Self-attention and generative adversarial networks for algae monitoring

Nhut Hai Huynh, Gordon Böer, and Hauke Schramm

ABSTRACT
Water is important for the natural environment and human health. Monitoring algae concentrations yield information on the water quality. Compared with in situ measurements of water quality parameters, which are often complex and expensive, remote sensing techniques, using hyperspectral data analysis, are fast and cost-effective. The objectives of this study are (1) to estimate the algae concentrations from hyperspectral data using deep learning techniques, (2) to investigate the applicability of attention mechanisms in the analysis of hyperspectral data, and (3) to augment the training data using generative adversarial networks (GANs). The results show that the accuracy of deep learning techniques is 7.6% higher than that of simpler artificial neural networks. Compared to noise injection and principal component analysis-based data augmentation, the use of a GAN-based data augmentation method significantly improves the accuracy of algae concentration estimates (>5%). In addition, models with added attention mechanisms yield an average 3.13% higher accuracy than those without attention techniques. The result demonstrates the improvement of spectral features of artificial hyperspectral data based on the self-attention approach, revealing the potential of attention techniques in hyperspectral remote sensing.

Introduction
Inland waters, such as lakes and rivers, play important roles in ecosystems and the biodiversity of freshwater species. In addition, these water resources are fundamental for human health. However, inland waters are contaminated with chemicals, wastewater, and plastics. Algae blooms pose a threat because they lead to oxygen deficiency and toxic emissions. Due to the dangerous effects of algae, the demand for an algae-growth monitoring system is high. Algae can be used as environmental indicators to assess the health or determine the content of water bodies such as inland waters because of their nutrient needs, rapid reproduction rates, and very short life cycles. Hence, the measurement of the algal spread with respect to species composition and densities reflects significant changes in the water chemistry and conditions. By monitoring the development of algae in inland waters, conclusions can be drawn about the water quality. For this study, four types of algae were selected because of their key effects on the environment: green algae, blue-green algae (cyanobacteria), diatoms, and cryptophyta. Green algae, one of the major algae types, supply food and oxygen sources to aquatic organisms; similarly, diatoms are the base of the food chain in lakes and rivers. Cryptophyta convert the energy of sunlight into sugar, which is an important food source for aquatic organisms in deep lakes or during the darkness of winter. Furthermore, the growth of blue-green algae reduces the water quality and can lead to mass fish death because these bacteria consume a large amount of oxygen.

A common approach to assess algae content involves identifying and counting algal cells using light microscopy, either manually or semi-automatically using image processing techniques (Coltelli et al., 2014). Based on algal photosynthetic pigments, flow cytometry is another approach that allows for a more rapid identification of dominant algae (Ormerod & Imrie, 1990). Remote sensing offers a viable alternative to these approaches, which have inherent limitations in terms of processing time and water sample preservation. Because the distinct spectral shape of energy released by algal photosynthesis depends on the algae type and pigment composition, spectral analysis provides insights into algae in inland waters (MacIntyre et al., 2010; Yentsch & Yentsch, 1979).

Several methods can be used to collect remote sensing data such as the use of spectrometers just above the water surface, airborne hyperspectral cameras, or multispectral sensors on Earth observation satellites. Satellite data covers huge areas, and many algorithms are applied for hyperspectral data analysis. In the last two decades, support vector machines became a popular method for the processing of this kind of data (Demir & Erturk, 2007; Gualtieri & Crompt, 1999). Some advanced algorithms, such as active
learning (Rajan et al., 2008), semi-supervised learning (Shi et al., 2013), and adversarial learning (Shi et al., 2020), leverage the powerful ability of machine learning for classification of hyperspectral images. Fang et al. (2019) and Xu et al. (2020) used attention mechanisms but did not explain the spectral-wise attention analysis of hyperspectral data. Recently, Guo et al. (2020) proposed a multitask classifier with attention techniques to reduce misclassification and improve robustness. Other algorithms can be used to estimate the algal concentration from spectra collected by above-water surface spectrometers. Recently, Maier and Keller (2019) applied several machine learning (ML) strategies to monitor algae and obtained promising results. This motivated us to investigate to what extent deep learning (DL), particularly convolutional neural networks (CNNs), and data augmentation methods can improve the prediction performance.

In addition to DL, the benefits of data augmentation methods for the estimation of the algae concentration were analyzed. A principal component analysis (PCA)-based method, noise injection, and generative adversarial network (GAN) frameworks were used in this study for data augmentation. Nalepa et al. (2019) proposed online and offline modes for data augmentation. In the offline mode, known as traditional data augmentation, synthesized samples are stored on a disk before training the prediction models. In online data augmentation, generative models produce data promptly. In this study, data augmentation was implemented based on the two above-mentioned procedures. Particularly, multitask GANs were utilized for online data augmentation. The generators and discriminators of the GANs are responsible for data augmentation and algae estimation, respectively.

Because the above-mentioned preliminary studies showed the ability to extract and select only task-relevant features, the authors of this work applied deep learning and attention mechanisms to utilize spectral information for the estimation of algal concentration. In addition, due to the lack of training data, this study will adopt data augmentation methods. This work proposes a single model with a multitask GAN-based framework and attention mechanisms with its variants to compare and evaluate the effectiveness of our proposed model.

- The major contributions of this work are summarized as follows: The study justifies the advantages of DL models in the estimation of algal concentration by showing that CNNs outperform the simple artificial neural network (ANN). It additionally proposes a multitask GAN-based model with attention mechanisms that achieves state of the art performance by using online data augmentation and a spectral-wise attention technique. In order to show the improvement obtained from training multitask GAN-based models, the study also compares the performance of models when using online and offline data augmentation. Finally, it is shown that attention-guided deep learning methods produce better results than regular deep learning methods. Besides that, the authors analyzed and visualized a spectral-wise attention technique for hyperspectral data.

Related Work

Traditionally, microscopy is used to identify phytoplankton at the species level based on morphological characteristics. Algae cells are first collected on filters, identified, and then counted. Despite the essential role microscopy plays in algae monitoring (Tomas, 1997), it is time-consuming and requires profound expertise regarding taxa discrimination. In addition, it is difficult to recognize very small phytoplankton (i.e., pico- and nano-sized) and monitoring is carried out on extremely small sampling volumes.

