Fully automatic detection and classification of phytoplankton specimens in digital microscopy images

David Rivas-Villar [a,b,*], José Rouco [a,b], Rafael Carballeira [c], Manuel G. Penedo [a,b], Jorge Novo [a,b],

[a] *Centro de investigacion CITIC, Universidade da Coruña, 15071 A Coruña, Spain*

[b] *Grupo VARPA, Instituto de Investigacion Biomédica de A Coruña (INIBIC), ˜ Universidade da Coruna, 15006 A Coruña, Spain*

[c] *Centro de Investigacions Científicas Avanzadas (CICA), Facultade de Ciencias, Universidade da Coruna, 15071 A Coruña, Spain*

Abstract

Background and objective The proliferation of toxin-producing phytoplankton species can compromise the quality of the water sources. This contamination is difficult to detect, and consequently to be neutralised, since normal water purification techniques are ineffective. Currently, the water analyses about phytoplankton are commonly performed by the specialists with manual routine analyses, which represents a major limitation. The adequate identification and classification of phytoplankton specimens requires intensive training and expertise. Additionally, the performed analysis involves a lengthy process that exhibits serious problems of reliability and repeatability as inter-expert agreement is not always reached. Considering all those factors, the automatization of these analyses is, therefore,

highly desirable to reduce the workload of the specialists and facilitate the process.

Methods This manuscript proposes a novel fully automatic methodology to perform phytoplankton analyses in digital microscopy images of water samples taken with a regular light microscope. In particular, we propose a method capable of analysing multi-specimen images acquired using a simplified systematic protocol. In contrast with prior approaches, this enables its use without the necessity of an expert taxonomist operating the microscope. The system is able to detect and segment the different existing phytoplankton specimens, with high variability in terms of visual appearances, and to merge them into colonies and sparse specimens when necessary. Moreover, the system is capable of differentiating them from other similar objects like zooplankton, detritus or mineral particles, among others, and then classify the specimens into defined target species of interest using a machine learning-based approach.

Results The proposed system provided satisfactory and accurate results in every step. The detection step provided a FNR of 0.4%. Phytoplankton detection, that is, differentiating true phytoplankton from similar objects (zooplankton, minerals, etc.), provided a result of 84.07% of precision at 90% of recall. The target species classification, reported an overall accuracy of 87.50%. The recall levels for each species are, 81.82% for W. naegeliana, 57.15% for A. spiroides, 85.71% for D. sociale and 95% for the "Other" group, a set of relevant toxic and interesting species widely spread over the samples.

Conclusions The proposed methodology provided accurate results in all the designed steps given the complexity of the problem, particularly in terms of specimen identification, phytoplankton differentiation as well as the classification of the defined target species. Therefore, this fully automatic system represents a robust and consistent tool to aid the specialists in the analysis of the quality of

the water sources and potability.

Key words: Microscope images; Phytoplankton detection; Colony merging; Gabor filters; Bag of Visual Words; Deep features

---

* Corresponding author. Centro de investigacion CITIC, Universidade da Coruña, 15071 A Coruña, Spain

Email addresses: david.rivas.villar@udc.es (David Rivas-Villar), jrouco@udc.es (José Rouco), r.carballeira@udc.es (Rafael Carballeira), mgpenedo@udc.es (Manuel G. Penedo), jnovo@udc.es (Jorge Novo).

1 Introduction

Phytoplankton is the basis of the food chain in all the aquatic ecosystems worldwide, either being fresh water such as rivers or saltwater like oceans. Phytoplankton carries out photosynthesis producing oxygen and acting as a carbon sink. It, therefore, holds a significant amount of scientific attention over those properties.

Phytoplankton is also relevant given that high concentrations of particular species, known as blooms, produce toxins that contaminate potable water, with the consequent drastic impact in the involved population. The effects of these toxins range from gastroenteritis to liver damage, cancer or death [1]. The main problem derived from these proliferations is that the produced toxins cannot simply be removed by purification techniques [2], thus needing continuous monitoring. Moreover, phytoplankton blooming is a phenomenon that has been linked to many factors but most importantly to global warming. This will, undoubtedly, increase the frequency and intensity of these dangerous proliferations in the coming years, increasing the severity of the impact and the consequences for the dependent populations.

Nowadays, the studies about phytoplankton and other types of aquatic microbes require exhaustive work by the specialists which results in a big bottleneck to the output and quality of these studies. Water samples must be collected in situ by the experts, in a time consuming and tedious process. After that, these samples need to be transported to the laboratory to be concentrated and treated for their observation under the microscope. Counting and classifying the existing specimens is commonly done by direct observation and manual routine, again, in a tedious and time consuming process. However,

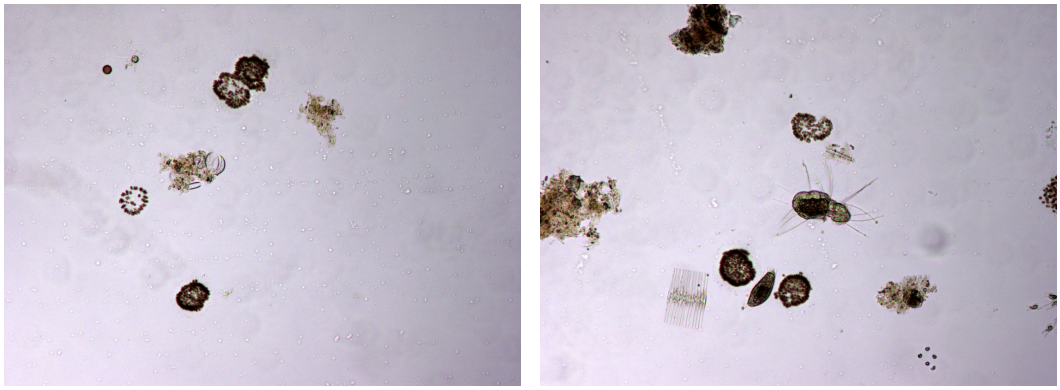
at this stage, it is possible to directly capture digital images of the samples for posterior analysis.

It has been shown that the assessment by the experts presents low recall rates, a 75%, in water with a high amount of debris [3]. The average of correct species identification rates below 73.7% [4]. Another issue related to the manual assessment of phytoplankton studies is the frequent lack of inter-expert agreement as well as discrepancies between criteria of the same expert in different time frames, species-wise, etc. It has been reported that the coefficient of variation between expert cell counts ranges from 8% to 57% depending on the target species [5]. The high disagreement on some species is explained by the varying complexity of the specimens, which is severely alleviated when objective protocols and criteria are established among experts. In other work [6], it was reported that trained specialists, or "book experts", can achieve a self-consistency of 67-83% in the species identification, and inter-expert agreement as low as 43%. However, when the same task was evaluated on experienced taxonomists, who routinely carry out field studies, the intra-expert agreement rises to 94-99%, and the agreement among peers rises to 84-95% [6]. Again, in this case, the performance variability depends on the complexity of the specific target species. These studies evidence, on one hand, that the phytoplankton identification and analysis is a complex task that requires extensive experience and knowledge. While, on the other hand, the use of systematic approaches helps both to improve the repeatability and the accuracy of the analyses. Thus, these analysis can certainly take advantage of automated computer-based approaches, that can be systematically applied without relying on experienced taxonomists. Moreover, the human performance is also influenced by psychological factors, like boredom and fatigue, derived from the repetitive and tedious manual work of identifying and sorting organisms under microscopes.

Therefore, plankton studies would benefit from available automatic approaches to simplify and alleviate the workload of the researchers along with the potential improvement of the results. Also, the continuous monitoring of water sources further motivates the necessity of consistent and reliable systems. The current dependency on experienced taxonomists for reliable manual analyses usually derives in an insufficient coverage of these monitoring studies to effectively prevent public health issues derived from toxic blooms [7].

Regarding computational proposals, partially automated analysis procedures have already been developed. We can classify the different developed approaches into two main separate groups: sample gathering and detection of specimens, and species identification approaches.

The sample gathering group is focused on building contraptions that are able to gather water data by themselves. The main disadvantage of these methods is their cost, size and the fact that most of them do not save the water samples, just images. The images also tend to be of less quality than those produced by a conventional microscope. Representative examples of these devices are Video Plankton Recorder (VPR) [8–10], Shadowed Image Particle Profiling and Evaluation Recorder (SIPPER) [11], FlowCytobot [12] and KRIP [13]. FlowCam [14] stands out as the most widely used device, working differently than the rest of the previously cited methods. In particular, it requires the gathering of the water as performed with the laboratory equipment. It produces higher quality images when compared with other related methods. This method simplifies the capturing process by passing specimens one at a time through a small tube. It, however, requires different flow cells depending on the size of the analysed specimens [15]. Overall, the algorithms that are used by these methods to detect the different specimens in the images are highly



(a) (b)

Fig. 1. Input images.

dependent on the specific acquisition features of each particular device and the specific target subset of phytoplankton species.

While microscopes are widely available, these automated imaging systems are expensive and often require other equipment to work, like a boat to tow them. Our proposed methodology uses, as input, microscopy images taken with a regular light microscope and, therefore, should be more affordable. Some examples of input images can be seen in Figure 1. These images include many specimens, as well as many spurious objects like mineral particles, organic debris or zooplankton. These objects may look very similar to phytoplankton specimens in many ways, so they are not easily discernible in many cases. Thus, the method intends on separating these spurious elements from the real phytoplankton specimens and later classifying the true specimens of the relevant species from the rest.

In terms of the detection step, there are few available approaches. Some of these approaches take regular microscopy images either containing a single or multiple specimens, as input [16,17] while other approaches, like the algorithm used in the automatic sampler KRIP [13], take the images from a gathering device. The output usually consists of segmentation masks for each

individual specimen. In the work of Zheng et al. [16], saliency object detection is used to distinguish specimens from the background in microscopy images containing only a single species. Verikas et al. [17] are able to detect round phytoplankton species through phase congruency and, then, refine these detections using ensemble classifiers. The algorithm used in the automatic sampler KRIP [13] applies a Sobel transformation as well as a threshold and morphological operators to the images captured by this device in order to separate the phytoplankton.

With respect to the species identification approaches, they are usually focused on classifying the single specimen images obtained by the previous automatic image acquisition methods. The input images of the existent approaches usually consist of the single-specimen images of marine phytoplankton available in several public datasets [18,19]. Thus the existent methods are focused on the particular features and characteristics of the images obtained using the automatic devices, as well as on those of the target species from marine phytoplankton. In this regard, it should be noticed that marine phytoplankton biodiversity is much lower than the fresh water one. Regardless of this, the prior approaches have explored the extraction of a wide variety of features from these images, as well as the use of several machine learning techniques [20–23]. Thus, as reference, the works in the state of the art have approached the issue with techniques like Support Vector Machines (SVM) [24–26], k Nearest Neighbour [20], Artificial Neural Networks (ANN) [27–29] or Deep Learning [21,30–33].

