applied sciences

MDPI

Article

A Low-Cost Automated Digital Microscopy Platform for Automatic Identification of Diatoms

Jesús Salido [1,*], Carlos Sánchez [2], Jesús Ruiz-Santaquiteria [1], Gabriel Cristóbal [2], Saul Blanco [3] and Gloria Bueno [1]

1 VISILAB, University of Castilla-La Mancha, Av. Camilo José Cela s/n, 13071 Ciudad Real, Spain; Jesus.RAlegre@uclm.es (J.R.-S.); gloria.bueno@uclm.es (G.B.)
2 Institute of Optics, Spanish National Research Council (CSIC), Serrano 121, 28006 Madrid, Spain; Carlos.SanchezBueno@uclm.es (C.S.); gabriel@optica.csic.es (G.C.)
3 Institute of the Environment, University of Leon, 24071 León, Spain; sblal@unileon.es
* Correspondence: jesus.salido@uclm.es; Tel.: +34-926-295-300

Received: 28 July 2020; Accepted: 25 August 2020; Published: 31 August 2020

Featured Application: Development of a fully operative low-cost automated digital microscope for the detection of diatoms by applying deep learning.

Abstract: Currently, microalgae (i.e., diatoms) constitute a generally accepted bioindicator of water quality and therefore provide an index of the status of biological ecosystems. Diatom detection for specimen counting and sample classification are two difficult time-consuming tasks for the few existing expert diatomists. To mitigate this challenge, in this work, we propose a fully operative low-cost automated microscope, integrating algorithms for: (1) stage and focus control, (2) image acquisition (slide scanning, stitching, contrast enhancement), and (3) diatom detection and a prospective specimen classification (among 80 taxa). Deep learning algorithms have been applied to overcome the difficult selection of image descriptors imposed by classical machine learning strategies. With respect to the mentioned strategies, the best results were obtained by deep neural networks with a maximum precision of 86% (with the YOLO network) for detection and 99.51% for classification, among 80 different species (with the AlexNet network). All the developed operational modules are integrated and controlled by the user from the developed graphical user interface running in the main controller. With the developed operative platform, it is noteworthy that this work provides a quite useful toolbox for phycologists in their daily challenging tasks to identify and classify diatoms.

Keywords: applied deep learning; digital microscopy; diatom identification; diatom classification; microscope automation

1. Introduction

1.1. What Is a Diatom and How Can We See It?

Diatoms are microscopic unicellular algae present on Earth since at least 180 millions years ago. They survive in any aquatic ecosystem with enough light—sea, lakes, rivers, even mud—where they live in the body of water as plankton or attached to plants, rocks, or small sand particles. Diatoms are the most diverse protists on the planet (i.e., unicellular organisms with a nucleus) with around 20,000 known species with 90% of them—according to estimates—remaining still unveiled. It is estimated that diatoms are responsible for 20% of total carbon fixation on Earth, being more productive than all the rainforests on the planet. Thus, they are essential to the whole ecosystem equilibrium of our planet [1–3].

Due to the microscopic size of diatoms—ranging from 2 µm to 2 mm in length for most species—they are invisible to the naked eye; hence, scientists use light microscopes (LM) or scanning electron microscopes (SEM) to reveal diatoms' secrets.

Diatoms are unique living beings whose cell walls are based on inorganic opaline silica, quite resistant to decay, heat, and acids. It is as if the cell were inside a Petri dish, called the frustule. The shell consists of two interlocking halves, the epitheca and the—slightly smaller—hypotheca (see Figure 1 for details). As shown in Figure 1, each thecae is formed by a structure called the valve—the main element—and several connected bands attached to it, also known as the girdle bands. Morphologically, diatoms can be separated in two types: centric—with radial symmetry—and pennate—with a bilateral symmetry.

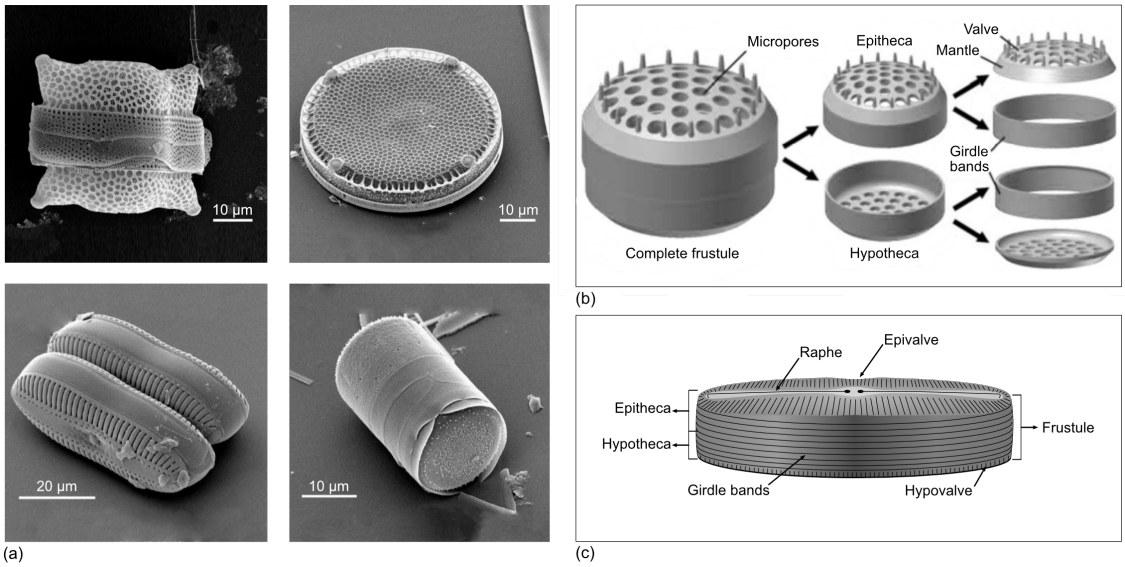


Figure 1. (**a**) Scanning electron micrographs (SEM) of four diatoms with diverse morphologies (images courtesy of Mary Ann Tifanny, San Diego Univ.); (**b**) parts of a circular frustule; and (**c**) parts of a pennate frustule.

Because of their microscopic size, diatoms are not collected individually, but in samples with an enormous number of specimens, even though they are not visible until the samples are processed. The preparation of samples by diatomists involves a careful cleaning process carried out with several oxidizing substances to eliminate all organic detritus in the sample [4,5]. All that is left after cleaning the sample are the siliceous skeletons of the diatoms, usually the disassociated elements of their frustules (i.e., valves and girdle elements; see Figure 1).

Currently, the identification and classification of diatoms is carried out using the optical microscope and digital cameras attached to it for image acquisition. With optical microscopy, it is possible to reach practical resolution values of about 0.25 µm, capable of distinguishing striae density, to 4/µm. It is worth noting that valves are transparent structures, so the image contrast from an optical microscope is a consequence of differences in the refractive index (RI) between the subject and its surroundings. The diatoms siliceous frustules have an RI of 1.4–1.43, and usually, they are mounted in a medium such as Naphrax—a synthetic resin for slide preparation—with an RI of about 1.6–1.7. Thus, what is seen of a diatom frustule depends on reflection, refraction, and diffraction at the frustule surface, and this is greatly affected by the focus over its three-dimensional morphology [6–8].

Magnification—usually ×100, ×60, ×40, and ×20 are used—is very important to capture the microscopic features of a valve, though illumination and focus are crucial to get images with good contrast. Brightfield and oblique illumination have proven their suitability for capturing high quality diatoms images. Both illumination and focus, rather than absolute choices, are relative choices with respect to the specimen under observation and its three-dimensional character. In recent studies,

multifocus and multiexposure fusion techniques have been applied to enhance digital images of diatoms [9–13].

The recording of diatom images has evolved since they were observed for the first time in 1703. Since then, drawing, photomicroscopy, and digital microscopy have provided a time-durable record of observations. Currently, besides the direct observation by a microscope, images are captured by digital cameras, with many advantages over previous methods [7,14]:

- Quick and simple capture, storage, and reproduction;
- Easy and reproducible ways to edit and enhance images by computer based image processing algorithms such as: erasing artifacts, noise reduction, multifocus, multiexposure fusion, and so on;
- Feasible enrichment of digital images with valuable information such as scale bars and other metadata;
- Processing of digital images by high level algorithms to extract features and obtain useful knowledge from the images (e.g., identification, classification).

1.2. How Diatom Are Species Distinguished?

Diatoms' identification and classification are not trivial tasks, even for the experienced diatomist. The main obstacles derive from the great number of species, the similarities between species—cryptic speciation—and the differences within species—mainly arising from the current stage of life cycle and others that are environmentally driven [15]. The main features used to identify and classify diatoms rely on the morphology (see Figure 2) and ornamentation of their valves [8,16,17], such as:

- Shape and symmetry of the valves;
- Number, density, and orientation of the striae formed by the alignment of several pores or areolas;
- Morphology of the central groove or raphe—involved in mobility—in the most pennate diatoms, especially its central area and terminal endings, and so on.