Flow cytometry is a faster approach than the traditional method in which phytoplankton cells in a water sample are forced through a small aperture into a light field. Based on optical features, the sensor measures the chlorophyll fluorescence and scattering, which can be used to identify the dominant algae (Olson et al., 1989). This approach is more advanced than the microscopic technique because the analysis is much faster and small algae can be identified. Advances in the ability of underwater measurement are promising with respect to automatic sampling (Dubelaar & Gerritzen, 2000) and data acquisition on algal dynamics (Sosik et al., 2003). However, this method has temporal and spatial limitations because of the small water sample volumes. In contrast to microscopy, flow cytometry can effectively analyze pico- and nanophytoplankton; however, it incorrectly recognizes microphytoplankton.

High-performance liquid chromatography (HPLC) is commonly used to monitor algae at the group level according to the identification and quantification of several pigments (Jeffrey & Vesk, 1997). The principle of this method is that several pigments, which are grouped in specific taxa of phytoplankton, so-called pigment fingerprints, are used to recognize and quantify algal groups. HPLC is renowned for its rapid analysis and useful interpretation at the class level since it eliminates the need for profound knowledge of algal taxonomy. However, in addition to cost limitations, a considerable barrier to the use of HPLC is the variability in pigment fingerprints across algal taxa (Nair et al., 2008).
Phytoplankton emits light energy during photosynthesis, which varies in spectral shape and intensity among algae groups and can be used to detect and characterize them in remote sensing approaches. The energy is recorded and then normalized to eliminate illumination variations, yielding the so-called remote-sensing reflectance.

The data recorded by satellites covers a relatively large area of inland waters. Many algorithms can be used to analyze hyperspectral images. Lin et al. (2013) applied stacked autoencoder and logistic regression to classify hyperspectral images. Due to the advent of DL, the performance of hyperspectral analysis has been maximized in recent studies. Hu et al. (2015) used a deep CNN to classify hyperspectral images. Zhu et al. (2018) pointed out that the problem of applying DL to hyperspectral data is the lack of training data and proposed an auxiliary classifier GAN (ACGAN) to generate augmented data and mitigate this problem. In addition, Nalepa et al. (2019) utilized several data augmentation methods on- and offline. To overcome the lack of labeled hyperspectral data, Lin et al. (2018) reused other preexisting data and associated interpretations while training a related problem and then applied low-level features obtained from the initial layers to the hyperspectral image problem. Recently, Fang et al. (2019) and Xu et al. (2020) improved the distinguishability of spectral features by using a spectral-wise attention technique and achieved excellent results. However, the aforementioned studies strongly relied on hyperspectral data from satellites, whereas this type of data is subject to very large uncertainties. These uncertainties are caused by adjacency and aerosol effects as well as sun and sky glint due to the large distance between the spacecraft and water surface. Hence, the data must be validated in accordance with the satellite missions by using just-above-water surface measurements.

The sensitivity to the environmental effects of just-above-water surface measurements recorded by spectrometers is generally less pronounced than that of data measured from space. In this work, we analyzed signals collected by spectrometers. Previously, several authors, such as Gitelson (1992), Schalles et al. (1998), Gitelson et al. (2007), and Zhou et al. (2013), analyzed the relation of algal occurrences to the ratio between the reflection minimum and maximum at ~700 nm to monitor algae in inland waters. Furthermore, Rundquist et al. (1996) relied on derivatives of the spectra for the estimation of the chlorophyll a concentration. Recently, Maier and Keller (2019) attempted to estimate the algal concentration via chlorophyll a from data recorded by spectrometers using several ML algorithms such as support vector machine, random forest, and neural network. However, they faced difficulties in terms of varying spectral features and the availability of training data because of weather conditions. Therefore, effective estimation and data augmentation models are needed.

Proposed Approach

Dataset

This study uses the dataset recorded by Maier and Keller (2019) at two locations (several sampling sites) in Karlsruhe, Germany. In total, 898 spectra were collected during the summer of 2018, consisting of 383 and 515 data points measured at the Physikerteich and Schlossgarten, respectively. The data consist of spectra recorded by a spectrometer and algal concentrations, which are used as input data and target values, respectively. The concentrations and corresponding spectra were matched based on their respective timestamps.

Hyperspectral data were collected by a reflectance box (RoX) spectrometer. The spectrometer was mounted on a tripod and oriented in vertical direction. The device includes a cosine receptor, which points to the zenith axis to measure incoming radiance, and an upwelling radiance receptor, which was aligned with the water surface. Finally, the ratio between the obtained incoming and upwelling radiance was determined, resulting in spectra with a spectral range of 341 to 1015 nm and an average sampling interval of 0.67 nm. The integration time of a measurement was 15 s. Maier and Keller (2019) prepared the data for algae estimation using the following procedure:

- Select the wavelength range to avoid sensor noise.
- Identify and eliminate outliers by comparing the intensities at any wavelength with the valid range with the centered median of the data points within the sampling period at a single point.
- Match hyperspectral data points to ensure that the differences between the timestamps of the recorded spectra and reference data are within a time span of one minute.

Figure 1 shows several example spectra collected at the Physikerteich with respect to various chlorophyll a concentrations. The water samples were collected as reference data every five minutes at a depth of
10 cm and sampling site close to the spectrometer. The water samples were kept out of sunlight and then evaluated with an algae lab analyser (ALA) device from bbe Moldaenke GmbH.

**Methodology**

The authors propose a DL model with multitask GAN-based framework and multihead self-attention for data augmentation and representation enhancement, respectively. This section introduces two key concepts of the above techniques: multitask GANs and multihead self-attention.

**Multitask GAN-based framework**

The GAN framework (Goodfellow et al., 2014) consists of a generator G and discriminator D, which are in competition with each other, that is, G tries to generate fake data, while D tries to classify data of unknown origin as fake or real data. Based on this competition, the two neural networks continuously and effectively improve. The optimization function of G and D can be expressed as follows:

$$\min_G \max_D E[\log P(\text{real}|X_{\text{real}})] + E[\log P(\text{fake}|X_{\text{fake}})]$$  (1)

where $X_{\text{real}}$ and $X_{\text{fake}}$ denote fake and real input data of discriminator D, respectively.

Odena et al. (2017) proposed the ACGAN, a multitask GAN, in which the discriminator identifies not only the origin of the data via source recognition loss minimization as in the vanilla GAN but also the class of data through classification loss optimization. The generator maximizes the genuinity loss and minimizes the auxiliary classification loss. Based on this principle, Zhu et al. (2018) utilized ACGAN to boost the hyperspectral image classification performance. This inspired us to apply ACGAN to both our online data augmentation and algae regression problems.