Specifically, among the Deep Learning methods used for this task, some approaches extract deep features using pretrained CNNs. In the work of Orenstein and Beijbom [30], the chosen CNN is AlexNet [34]. Several configura-

tions are tested, including using the pretrained network as is, fine-tunning it, or training from scratch. In the work of González et al. [33], the selected network is resNet [35], it was only tested with pretraining. In both works, the pretraining is achieved using the ImageNet dataset [36].

There are other techniques, like those based on chlorophyll fluorescence analysis, following a different approach to obtain an approximate count estimation of phytoplanktae. In this case, they cannot classify species as they only detect the amount of chlorophyll in water through reflected light. These methods are not particularly accurate as the readings can be misleading since many factors are able to affect them. Some of these factors include varying pigment cell content, cell size, cell agglomeration, water turbidity, temperature, etc. [1]. However, they prove to be useful for water potability testing despite their numerous flaws [37]. Thus, fluorescence can be used as a complement to other imaging techniques [16,38,39], making the differentiation of phytoplankton and non-phytoplankton specimens easier [40]. That is, phytoplankton contains chlorophyll, thus reflecting light while the rest of objects and artefacts do not [22].

Overall, the accurate detection and segmentation of phytoplankton specimens in images represents a challenging task. Phytoplankton presents significant intra-species morphological variance, that is, vast differences between organisms of the same species. This, coupled with the high degree of morphological overlapping between species (specimens from different species are very similar despite belonging to separate species), makes the classification task very challenging. Morphological differences between individuals may not only be in shape but also in colour or texture, depending on the many variables that can affect the different species. Thus, many of the systems that were developed

over the years perform partial and reduced analysis as each one is specialised in a subset of plankton. The changes in size and characteristics precludes, currently, from developing a generalist system that can classify all the plankton types [41]. Most of the current phytoplankton species classification systems are centred on target marine species since they are of special interest not only for the environment but also for food production as red tides can contaminate seafood. Despite its importance, freshwater phytoplankton analysis is still unexplored and many places still lack the capability to carry out such analysis.

To date, regular microscopy images are not usually considered for automatic analysis approaches due to the requirement of a biologist to manually capture them. These images, however, offer much higher quality than most of the automatic approaches. These images are also more affordable, as regular microscopes equipped with digital cameras are available in most of the laboratories. To the best of our knowledge, only some few works in the state of the art have focused on the automatic analysis of regular microscope images [40,42,43]. Planktovision [40] and PLASA [42] propose to use fluorescence imaging in combination with brightfield images of multiple focal points. Furthermore, PLASA included input images with multiple magnifications as well. In order to successfully apply this complex imaging protocol, these two methods propose to use computerised microscopes during the imaging process. In [43], instead, a manually adjusted imaging procedure is used to ensure that each specimen is imaged with the appropriate magnification and focal point. However, this requires that an experienced taxonomist operates the microscope. While this approach may be suitable for automating the tedious cell counting and specimen measurement procedures, it is not suitable for releasing the expert from the imaging, detection and identification tasks.

In contrast to these previous approaches, in this work, we propose to use a systematic imaging approach using a fixed focal point and magnification level. This ensures that the imaging can be performed by a laboratory technician, not necessarily experienced in phytoplankton identification, as the detection and identification tasks are completely delegated to the automatic system. Additionally, the proposed system aims at achieving a fully automatic detection and species classification of freshwater phytoplankton specimens using these regular microscopy images. In particular, the novel method is able to detect and, posteriorly, segment the existent phytoplankton specimens in multi-specimen images and it is able to merge sparse appearance specimens, as these are common for some freshwater phytoplankton species. Furthermore, it differentiates true phytoplankton from other similar objects like minerals, bubbles, zooplankton or detritus of other kinds. Finally, the system is capable of classifying the identified phytoplankton species into specific target species of interest.

It should be noted that, contrary to most of the state of the art works, our system targets fresh water phytoplankton. When compared to salt water, fresh water has both more biodiversity in terms of species and a higher concentration of debris, significantly complicating both the detection and classification. On the same note, state of the art approaches usually focus on single-specimen-images captured by in-flow cytometry. Instead, we use regular microscope images, with fixed focal point and magnification, in which several specimens and debris appear, without needing an expert to select the relevant images and fine-tune the acquisition. In this sense, the proposed method has been specifically designed for microscopy images taking into account the particular features and caveats of these images.

This manuscript is organised as follows: Section 2 presents the proposed methodology and details its main features. Next, Section 3 exposes the testing methodology, describing the used dataset and the different combinations of classifiers and features. Section 4 describes and discusses the results that were obtained and the main challenges of this work. Finally, Section 5 includes the conclusions about the proposed system as well as some potential future lines of work.

2 Methodology

This proposed system is divided into several steps, each one with an specific purpose, graphically illustrated in Figure 2. First, the foreground-background stage binarizes the input image into two separate classes corresponding to the foreground and the background. The objective is that foreground class includes all the potential target specimens. Next, every possible candidate region from the foreground is selected and analysed, discarding some artefacts through the use of domain-related criteria. To improve the results, a merging algorithm is proposed to fuse sparse individuals as well as colonies into single detections. This is a necessary step because several species present a sparse appearance, as the physical links between their parts are not visually evident in the microscopy images. Subsequently, the remaining detections are differentiated as true phytoplankton specimens and other objects using a learning-based approach that exploits representative complementary features. Finally, a similar classification strategy is used to identify several target species among the rest of phytoplankton.

It should be noted that every component in this pipeline has been included

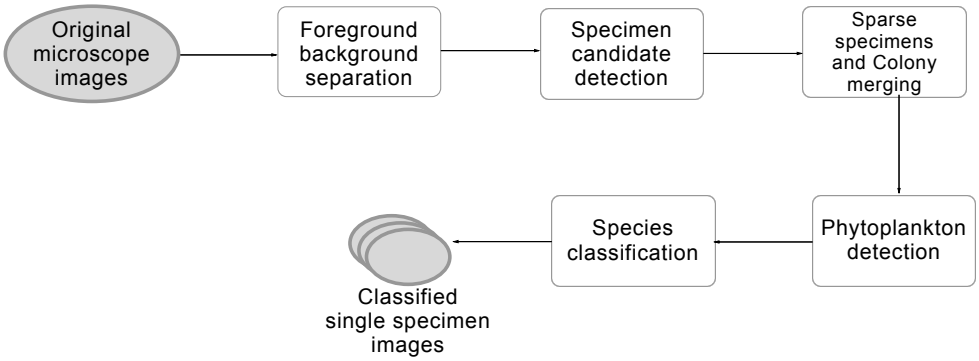
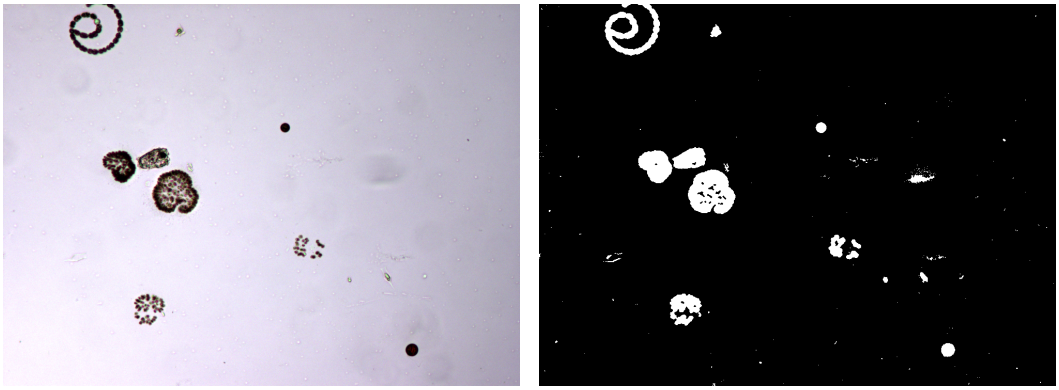


Fig. 2. Main steps of the proposed methodology.

and designed according to the intrinsic features of the microscopy images. For instance, methods in the foreground/background separation are specifically tuned for high recall rates. Another example of this is the colony merging step, also created with high recall as the objective, which addresses the common issue of sparse appearance of specimens from several species of freshwater phytoplankton. Finally, the explored learning strategies and features are intended to capture the appearance and variability, inter- and intra-class, of the target phytoplankton taxa, as well as the diversity of non target objects in the image, also with high recall and precision in mind.

2.1 Foreground-background separation

This first step takes the input microscopy images and separates the foreground from the water background. This step is designed to provide a high recall. In this sense, the objective is to retrieve all the relevant foreground, even if some spurious elements like debris or artefacts are included in it. In particular, the proposed algorithm splits the image into its three RGB channels, thresholding each one of them separately. The results of each one of the channels are later combined through an OR operator to preserve the highest amount of



(a) (b)

Fig. 3. Foreground-background separation. (a) Original image. (b) Binary foreground-background separation map.

information among the three channels.

In this case, an adaptive Gaussian threshold [44] is used as it allows to deal with the uneven illumination of the images. This algorithm computes the threshold for each image position using a Gaussian-weighted local average minus a fixed value C. The spread of the Gaussian windows was set to $\sigma = 75\mu m$, which is a larger size than all the target specimens. This way, spurious tiny irregularities are not enhanced during the thresholding. The offset is empirically set to $C = 8\%$ of the dynamic range. The output of this step, illustrated in the example of Figure 3, is a binary map which separates the target foreground from the rest of the background.

2.2 Specimen candidate detection

Every connected component from the binary foreground map is identified and analysed separately. To do so, we employ the Suzuki and Abe's algorithm [45]. This algorithm is able to detect the individual connected regions by tracing their contours. Each separated contour is considered a specimen candidate but only their external contour is taken into account, that is, their internal holes

are filled in.

Once the primary selection of the specimen candidate detections is done, the results are cleaned up, eliminating the candidates that are too small to be phytoplankton. We consider expert knowledge about the phytoplankton size to perform this filtering. Therefore, all the specimens smaller than $5\mu m^2$ are discarded. The parameters of the algorithm are derived accordingly to the magnification and resolution of the images, which is assumed to be fixed and known beforehand. The specimens that are in contact with the borders of the image are also discarded. This is based on the notion that the experts commonly ignore partial specimens during the manual analysis.