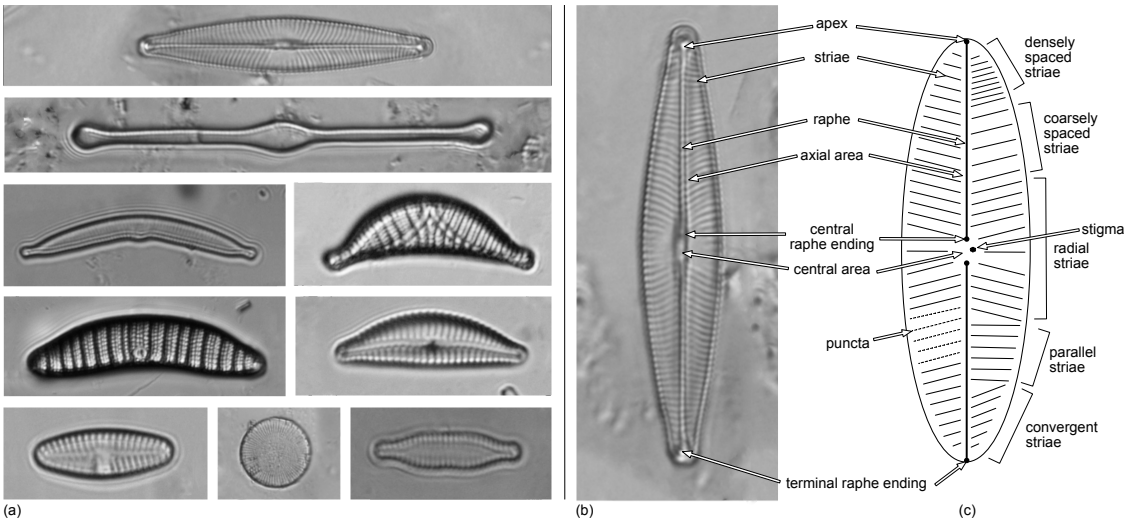


Figure 2. (**a**) Examples of the shape diversity of diatom valves (not at the same scale); (**b**,**c**) features of a pennate valve (source: AQUALITASpublic database [18]).

Diatomists identify diatoms by following two methods [8]:

1. Matching the specimen under study with a picture of a known diatom. The correct identification relies on a trained eye to interpret the subtle differences between the specimen and the reference picture.
2. Working through a decision tree about the presence/absence of morphological features. In this case, the identification concludes with the best match of features considered by the classification

key. Overall, this is the most effective approach when multiple access keys are used, where each feature is numbered in a table with a list of the coded status of each feature.

1.3. Why Is Research of Diatoms Relevant?

Beyond their contribution to the planet's ecology and even their intrinsic appeal, there are some aspects of diatoms' biology that make them important to researchers:

- The shape and decoration of the frustule—especially the valves—are very particular for each species, and most diatoms can be identified based on their shape and ornaments.
- To survive, each diatom species is adapted to a relatively specific ecosystem. Thus, the quality of aquatic ecosystems can be inferred by obtaining indexes from the diatom species present [19–22].
- They are livings beings with a siliceous inorganic cellular skeleton that is very resistant to decay (i.e., putrefaction), heat, and acids. Hence, they can be collected from seabed and lake sediments even millions of years after the cells have died.

Therefore, the study of diatoms can be applied to ecological monitoring (e.g., freshwater quality) and recreating past environments (e.g., in archaeological, geological, and forensic research) [23]. All the applied analytical research depends on the correct identification of diatom species based on a purely systematic classification and evolutionary studies carried out by a few specialized diatomists in the world.

In phycology studies, identification answers the question of which species correspond to an exemplar, and classification provides the features that allow differentiating species [3,6]. This distinction is important since in image pattern recognition jargon, identification determines in the image if a diatom "is present" and "where", whereas classification answers the question of "what diatom species" is in the image.

1.4. Automated Tools for Productive Research with Diatoms

Currently, digital microscopy provides several automation methods with many advantages over traditional ones, particularly in applied studies where the identification, counting, and classification—in the sense of pattern recognition—of diatoms are required. In this context, the tasks subject to being automated are related to: (1-DIG) image capture and digitalization; (2-PRE)image preprocessing to improve image quality by eliminating artifacts [14]; (3-SAV) the storage, recovery, and display of digital images; and (4-ADV) advanced image processing algorithms. Next, we enumerate the tasks for each mentioned category:

- (1-DIG) Programmable illumination: To obtain better image contrast, it is possible to choose different illumination parameters such as intensity, modality (e.g., brightfield, darkfield, oblique, etc.), wavelength (e.g., white, UV, etc.), and so on [24–26].
- (1-DIG) Scanning automation: When the sample under observation is not covered by the field of view (FOV), the slide should be shifted below the objective—by stage motion—to obtain information from the whole slide [27].
- (1-DIG) Stitching: If a slide requires several image captures (i.e., patches) to cover the whole preparation, the stitching process produces a new image generated by joining the individual patches [28,29].
- (1-DIG) Multifocus and multiexposure fusion: When several images are captured at different focal planes, the inherent three-dimensional morphology of diatom valves is recorded and could be fused in a new complete in focus image. Similarly, multiexposure fusion is used to obtain the right exposure over all regions in the fused image from individual images with underexposed and overexposed regions [9,10,13,30–36].

- (2-PRE) Noise reduction and artifact cleaning: Digital images captured by any microscope are affected by several sources of noise and artifacts—derived from imperfect illumination, optical systems, debris on the slide, etc.—the elimination of which is desired [14,37–39].
- (2-PRE) Contrast enhancement: Contrast is related to the different levels of intensity present in the image. The more levels are in the image, the easier it is to distinguish the details [14,38].
- (3-SAV) Adding scale bars and metadata: Useful information should be recorded and linked to each image for enrichment with additional information—not visible in the raw image—for easy recovery as needed.
- (4-ADV) Detection and identification: This task consists of determining the presence of diatoms in an image and the information about their location and quantity [8,16,40–44].
- (4-ADV) Classification: If diatoms are present in the image, classification answers the question about their taxa [45–53].

1.5. Objective

The significance of applied diatom research and interest in tools to facilitate the most cumbersome and time-consuming tasks are well accepted [3,7,19–21,23]. This work aims to integrate a fully operative set of tools to cope with the main workflow of digital microscopy with diatoms. In such a direction, the authors want to demonstrate the feasibility of a low-cost digital automated microscopy platform to deal with the detection and classification involved in applied research with diatoms. This platform will materialize a test-bed for algorithms developed by the authors and others in a fully operative environment.

Since classical machine learning algorithms—relying on the selection of handcrafted descriptors—have been applied to detection and classification for years, this work harnesses the new trends based on deep learning where the machine learns the most representative features from scratch.

2. Materials and Methods

Figure 3 shows an overview of the proposed system with its constituent components. In the following subsections, a more detailed description of the components, functionalities, and methods is provided.

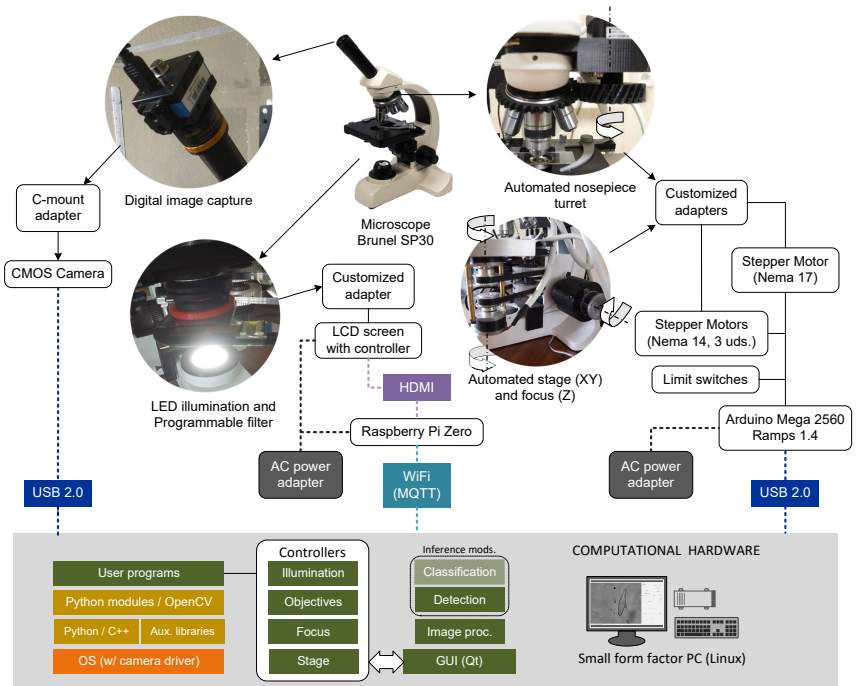


Figure 3. System overview and components.