In order to promote the diversity-oriented quality of artificial spectra, the combination between real and reference spectra is the input of our proposed GANs, yielding the so-called auxiliary regression with reference GAN (rGAN). A reference spectrum is defined as a “neighboring” spectrum belonging to the same group as the primary spectrum. To identify the group of spectra, mean-shift clustering was used in the preprocessing phase to cluster adjacent data points with comparable phytoplankton concentrations, that is, the algal concentrations between data points are similar. In fact, the combination of the original and reference spectra diversifies genuine spectra, while real–fake and regression losses enhance the realistic and augmentation-related quality of fake spectra, respectively. Figure 2 visualizes the scheme of rGAN.

The discriminator receives either a fake or real spectrum as input and yields the source probability of the input spectrum and the regressed algal concentrations. Although the source loss is the same as that in ACGAN, the classification loss is replaced by regression loss, $L_R$, that is, mean absolute difference (MAD), as shown in Equation 2, to consider the regression task.

$$\text{MAD} = \frac{1}{N} \sum_{i=1}^{N} |C_i - \hat{C}_i|$$  (2)

where $C_i$ and $\hat{C}_i$ denote the true and predicted algae concentrations of sample $i$, respectively, and $N$ is the number of samples. Equation 3 presents the source

![Figure 1. Examples of spectra collected at the Physikerteich with respect to various chlorophyll a concentrations.](image-url)
loss $L_S$ of rGAN. Equations 4 and 5 are the optimization functions of the generator $G$ and discriminator $D$ in rGAN, respectively.

$$L_S = E[\log P(\text{real}|X_{\text{origin}} + X_{\text{ref}})] + E[\log P(\text{fake}|X_{\text{fake}})]$$

(3)

where $X_{\text{fake}}$, $X_{\text{origin}}$ and $X_{\text{ref}}$ are the artificial, original and reference spectra, respectively.

$$\min_G (L_S + L_R) = \min_G (L_S + \text{MAD})$$

(4)

$$\max_D (L_S - L_R) = \max_D (L_S - \text{MAD})$$

(5)

**Multihead self-attention**

Sathyendranath et al. (1989), Wernand et al. (1997), and Lee et al. (2014) reported the strong relationships of reflective values among spectral bands. In the present work, the attention mechanism (Cheng et al., 2016) was utilized to analyze intraspectral relationships and their respective significance for the estimation of the algae content. This information can be used to emphasize important combinations and to reduce the influence of superfluous ones. In addition, multi-head self-attention (Vaswani et al., 2017) was used for ensemble information of various attention representations from different subspaces. All products of independent single-head attentions were concatenated and linearly transformed to obtain new representations. In this work, multihead self-attention is considered as a layer and added to CNN and GANs, yielding the so-called attention CNN (aCNN) and a GAN-based model with self-attention, respectively.

**Combination of self-attention and GANs**

Based on Lee et al. (2014), the interdependence of reflectance between wavebands was used for hyperspectral reconstruction. In the present work, self-attention was applied to a GAN generator to analyze the spectral relationships instead of the Pearson correlation coefficient, as used in Lee et al. (2014). Zhang et al. (2018) combined self-attention and GAN, resulting in the so-called self-attention GAN (SAGAN) and achieved a good performance. The proposed model with combination of self-attention and GANs in the present work is an rGAN with five-head attention layers added to the generator and discriminator. In order to distinguish this proposed model from its variant without self-attention (rGAN), the authors name it as self-attention with reference GAN (SArGAN). Similar to the original SAGAN, spectral normalization (SN; Miyato et al., 2018) was applied to each layer of our SArGAN.

**Experiments**

The experiments were carried out in two phases: (1) training phase including hyperparameter tuning and model optimization and (2) testing phase to evaluate the model performance. During the training phase,
a random search was applied to tune the hyperparameters of the models such as the architecture, learning rates, and dropout rates. This was done by applying a five-fold cross-validation to the training data of each model. The configuration of each model that minimizes the root-mean-square error (RMSEs) was chosen. During the testing phase, each regression model estimated the chlorophyll a concentration based on testing data. The performance of each model was evaluated based on the coefficient of determination ($R^2$) and mean absolute percentage error (MAPE).

In general, the input data of both generative and regression models, i.e., hyperspectral data, are presented as a sequence of reflectivity values or a one-dimensional (1D) array. Hence, the outputs of the generators in both on- and offline data augmentation methods are 1D arrays of reflectivity values. The predictor (in offline mode) and estimator (in online mode) generate four outputs, corresponding to four predicted algae concentrations.

There are three kinds of comparisons corresponding to three main objectives of this study: investigating DL applicability, online data augmentation effectiveness, and attention robustness.

**DL applicability**

To explore the applicability of DL models, the authors compared CNN to ANN. Because an ANN achieves a good performance in monitoring algae according to Maier and Keller (2019), a simple ANN was applied in this study as a baseline model. The optimal architecture of the simple ANN, shown in Table 1, includes five fully connected (FC) layers with a dropout rate of 0.3.

The reason for the application of DL models, especially CNNs, is that they are renowned for noise tolerance and the extraction of important information. Because raw spectral data are noisy, convolutional filters are used to effectively smooth signals via filtering–parameter learning from mistakes. Furthermore, Gitelson (1992) and Dekker (1993) indicated that the peak at ~700 nm can be used to estimate algae; however, the peak spans a range of 40 nm between 685 and 725 nm. Because convolutional filters have the ability to compress the aforementioned wavelength range to a single value, CNN-based models can extract key information from this peak. In practice, a proven CNN comprises three 1D-convolutional layers (Conv1D) with a rectified linear unit (ReLU) as activation function, each of which is followed by batch normalization (BN) and a max pooling layer (Pool1D) for feature extraction and two FC layers for regression. Table 2 summarizes the architecture of the CNN.