2.3 Sparse specimen and colony merging

The previous steps are able to produce accurate segmentations for most of the specimens in the analysed images. However, sparse specimens, which physical links are not visible in the microscopy images, are usually detected as separate specimens. The same happens with the colonies, which are formed by specimens of the same species, but grouped together. To address these situations, a colony merging step, that fuses nearby detections with similar colour, is proposed. An example of sparse specimen being detected as several independent candidates is show in Figure 4.

The proposed algorithm is based on the similarity of the candidates. To analyse this similarity between neighbouring detections, we first create a graph of all the detections employing a Delaunay Triangulation [46,47] of the candidate detection centroids. An example of the result of a Delaunay Triangulation can be seen in Figure 5. Next, the method measures the similarity between

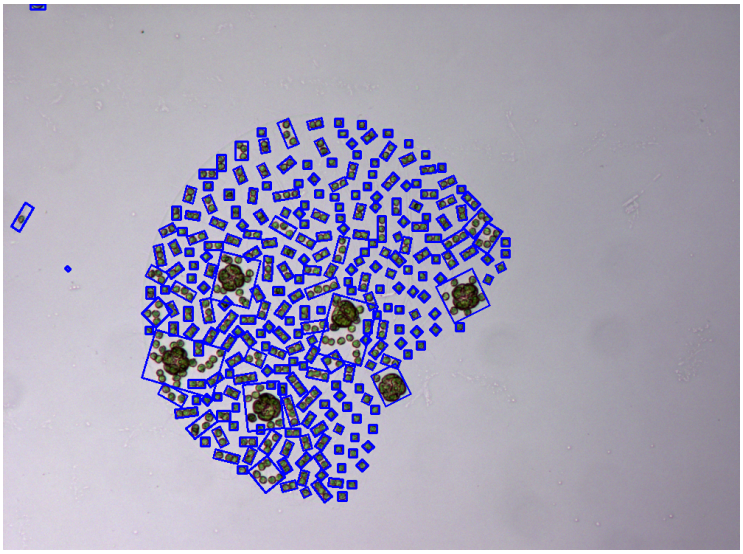


Fig. 4. Sparse specimen detected in fragments.

neighbouring detections based on their colour and distance. If the Euclidean distance between the detection or the distance between the colours of the candidates are above a certain threshold, the link is pruned. Those thresholds are empirically selected corresponding to 105 μm in distance and 15% of dynamic range per RGB channel. The colour values for the detections are computed using the average of the RGB pixels inside each candidate segmentation mask. This prevents the colour distribution from being skewed by the background. Once all the links have been explored, the nodes that remain connected are fused into single candidates. Figure 6 shows the final detection result of the fusion step over the previous presented example of Figures 4 and 5.

2.4 Classification stages

The two last steps of the methodology share the objective of separating the previously identified specimens in different groups. Firstly, the phytoplankton detection stage separates true phytoplankton specimens from the non-phytoplankton elements that are present in the images. These non-phytoplankton

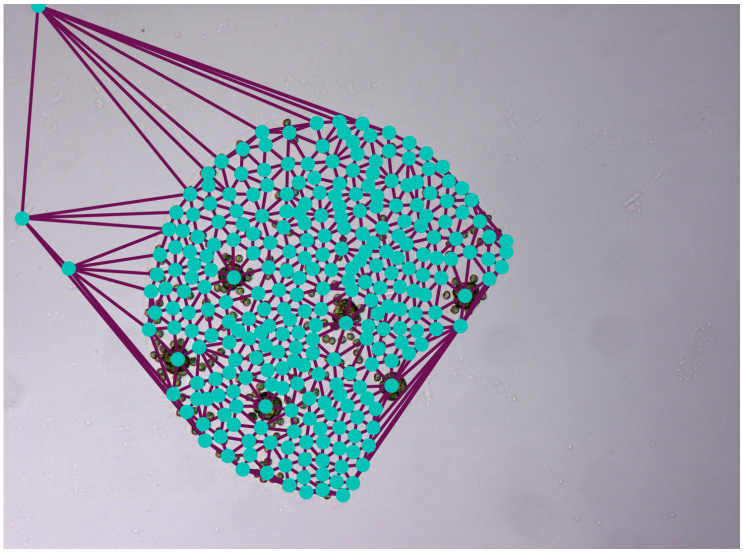


Fig. 5. Example of a Delaunay Triangulation linking detections.

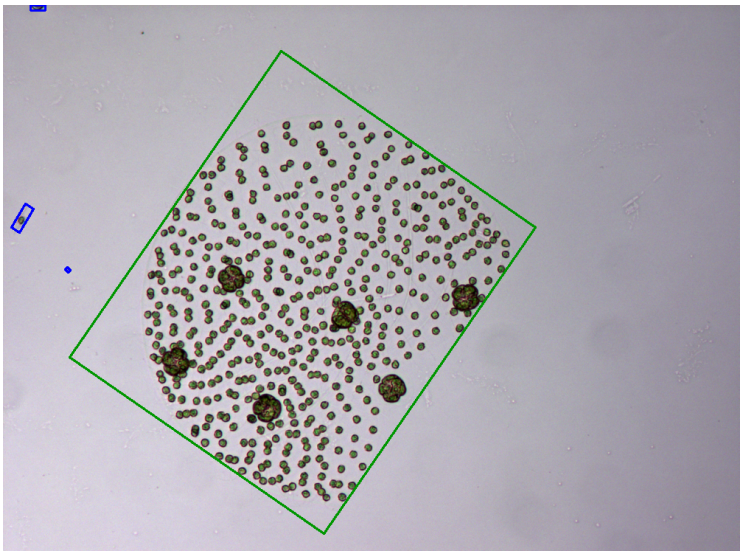
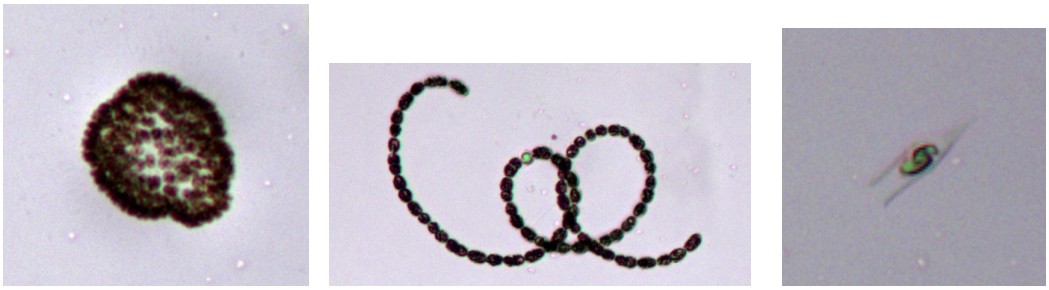


Fig. 6. Result of the fusion step over sparse specimen from Figures 4 and 5.

elements include, but are not limited to, zooplankton, garbage and minerals, which can be, in many cases, very similar to phytoplankton. Next, from the filtered true phytoplankton specimens, the target species classification stage groups the candidates belonging to the species of interest.

Given the huge amount of potential species to be identified, in this work, we targeted a set of relevant and representative ones. In particular, we selected a couple of frequent species that produce toxins and a harmless one. The



(a) (b) (c)

Fig. 7. Examples of relevant species that are classified separately. Images not to scale. (a) Woronichinia naegeliana. (b) Anabaena spiroides. (c) Dinobryon sp.

selected species subset is representative of the complexity of phytoplankton species identification. The chosen species, with examples depicted in Figure 7, are:

- ***Woronichinia naegeliana***: The most representative species that produces toxins, which is very frequent in the analysed samples. It represents the majority of the biomass of toxin-producing cyanobacteria.
- ***Anabaena spiroides***: Another toxic species, characterised by their elongated shape and differences in size. It is also numerous in the water samples and has several challenging image features to analyse.
- ***Dinobryon sociale***: This is a harmless species, that is, it does not produce toxins. However, due to its characteristics like very wide morphological differences between specimens as well as the fact that it can possess a transparent capsule make this species ideal to benchmark the system.

Both of these steps share the same characteristics as the phytoplankton can be distinguished by its texture, colour or shape. Therefore, we designed both steps taking it into account. In order to asses how each image feature is able to represent the different groupings and impact the overall classification we designed both steps using the same features. In particular, we use texture and colour features using a Bag of Visual Words (BoVW) model [48–50] as well

as deep features from a pre-trained deep convolutional neural network [51]. These features exploit the most representative differences among the different phytoplankton species as well as the spurious objects.

The features are extracted from the image positions obtained in the previous steps. The bounding boxes are, however, enlarged by 10% in each direction to enclose more contextual information. It should also be noted that, for each feature, a masked and unmasked version are tested. The unmasked version includes all the information within the bounding box whereas the masked version uses the segmentation map to extract only the information belonging to the specimen, discarding the background. While the masked approach may facilitate the work of a classifier by only providing the internal specimen appearance information, it avoids some potentially relevant information derived from the external context (like shape information), therefore testing is needed to see which hypothesis is correct.

The previously described features are combined under different configurations and settings to measure their suitability and impact in the classification process of both designed steps. Each feature is tested alone and, then, each classical feature is tested in conjunction with deep features separately. Finally the three separate features are tested together. For this two different methods will be tested, using a single bag of visual words for each, colour and texture descriptors, and using separate bags. Separate bags of words will result in more features and therefore more information. On the other hand, using a single bag would create a feature vector containing information about both characteristics which could prove more information-rich than each one separately, as they are complementary.