2.1. Microscope and Accessories

- Compound transmission microscope Brunel Westbury SP30, monocular, with quadruple revolving nosepiece equipped with ×4, ×10, ×40, ×100 (oil) DIN standard parfocal objectives and eyepiece with ×10 magnification for reaching a magnification range from ×40 to ×1000. It includes: (A) a mechanical stage with motion along the X-Y axes and Vernier reference scales; (B) a coarse and fine focus mechanism consisting of a rack and pinion for motion along the Z axis with focus stop; (C) a substage Abbe achromatic condenser with independent focus control (N.A. 1.25); (D) an iris diaphragm and filter carrier; and (E) controlled LED illumination.
- Additional objectives with magnifications of ×20 and ×60 (N.A. 0.85) to provide the required value range. These objectives substitute respectively the ×10 and ×100 (oil) objectives originally included.
- Digicam CCTV Adapter. This is mounted in the eyepiece to provide a secure attachment to an external digital camera by a C-mount adapter (see the details in Figure 3).
- Digital camera coupled to the microscope eyepiece to capture digital images of diatoms. Three models were evaluated: (1) UI-1240LE-C-HG (by IDS) CMOS USB2.0 color camera with a resolution of 1280 × 1024 pixels (1.3 Mega pixels), 25 fps (frames per second), and 5 µm pixel size; (2) Toupcam LCMOS LP605100A (by ToupTek Europe) CMOS USB2.0 color, res. $w \times h$ of 2592 × 1944 pixels (5 Mpx) and 2.2 µm pixel size; and (3) DMK 72BUC02 (by The ImagingSource) CMOS USB2.0 monochrome with res. $w \times h$ of 2592 × 1944 pixels (5 Mpx) and 2.2 µm pixel size.

  The digital camera should be selected carefully having in mind whether it will be used with the provided proprietary software or it is necessary to develop customized software. In the latter case—as our case—the camera should provide an open driver and/or an SDK (software development kit) for the target operating system (e.g., Linux) where the user programs must run.

2.2. Mechanical System

The microscope mechanical system is constituted by a combination of racks and pinions to achieve the stage motion on the X-Y axes and coarse/fine focus on the Z axis. These motions are manually applied by four knobs. In addition, the rotating nosepiece turret allows changing the four objectives connected to it. In order to accomplish the automatic movement of the three mentioned systems—X-Y stage, Z focus, and nosepiece turret—several elements are employed (see the details in Figure 3):

- Bipolar stepper motors NEMA 14 (3 uds.) with step angle 0.9º/400, 17 oz-in(torque), and 650 mA (max. current) for movements following the X-Y/Z axes. These are coupled to the knobs and stage by customized adapters: (A) a supporting plate attached to the stage, made of thermoplastic PLA (polylactic acid) filament by 3D printing; (B) a pulley–belt for the X-Y stage coordinates; and (C) a direct drive for the focus Z coordinate.
- Bipolar stepper motor NEMA 17 with step angle 1.8º/200, 44 oz-in (torque), and 1200 mA (max. current) for the nosepiece turret objective changer. In this case, the customized adapter consists of a cogwheel pair—made by 3D printing.
- Optical limit switch for controlling the motors' homing motions.

The four stepper motors are controlled by an Arduino Mega 2560 board, with a Ramps 1.4 shield for the POLOLU A4988 current driver, connected by a USB2.0 interface to the main microscope PC controller.

An important limitation of the mechanical system is derived from backlash effects in the transmission elements (see Figure 4). This effect appears during X-Y and Z motions with pernicious consequences on precision, but inherent to the use of low-cost components. In our system, backlash is more severe in the pulley–belt transmissions (i.e., X-Y stage motion). In the Z focus axis, backlash has less magnitude because a direct transmission is employed. Backlash represents a truly tough problem

with a difficult solution. For mitigating its effect, the more precise controlled movements should be done in the same direction whenever possible to standardize the derived position bias.

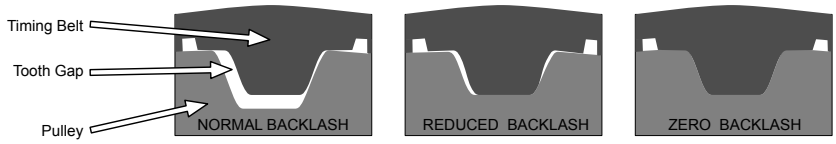


Figure 4. Backlash derived from the gap between the belt and pulley.

2.3. Programmable Illumination

Brightfield illumination of diatom specimens does not allow revealing their whole structures because they are transparent. Thus, what it is seen are patterns derived from optical reflection, refraction, and diffraction phenomena over the three-dimensional structure of a diatom frustule. It is possible to improve the lateral spatial resolution and accentuate the contrast of three-dimensional structures—obtaining a pseudo-relief effect—by applying different modalities of brightfield illumination [24–26].

In order to achieve different brightfield illumination modalities, a light filter could be located in the substage back focal plane of the condenser (see the details in Figure 3). The mechanical link is achieved by a customized coupler made by 3D printing. Applying the appropriate mask through the filter, it is possible to modify the direction of incident light to the objective. To customize the desired mask in a very flexible way, a projector LCD color screen is used. The LCD screen has a resolution of 320 × 240 pixels, and it is connected through an HDMI interface to a Raspberry Pi Zero based controller for programming the desired mask filters. In this configuration, the filter mask consists of an image displayed on the LCD screen sent from the Raspberry Pi Zero. In addition to the HDMI interface, the Raspberry controller must have a supplementary connection to the central computational hardware where the user programs run. For this link, we adopted an MQTT (Message Queuing Telemetry Transport) protocol, a lightweight standard open protocol widely available for wireless connection with a small code footprint conceived of for resource-constrained devices and low bandwidth.

Figure 5 shows three images to be displayed in the projector LCD screen corresponding to the respective modalities of illumination:

- Brightfield: This is the classical illumination modality used for diatom observation. The light is transmitted in a parallel direction from the source to the objective through the specimen under observation. The structures that absorb enough light are observed as darker than the bright background.
- Concentric oblique brightfield: In this illumination modality, the light cone is altered, so the central area is masked. Therefore, the specimen is illuminated only by a concentric oblique annular light.
- Eccentric oblique brightfield: When the illumination has a definite direction, the lateral resolution is improved. Furthermore, the three-dimensional nature of specimens is accentuated.

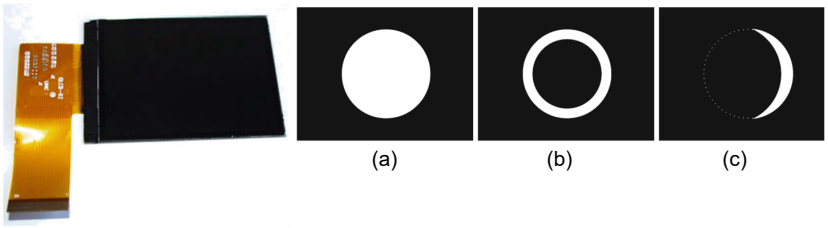


Figure 5. Projector LCD screen to obtain programmable masks for light filtering. Examples of filters for three different modalities of illumination: (**a**) brightfield, (**b**) concentric oblique brightfield, and (**c**) eccentric oblique brightfield.

2.4. Computational Hardware

The whole system with the individual controllers—mechanical subsystem and programmable illumination—must be coordinated by a main controller (see the overview in Figure 3). This main controller is constituted by a compact small form factor computer, Intel NUC (mod. DN2820FYKH) with an Intel Celeron processor (mod. N2830), 8 GB (RAM), 2 TB (HDD), and OS Linux Ubuntu 14.04 LTS (64 bits). It runs the main controller and user required functions launched from the developed GUI (graphical user interface). For instance, it will execute the image processing tasks during automatic slide scanning and the inference modules for live automatic detection of diatoms. Moreover, it is connected to an external monitor for visualization purposes.

2.5. Automatic Slide Scanning

In many studies where diatoms are involved, counting the number of specimens in a sample represents a very time-consuming task. For instance, the European Standard for quality assessment based on diatoms indexes requires at least counting 400 valves in a sample [20]. The counting task cannot be fully achieved by exploring a single field of view (FOV) captured from the slide; hence, a complete slide observation should be carried out. At times when the distribution of diatoms in the sample is approximately uniform, some diatomists suggest a random subsampling of the slide to reduce the time for counting. In any case, this tedious task can be alleviated by the automatic scanning of the slide to get a complete image with the interesting specimens in the sample. In order to cope with the automatic slide scanning task, several steps should be carried out:

- Motor calibration: Stage and focus motions are achieved by controlling the stepper motors linked to the X-Y knobs in the stage and the focus Z axis. An essential aspect for the microscope's automation is to calibrate these motors to establish the rate of X-Y and Z displacement corresponding to a single turn step of each motor.
- Image calibration: This establishes the relation between image pixels and the true distance (i.e., pixels/μm), also known as image resolution. This calibration is carried out by using a scaled slide with a known distance between consecutive divisions. Image calibration is used later to add a ruler on the image corresponding to the scale of the displayed specimens in the slide. Table 1 shows how image spatial resolution and FOV size are affected by the objective magnification.