**Online data augmentation effectiveness**

The comparison of online and offline data augmentation in this study shows the effectiveness of a multitask GAN for data augmentation. For offline data augmentation, noise injection and PCA- and single-task-GAN-based data generation methods were used to synthesize the spectra, which were saved on disk before training the prediction models. These augmented data are then fed to several single-task predictors (ANN, CNN, and aCNN) to estimate algae concentration. To add noise to raw data, the spectra were multiplied by a random number, which was selected from a range of 0.97 to 1.03. This range was chosen after parameter tuning. Nalepa et al. (2019) introduced data augmentation using PCA, where spectra are first decomposed into components. Subsequently, the first components are multiplied by a random number produced by a uniform distribution. Finally, the spectra are reconstructed by a PCA-based inverse transformation. Following this methodology, the first component was multiplied by a random number, which was selected from the range of 0.9 to 1.1 in this study to

---

**Table 3. The architecture of off-line GAN-based data augmentation.**

| Layer | Generator Architecture | Concat. | Layer | Discriminator Architecture |
|-------|------------------------|---------|-------|---------------------------|
| 1     | Conv1D(32x3) – BN – ReLU | None    | 1     | Conv1D(32x3) – BN – ReLU  |
| 2     | Conv1D(32x3) – BN – ReLU | None    | 2     | Conv1D(32x3) – BN – ReLU  |
| 3     | Conv1D(64x3) – BN – ReLU | None    | 3     | Conv1D(64x3) – BN – ReLU  |
| 4     | Conv1D(64x3) – BN – ReLU | None    | 4     | Conv1D(64x3) – BN – ReLU  |
| 5     | Conv1D(128x3) – BN – ReLU | None    | 5     | Conv1D(128x3) – BN – ReLU  |
| 6     | Trans-Conv1D(64x3) – layer 4 | FC(1024, drop: 0.3) | 6     | FC(1) – Sigmoid |
| 7     | Trans-Conv1D(64x3) – BN – ReLU | layer 3 | 7     | FC(1) – Sigmoid |
| 8     | Trans-Conv1D(32x3) – BN – ReLU | layer 2 | 8     | FC(1) – Sigmoid |
| 9     | Trans-Conv1D(32x3) – BN – ReLU | layer 1 | 9     | FC(1) – Sigmoid |
| 10    | Trans-Conv1D(16x3) – BN – ReLU | None    | 10    | FC(1) – Sigmoid |
| 11    | Conv1D(1x3, stride:1) – Tanh | None    | 11    | FC(1) – Sigmoid |
generate augmented data. In offline GAN-based data augmentation, the generator, supplied with the real spectrum as input, generates a virtual spectrum as output. Table 3 presents the GAN architecture used in this study. The generator G has an architecture similar to that of the U-Net (Ronneberger et al., 2015); it consists of five convolutional and five transposed-convolutional (Trans-Conv1D) layers. The discriminator D predicts the source probability of the spectral input. This network consists of five convolutional layers and two FC layers. The activation functions of earlier and intermediate layers in G and D are ReLU and leaky ReLU (lReLU), respectively. The hyperbolic tangent (Tanh) and sigmoid functions in the last layers are the activation functions of the G and D, respectively.

For online data augmentation, this study applies a multitask GAN in which the generator and discriminator function as data augmentation and algae concentration estimator, respectively. Table 4 displays the architecture of rGAN. The shape of the generator in rGAN is consistent with that of offline GAN-based data augmentation. The architecture of the five convolutional layers of D in rGAN is same as that used for offline GAN-based data augmentation. The next layer is a dense layer with 1024 units. Its output is fed into two paths corresponding to two tasks: source-recognition and concentration estimation. Table 5 presents SArGAN architecture. The SArGAN, used in this work, is shaped with respect to rGAN; however, five-head attention layers are added to the generator and discriminator before the second last transposed-convolutional layer and third convolutional layer, respectively.

**Attention robustness**

Regarding this objective, five-head self-attention was added to CNN and rGAN to form aCNN and SArGAN, respectively. In practice, better results can be achieved when the attention layer is applied to intermediate or the last layers. In our work, the attention layer was added to the intermediate layer, as shown in Tables 6 and 5. The authors then compared models with attention layer to their variants without attention layer (CNN vs. aCNN, rGAN vs. SArGAN).

In summary, the authors compared DL model to the baseline model and the proposed model to its variants to investigate the key objectives. Table 7 shows the main differences between models corresponding to their purposes.

**Results and Discussion**

**Spectra Generated by Data Augmentation**

Figure 3 shows the spectra synthesized by noise injection and PCA- and GAN-based data augmentation methods. It can be seen that the spectrum generated by offline data augmentation insignificantly differs from the original one.

The spectra generated by multitask GANs (rGAN and SArGAN) based on the combination of primary and reference spectra are depicted in Figure 4. These

---

**Table 3. The architecture of SArGAN.**

| Layer | Architecture | Concat. | Layer | Architecture |
|-------|--------------|---------|-------|--------------|
| 1     | Conv1D(32x3) – BN – RelU. | None    | 1     | Conv1D(32x3) – BN – RelU. |
| 2     | Conv1D(32x3) – BN – RelU. | None    | 2     | Conv1D(32x3) – BN – RelU. |
| 3     | Conv1D(64x3) – BN – RelU. | None    | 3     | Conv1D(64x3) – BN – RelU. |
| 4     | Conv1D(64x3) – BN – RelU. | None    | 4     | Conv1D(64x3) – BN – RelU. |
| 5     | Conv1D(128x3) – BN – RelU. | None    | 5     | Conv1D(128x3) – BN – RelU. |
| 6     | Trans-Conv1D(64x3) – BN – RelU. | layer 4 | 6     | FC(1024, drop: 0.3) |
| 7     | Trans-Conv1D(64x3) – BN – RelU. | layer 3 | 7     | FC(1) – Sigmoid |
| 8     | Trans-Conv1D(128x3) – BN – RelU. | layer 2 | 8     | FC(128, drop: 0.3) – FC(4) |
| 9     | Trans-Conv1D(32x3) – BN – RelU. | layer 1 | 9     | FC(4) |
| 10    | Trans-Conv1D(16x3) – BN – RelU. | None    | 10    | FC(128, drop: 0.3) – FC(4) |
| 11    | Conv1D(1x3, stride:1) – Tanh. | None    | 11    | FC(4) |