These different combinations are tested for both masked and unmasked fea-

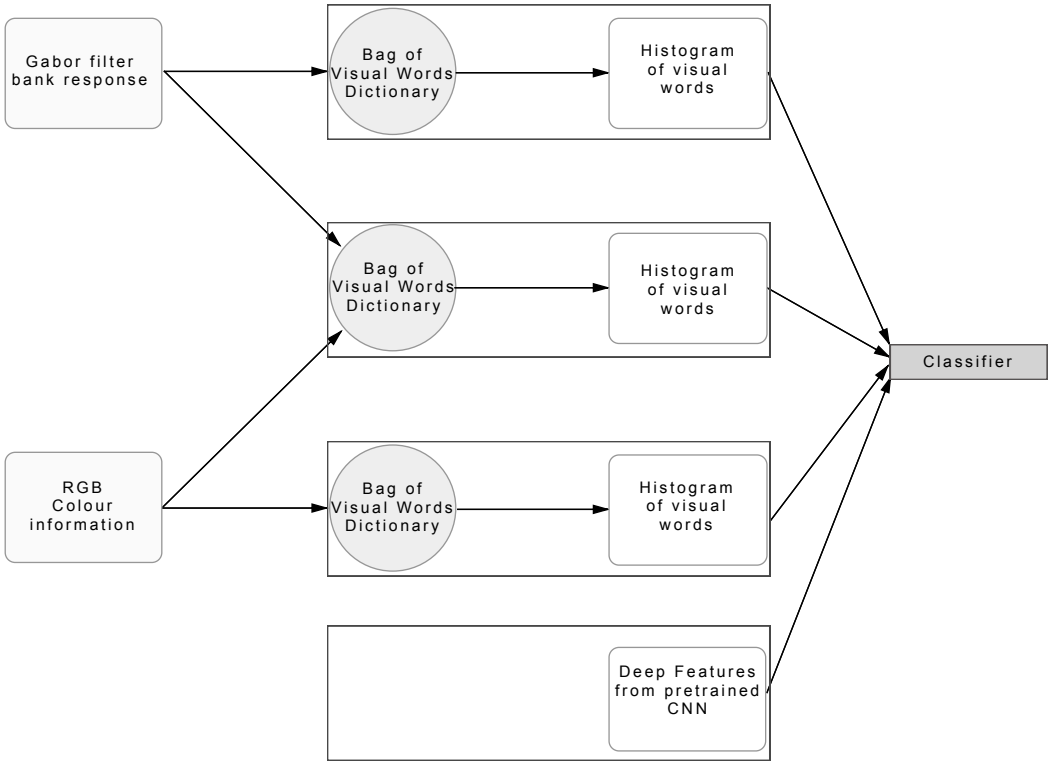


Fig. 8. Pipeline used for the classification steps showing the different possible combination of features.

tures. It should also be noted that, several free parameters are left to be selected empirically during the experimentation, especially regarding the Gabor filters used in texture descriptors. A visual representation of the novel proposed pipeline can be seen in Figure 8, representing how the different features are specifically combined.

In terms of classifiers, we test each feature combination with Random Forest (RF) and Support Vector Machine (SVM) classifiers.

2.4.1 Colour Features

The colour features are obtained using a BoVW model. This model is able to represent the contents of an image region through histograms of visual patterns, called words. These visual words are local patterns defined by specific

combinations of reference visual features. The specific combinations that define the visual words are learned using a clustering algorithm, k-means [52] in this work, that are later stored into a dictionary. This dictionary is used as reference to compute the histograms.

The colour features are obtained from the RGB values of each image position. The RGB values from all the training set boxes are gathered together and clustered, in this case with $k-$means, using a specific number of cluster centres, K_c. The number of centres is left as a free parameter during our experimentation to be empirically selected. The resulting histogram over the visual words in the dictionary would represent, in this case, the distribution of colours within the bounding box. It is important to mention that the colours can also include the background, therefore some implicit shape features should also be present as the histograms represent the foreground-background proportions of each box.

2.4.2 Texture Features

The base texture features are extracted using the complex responses to a Gabor filter [53,54] bank for each image position in combination with the BoVW model described above. The visual words produced by the BoVW model would represent in this case local shape patterns. The histograms of the patterns for the whole image regions represent texture in terms of the distribution of local shape. The used filter bank consists of a series of Gabor filters of a single scale (central frequency) and bandwidth with a varying number of orientations. The real and imaginary responses of these filters are taken into account. The complex response for each filter, given its orientation θ, is

$$g(x, y; \theta) = \frac{f_c^2}{\pi\sigma^2} e^{f_c(\frac{x'^2 y'^2}{\sigma^2})} e^{-j2\pi f_c x'}, \tag{1}$$

where f_c denotes the central frequency of the filters, j denotes the imaginary unit and σ controls the spread of the isotropic envelope. The x' and y' coordinates are a rotated space according to

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta, \\ y' &= -x\sin\theta + y\cos\theta, \end{aligned} \tag{2}$$

where θ denotes the rotation angle. The σ value is computed depending on the central frequency f_c according to

$$\sigma = \frac{1}{\pi f_c} \cdot \sqrt{\frac{\ln 2}{2}} \cdot \frac{2^B + 1}{2^B - 1} \tag{3}$$

where B denotes the filters bandwidth in octaves.

The bank of filters is composed of N_o filters with $\theta \in \{\theta, \pi/N_o, \ldots, \pi(N_o - 1)/N_o)\}$, corresponding to N_o evenly distributed orientations. This results in a total of $2 \cdot N_o$ responses for each image position, as we consider both real and imaginary responses for the filters.

The responses to the filter banks are processed, like the colour features, by the BoVW model. It should be noted that the image is converted to grayscale for this procedure and the Gabor filer responses are computed over the whole image rather than by individual bounding boxes to minimise potential aliasing effects due to the bounding box boundaries.

The central frequency f_c, the frequency bandwidth B, the number of orientations N_o and the number of k-means clusters k_t are left as free parameters and their values are selected empirically during the experiments.

2.4.3 Deep Features

To obtain Deep Features, CNNs are employed bypassing their classification layers. These features can then be used with some classifier to produce the desired classification. The main advantage of this approach is the use of pretrained networks, which are widely available. These networks have already learned general features without the need to expend a long time training them. In this work, we use a resNet [35], specifically the resNet50 version, in order to extract the features from each image position. The particular network that we used was pretrained in ImageNet [36] and was not finetuned any further. This pretraining should allow the network to obtain relevant features from our set of images despite having never seen it. In this work, we take the deep features from the global average pooling layer, just before the classification layers of the network. This results in a vector with 2048 components for each specimen image.

2.4.4 Classifiers

Random Forest is a suitable classifier for this task since it performs well when there are many features and it is also able to develop complex classification boundaries. This method combines multiple decision trees in a process called bagging. This process creates trees independently and combines their classification through different methods like voting. Each tree only receives a randomly selected subset of features. The combination of these trees allows to

create a strong classifier from many weaker ones.

Support Vectors Machines separate the data arranged in an n-dimensional space (each dimension is component of the feature vector) by using hyperplanes that maximise the classification margin. This method is suitable for our classification since varying the different parameters of each kernel allows to obtain complex classification boundaries which may be needed for this task. SVMs perform well for tasks involving big dimensional spaces although they can also suffer the curse of dimensionality [55].

3 Experimental Setup

3.1 Dataset

The proposed system is trained, validated and tested using 293 multi-specimen microscopy images captured from several water samples belonging to different periods of time. This is done with the intention of capturing a representative amount of species as their concentration varies influenced by many factors like seasonality.

The images produced by the digital microscope were captured at $10\times$ of magnification, which provided a resolution of approximately 0.67 μm per pixel.

Representative examples of images can be seen in Figure 1. The used dataset consists in 293 images that contain 1611 phytoplankton specimens from a large variety of species. The images also contain a significant amount of spurious objects like zooplankton, minerals or garbage.

For the first steps, detection and merging of colonies, we selected random

subset of 50 images to act as the training set. The rest are left in the test set to evaluate the algorithms.

In the classification steps, the multi-specimen images were randomly split 80-20, with 80% of the images being dedicated to the training and validation using a 10-fold crossvalidation with grid search to obtain the best parameters with which the system is then retrained. The remaining 20% of the images was used as the test set.

The ground truth consists of bounding boxes that enclose the phytoplankton specimens, with an associated label identifying the species. An experienced taxonomist manually marked and identified all the target specimens in the image dataset. The expert consulted other peers in difficult or challenging cases in order to ensure high quality tagging.

During the species classification step, the chosen set of target species are classified against a general group, "others", which encompasses the rest of species that also appear in the dataset but are not specifically considered for classification. Therefore, the resulting classifier has four different classes in which to group the images. Overall, the dataset employed consists of 175 specimens of W. naegeliana, 41 of A. spiroides and 202 of D. sociale. The rest, 425 specimens, constitute the species grouped under the Other category.

3.2 Evaluation metrics

3.2.1 Specimen detection and merging

In order to evaluate the performance of the first steps, including specimen detection and colony merging we need to observe the overlapping between the

bounding boxes marked by the expert and the ones detected by our system. As gold standard, to count as positives, the detected bounding boxes should enclose at least 50% of the specimens' area. The metric used to evaluate the specimen detection step is False Negative Rate (FNR) as the most important goal of this step is to capture as many specimens as possible.

The sparse specimen and colony merging step is evaluated by comparing the amounts of over and under-segmentation before and after its application.

3.2.2 Phytoplankton identification

The phytoplankton identification is evaluated with precision at 90% or 95% of recall. This metric is chosen because our intention is to preserve as many true specimens as possible even if that means including some false positives. Precision Recall (PR) curves are also used to illustrate the performance of this step. PR curves are used instead of Receiver Operating Characteristic (ROC) curves due to the class imbalance and the specific relevance of the positive class in the system's goals.

3.2.3 Target species classification

This step is proposed as a multi-class classification problem. The main metrics to evaluate the performance of this step are the overall classification accuracy and the F1-Score. Additionally, precision, recall, F1-Score and accuracy are computed for every class. These metrics allow to quantify how the system performs individually for each species and, therefore, measure the suitability for each one of them separately.

Parameter	Values
# of estimators	10, 50, 100, 500, 1000
Split Criterion	gini, entropy
min. samples per split	2, 5, 8, 10
min. samples per leaf	1, 2, 4
max_features	# of input features

Table 1
Tested hyperparameters during Grid Search for RF.

Parameter	Values
C	0.001, 0.01, 0.1, 1, 5, 10
Kernel	linear, polynomial, rbf
Gamma (only polynomial)	0.1, 0.01, 0.001, 0.0001

Table 2
Tested hyperparameters during Grid Search for SVM.

3.3 Classifier parameters

The SVM and Random Forest hyperparameters are optimised using a grid search each of the classification cases.

For the Random Forest Grid search the candidate parameters to be tested and chosen by the system are represented in Table 1.

For the SVM Grid search the candidate parameters to be tested and chosen by the system are represented in Table 2.

3.4 Feature parameters

Both the Gabor filter bank and colour features are subject to the tuning of several parameters. The best set of parameters for these features will be selected employing an exhaustive grid search over all the possibilities using the

Parameter	Values
Bandwidth	2, 1.5 ,1 , 0.5
Frequency	0.5, 0.3535, 0.25, 0.177, 0.125, 0.088, 0.0625, 0.0442, 0.03125, 0.0221, 0.0156, 0.01105
Orientations	4, 8

Table 3
Tested parameters for the Gabor filter bank.

phytoplankton identification set, that is, the one where true phytoplankton specimens are separated from spurious objects.

The parameter for the set of Gabor filter banks to be tested is represented in Table 3.