Table 1. Image resolution and FOV size depending on the objective magnification—using the UI-1240LE-C-HG camera by IDS.

Objective	Res. (pixels/μm)	FOV Size (mm^2)
4×	0.54	4.485
20×	2.66	0.179
40×	5.39	0.042
60×	7.91	0.018

- Scanning: Since it is impossible to observe the whole slide in a single FOV, it is imperative to look over the entire slide, taking successive captures (i.e., scanning the slide). There are several strategies to deal with the scanning of the slide depending on: (a) the position of successive captures at each FOV center and (b) the path followed to reach consecutive points for image acquisition. Figure 6 illustrates three possible strategies. In Figure 6a, the whole slide is divided into an array of FOVs covered by a snake-by-row path. Figure 6b shows a row-by-row path with image acquisition in the same direction. Finally, Figure 6c shows scanning based on regions of interest (ROIs) covered by a heuristic path driven by the closest uncovered FOV. This third strategy could be appropriate in cases with very few ROIs where their discovery is compensated by the time saved in acquiring fewer images. The design of an optimal scanning

strategy should consider the extra time needed for the steady stop of the stage to avoid blurry images. It is important to notice the inherent difficulty in controlling the two stepper motors simultaneously—corresponding to X-Y motions—to obtain the desired trajectory.

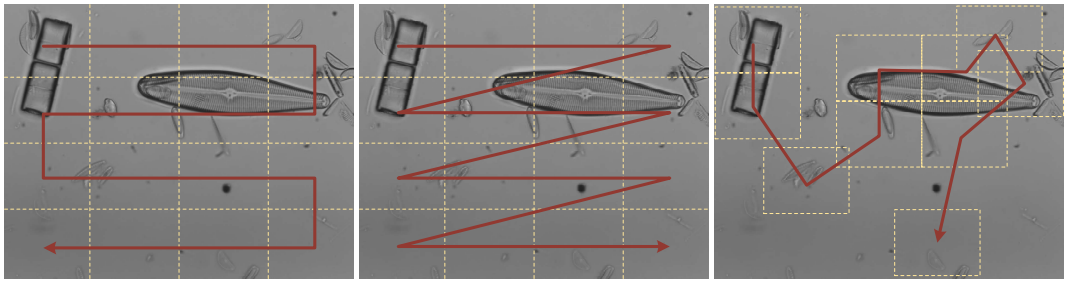


Figure 6. Different sequential scanning strategies: (**a**) following a snake-by-row path, (**b**) row-by-row path, and (**c**) based on ROIs following the next closest path.

- Autofocus: Automatic scanning relies on capturing multiple images over one single slide. As the FOV changes from one capture to the other, dynamic focus adaptation is eventually needed. There are different methods to measure the focus level in an image based on calculating gradients (i.e., derivative operators) over the image. The idea underlying autofocus algorithms based on derivative operators is that focused images have fewer levels of grey variability (i.e., sharper edges) and therefore higher frequency components than unfocused images.

  The variance of Laplacian has demonstrated its feasibility to validate in-focus diatom observation in such a way that the focus is much better when its calculated value is higher for the image [31,33,34,54–56]. The Laplacian operator Δ is calculated by convoluting the 3×3 kernel $\mathcal{L}$ over the image $I(m,n)$ (with m,n being the width and height in pixels of the image I). Then:

$$\Delta(I) = I(m,n) * \mathcal{L} \tag{1}$$

with $\mathcal{L}$ being:

$$\mathcal{L} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{2}$$

After applying the operator Δ to an $m \times n$ image, a new array is obtained. Then, the value of the variance Φ is computed for this new array by using the following equations:

$$\Phi_{m,n} = \sum_{i}^{m} \sum_{j}^{n} \left[|\Delta I(i,j)| - \overline{\Delta I} \right]^2 \tag{3}$$

with $\overline{\Delta I}$ being the average of the Laplacian:

$$\overline{\Delta I} = \frac{1}{m \times n} \sum_{i}^{m} \sum_{j}^{n} |\Delta I| \tag{4}$$

Based on the focus value (e.g., variance of Laplacian), it is necessary to define a strategy to reach the optimal focal position (Z) after an X-Y stage motion. The first type of strategies is based in a global search over a stack of images focused at different distances to select the image with the maximum focus value (i.e., the best focused image). These strategies find the best focused image and also allow fusing a selected sub-stack to obtain a final all-in-focus image. This possibility is

very interesting in microscopy to obtain better quality on images with several acquired ROIs at different focal planes [9,13,35]. However, these strategies are not appropriate when the response is time critical and the computing resources are quite limited. For automatic sequential scanning, this constitutes a valid approach since usually, there is very limited defocusing between consecutive patches; hence, a limited focal stack size usually includes the interesting images.

Other types of strategies are motivated by reducing the time and stack size to find the image with the best focus (e.g., maximizing the number of ROIs in focus). These strategies are based on optimization methods to estimate some extreme point of a function (e.g., maximum of a focus function) [32,35,36,57].

- Stitching: This is the process carried out to obtain an image composed by several images with overlapping FOVs. In digital microscopy, the automatic scanning is followed by the stitching of individual patches [28,29]. This combined process is accomplished in three stages: (1) feature extraction, (2) image registration, and (3) blending of patches.
- Preprocessing: This step consists of a set of operations applied over the image in order to mitigate the undesired effects of different sources of noise (e.g., thermal, electronic, etc.) affecting the digital image sensor and the appearance of undesirable artifacts in the image due to illumination, dust, debris, and so on. Usually, there are several types of imperfections with negative effects on microscopic images. The first type is independent of the observed specimens, and they are related to: the optical systems applied, illumination (e.g., dust particles on surfaces, uneven illumination), and noise affecting the image sensors. To remove the mentioned artifacts, the most typical operations carried out are those of [7,14]:

  (A) Noise reduction: This is usually achieved by Gaussian filters when noise can be modeled with a Gaussian distribution. Multiple types of denoising filters have been developed depending on the necessity to preserve specific image features (e.g., edges) [37,39].

  (B) Background correction: Some other effects have a homogeneous influence on all the images, and they can be removed easily by combining two images captured for the same specimen, one in focus and the other completely defocused. The image with the specimen in focus is divided by the second one, getting a resultant image free of all common imperfections. After division, the intensity of each pixel is normalized—an operation also known as contrast stretching—to obtain a better dynamic range.

  (C) Contrast enhancement: Its main purpose is to improve the perception of the information present in the image by enhancing the difference of luminance and colour, also known as contrast. There is a direct relationship between contrast and the image histogram; thus, contrast enhancement relies on operating over the histogram. Several strategies have been developed to achieve histogram equalization (HE). The underlying idea in HE is to stretch out the histogram in such a way that accumulated frequencies are approximate to a linear function. To avoid a homogeneous transformation over the whole image, some adaptive algorithms such as contrast limit adaptive histogram equalization (CLAHE) have been proposed, thus improving contrast without amplifying the noise effect [38].

With the purpose of quantifying the final image quality obtained by applying distinct illumination modalities and processing algorithms, several test have been carried out using the 1951 USAF resolution test chart, also employed to analyse the frequency response by the MTF (modulation transfer function).

2.6. Deep Leaning Applied

After slide scanning and image preprocessing have finished, higher level tasks should be carried out to understand the information present in the image. In research with diatoms, some of the most interesting high level tasks deal with answering the questions [8]:

- "Where and how many diatoms are in the image?" (i.e., detection)
- "To what taxon does each diatom belong?" (i.e., classification)

The automation of the mentioned tasks is truly worthwhile because detection is time-consuming, and classification is carried out by scarce diatomists able to distinguish specimens' taxa with subtle invisible differences to the untrained eye [58,59]. Automatic detection and classification have been the subject of study for many decades, being addressed by diverse machine learning algorithms. However, the more recent trend is to apply deep neural networks (DNNs) to these topics considering their promising results and the unnecessary handcrafted selection of features required by other approaches.

In the next sections, it will be explained how DNNs have been applied to the detection and classification of diatoms in our system.

2.6.1. Automatic Detection

According to European Union directives, the assessment of water quality requires counting about 400 diatom valves per sample [20], this being a time-consuming task. Therefore, the automatic detection of diatoms would alleviate this workload and constitutes a previous step to diatom counting.