**Table 4. The architecture of rGAN.**

| Layer | Architecture | Concat. | Layer | Architecture |
|-------|--------------|---------|-------|--------------|
| 1     | Conv1D(32x3) – BN – ReLU. | None    | 1     | Conv1D(32x3) – BN – ReLU. |
| 2     | Conv1D(32x3) – BN – ReLU. | None    | 2     | Conv1D(32x3) – BN – ReLU. |
| 3     | Conv1D(64x3) – BN – ReLU. | None    | 3     | Conv1D(64x3) – BN – ReLU. |
| 4     | Conv1D(64x3) – BN – ReLU. | None    | 4     | Conv1D(64x3) – BN – ReLU. |
| 5     | Conv1D(128x3) – BN – ReLU. | None    | 5     | Conv1D(128x3) – BN – ReLU. |
| 6     | Trans-Conv1D(64x3) – BN – ReLU. | layer 4 | 6     | FC(1024, drop: 0.3) |
| 7     | Trans-Conv1D(64x3) – BN – ReLU. | layer 3 | 7     | FC(1) – Sigmoid |
| 8     | Trans-Conv1D(128x3) – BN – ReLU. | layer 2 | 8     | FC(128, drop: 0.3) – FC(4) |
| 9     | Trans-Conv1D(32x3) – BN – ReLU. | layer 1 | 9     | FC(4) |
| 10    | Trans-Conv1D(16x3) – BN – ReLU. | None    | 10    | FC(128, drop: 0.3) – FC(4) |
| 11    | Conv1D(1x3, stride:1) – Tanh. | None    | 11    | FC(4) |

**Table 5. The architecture of SArGAN.**

| Layer | Architecture | Concat. | Layer | Architecture |
|-------|--------------|---------|-------|--------------|
| 1     | Conv1D(32x3) – SN – BN – RelU. | None    | 1     | Conv1D(32x3) – SN – BN – RelU. |
| 2     | Conv1D(32x3) – SN – BN – ReLU. | None    | 2     | Conv1D(32x3) – SN – BN – ReLU. |
| 3     | Conv1D(64x3) – SN – BN – ReLU. | None    | 3     | Conv1D(64x3) – SN – BN – ReLU. |
| 4     | Conv1D(64x3) – SN – BN – ReLU. | None    | 4     | Conv1D(64x3) – SN – BN – ReLU. |
| 5     | Conv1D(128x3) – SN – BN – ReLU. | None    | 5     | Conv1D(128x3) – SN – BN – ReLU. |
| 6     | Trans-Conv1D(64x3) – SN – BN – ReLU. | layer 4 | 6     | FC(1024, drop: 0.3) |
| 7     | Trans-Conv1D(64x3) – SN – BN – ReLU. | layer 3 | 7     | FC(1) – Sigmoid |
| 8     | Trans-Conv1D(128x3) – SN – BN – ReLU. | layer 2 | 8     | FC(128, drop: 0.3) – FC(4) |
| 9     | Trans-Conv1D(32x3) – SN – BN – ReLU. | layer 1 | 9     | FC(4) |
| 10    | Trans-Conv1D(16x3) – SN – BN – ReLU. | None    | 10    | FC(128, drop: 0.3) – FC(4) |
| 11    | Conv1D(1x3, stride:1) – SN – Tanh. | None    | 11    | FC(4) |
spectra are much more diverse in terms of intensity and shape, while the objective functions of rGAN and SArGAN ensure the quality of this fake spectrum.

Correlation Maps of Attention Mechanisms

Figure 5 presents the correlation map (third head) of the intensity at the 109th position in the feature map with intensities at other positions. This correlation map was produced by the attention layer in the third layer of the discriminator. The 109th position in the feature map has a receptive field covering the wavelengths 669.572, 670.244, 670.915, and 671.586 nm. The peak at the 122nd position in the feature map, which corresponds to the wavelengths of 703.580, 704.242, 704.904, and 705.566 nm, illustrates the high correlation between intensities at 670 and 705 nm, which agrees with previous findings (Dekker, 1993; Gitelson, 1992; Gitelson & Kondratyev, 1991). Similarly, the attention map of the 111th position in the feature map (Figure 6), reflecting the intensities of the wavelengths of 674.938, 675.608, 676.278, and 676.948 nm, shows an intense peak at the 129th position, which has a receptive field

| Model  | Architecture | Purpose |
|--------|--------------|---------|
| ANN    | Only fully connected layers | Baseline model to evaluate the applicability of other DL |
| CNN    | Convolutional layers Single-task predictor for algal concentration estimation No attention layer | DL model for DL applicability investigation Usage of offline data augmentation to compare against online data augmentation effectiveness Baseline model to investigate the robustness of the attention method |
| aCNN   | Single-task predictor for algal concentration estimation Attention layer | Usage of offline data augmentation to compare against online data augmentation effectiveness Investigate the robustness of an attention-guided model |
| rGAN   | Generator for generating spectra and multitask discriminator for genuinity and algal concentration prediction No attention layer | Usage of online data augmentation to evaluate its effectiveness Baseline model to investigate the robustness of the attention method Usage of online data augmentation to evaluate its effectiveness Investigate the robustness of the attention method |
| SArGAN | Generator for generating spectra and multitask discriminator for genuinity and algal concentration prediction Attention layer | |

Table 8. Comparisons of MAPE between on- and off-line data augmentation models as well as deep-learning-based approaches to baseline (ANN).

| Algae    | Predictor | Raw | Noise | PCA | GAN | AE | rGAN | SArGAN |
|----------|-----------|-----|-------|-----|-----|----|------|--------|
| Green    | ANN       | 13.56 | 10.19 | 8.37 | 12.55 | 11.17 | 8.07 | 5.79 |
|          | CNN       | 8.80  | 10.49 | 10.53 | 9.31 | 9.78 | 7.68 |
|          | aCNN      | 10.55 | 9.86  | 8.82 | 7.75 | 9.25 | 8.73 |
| BlueGreen| ANN       | 29.77 | 19.53 | 18.81 | 22.53 | 22.66 | 13.68 | 9.43 |
|          | CNN       | 25.29 | 12.47 | 14.32 | 12.17 | 12.06 | 12.47 | 9.31 |
|          | aCNN      | 11.85 | 7.86  | 8.80 | 8.99 | 9.38 | 7.68 |
| Diatoms  | ANN       | 10.83 | 10.35 | 9.87 | 10.66 | 10.43 | 8.97 | 6.91 |
|          | CNN       | 25.77 | 10.27 | 9.84 | 8.85 | 8.85 | 7.68 |
|          | aCNN      | 9.21  | 10.57 | 10.31 | 10.42 | 10.13 | 9.45 |
| Cryptophyta | ANN     | 47.80 | 79.12 | 75.95 | 79.42 | 80.57 | 30.96 | 24.85 |
|          | CNN       | 49.68 | 67.35 | 60.11 | 62.35 | 59.87 | 30.96 | 24.85 |
|          | aCNN      | 44.48 | 32.11 | 36.06 | 40.69 | 38.34 | 30.96 | 24.85 |

