woebbecke, G. E. Meyer, K. Von Bargen, and D. A. Mortensen, "Color indices for weed identification under various soil, residue, and lighting conditions," *Trans. ASAE*, vol. 38, no. 1, pp. 259–269, 1995.
[34] *Images Taken By Aircraft*. Accessed: Feb. 12, 2019. [Online]. Available: https://www.water.or.kr/disaster/live/airPhotoMain.do

CHEOL WOO PARK received the B.S. degree in physics from Pusan National University, South Korea, in 2018, where he is currently pursuing the M.S. degree in electronics engineering. His research interests include computer vision and image processing.

JONG JU JEON received the B.S. degree in electronics engineering from Pusan National University, South Korea, in 2014, where he is currently pursuing the Ph.D. degree in electronics engineering. His research interests include digital image forensic, image processing, and computer vision.

YONG HO MOON received the B.S., M.S., and Ph.D. degrees in electronics engineering from Pusan National University, South Korea, in 1992, 1994, and 1998, respectively. He has been with the Faculty of Gyeongsang National University, since 2007, where he is currently a Professor with the Department of Aerospace and Software Engineering. His research interests include image processing, SoC, embedded systems, and digital forensics.

IL KYU EOM received the B.S., M.S., and Ph.D. degrees in electronics engineering from Pusan National University, South Korea, in 1990, 1992, and 1998, respectively. He has been with the Faculty of Pusan National University, since 1997, where he is currently a Professor with the Department of Electronics Engineering. His research interests include image forensic and image processing, computer vision, and machine learning.