Until now, most of the methods for diatom detection relied on classical techniques, the most recent ones [53] being based on: active contours [17,42,60,61], region segmentation [43,50], and filtering [62,63]. However, fewer studies applied DNNs to diatom segmentation. The reported average accuracy for the classical methods ranged from 88% to 95%, although with evidence of shortcomings such as: manual setting of the initial curve—in the case of active contours—and sensitivity to noise—in region segmentation methods. Moreover, all these techniques have been tested on images for the detection of a single diatom taxon.

To overcome the shortcomings of previous classical works and to provide a feasible tool to help diatomists, several methods for image segmentation are compared in order to choose a valuable one to be integrated in our system. For the mentioned comparison, the dataset is constituted by 126 images acquired at 60× magnification with a resolution of 2592 × 1944 pixels, with 10 possible diatom taxa present in the images—usually with two or three different taxa present simultaneously in the same image, at most. The chosen taxa are representative of diatoms' diversity of morphological features (i.e., external shape and internal structures). In order to prepare the 126 images of the dataset, an expert diatomist labeled 1446 diatom valves in total, with 144 average ROIs per each of the 10 taxa, using the VGG Image Annotator (VIA) [64]. The isolated ROIs are included in the publicly available AQUALITAS dataset [18].

Three DNN flavors were compared, then against classical approaches unrelated to DNNs. For the training and assessment of the three methods, the dataset was divided into two datasets for training and testing, with 105 and 21 images, respectively. The test dataset was randomly selected from the available data and held out until the final comparison of the different approaches. The comparison was achieved using three metrics calculated for the test dataset after the testing detection of the 10 possible taxa present in the sample images. For each of the three selected DNNs, a validation dataset was chosen randomly to evaluate the model performance at the end of each epoch. The evolution of the loss function for the training and validation sets constitutes an adequate tool for tuning the algorithms and implementing an early stopping strategy when overfitting to the training data occurs—the loss function increasing for the validation set while remaining steady for the training set.

Therefore, in our work, the following DNNs were chosen to solve the problem of diatom detection by image segmentation:

1. Fully convolutional fast/faster region based convolutional network: A successful evolution of R-CNN based on fully convolutional networks (FCN) is YOLO ("You Only Look Once") [65,66]. Rather than using a model for different regions, orientations, and scales to detect objects, YOLO accomplishes its task at once, with a single network for the whole image. Previous region based segmentation relied on additional algorithms that generate candidate regions.

However, the YOLO framework is able to generate a multitude of candidates—most of them with little confidence—filtered by using a suitable threshold. As the training dataset is quite small (105 images), a pretrained YOLO implementation was used [65,66]. A further training—also known as fine-tuning—was carried out for 10,000 epochs with the learning rate set to 0.001 and the optimizer based on stochastic gradient descent (SGD) with a mini-batch size of 4 images and a momentum coefficient of 0.9.

2. CNN for semantic segmentation: The chosen architecture was SegNet [67] because it was conceived of to be fast at inference (e.g., autonomous driving). This architecture tries to cope with the loss of spatial information derived from pixel-level classification caused by semantic segmentation. SegNet is constituted by: (A) one encoder network, (B) the corresponding decoder network, and (C) one final pixel-level classification layer. The encoder network consists of the first thirteen layers of the VGG16 network [68], pretrained with the COCO dataset [69]. These layers carry out operations such as: convolution, batch normalization, function activation by rectified linear units (ReLU), and max-pooling. In SegNet, the fully connected layer of VGG16 is replaced by a decoder network that recovers the feature maps' resolution prior to the max-pooling operation depending on the position at which each feature reaches the maximum value. A final softmax classifier is fed with the output of the previous decoder network. The predicted segmentation corresponds to the taxa with the maximum likelihood at each pixel. Unlike YOLO, fine-tuning is reduced to 100 epochs with the learning rate set to 0.05.
3. CNN for instance segmentation (Mask-R-CNN): This framework was created as a combination of object detection and semantic segmentation. Mask-R-CNN [44,70] is a modified version of the Faster-R-CNN object detection framework with the addition of achieving the segmentation of detected ROIs. Firstly, a CNN generates a feature map from an input image. Then, a region proposal network (RPN) produces a candidate bounding box for each object. The RPN generates the bounding box coordinates and the likelihood of being an object. The main difference between Mask-R-CNN and Faster-R-CNN consists of the layer that obtains the individual ROI feature maps using the proposed bounding boxes by the RPN. In Mask-R-CNN, that layer aligns the feature maps with the bounding boxes using continuous bins rather than quantized ones and bilinear interpolation for better spatial correspondence. A fully connected layer predicts simultaneously the likelihood of a class—using a softmax function—and the object boundaries by bounding box regression. Moreover, Mask-R-CNN adds a parallel branch with mask prediction to achieve ROI segmentation. At this step, a fully convolutional network (FCN) performs a pixel-level classification for each ROI and class.

For comparison purposes, two baseline algorithms were implemented: (A) Viola–Jones (VJ) [71]—a classical algorithm for detection based on a sliding window and a bank of simple filters—and the scale and curvature invariant ridge detector (SCIRD) [63,72]. SCIRD is based on a bank of Gaussian non-linear filters convoluted with the image. Each filter in the bank is specialized to detect a particular shape, and its parameters are adjusted according to contrast, detritus density, and low noise levels.

After training, each of the four methods to be compared produces a set of bounding boxes for the diatoms present in the images of the test dataset. The results of each method can be checked at the pixel level by checking if a predicted pixel belonging to a diatom was classified correctly or not. Three metrics were used to compare the results of the baseline algorithms (VJ and SCIRD) and those based on DNNs (YOLO and SegNet) for the detection of 10 diatom taxa in the test images. The chosen metrics are:

- Sensitivity (also called recall or true positive rate): This measures the percentage of pixels detected correctly over the whole number of pixels belonging to a diatom:

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \tag{5}$$

- Specificity (also called the true negative rate):

$$\text{Specificity} = \frac{TN}{(TN + FP)} \tag{6}$$

- Precision (proportion correctly detected):

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{7}$$

with TP = true positives (i.e., pixels correctly classified as belonging to a diatom), TN = true negatives (i.e., pixels correctly classified as not belonging to a diatom), FP = false positives (i.e., pixels incorrectly classified as belonging to a diatom), and FN = false negatives (i.e., pixels incorrectly classified as not belonging to a diatom).

2.6.2. Automatic Classification

Diatoms' classification—deciding the specific taxon for a diatom—constitutes a big challenge even for expert diatomists. The countless number of unknown species and the subtle differences among specimens belonging to the same taxon make the task of classifying diatoms a bottleneck in many research endeavors that depend on it [58]. Thus, automatic classification becomes a very interesting topic with a worthwhile reward.

The application of machine learning algorithms like Decision tree ensembles has produced very promising results [45] with average accuracies ranging from 96.17%—in studies with 55 taxa distributed into 1098 samples—to 97.97%—for 38 taxa distributed into 837 samples. In works carried out by the authors [52], these results were improved to 98.11%—for 80 taxa distributed into 24,000 samples—after an exhaustive study of 273 features including morphological, statistical, textural, and space-frequency descriptors. As mentioned in the results of these previous studies, the performance of classification relies strongly on a convenient handcrafted selection of image descriptors followed by a smart reduction to the most discriminant ones.

One of the great advantages of applying neural networks to classification tasks derives from their capacity to discover features by learning (i.e., generalization). Thus, we decided to carry out a study using convolutional neural networks (CNNs) to classify diatoms and to compare its performance against the noted classical methods that require the selection of handcrafted descriptors.

The dataset was obtained by the annotation of sample images—labeled with locations and corresponding diatom taxa also used for automatic detection—using the VGG Image Annotator (VIA) [64]. VIA is an image annotator tool with a web interface that generates a JSON file for modelling the ground truth (GT) image set. After image annotation, every single diatom specimen is cropped from its original image to obtain the 8000 annotated images—with 100 images on average per each of the 80 included taxa—that constitute the public AQUALITAS dataset [18]. Several data augmentation techniques (e.g., rotation and flipping) have been applied to get bigger datasets and to analyze the influence of samples per class—with 300, 700, and 1000 images per class. Moreover, other influences have been tested (e.g., illumination conditions) including diversity in the dataset by:

- Image segmentation for background elimination and
- Histogram normalization by histogram matching based on specific samples with good contrast.

Table 2 summarizes the 80 diatom taxa classified, including the number of samples for each taxon in the initial (i.e., original) dataset.

Therefore, several datasets were obtained to perform the experiments and to study the classification performance reached under the influence of the number of samples per class—images per taxon—in different datasets: (A) original, (B) segmented, (C) normalized, and (D) original plus normalized. Note that each dataset was augmented to obtain several distributions of images per class.