For the bag of visual words, the free parameter to be selected during experimentation is the number of bins of the histograms and the number of words to be clustered using k-means. Both of these are made to coincide so, overall, the BoVW model only has a single free parameter. The values tested for this parameter are: 100, 50, 20, 10, 8, 5, 3, 2.

The best parameters obtained for these features in the phytoplankton identification task are used in the species classification step. This means that the hand-crafted features will be shared between tests as only a single set of parameters will be used in both classification steps.

The best results for each of these two features can then be combined with Deep Features to asses their impact on the robust features from a CNN. This would provide us with results on whether hand-crafted classical features can improve the results of a CNN's deep features.

4 Results and Discussion

4.1 Specimen detection

The proposed method produced a FNR of 0.4% in the test set. This corresponds to just 5 missed specimens among the 1249 that are included in the test set for this step. The low FNR meant that several false positives are included, a total of 1136, which would be later discarded. Considering the high degree of complexity, we can consider the phytoplankton detection to be accurate despite the false positive rates that it detects. This is because the system was designed at this stage to provide the highest possible phytoplankton recall, allowing the following steps to reduce the false positives.

4.2 Colony merging

Colony merging is evaluated comparing over and undersegmentation metrics before and after the merging step. Before the colony merging algorithm, 21.93% of candidates are oversegmented, mainly due to the presence of sparse specimens. On the other hand, 8.57% of candidates are undersegmented, due to the overlapping of specimens. After the merging stage, the oversegmentation was reduced to a 3.76% while the undersegmentation increased to a 17.77%. Therefore, the merging process presents an adequate impact in the ratios as it improves the oversegmentation reducing it by a 18.17%, while only increasing the undersegmentation by a 9.2%.

The best set of parameters for the texture features using Random Forest are 100 bins for the BoVW model and a bandwidth of 1.5 octaves, a frequency of 0.25 px^{-1} and 8 orientations for the Gabor filter bank. The best result for these features is a 77.35% of precision at 90% of recall, which was obtained using masked features. Using all the information, that is, unmasked the result is 75.34%. The best parameters for colour features using RF are 100 bins for the BoVW model, which results in a top precision of 77.86% at 90% of recall with, again, masked features. The unmasked version reported a significantly lower precision, 70.32% at 90% of recall.

The best set of parameters for texture features using SVM as the classifier are 3 bins of the BoVW and, for the Gabor filter bank a bandwidth of 1.5, a frequency of 0.0312 px^{-1} and 4 orientations. This combination of parameters for the texture features, in the case of SVM, report a precision of 69.89% at 90% of recall while unmasked. For this combination, the masked version resulted in 67.46% of precision at 90% of recall. In terms of the colour features for the SVM, the number of bins for the BoVW is 10, which results in top performance 68.25% of precision at 90% of recall, in this case masking the features. Using the same feature parameters for colour, the unmasked version reports a precision of 62.38% at 90% of recall.

These results illustrate the differences in the amount of features that each classifier is able to extract the most information from. Random Forest is most robust when using many features, in this case, the maximum for both colour and texture as they both use 100 bins. On the other hand, SVM tends to use lower amounts of features, in this case 3 bins in texture and 10 in colour, both

	RF		SVM	
	Prec. 90% Recall	Prec. 95% Recall	Prec. 90% Recall	Prec. 95% Recall
Gabor Filters	75.34%	65.34%	69.89%	63.65%
Colour Features	70.32%	62.99%	62.38%	46.32%
Deep Features	73.43%	67.13%	**84.07%**	**71.02%**
Gabor + DF	77.86%	75.58%	**84.07%**	**71.02%**
Colour + DF	81.13%	77.78%	**84.07%**	**71.02%**
Gabor + Colour +DF	**81.70%**	**78.78%**	**84.07%**	**71.02%**
Gabor+Colour(same bag)+DF	77.86%	73.58%	**84.07%**	**71.02%**

Table 4
Comparison of the different unmasked feature combinations for both classifiers. Best results are highlighted in bold.

	RF		SVM	
	Prec. 90% Recall	Prec. 95% Recall	Prec. 90% Recall	Prec. 95% Recall
Gabor Filters	77.35%	72.28%	67.46%	59.85%
Colour Features	77.86%	62.03%	68.25%	61.11%
Deep Features	67.46%	57.38%	69.06%	54.82%
Gabor + DF	78.92%	71.02%	70.75%	61.72%
Colour + DF	**81.13%**	**77.78%**	**71.18%**	**61.72%**
Gabor + Colour +DF	78.92%	67.87%	**71.18%**	**61.72%**
Gabor+Colour(same bag)+DF	78.93%	75.86%	70.32%	61.72%

Table 5
Comparison of the different masked feature combinations for both classifiers. Best results are highlighted in bold.

notably smaller than the ones used in RF. A similar point can be made about Gabor filter orientations as the SVM prefer just 4 when compared to RF that performs better with 8.

We can see that, in these results, most of the overall best results correspond to masked features which reveal the importance of giving the classifier only the relevant information, not skewed by the background. While masking the features can limit some information, it proved to be beneficial. Except texture only in the SVM classifier, the rest of the top results use masked features.

The results of the identification of phytoplankton, combining Deep Features with BoVW features are illustrated in the Tables 4 and 5, complementarily illustrated in Figures 9 and 10. A comparison between the PR curves for the best results with each classifier is shown in Figure 11. Finally, representative examples of this step are shown in Figure 12. These examples were obtained from the test set and were classified using the best model obtained.

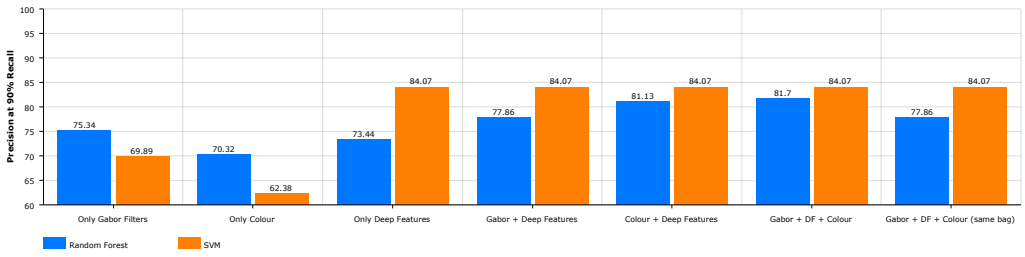


Fig. 9. Comparison between different classifiers and combinations of unmasked features in phytoplankton identification.

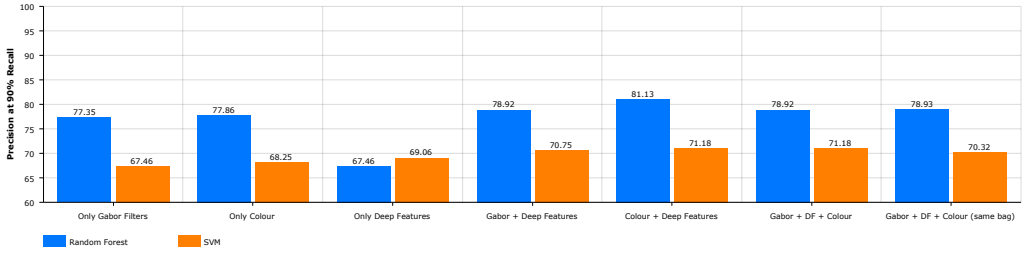


Fig. 10. Comparison between different classifiers and combinations of masked features in phytoplankton identification.

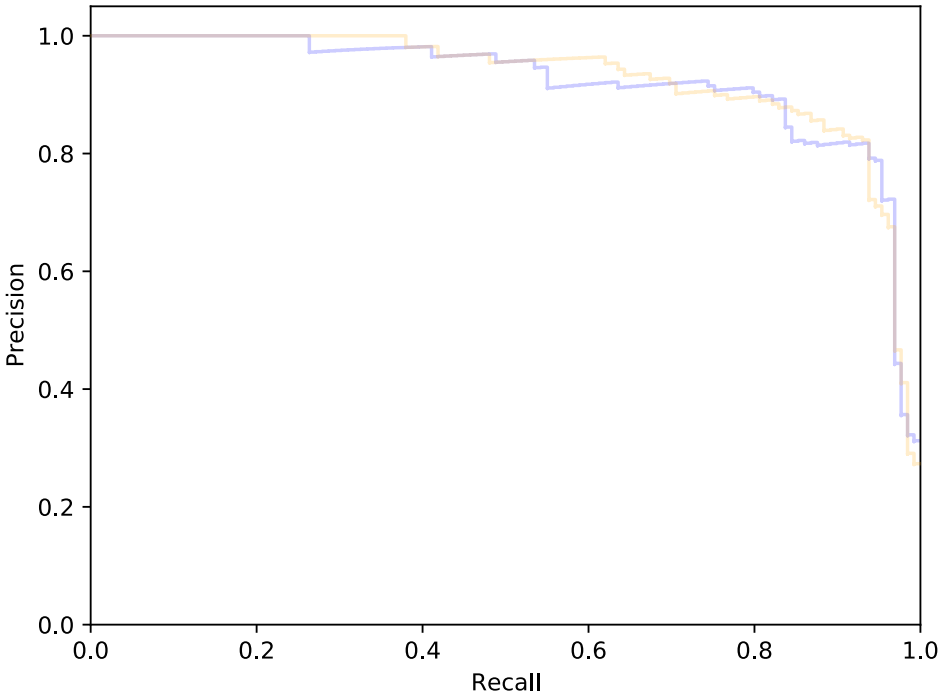
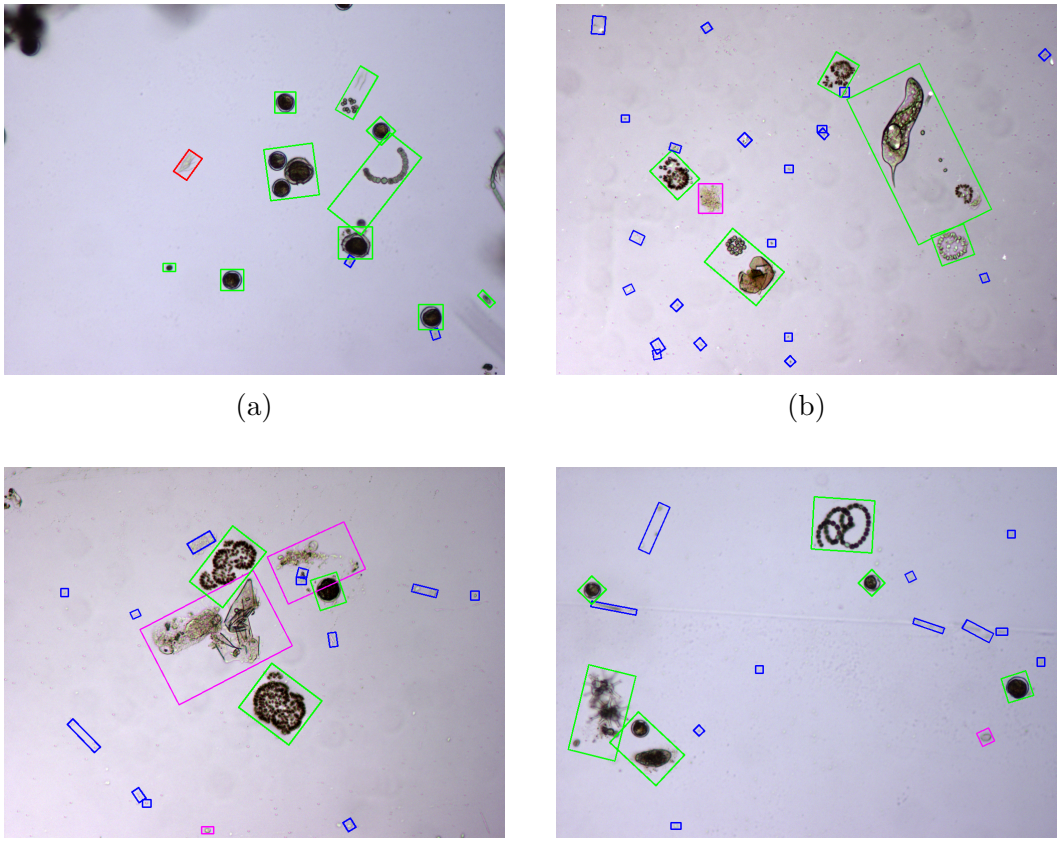