Table 9. Comparisons of $R^2$ between on- and off-line data augmentation models as well as deep-learning-based approaches to baseline (ANN).

| Algae    | Predictor | Raw | Noise | PCA | GAN | AA | rGAN | SArGAN |
|----------|-----------|-----|-------|-----|-----|----|------|--------|
| Green    | ANN       | 93.36 | 95.63 | 96.80 | 94.33 | 95.03 | 97.82 | 98.46 |
|          | CNN       | 95.93 | 96.64 | 95.86 | 96.63 | 96.27 | 97.82 | 98.46 |
|          | aCNN      | 94.74 | 96.77 | 97.27 | 97.31 | 96.52 | 97.82 | 98.46 |
| BlueGreen| ANN       | 80.66 | 91.60 | 89.48 | 82.77 | 86.13 | 98.45 | 98.71 |
|          | CNN       | 90.10 | 94.23 | 92.69 | 92.76 | 92.45 | 97.77 | 98.71 |
|          | aCNN      | 98.48 | 98.16 | 97.01 | 96.64 | 97.64 | 98.45 | 98.71 |
| Diatoms  | ANN       | 94.62 | 95.23 | 95.04 | 95.05 | 94.99 | 95.84 | 94.71 |
|          | CNN       | 92.45 | 92.87 | 94.58 | 94.72 | 93.66 | 94.71 | 94.71 |
|          | aCNN      | 95.21 | 93.01 | 92.83 | 93.82 | 93.72 | 94.71 | 94.71 |
| Cryptophyta| ANN     | 37.61 | 37.75 | 43.58 | 39.90 | 39.71 | 77.09 | 87.16 |
|          | CNN       | 57.65 | 68.83 | 62.81 | 66.41 | 63.93 | 87.16 | 87.16 |
|          | aCNN      | 64.75 | 76.54 | 74.87 | 78.68 | 73.71 | 87.16 | 87.16 |
covering the wavelengths of 724.672, 725.328, 725.984, and 726.640 nm. This explains the close relationship between the intensities at 675 and 725 nm, which also agrees with the results of previous studies (Hoge et al., 1987; Pierson & Strömbeck, 2000; Ruddick et al., 2001; Yacobi et al., 1995). The above-mentioned analogy between the spectral correlation based on self-attention and the previously investigated channel relationship proves that the attention mechanism has the ability to use decisive spectral connections between wavelengths for algal concentration estimation. Note that the horizontal axes in Figures 5 and 6 were converted from the positions in the feature map to approximate wavelengths.

**Comparison of the Model Performance**

The coefficient of determination ($R^2$) and mean absolute percentage error (MAPE) are used in this work to express the algae prediction performance. Equations 6 and 7 present $R^2$ and MAPE, respectively.
where $C_i$ and $\tilde{C}_i$ denote the true and predicted algae concentrations of sample $i$, respectively. $\bar{C}$ is the mean of true algae concentrations.

Tables 8 and 9 display the comparisons of the performances of on- and offline data augmentation models as well as DL-based approaches to baseline (ANN) in terms of the MAPE and $R^2$, respectively. The rows present the predictors in the offline mode and estimators of multitask GAN-based models in the online mode. The columns illustrate the origin of the data: only raw data or data with supplementations by the offline augmentation technique or generator of multitask GANs in online mode.

Regarding the comparison of the prediction models, the average accuracy (AA) and average error (AE) of the CNN are higher and lower than those of the ANN for most algae, respectively, except for diatoms. In addition, the performances of the CNN improve by adding an attention technique. For instance, the AE of aCNN is 6.68% lower than that of the CNN and the AA of aCNN is 5.12% higher than that of the CNN for Blue-Green estimation. Similar to the predictors in offline data augmentation, the outcomes of the estimators using a multitask GAN for data augmentation are enhanced by applying an attention technique. The model trained with data augmented using SArGAN yield a $\sim$ 10% higher accuracy than that treated with rGAN.

Data augmentation methods leverage the prediction performance. Compared to offline data augmentation models, the models utilizing multitask GAN for the prompt generation of virtual spectra yield a better performance. The accuracy
of the model for which SArGAN was used as data augmentation for cryptophyta is more than 25% higher than that of the models to which offline data augmentation was applied.

Regarding the performance of algae prediction, the error and accuracy for cryptophyta are mediocre compared with those of other algae. The minimum MAPE for cryptophyta is 24.85% compared to 5.79% to 7.43% obtained for other algae.

In summary, the CNN has a 7.61% higher accuracy than the ANN, indicating the potential of the application of DL for phytoplankton concentration estimation. The accuracy of online GAN-based data augmentation methods is 5.94% and 5.57% higher than that of models using augmented data from noise injection and PCA-based data augmentation, respectively. The comparison of models with and without attention (aCNN vs. CNN, SArGAN vs. rGAN) showed that the accuracy improves by 3.81% and 2.46%, respectively, when attention mechanisms were used.

Conclusions

This study has shown that DL is applicable for the estimation of the phytoplankton concentration from spectra obtained by spectrometer measurements. DL models were found to give better results than those using a simple ANN to estimate phytoplankton concentration. The authors proposed a model with a multitask GAN-based framework and multihead self-attention mechanisms for data augmentation and representation enhancement, respectively. By experimentally comparing this model with its variants without either of the above techniques (aCNN and rGAN), the effectiveness of the proposed model was demonstrated. In particular, the multitask GAN framework outperforms models using only raw data, or virtual spectra, generated by noise injection or PCA. In addition, it could be shown that multihead self-attention mechanisms enhance the representation and significantly improve the algae prediction performance.

Further analysis of attention mechanisms expresses that the spectral correlation based on self-attention agrees with the earlier investigated channel relationship.

However, the estimation of the cryptophyta concentration remains difficult and further research is required. The limitations are caused by the variation of the environment, such as sunlight, solar position, and geometry, during the measurements. Therefore, we plan to deploy a sun-following system that keeps spectrometers at the optimal position against the Sun as Mobley (1999) suggested. A viewing direction of 40° from the nadir and 135° away from the Sun would mitigate the effects of sun- and sky-glint. In addition, environmental information will be embedded in hyperspectral data to reduce the variation caused by the environment.