To achieve the automatic classification of diatoms over the 80 possible taxa in the previously mentioned datasets, the CNN architecture AlexNet [73] was selected. This network is well established as a baseline for experimentation in the deep learning field [51]. Moreover, it provides the advantageous possibility of using transfer leaning—starting with a pretrained net—to drastically reduce the time for the training stage, also known in this case as fine-tuning the CNN. Since AlexNet was pretrained with ImageNet [74], it was tuned by training with an initial leaning rate of 0.001, decreasing by a factor of 0.1 and a period of eight. The selected optimizer was SGD with L2-regularization of 0.004 to prevent overfitting. Moreover, AlexNet allows additional strategies to cope with overfitting like dropout and the previously mentioned data augmentation. The final layer of AlexNet has a node corresponding to each class. The output for each such node computes the likelihood of the input image belonging to the corresponding class. This calculation is reached by a softmax function that normalizes the total sum of the output for the final nodes in the range of 0–1.

Table 2. List of the 80 species for automatic classification with DNN [51].

Class Number	Species (Number of Images)	Class Number	Species (Number of Images)
1.	*Achnanthes subhudsonis* (123)	2.	*Achnanthidium atomoides* (129)
3.	*Achnanthidium caravelense* (59)	4.	*Achnanthidium catenatum* (187)
5.	*Achnanthidium druartii* (93)	6.	*Achnanthidium eutrophilum* (97)
7.	*Achnanthidium exile* (98)	8.	*Achnanthidium jackii* (125)
9.	*Achnanthidium rivulare* (305)	10.	*Amphora pediculus* (117)
11.	*Aulacoseira subarctica* (113)	12.	*Cocconeis lineata* (81)
13.	*Cocconeis pediculus* (49)	14.	*Cocconeis placentula var euglypta* (117)
15.	*Craticula accomoda* (86)	16.	*Cyclostephanos dubius* (85)
17.	*Cyclotella atomus* (99)	18.	*Cyclotella meneghiniana* (103)
19.	*Cymbella excisa var angusta* (79)	20.	*Cymbella excisa var excisa* (241)
21.	*Cymbella excisiformis var excisiformis* (142)	22.	*Cymbella parva* (177)
23.	*Denticula tenuis* (181)	24.	*Diatoma mesodon* (115)
25.	*Diatoma moniliformis* (134)	26.	*Diatoma vulgaris* (88)
27.	*Discostella pseudostelligera* (82)	28.	*Encyonema minutum* (120)
29.	*Encyonema reichardtii* (152)	30.	*Encyonema silesiacum* (108)
31.	*Encyonema ventricosum* (101)	32.	*Encyonopsis alpina* (106)
33.	*Encyonopsis minuta* (89)	34.	*Eolimna minima* (174)
35.	*Eolimna rhombelliptica* (132)	36.	*Eolimna subminuscula* (94)
37.	*Epithemia adnata* (72)	38.	*Epithemia sorex* (85)
39.	*Epithemia turgida* (93)	40.	*Fragilaria arcus* (93)
41.	*Fragilaria gracilis* (54)	42.	*Fragilaria pararumpens* (74)
43.	*Fragilaria perminuta* (89)	44.	*Fragilaria rumpens* (49)
45.	*Fragilaria vaucheriae* (82)	46.	*Gomphonema angustatum* (86)
47.	*Gomphonema angustivalva* (55)	48.	*Gomphonema insigniforme* (90)
49.	*Gomphonema micropumilum* (89)	50.	*Gomphonema micropus* (117)
51.	*Gomphonema minusculum* (158)	52.	*Gomphonema minutum* (93)
53.	*Gomphonema parvulum saprophilum* (52)	54.	*Gomphonema pumilum var elegans* (128)
55.	*Gomphonema rhombicum* (64)	56.	*Humidophila contenta* (105)
57.	*Karayevia clevei varclevei* (84)	58.	*Luticola goeppertiana* (136)
59.	*Mayamaea permitis* (40)	60.	*Melosira varians* (146)
61.	*Navicula cryptotenella* (136)	62.	*Navicula cryptotenelloides* (107)
63.	*Navicula gregaria* (50)	64.	*Navicula lanceolata* (77)
65.	*Navicula tripunctata* (99)	66.	*Nitzschia amphibia* (124)
67.	*Nitzschia capitellata* (123)	68.	*Nitzschia costei* (72)
69.	*Nitzschia desertorum* (71)	70.	*Nitzschia dissipata var media* (81)
71.	*Nitzschia fossilis* (76)	72.	*Nitzschia frustulum var frustulum* (226)
73.	*Nitzschia inconspicua* (255)	74.	*Nitzschia tropica* (65)
75.	*Nitzschia umbonata* (91)	76.	*Rhoicosphenia abbreviata* (94)
77.	*Skeletonema potamos* (155)	78.	*Staurosira binodis* (94)
79.	*Staurosira venter* (87)	80.	*Thalassiosira pseudonana* (70)

In order to evaluate the classification performance reached with each dataset, a 10-fold cross-validation scheme was followed. The comparison of the results was achieved by inspection of the confusion matrix and the metrics derived from it. In our case, the most informative metric to compare the performance of classification with each dataset variant was accuracy—calculated as the sum of the average number of correct predictions for all classes.

3. Results and Discussion

As was established in Section 1.5, we developed an automated low-cost digital microscopy platform—using available components on the market—that exhibits fully operational automation of many interesting integrated tasks for research with diatoms. In this section, a detailed description—ordered into subsections—of the most remarkable results of the work is given with a discussion of the most relevant related issues.

3.1. Programmable Illumination

A very flexible programmable illumination system was achieved by controlling a projector LCD screen that allows various illumination modes: (1) brightfield, (2) concentric oblique brightfield, and (3) eccentric oblique brightfield. Depending on the specimen under observation, specific illumination modalities could dramatically improve the image contrast. Figure 7 shows the effect on microscopic images captured using three of the filter masks tested. In the image on the right (Figure 7c), the enhancement of the contrast and relief of the diatom valve can be appreciated.

Figure 7. Effect on a microscopy image captured (at 60×, res. 1280 × 1024 pix.) with different filters masks: (**a**) circular, (**b**) concentric, and (**c**) eccentric.

After the experiment with different filter masks, we included in the system two illumination modes: brightfield—default mode—and eccentric oblique brightfield selectable through the GUI with the dark-field check box in the lateral configuration panel (see Figure 9a).

3.2. Automatic Slide Sequential Scanning

In our microscope, a sequential scanning process was implemented following a row-by-row path—to standardize backlash effects—depending on the FOVs' array size (i.e., X rows and Y columns) adjusted by the user (see Figure 9b). Moreover, this task implies several image processing steps such as: denoising, background correction, contrast enhancement, autofocus, focal stacking, and multifocus fusion.

In addition to sequential scanning, a mode of random scanning was included to allow aleatory subsampling of slides in order to reduce scanning times, especially if a uniform distribution of diatoms in the sample may be assumed. In random scanning, a certain number of cells' FOVs are selected aleatory from the array that cover the inspected region on the slide. Random scanning is selected by selecting the random fields button on the GUI control panel (see Figure 9b).

Table 3 summarizes the time—in seconds—for image acquisition (Acq. time), image processing (Proc. time), and the total time (Acq. time + Proc. time) obtained with different scanning sizes (i.e., FOV array size). Note that the image resolution is obtained by stitching individual patches of 1280 × 1024 pixels—with overlap among adjacent patches.

Table 3. Summary of the results for automatic scanning with different FOV array sizes. Acq., acquisition; Proc., processing.

FOV Array Size	Image res. ($w \times h$)	Tot.Size (mm^2)	Acq. Time (s)	Proc. Time (s)	Tot. Time (s)
3 × 3	3660 × 2928	0.17	5.719	9.678	15.397
5 × 5	6040 × 4832	0.46	13.783	24.074	37.858
7 × 7	8420 × 6736	0.90	25.916	63.642	89.557
10 × 10	11,990 × 9592	2.07	50.113	116.960	167.073

For the sequential scanning, the whole scanned area ranged from 0.17 mm^2 (for a 3 × 3 FOV array size) to 2.07 mm^2 (for a 10 × 10 FOV array size) with the respective total $w \times h$ resolution from 3660 × 2928 to 12,800 × 10,240 pixels and total time of operation (i.e., acquisition image plus image processing) from around 15 s to less than 3 min.

The acquisition time is strongly dependent on the speed of the stepper motors and the time to settle at each stop position to avoid blurry images. In order to reduce the required time, it is possible to investigate future improvements for the stepper motors' control strategy to achieve: (**A**) higher speeds without losing steps and shorter stopping times and (**B**) simultaneous X–Y motion to follow the optimum paths to mitigate the backlash effects.