Fig. 11. Comparison between the best results for RF (blue) and SVM (orange).

We can see in the provided results that the classification separating phytoplankton and spurious objects is successful, despite the complexity of the step



(a) (b)

(c) (d)

Fig. 12. Representative examples of phytoplankton identification extracted from the test set. True positives in green, true negatives in blue, false negatives in red and false positives in magenta.

given the large variability of the large amount of species as well as the heterogeneity of the negative class (zooplankton, garbage, minerals, etc.). The top result using precision at 90% of recall is a SVM that makes use only of Deep features, resulting in a precision of 84.07%. It should be noted that, in the case of the SVM, neither Gabor nor colour features increased the precision any further as the top results stays the same. However, Random Forest did improve when using both Deep features and classical features, although it did not reach the performance of the SVM. This is likely due to the low number of BoVW features when compared to the deep features. While using only colour features or texture features SVM favoured lower amount of bins than RF. This means that, when combined with the Deep Features, just 3 components

	RF		SVM	
	Prec. 90% Recall	Prec. 95% Recall	Prec. 90% Recall	Prec. 95% Recall
Gabor+Colour (masked) and DF (unmasked)	81.13%	77.29%	84.07%	77.02%
Gabor+Colour (same bag and masked) and DF (unmasked)	79.46%	74.75%	84.07%	77.02%

Table 6
Combination of best performing features, mixing masked and unmasked.

for texture and/or 10 for colour belong to BoVW features as opposed to 2048 from deep features. This results in these features becoming mostly insignificant with the SVM classifier due to their low amount when compared with the deep features.

Considering precision at 95% of recall we obtain a similar high precision, 78.78%, this time using a Random Forest classifier which uses Gabor filter banks, Deep Features and colour information features. In this case, the texture and colour features are separated into two bags.

The results also show that masking the deep features reduces the performance of the system, dropping around a 7% precision for RF and a 15% for SVM. On the contrary, the BoVW features, often times, perform better being masked. This motivated some extra tests in which the classical features are masked while the deep features are left unmasked. They would take the highest performing configuration of each and join them. These results can be seen in Table 6. However, the results do not improve on the previously presented ones, although they come very close in the case of RF and tie in the case of SVM due to previously explained phenomenon about the irrelevance of hand-crafted features.

Similarly to the feature masking, joining texture and colour features in the same BoVW does not improve, most of the times, the results for this phytoplankton identification step. The only instance where this happens is using

RF and masked features, in this case the main difference is at 95% of recall where both features joined obtain around 7% higher precision. In the rest of the cases the results were equal or worse combining features.

Overall, we observed that this step is successful in the issue of eliminating spurious elements from the set of extracted specimens. We have to consider that the phytoplankton specimens represent a heterogeneous set, complicating its agglomeration as a single group, and, in many cases, they can be very similar to the spurious objects. Nevertheless, despite the complexity, this step showed accurate results even at high levels of recall, which implies few false negatives, which in turn is the objective of the step.

4.4 Target species classification

The results for the classification of target species are represented in Table 7. The comparison of results between features can be observed in the graphs provided in Figures 13 and 14. The overall best results for the classification of the target species step can be observed in Table 8 and the complete confusion matrix is shown in Figure 15.

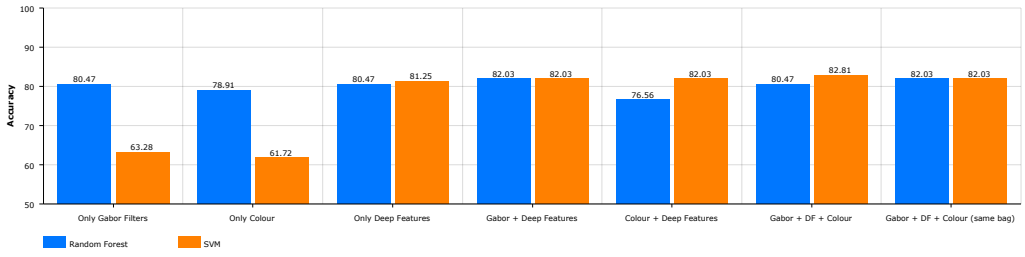


Fig. 13. Comparison between different classifiers and combinations of unmasked features in the phytoplankton classification.

In the results for the species classification, we obtain an overall accuracy of 87.50% for all the 4 classes while using a RF classifier with colour and deep

		RF		SVM	
		ACC	F1-Score	ACC	F1-score
Unmasked	Only Gabor	80.47%	80.84%	63.28%	61.59%
	Only Colour	78.91%	79.73%	61.72%	60.44%
	Only Deep Features	80.47%	80.71%	81.25%	81.36%
	Gabor + DF	82.03%	82.47%	82.03%	82.13%
	Colour + DF	76.56%	77.07%	82.03%	82.15%
	Gabor + colour + DF	80.47%	80.84%	82.81%	82.91%
	Gabor/colour S.B + DF	82.03%	82.48%	82.03%	82.13%
Masked	Only Gabor	78.12%	78.12%	59.38%	57.74%
	Only Colour	79.69%	80.00%	60.16%	58.57%
	Only Deep Features	83.59%	80.00%	**85.94%**	**86.15%**
	Gabor + DF	76.56%	77.34%	82.03%	82.50%
	Colour + DF	**87.50%**	**87.99%**	84.38%	84.75%
	Gabor + colour + DF	78.91%	79.80%	82.03%	82.50%
	Gabor/colour S.B + DF	78.91%	79.80%	82.03%	82.50%

Table 7
Results for the tests in the target phytoplankton species classification expressed in overall classification accuracy and F1-Score (weighted). S.B. means same bag. Best result highlighted in bold.

	Accuracy	Precision	Recall	F1-Score
W. naegeliana	94.53%	96.43%	81.82%	88.53%
A. spiroides	97.66%	100.0%	57.15%	72.73%
D. sociale	94.53%	88.89%	85.71%	87.27%
Others	88.28%	82.61%	95.00%	88.37%

Table 8
Results of the target species classification step.

features. Adding texture information significantly worsened the result, in spite of the BoVW configuration. That is, using a single BoVW or two separate ones did not influence the results positively as both configurations worsen then results. The top performance with the SVM is obtained with just deep

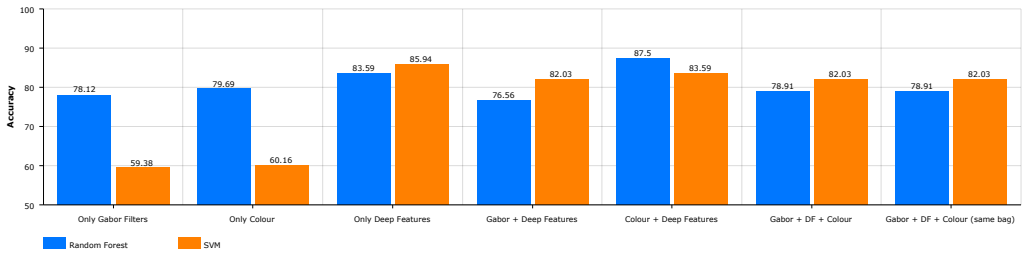


Fig. 14. Comparison between different classifiers and combinations of masked features in the phytoplankton classification.

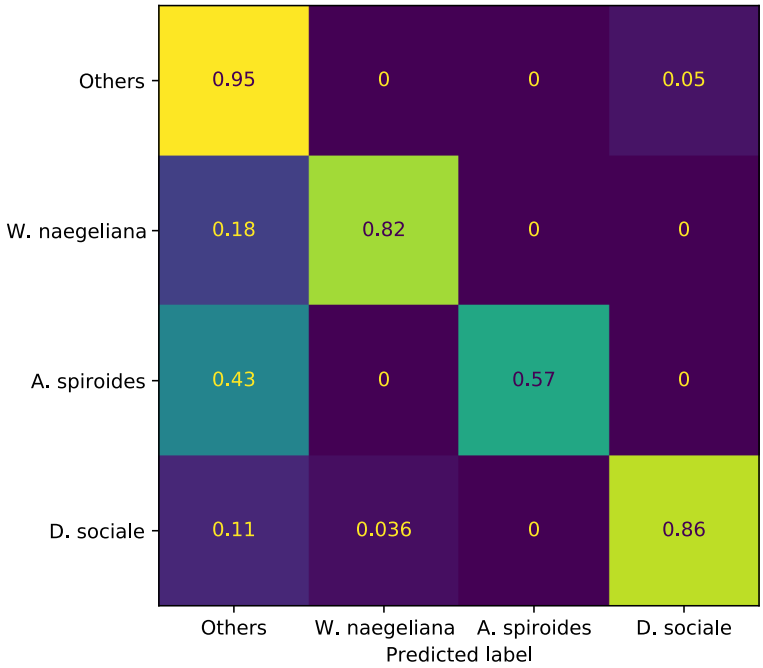


Fig. 15. Confusion matrix corresponding to the best results obtained in the target classification step.

features which results in 85.94% of overall classification accuracy.