Acknowledgments

The research is part of the CIWAWA project. The authors would like to thank project partners Philipp Maier, Christian Moldaenke and Andre Zaake for providing us hyperspectral data and algal concentrations.

Data availability statement

The data that support the findings of this study are available from Maier and Keller (2019). Restrictions apply to the availability of these data, which were used under license for this study. Data are available at https://datapub.gfz-potsdam.de/download/10.5880.FIDGEO.2020.036gr/ with the permission of Maier and Keller (2019).

Disclosure statement

This research is sponsored by bbe Moldaenke GmbH and may lead to the development of products that may be licensed to bbe Moldaenke GmbH, in which I have a business and/or financial interest. In addition, the research is supervised by Kiel University of Applied Sciences and used data from Maier and Keller (2019). I have disclosed those interests fully to Taylor & Francis and have in place an approved plan for managing any potential conflicts arising from this arrangement.

ORCID

Nhu Huy Ho \(\text{http://orcid.org/0000-0003-0622-100X}\)

Gordon Böer \(\text{http://orcid.org/0000-0002-6196-9558}\)

References

Cheng, J., Dong, L., & Lapata, M. (2016). Long short-term memory-networks for machine reading. In Empirical methods in natural language processing (Association for Computational Linguistics) (pp. 551–561). arXiv preprint arXiv:1601.06733.

Coltelli, P., Barsanti, L., Evangelista, V., Frassanito, A. M., & Gualtieri, P. (2014). Water monitoring: Automated and real time identification and classification of algae using digital microscopy. Environmental Science: Processes & Impacts, 16(11), 2656–2665. https://doi.org/10.1039/C4EM00451E

Dekker, A. G. (1993). Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing [Doctoral dissertation, Vrije Universiteit Amsterdam]. Free Universit. https://research.vu.nl/en/publications/detection-of-optical-water-quality-parameters-for-eutrophic-water

Demir, B., & Erturk, S. (2007). Hyperspectral image classification using relevance vector machines. IEEE Geoscience and Remote Sensing Letters, 4(4), 586–590. https://doi.org/10.1109/LGRS.2007.903069
Dubelaar, G. B. J., & Gerritsen, P. L. (2000). CytoBuoy: A step forward towards using flow cytometry in operational oceanography. *Scientia Marina, 64*(2), 255–265. https://doi.org/10.3989/scimar.2000.64n2255

Fang, B., Li, Y., Zhang, H., & Chan, J. C. W. (2019). Hyperspectral images classification based on dense convolutional networks with spectral-wise attention mechanism. *Remote Sensing, 11*(2), 159. https://doi.org/10.3390/rs11020159

Gitelson, A. A., & Kondratyev, K. Y. (1991). Optical models of mesotrophic and eutrophic water bodies. *International Journal of Remote Sensing, 12*(3), 373–385. https://doi.org/10.1080/01431169108929659

Gitelson, A. A., Schalles, J. F., & Hladik, C. M. (2007). Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment, 109*(4), 464–472. https://doi.org/10.1016/j.rse.2007.01.016

Gitelson, A. A. (1992). The peak near 700 nm on radiance spectra of algae and water: Relationships of its magnitude and position with chlorophyll concentration. *International Journal of Remote Sensing, 13*(17), 3367–3373. https://doi.org/10.1080/01431169208904125

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 27, pp. 2672–2680). Curran Associates, Inc. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

Guatieri, I. A., & Crump, R. F. (1999). Support vector machines for hyperspectral remote sensing classification. In *27th AIPR workshop: Advances in computer-assisted recognition* (Vol. 3584, pp. 221–232). International Society for Optics and Photonics.

Guo, H., Shi, Q., Du, B., Zhang, L., Wang, D., & Ding, H. (2020). Scene-driven multitask parallel attention network for building extraction in high-resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing 59* 4 4287–4406 . https://doi.org/10.1109/TGRS.2020.3014312

Hoge, F. E., Wright, C. W., & Swift, R. N. (1987). Radiance-ratio algorithm wavelengths for remote oceanic chlorophyll determination. *Applied Optics, 26*(11), 2082–2094. https://doi.org/10.1364/ACO.26.002082

Hu, W., Huang, Y., Wei, L., Zhang, F., & Li, H. (2015). Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors, 2015* Special Issue , 1–12 1687–725X doi:10.1155/2015/258619 .

Jeffrey, S. W., & Veski, M. (1997). Introduction to marine phytoplankton and their pigment signatures. In S. W. Jeffrey, R.F.C. Mantoura, and S.W. Wright (Eds.), *Phytoplankton pigments in oceanography*. UNESCO Publishing 37–84.

Lee, Z., Shang, S., Hu, C., & Zibordi, G. (2014). Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color satellite sensors. *Applied Optics, 53*(15), 3301–3310. https://doi.org/10.1364/AO.53.003301

Lin, J., Ward, R., & Wang, Z. J. (2018). Deep transfer learning for hyperspectral image classification. In *20th International Workshop on Multimedia Signal Processing (MMSp)* Vancouver, BC, Canada (pp. 1–5). IEEE. https://doi.org/10.1109/MMSP.2018.8547139

Lin, Z., Chen, Y., Zhao, X., & Wang, G. (2013). Spectral-spatial classification of hyperspectral image using autoencoders. In 9th International Conference on Information, Communications & Signal Processing Tainan (pp. 1–5). IEEE. https://doi.org/10.1109/ICICS.2013.6782778

MacIntyre, H. L., Lawrenz, E., & Richardson, T. L. (2010). Taxonomic discrimination of phytoplankton by spectral fluorescence. In D.J. Suggett, O. Prâslî, & M.A. Borowitzka (Eds.), *Chlorophyll a fluorescence in aquatic sciences: Methods and applications* (pp. 129–169). Springer.

Maier, P. M., & Keller, S. (2019). Estimating chlorophyll a concentrations of several inland waters with hyperspectral data and machine learning models Vosselman, G. In *ISPRS annals of photogrammetry, remote sensing and spatial information sciences* (ISPRS) (pp. 609–614). arXiv preprint arXiv:1904.02052.