3.3. Live Diatom Detection

The objective of diatom detection is to count the specimens in a slide under observation with the presence of 10 different possible species—taxa. To decide which method is the most appropriate to implement a live diatom detection function for our automated microscope, four methods were finally compared—based on three metrics. Two of them (VJ and SCIRD) are dependent on the classical careful selection of handcrafted features, and the other two are based on deep neural networks (YOLO and SegNet) capable of extracting the relevant features by themselves.

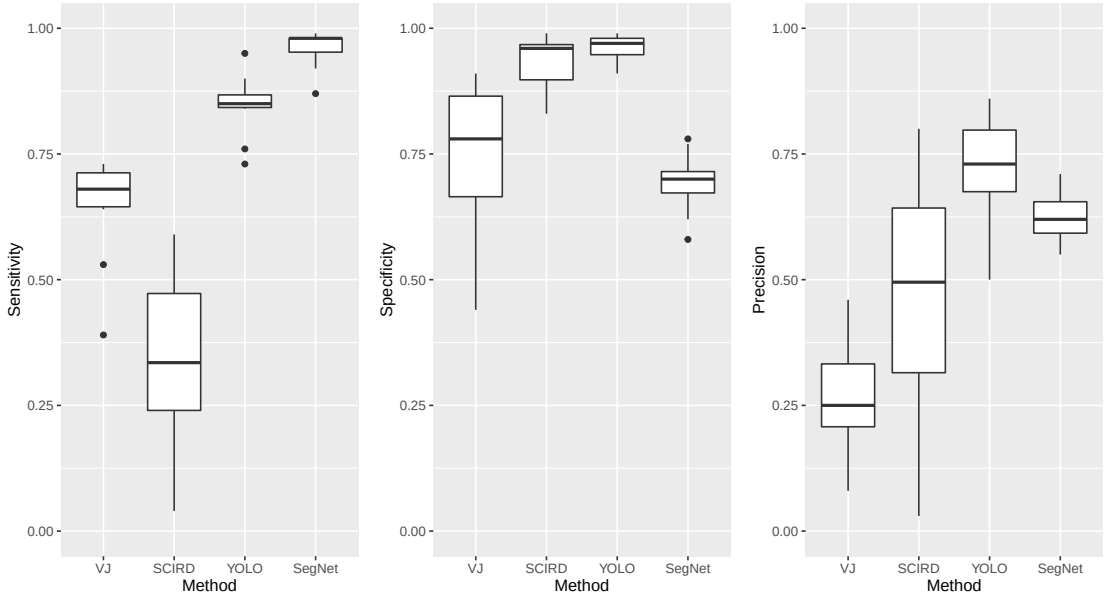
Tables 4 and 5 and Figure 8 show the results for testing the four methods for the 21 test images. Table 4 highlights the method with the best score for each metric calculated for the diatom taxa being detected. Table 5 and Figure 8 group the results for all the taxa to show the most global perspective.

Table 4. Summary of the best method for each taxon detection based on different metrics (the best values are highlighted in boldface).

Species	Sensitivity	Specificity	Precision
1. *Gomphonema rhombicum*	SegNet (0.98)	YOLO (0.93)	YOLO (0.86)
2. *Nitzschia palea*	SegNet (0.87)	SCIRD (0.97)	YOLO (0.73)
3. *Skeletonema potamos*	SegNet (0.92)	YOLO (0.97)	SegNet (0.63)
4. *Eolimna minima*	SegNet (0.98)	YOLO (0.94)	YOLO (0.64)
5. *Achnanthes subhudsonis*	SegNet (0.98)	YOLO (0.98)	YOLO (0.84)
6. *Staurosira venter*	SegNet (0.96)	YOLO (0.99)	YOLO (0.73)
7. *Nitzschia capitellata*	SegNet (0.99)	SCIRD (0.99)	SCIRD (0.80)
8. *Eolimna rhombelliptica*	SegNet (0.95)	YOLO (0.98)	YOLO (0.72)
9. *Nitzschia inconspicua*	SegNet (0.99)	YOLO (0.98)	YOLO (0.79)
10. *Nitzschia frustulum*	SegNet (0.98)	YOLO (0.97)	YOLO (0.80)

Table 5. Summary of the results for detection with several methods (the best values are highlighted in boldface): Viola–Jones (VJ), scale and curvature invariant ridge detector (SCIRD), YOLO, and SegNet.

Method	Statistics	Sensitivity	Specificity	Precision
VJ	*mean*	0.645	0.754	0.259
	std. dev	0.107	0.159	0.120
SCIRD	*mean*	0.334	0.934	0.451
	std. dev	0.179	0.056	0.254
YOLO	*mean*	0.846	0.962	0.727
	std. dev	0.063	0.026	0.107
SegNet	*mean*	0.960	0.692	0.628
	std. dev	0.038	0.061	0.051



(**a**) (**b**) (**c**)

Figure 8. Boxplots comparing the obtained results for three metrics (**a**) sensitivity, (**b**) specificity, and (**c**) precision; comparing each method (VJ, SCIRD, YOLO, and SegNet).

As is shown by these results, YOLO and semantic segmentation with SegNet overcame the classical methods not based on DNNs, although SegNet's specificity was quite poor against SCIRD. SegNet (96% sensitivity) and YOLO (96% specificity and 72.7% precision) scored the best for the metrics employed for comparison among the four strategies. Although YOLO and SegNet had an acceptable performance and were better than SCIRD (i.e., sensitivity and specificity), they produced false negatives—missing some diatoms—and false positives, even segmenting some detritus instead of diatoms.

In general, sensitivity was enhanced with semantic segmentation (SegNet) with room for improvement in precision. An observed weakness in semantic segmentation derives from the incorrect separation of ROIs when overlapping between specimens occurs. This problem may be solved with instance segmentation (Mask-R-CNN), but it has been proven only with 10 different species [44].

As a trade-off between performance and computational requirements, an inference module based on YOLO was integrated into the microscope software to provide a live diatom detection. The implementation finally integrated was adapted from the public code at the Darknet GitHub Repository [75].

3.4. Classification with CNNs

Several experiments were carried out to choose the best strategy for preparing the dataset that allowed fine-tuning AlexNet, to classify diatoms among 80 possible taxa (i.e., classes). The experiments considered the original dataset and three more:

- Segmented dataset obtained by ROI segmentation to remove the background,
- Normalized dataset produced by histogram matching, and
- Original dataset augmented with the normalized one.

Each dataset was used with three versions depending on the images/their class (300, 700, and 1000). Table 6 summarizes the results obtained by 10–fold cross-validation for the assessment of the classification with AlexNet trained with each dataset—and their three versions.

Table 6. Summary of the AlexNet results after fine-tuning with different datasets.

Dataset	Images/Class	Average Accuracy (%)	Standard Deviation
Original	1000	99.24	0.09
Segmented	1000	98.81	0.15
Normalized	1000	98.84	0.15
Original+Normalized	2000	99.51	0.048

After experimentation and parameter adjustment, the best result of 99.51% average accuracy (with a standard deviation of 0.048) was obtained with the augmentation of the original dataset with the normalized one. For this model, the sensitivity in the top 30 error-prone classes—the most difficult to classify—never fell under 93%. Moreover, the best performance was obtained always with the higher number of images per class.

In a computer equipped with a GPU NVIDIA GTX 960 with 6 GB of VRAM, the training process lasted for an hour, while the inference to classify a single image took 7 ms on the same computer. The implementation tested for AlexNet was the development deployed with the MATLAB Deep Learning Toolbox.

Although very good results were achieved, the current main controller (Intel NUC) does not provide enough computing power to deal with training and inference with AlexNet. Thus, automated classification has not been integrated in our system yet. We consider it future work to substitute the main controller with an NVIDIA JetSon module, an embedded AI computing platform capable of high-performance, low-power computation for deep learning and computer vision applications.

3.5. Software Integration and User Interface

Python was the main program language chosen for software development because of its versatility, the availability for many operating systems and platforms, and the numerous publicly available modules and libraries. Moreover, MATLAB was employed for the experiments and C++ when better code performance was preferred (e.g., the YOLO network). The most important modules and libraries used with Python were: OpenCV (for image processing), PyQt (for GUI development), NumPy (for scientific computing), and pySerial (for communications).

A graphical user interface (GUI) was developed as a friendly way to provide access to our automated microscopy platform and its functionality. The GUI provides an intuitive way to set the main configurations and microscope functions through three panels with the following controls (see Figure 9):

- Configuration panel: enables camera view (show button), setup for home position (setup and home buttons), switch between color and grey scale images (RGB check box), switch between brightfield and eccentric oblique illumination (dark-field check box), display a scale bar (scale bar check box),

autofocus (autofocus button), background correction (take background image button and background correction check box), live diatom detection (live detection check box), merge detected bounding boxes (merge box check box), and save image (save button).