In this stage, the results of masking the features are similar to the ones obtained in the phytoplankton detection step. While in that step both classical features did benefit from using the masks, in this case the results are not equal. Gabor filters perform better without masking and the same goes for colour with the SVM. The only instance of the opposite is colour with RF which, masked, improves around 1%. On the other hand, the situation of deep features

is reversed, while previously they always performed better being unmasked, now the better results are obtained by using the segmentation masks.

Like in the previous step, we obtain mixed results in terms of mixing colour and texture features in the same bag or using two separate ones. While with masked features both tie, using unmasked features RF performs better using a single BoVW for both types of data while SVM obtains better results with two separate bags.

We can see that W. naegeliana provided the best results among the two dangerous species. A. spiroides, on the contrary, shows the lowest recall metric. This may be due to the fact that while Woronichinias generally present constant shape and colours in the different specimens, while A. spiroides shows large variations in size, shape and colour. However, despite the larger intraclass variation of A. spiroides there is a lower number of specimens in the training dataset.

It is important to note the accurate metrics obtained by D. sociale (85.71% of recall and 88.89% of precision) despite that they also present a large intraclass variation. This species can appear alone or in groups, with or without transparent capsule, and its nucleus despite often being oval, can take many shapes. However, D. sociale is the single most abundant species in the dataset. The results for this species illustrate the generalisation capabilities of the system, being able to recognise diverse appearances of the target species if sufficiently represented in the training dataset.

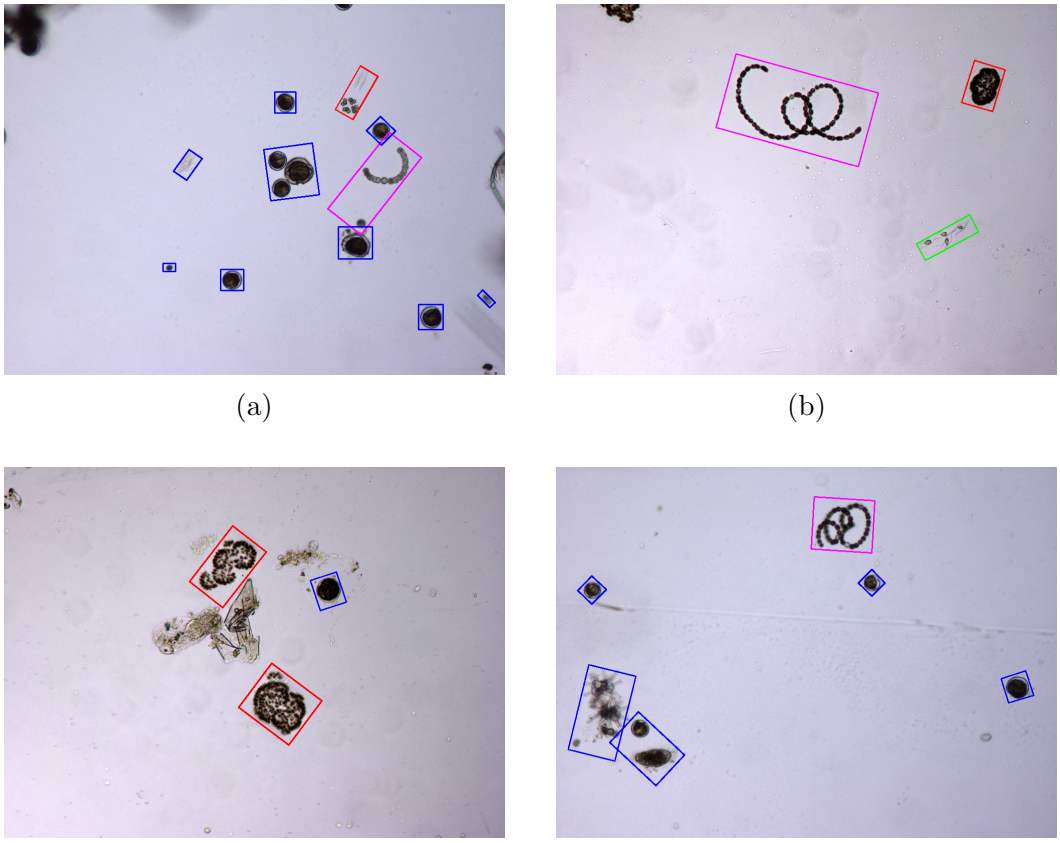
Overall, the performance for the different species is adequate given the high complexity of the problem and the variety of characteristics (size, shapes, etc.) that the specimens typically present. This is specially relevant if we take into

account the amount of work and experience a taxonomist has to endure in order to obtain comparable results. Some representative examples of the final classification stage for several images in the test set can be seen in Figure 16.

Finally, we note that one possible limitation of the presented study is that it does not include a comparison with prior detection and species identification computational approaches. This, on one hand, is due to the fact that previous works do not use an imaging protocol that is similar to ours, therefore making a fair comparison impossible. As reference, regarding the imaging conditions, Baek et al. [43] use varying magnification levels and focal points depending on the imaged species and on a specimen basis. This, along with not releasing the experienced experts from operating the microscope, simplifies the challenges to be faced by the automatic methods. Instead, Planktovision [40] and PLASA [42] use input images of multiple focal points, as well as multiple magnification levels in the case of PLASA, along with fluorescence imaging as complementary input to simplify the phytoplankton detection methods. On the other hand, each work considers different target species, which appearance may further vary depending on the climatic, seasonal and stage of development conditions. Therefore, even in the case of similar imaging conditions, the direct comparison of results from different works that do not share the exact same dataset may result in misleading conclusions. Thus, due to the aforementioned reasons, we could not include comparison results.

5 Conclusions

In this work, we propose a novel fully automatic methodology for the detection and taxonomic identification of phytoplankton specimens in regular mi-



(a) (b)

(c) (d)

Fig. 16. Representative examples of phytoplankton classification into target species extracted from the test set. Specimens classified as W. naegeliana in red, A. spiroides in magenta, D. sociale in green and Others in blue.

croscopy images. The main aim is to take advantage of the automatic image analysis to improve the repeatability and accuracy of freshwater phytoplankton studies, and releasing the experts from tedious and repetitive work, while using standard laboratory equipment. To that end, the work herein described approaches the image acquisition with regular microscopes, using a systematic protocol that fixes focal point and magnification for the whole analysis. This, in contrast to prior approaches, can also release the expert taxonomists from the image acquisition process, but imposes additional challenges to the computational analysis of the obtained images.

The proposed computational analysis methodology is a complete processing

pipeline. It takes as input multi-specimen microscope images, then detects, segments and identifies the specimens that are present on said images. Thus, this work presents a fully automatic system that processes microscope images, detecting the various phytoplankton specimens contained in them and classifying the target species.

The proposed methods include a novel algorithm, proposed to improve the detections of sparse specimens and colonies, often detected in parts due to the absence of visually evident links. Also, the classification stages use flexible feature extraction techniques, based on Bag of Visual Words and deep features which allowed to be adapted to the various objectives of the pipeline; i.e. phytoplankton detection and identification.

The experimental results show that the proposed system offers an accurate and satisfactory performance in the test set for the three main steps which are segmentation, phytoplankton detection and species classification. From the total of 1249 phytoplankton specimens in the test set, only 5 are missed. These detections presented a characteristic oversegmentation due to the appearance of the sparse specimens. The proposed merging algorithm allowed to successfully address this issue. Phytoplankton detection, that is, separating true phytoplankton specimens, yields a performance of 84.07% precision at 90% of recall. Finally, the target species classification results obtains an overall accuracy of 87.50%. The levels of recall for each class or species are of 81.82% for W. naegeliana, 57.15% for A. spiroides, 85.71% for D. sociale and 95% for the Other species. These results were achieved thanks to an exhaustive comparison of several automated feature extraction procedures, which could be easily adjusted for future goals.

The results prove to be satisfactory in every step, demonstrating good perfor-

mance despite the inherent difficulties of the process like similarities between species, high amounts of phytoplankton-like garbage or overlapping specimens. Overall, we can conclude that the system is able to successfully perform its task.

Future work includes the extension of the methodology with more sophisticated procedures, like deep learning methods. The current study using classical machine learning methods allowed to identify the specific requirements for the detection stage in order to achieve a high precision at high recall, along with the evidence that the high variability of phytoplankton and non-target object appearances can take advantage of advanced texture representations and deep features in the classification stages. Thus, the application of deep learning methods represent the logical next step for this work. Deep learning has had success in similar biomedical domains, therefore it should allow to improve the current methodology. Moreover, it is planned to extend the study to more target species (even considering marine phytoplankton), and to include a larger number of water samples labelled by multiple experts to achieve more reliable performance estimations. This would allow to further improve the system and extend its ability to detect a wider variety of dangerous species of phytoplankton with high reliability.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

This work is supported by the European Regional Development Fund (ERDF) of the European Union and Xunta de Galicia through Centro de Investigación del Sistema Universitario de Galicia, ref. ED431G 2019/01.

References

[1] A. Zamyadi, F. Choo, G. Newcombe, R. Stuetz, R. K. Henderson, A review of monitoring technologies for real-time management of cyanobacteria: Recent advances and future direction, TrAC Trends in Analytical Chemistry 85 (2016) 83–96.

[2] H. W. Paerl, V. J. Paul, Climate change: Links to global expansion of harmful cyanobacteria, Water Research 46 (5) (2012) 1349–1363.

[3] M. R. First, L. A. Drake, Performance of the human "counting machine": Evaluation of manual microscopy for enumerating plankton, Journal of Plankton Research 34 (12) (2012) 1028–1041.

[4] M. G. Kelly, M. M. Bayer, J. Hürlimann, R. J. Telford, Human error and quality assurance in diatom analysis, in: H. du Buf, M. M. Bayer (Eds.), Automatic Diatom Identification, Vol. 51 of Series in Machine Perception and Artificial Intelligence, World Scientific, 2002, pp. 75–91.

[5] K. Vuorio, L. Lepistoe, A.-L. Holopainen, Intercalibrations of freshwater phytoplankton analyses, Boreal Environment Research 12 (2007) 561–569.

[6] P. Culverhouse, W. R, B. Reguera, H. V, S. González-Gil, Do experts make mistakes? A comparison of human and machine identification of dinoflagellates, Marine Ecology Progress Series 247 (2003) 17–25.