Miyato, T., Kataoka, T., Koyama, M., & Yoshida, Y. (2018). Spectral normalization for generative adversarial networks. In *International Conference on Learning Representations* Vancouver, BC, Canada. arXiv preprint arXiv:1802.05957

Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Applied Optics, 38* (36), 7424–7455. https://doi.org/10.1364/AO.38.007442

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M. H., Devred, E., & Bouman, H. (2008). Remote sensing of phytoplankton functional types. *Remote Sensing of Environment, 112*(8), 3366–3375. https://doi.org/10.1016/j.rse.2008.01.021

Nalepa, J., Myller, M., & Kauhalo, M. (2019). Hyperspectral data augmentation. *IEEE Geoscience and Remote Sensing Letters, 16*(5), 806–810. arXiv preprint arXiv:1903.05580. https://doi.org/10.1109/LGRS.2018.2881102

Odena, A., Olah, C., & Shlens, J. (2017). Conditional image synthesis with auxiliary classifier GANs. In D. Precup, and Y.W. Teh (Eds.), *Proceedings of the 34th International Conference on Machine Learning* Sydney, NSW, Australia (Vol.70, pp. 2642–2651). JMLR.

Olson, R. J., Zettler, E. R., & Anderson, O. K. (1989). Discrimination of eukaryotic phytoplankton cell types from light scatter and autofluorescence properties measured by flow cytometry. *Cytometry: The Journal of the International Society for Analytical Cytology, 10*(5), 636–643. https://doi.org/10.1002/cyto.990100520

Ormerod, M. G., & Imrie, P. R. (1990). Flow cytometry Walker, J. M., and Pollard, J. W. in *Animal Cell Culture* (pp. 543–558). Humana Press.https://doi.org/10.1385/0-89603-150-0:543

Pierson, D. C., & Strömbeck, N. (2000). A modelling approach to evaluate preliminary remote sensing algorithms: Use of water quality data from swedish great lakes. *Geophysica, 36*(1–2), 177–202 https://www.geophysica.fi/pdf/geophysica_2000_36_1-2_177_pierson.pdf .

Rajan, S., Ghosh, J., & Crawford, M. M. (2008). An active learning approach to hyperspectral data classification. *IEEE Transactions on Geoscience and Remote Sensing, 46*(4), 1231–1242. https://doi.org/10.1109/TGRS.2007.910220

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* Munich, Germany (pp. 234–241). Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28
Vaswani, K. G., Gons, H. J., Rijkeboer, M., & Tilstone, G. (2001). Optical remote sensing of chlorophyll a in case 2 waters by use of an adaptive two-band algorithm with optimal error properties. Applied Optics, 40(21), 3575–3585. https://doi.org/10.1364/AO.40.003575

Rundquist, D. C., Han, L., Schalles, J. F., & Peake, J. S. (1996). Remote measurement of algal chlorophyll in surface waters: The case for the first derivative of reflectance near 690 nm. Photogrammetric Engineering and Remote Sensing, 62(2), 195–200. https://www.asprs.org/wp-content/uploads/pers/1996journal/feb/1996_feb_195-200.pdf

Sathyendranath, S., Prieur, L., & Morel, A. (1989). A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal waters. International Journal of Remote Sensing, 10(8), 1373–1394. https://doi.org/10.1080/01431168908903974

Schalles, J. F., Gitelson, A. A., Yacobi, Y. Z., & Kroenke, A. E. (1998). Estimation of chlorophyll a from time series measurements of high spectral resolution reflectance in an eutrophic lake. Journal of Physiology, 34(2), 383–390. https://doi.org/10.1046/j.1529-8817.1998.340383.x

Shi, Q., Liu, M., Liu, X., Liu, P., Zhang, P., Yang, J., & Li, X. (2020). Domain adaption for fine-grained urban village extraction from satellite images. IEEE Geoscience and Remote Sensing Letters, 17(8), 1430–1434. https://doi.org/10.1109/LGRS.2019.2947473

Shi, Q., Zhang, L., & Du, B. (2013). Semisupervised discriminative locally enhanced alignment for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 51(9), 4800–4815. https://doi.org/10.1109/TGRS.2012.2230445

Sosik, H. M., Olson, R. J., Neubert, M. G., Shalapyonok, A., & Solow, A. R. (2003). Growth rates of coastal phytoplankton from time-series measurements with a submersible flow cytometer. Limnology and Oceanography, 48(5), 1756–1765. https://doi.org/10.4319/lo.2003.48.5.1756

Tomas, C. R. (Ed.). (1997). Identifying marine phytoplankton. Elsevier.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in neural information processing systems (Vol. 30, pp. 5998–6008). Curran Associates, Inc. http://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf

Wernand, M. R., Shimwell, S. J., & De Munck, J. C. (1997). A simple method of full spectrum reconstruction by a five-band approach for ocean colour applications. International Journal of Remote Sensing, 18(9), 1977–1986. https://doi.org/10.1080/014311697217981

Xu, Q., Xiao, Y., Wang, D., & Luo, B. (2020). CSA-MSO3DCNN: Multiscale octave 3D CNN with channel and spatial attention for hyperspectral image classification. Remote Sensing, 12(1), 188. https://doi.org/10.3390/rs12010188

Yacobi, Y. Z., Gitelson, A., & Mayo, M. (1995). Remote sensing of chlorophyll in Lake Kinneret using high-spectral-resolution radiometer and Landsat TM: Spectral features of reflectance and algorithm development. Journal of Plankton Research, 17(11), 2155–2173. https://doi.org/10.1093/plankt/17.11.2155

Yentsch, C. S., & Yentsch, C. M. (1979). Fluorescence spectral signatures: The characterization of phytoplankton populations by the use of excitation and emission spectra. J. Mar. Res, 37(3), 471–483. https://images.peabody.yale.edu/publications/jmr/jmr37-03-06.pdf.

Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2018). Self-attention generative adversarial networks. In K. Chaudhuri, and R. Salakhutdinov (Eds.), Proceedings of the 36th International Conference on Machine Learning California, USA (Vol.97, pp. 7354–7363). PMLR. arXiv preprint arXiv:1805.08318

Zhou, L., Xu, B., Ma, W., Zhao, B., Li, L., & Huai, H. (2013). Evaluation of hyperspectral multi-band indices to estimate chlorophyll-A concentration using field spectral measurements and satellite data in Dianshan Lake, China. Water, 5 (2), 525–539. https://doi.org/10.3390/w5020525

Zhu, L., Chen, Y., Ghamisi, P., & Benidiktsson, J. A. (2018). Generative adversarial networks for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing, 56(9), 5046–5063. https://doi.org/10.1109/TGRS.2018.2805286