[7] T. Cuellar-Martinez, A. C. Ruiz-Fernández, C. Alonso-Hernández, O. Amaya-Monterrosa, R. Quintanilla, H. L. Carrillo-Ovalle, N. Arbeláez M, L. Díaz-Asencio, S. M. Méndez, M. Vargas, N. F. Chow-Wong, L. R. Valerio-Gonzalez, H. Enevoldsen, M.-Y. Dechraoui Bottein, Addressing the problem of harmful algal blooms in latin america and the caribbean- a regional network for early warning and response, Frontiers in Marine Science 5 (2018) 409.

[8] C. Davis, S. Gallager, M. Berman, L. Haury, J. R. Strickler, The video plankton recorder (vpr): Design and initial results, Arch. Hydrobiol. Beih. Ergebn. Limnol. 36 (1992) 67–81.

[9] C. S. Davis, S. M. Gallager, M. Marra, W. Kenneth Stewart, Rapid visualization of plankton abundance and taxonomic composition using the Video Plankton Recorder, Deep Sea Research Part II: Topical Studies in Oceanography 43 (7) (1996) 1947–1970.

[10] C. S. Davis, S. M. Gallager, A. R. Solow, Microaggregations of Oceanic Plankton Observed by Towed Video Microscopy, Science 257 (5067) (1992) 230–232.

[11] S. Samson, T. Hopkins, A. Remsen, L. Langebrake, T. Sutton, J. Patten, A system for high-resolution zooplankton imaging, Oceanic Engineering, IEEE Journal of 26 (2001) 671 – 676.

[12] R. Olson, A. Shalapyonok, H. Sosik, An automated submersible flow cytometer for analyzing pico- and nanophytoplankton: Flowcytobot, Deep Sea Research Part I: Oceanographic Research Papers 50 (2003) 301–315.

[13] Y. Nagashima, Y. Matsumoto, H. Kondo, H. Yamazaki, S. Gallager, Development of a realtime plankton image archiver for AUVs, in: 2014 IEEE/OES Autonomous Underwater Vehicles (AUV), 2014, pp. 1–6.

[14] C. Sieracki, M. Sieracki, C. Yentsch, An imaging-in-flow system for automated analysis of marine microplankton, Marine Ecology-progress Series - MAR ECOL-PROGR SER 168 (1998) 285–296.

[15] N. Barteneva, I. Vorobjev, D. Basiji, A. Lau, T. T.W.Wong, H. C. Shum, K. Wong, K. Tsia, M. Hildebrand, A. Davis, R. Abbriano, H. Pugsley, J. Traller, S. Smith, R. Shrestha, O. Cook, E. Sanchez-Alvarez, K. Manandhar-Shrestha, B. Alderete, Imaging Flow Cytometry: Methods and Protocols, Methods in Molecular Biology, vol. 1389, Springer, New York, 2016.

[16] H. Zheng, N. Wang, Z. Yu, Z. Gu, B. Zheng, Robust and automatic cell detection and segmentation from microscopic images of non-setae phytoplankton species, IET Image Processing 11 (11) (2017) 1077–1085.

[17] A. Verikas, A. Gelzinis, M. Bacauskiene, I. Olenina, S. Olenin, E. Vaiciukynas, Phase congruency-based detection of circular objects applied to analysis of phytoplankton images, Pattern Recognition 45 (4) (2012) 1659–1670.

[18] Woods hole oceanographic institution plankton, `https://darchive.mblwhoilibrary.org/handle/1912/7341`, accessed: 2020-07-29.

[19] Kaggle national data science bowl, `https://www.kaggle.com/c/datasciencebowl/data`, accessed: 2020-07-29.

[20] I. Corrêa, P. Drews, M. S. d. Souza, V. M. Tavano, Supervised Microalgae Classification in Imbalanced Dataset, in: 2016 5th Brazilian Conference on Intelligent Systems (BRACIS), 2016, pp. 49–54.

[21] I. Correa, P. Drews, S. Botelho, M. S. d. Souza, V. M. Tavano, Deep Learning for Microalgae Classification, in: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017, pp. 20–25.

[22] E. Álvarez, A. López-Urrutia, E. Nogueira, Improvement of plankton biovolume estimates derived from image-based automatic sampling devices: application to FlowCAM, Journal of Plankton Research 34 (6) (2012) 454–469.

[23] Tong Luo, K. Kramer, D. B. Goldgof, L. O. Hall, S. Samson, A. Remsen, T. Hopkins, Recognizing plankton images from the shadow image particle

profiling evaluation recorder, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 34 (4) (2004) 1753–1762.

[24] D. A. Lisin, M. A. Mattar, M. B. Blaschko, E. G. Learned-Miller, M. C. Benfield, Combining Local and Global Image Features for Object Class Recognition, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops, 2005, pp. 47–47.

[25] J. Zhao, H. Guo, X. Sun, A Research on the Recognition of Chironomid Larvae Based on SVM, in: 2009 Pacific-Asia Conference on Circuits, Communications and Systems, 2009, pp. 610–613.

[26] Q. Hu, C. Davis, Automatic plankton image recognition with co-occurrence matrices and Support Vector Machine, Marine Ecology Progress Series 295 (2005) 21–31.

[27] L. Boddy, C. Morris, M. Wilkins, L. Al-Haddad, G. Tarran, R. Jonker, P. Burkill, Identification of 72 phytoplankton species by radial basis function neural network analysis of flow cytometric data, Marine Ecology Progress Series 195 (2000) 47–59.

[28] C. e. a. Pf, Automatic classification of field-collected dinoflagellates by artificial neural network, Marine Ecology Progress Series 139 (1996) 281–287.

[29] K. V. Embleton, C. E. Gibson, S. I. Heaney, Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method, Journal of Plankton Research 25 (6) (2003) 669–681.

[30] E. C. Orenstein, O. Beijbom, Transfer Learning and Deep Feature Extraction for Planktonic Image Data Sets, in: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 2017, pp. 1082–1088.

[31] H. Lee, M. Park, J. Kim, Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning, in: 2016

IEEE International Conference on Image Processing (ICIP), 2016, pp. 3713–3717.

[32] E. C. Orenstein, O. Beijbom, Transfer learning and deep feature extraction for planktonic image data sets, in: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 2017, pp. 1082–1088.

[33] P. González, A. Castaño, E. E. Peacock, J. Díez, J. J. Del Coz, H. M. Sosik, Automatic plankton quantification using deep features, Journal of Plankton Research 41 (4) (2019) 449–463.

[34] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25, Curran Associates, Inc., 2012, pp. 1097–1105.

[35] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[36] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, Li Fei-Fei, Imagenet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009, pp. 248–255.

[37] K. Rodenacker, B. Hense, U. Jütting, P. Gais, Automatic analysis of aqueous specimens for phytoplankton structure recognition and population estimation, Microscopy Research and Technique 69 (9) (2006) 708–720.

[38] A. Gelzinis, A. Verikas, E. Vaiciukynas, M. Bacauskiene, A novel technique to extract accurate cell contours applied for segmentation of phytoplankton images, Machine Vision and Applications 26 (2) (2015) 305–315.

[39] E. C. Orenstein, O. Beijbom, E. E. Peacock, H. M. Sosik, WHOI-Plankton- A Large Scale Fine Grained Visual Recognition Benchmark Dataset for Plankton Classification, arXiv:1510.00745 [cs]ArXiv: 1510.00745.

[40] K. Schulze, U. M. Tillich, T. Dandekar, M. Frohme, PlanktoVision - an automated analysis system for the identification of phytoplankton, BMC Bioinformatics 14 (1) (2013) 115.

[41] A. McQuatters-Gollop, D. Johns, E. Bresnan, J. Skinner, I. Rombouts, R. Stern, A. Aubert, M. Johansen, J. Bedford, A. Knights, From microscope to management: The critical value of plankton taxonomy to marine policy and biodiversity conservation, Marine Policy 83 (2017) 1–10.

[42] K. Rodenacker, B. Hense, U. Jütting, P. Gais, Automatic analysis of aqueous specimens for phytoplankton structure recognition and population estimation, Microscopy research and technique 69 (2006) 708–20.

[43] S. Baek, J. Pyo, Y. Pachepsky, Y. Park, M. Ligaray, C.-Y. Ahn, Y.-H. Kim, J. A. Chun, K. Cho, Identification and enumeration of cyanobacteria species using a deep neural network, Ecological Indicators 115 (2020) 106395.

[44] D. Bradley, G. Roth, Adaptive thresholding using the integral image, J. Graphics Tools 12 (2007) 13–21.

[45] S. Suzuki, K. Abe, Topological structural analysis of digitized binary images by border following, Computer Vision, Graphics, and Image Processing 30 (1) (1985) 32–46.

[46] B. Delaunay, Sur la sphère vide, Bulletin de l'Académie des Sciences de l'URSS, Classe des Sciences Mathématiques et Naturelles 6 (1934) 793–800.

[47] M. d. Berg, O. Cheong, M. v. Kreveld, M. Overmars, Computational Geometry: Algorithms and Applications, 3rd Edition, Springer-Verlag TELOS, Santa Clara, CA, USA, 2008.

[48] J. Winn, A. Criminisi, T. Minka, Object categorization by learned universal visual dictionary, in: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, Vol. 2, 2005, pp. 1800–1807 Vol. 2.

[49] R. Shekhar, C. V. Jawahar, Word image retrieval using bag of visual words, in: 2012 10th IAPR International Workshop on Document Analysis Systems, 2012, pp. 297–301.

[50] K. Kesorn, S. Poslad, An enhanced bag-of-visual word vector space model to represent visual content in athletics images, IEEE Transactions on Multimedia 14 (1) (2012) 211–222.

[51] J. d. Moura, J. Novo, M. Ortega, Deep Feature Analysis in a Transfer Learning-based Approach for the Automatic Identification of Diabetic Macular Edema, in: 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8, iSSN: 2161-4393.

[52] C. Ma, J. Wu, Data Clustering: Theory, Algorithms, and Applications, Vol. 20, SIAM, 2007.

[53] M. R. Turner, Texture discrimination by gabor functions, Biological Cybernetics 55 (2) (1986) 71–82.

[54] M. Idrissa, M. Acheroy, Texture classification using Gabor filters, Pattern Recognition Letters 23 (9) (2002) 1095–1102.

[55] F. G. Bersimis, I. Varlamis, Chapter 2 - use of health-related indices and classification methods in medical data, in: N. Dey (Ed.), Classification Techniques for Medical Image Analysis and Computer Aided Diagnosis, Advances in ubiquitous sensing applications for healthcare, Academic Press, 2019, pp. 31 – 66.