- Scanning and processing settings panel: objective selection (objective radio buttons), color normalization (Reinhard check box), random scanning (random fields button), sequential scanning (X-size, Y-size dialog boxes and scan button), stitching (stitching button), focal stack acquisition (# images and step dialog boxes and take stack button), and multifocus fusion based on the extended depth of field technique [9] (EDF button).
- Motorized stage control panel: step size of motion (step size radio button), unit of motion (select unit radio button), XYZ motion (XYZ dialog boxes and move button), Z motion (Z− and Z+ buttons), and X-Y motion (arrow buttons).

Figure 9. Main GUI window showing the camera view and the control panels: (**a**) configuration panel, (**b**) scanning and processing settings panel, and (**c**) motorized stage control panel.

4. Conclusions

This work demonstrates the feasibility of a low-cost automatic platform for digital microscopy to assist expert diatomists in time-consuming tasks such as diatom counting and classification. The most remarkable functions integrated in the developed system are:

- Stage X-Y automatic motion control.
- Automated autofocus.
- Programmable illumination modes by the automatic generation of filter masks for image contrast enhancement (see Figure 7c).
- Automatic sequential scanning covering a slide area ranging from 0.17 to 2.07 mm^2 with the corresponding total time for acquisitions and processing from 15 s to 3 min (see Table 3); in addition, a derived random scanning scheme is included.
- Live detection of diatoms for faster diatom counting by using YOLO for on-time inferences with an average sensitivity of 84,6%, specificity of 96.2%, and precision of 72.7% (see Table 5).
- Focal stack acquisition and multifocus fusion by EDF.
- Full integration of operation and control from a user-friendly GUI.

This work shows a path of science transference to the applied field of digital microscopy with planned future improvements to increase the capabilities of the system and its possible adoption by the market. Some future work for improving the system are:

- Substitution of the small form factor PC by a computational unit with a GPU for embedded systems, capable of fast deep learning inference even for very demanding machine learning algorithms; in this line, a preliminary prototype is being testing with the NVIDIA JetSon platform.
- More precise segmentation algorithms to reduce false positives and false negatives.
- Better modularity of mechanical elements, for fast de/coupling to standard microscopes.
- Faster response times for sequential scanning, applying more complex motor control strategies, minimizing backlash effects, and optimizing followed paths.

Author Contributions: All the authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Spanish Government under the AQUALITAS-RETOS project with Ref. CTM2014-51907-C2-2-R-MINECO.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
CLAHE	Contrast limit adaptive histogram equalization
CNN	Convolutional neural network
DNN	Deep neural network
EDF	Extended depth of field
FOV	Field of view
FCN	Fully convolutional network
GUI	Graphical user interface
GPU	Graphics processing unit
GT	Ground truth
HDMI	High-definition multimedia interface
HE	Histogram equalization
JSON	JavaScript object notation
LCD	Liquid crystal display
LM	Light microscope
MQTT	Message Queuing Telemetry Transport
MTF	Modulation transfer function
N.A.	Numerical aperture
PLA	Polylactic acid
ReLU	Rectifier linear unit
RI	Refractive index
ROI	Region of interest
RPN	Region proposal network
SCIRD	Scale and curvature invariant ridge detector
SDK	Software development kit
SEM	Scanning electron microscope
SGD	Stochastic gradient descent
UV	Ultraviolet
VJ	Viola–Jones

References

1. Mann, D.G. Diatoms. Version 07. In *The Tree of Life Web Project*; 2010. Available online: http://tolweb.org/Diatoms/21810/ (accessed on 14 July 2020).
2. Guiry, M.D. How many species of algae are there? *J. Phycol.* **2012**, *48*, 1057–1063. [CrossRef] [PubMed]

3. Cristóbal, G.; Blanco, S.; Bueno, G. Overview: Antecedents, Motivation and Necessity. In *Modern Trends in Diatom Identification: Fundamentals and Applications*; Cristóbal, G., Blanco, S., Bueno, G., Eds.; Springer, Nature: Cham, Switzerland, 2020; Chapter 1, pp. 3–10. [CrossRef]
4. Carr, J.M.; Hergenrader, G.L.; Troelstrup, N.H. A Simple, Inexpensive Method for Cleaning Diatoms. *Trans. Am. Microsc. Soc.* **1986**, *105*, 152–157. [CrossRef]
5. Franchini, W. The Collecting, Cleaning, and Mounting of Diatoms. "How To" Tutorial Series in Modern Microscopy Journal (art. 107). 2013. Available online: https://www.mccrone.com/mm/the-collecting-cleaning-and-mounting-of-diatoms/ (accessed on 14 August 2020).
6. du Buf, H.; Bayer, M.; Droop, S.; Head, R.; Juggins, S.; Fischer, S.; Bunke, H.; Wilkinson, M.; Roerdink, J.; Pech-Pacheco, J.; et al. Diatom identification: A double challenge called ADIAC. In Proceedings of the 10th International Conference on Image Analysis and Processing, Venice, Italy, 27–29 September 1999; pp. 734–739. [CrossRef]
7. Bayer, M.M.; Droop, S.J.M.; Mann, D.G. Digital microscopy in phycological research, with special reference to microalgae. *Phycol. Res.* **2001**, *49*, 263–274. [CrossRef]
8. du Buf, H. *Automatic Diatom Identification*; World Scientific: Singapore, 2002; Volume 51. [CrossRef]
9. Forster, B.; Van De Ville, D.; Berent, J.; Sage, D.; Unser, M. Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images. *Microsc. Res. Tech.* **2004**, *65*, 33–42. [CrossRef]
10. Mertens, T.; Kautz, J.; Van Reeth, F. Exposure fusion: A simple and practical alternative to high dynamic range photography. In *Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2009; Volume 28, pp. 161–171. [CrossRef]
11. Li, S.; Kang, X.; Hu, J. Image fusion with guided filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875. [CrossRef]
12. Liu, Y.; Jin, J.; Wang, Q.; Shen, Y.; Dong, X. Region level based multi-focus image fusion using quaternion wavelet and normalized cut. *Signal Process.* **2014**, *97*, 9–30. [CrossRef]
13. Singh, H.; Cristóbal, G.; Kumar, V. Multifocus and Multiexposure Techniques. In *Modern Trends in Diatom Identification: Fundamentals and Applications*; Cristóbal, G., Blanco, S., Bueno, G., Eds.; Springer, Nature: Cham, Switzerland, 2020; pp. 165–181. [CrossRef]
14. Wu, Q.; Merchant, F.; Castleman, K. *Microscope Image Processing*; Elsevier: Amsterdam, The Netherlands, 2010.
15. Olenici, A.; Baciu, C.; Blanco, S.; Morin, S. Naturally and Environmentally Driven Variations in Diatom Morphology: Implications for Diatom–Based Assessment of Water Quality. In *Modern Trends in Diatom Identification: Fundamentals and Applications*; Cristóbal, G., Blanco, S., Bueno, G., Eds.; Springer Nature: Cham, Switzerland, 2020; pp. 39–50. [CrossRef]
16. Droop, S. Introduction to Diatom Identification. 2003. Available online: https://rbg-web2.rbge.org.uk/ADIAC/intro/intro.htm (accessed on 14 June 2020).
17. Hicks, Y.A.; Marshall, D.; Rosin, P.L.; Martin, R.R.; Mann, D.G.; Droop, S.J.M. A model of diatom shape and texture for analysis, synthesis and identification. *Mach. Vis. Appl.* **2006**, *17*, 297–307. [CrossRef]
18. Sanchez Bueno, C.; Blanco, S.; Bueno, G.; Borrego-Ramos, M.; Cristobal, G. *Aqualitas Database (Full Release)*; Aqualitas Inc.: Bedford, NS, Canada, 2020. [CrossRef]
19. John, D.M. Use of Algae for Monitoring Rivers III, edited by J. Prygiel, B. A. Whitton and J. Bukowska (eds). *J. Appl. Phycol.* **1999**, *11*, 596–597. [CrossRef]
20. European Parliament and Council of the European Union. Water Framework Directive 2000/60/EC establishing a framework for community action in the field of water policy. *Off. J. Eur. Communities* **2000**, *327*, 1–73.
21. Blanco, S.; Bécares, E. Are biotic indices sensitive to river toxicants? A comparison of metrics based on diatoms and macro-invertebrates. *Chemosphere* **2010**, *79*, 18–25. [CrossRef]
22. Wu, N.; Dong, X.; Liu, Y.; Wang, C.; Baattrup-Pedersen, A.; Riis, T. Using river microalgae as indicators for freshwater biomonitoring: Review of published research and future directions. *Ecol. Indic.* **2017**, *81*, 124–131. [CrossRef]
23. Smol, J.P.; Stoermer, E.F. *The Diatoms: Applications for the Environmental and Earth Sciences*; Cambridge University Press: Cambridge, UK, 2010.
24. Piper, J. A review of high-grade imaging of diatoms and radiolarians in light microscopy optical–and software–based techniques. *Diatom Res.* **2011**, *26*, 57–72. [CrossRef]

25. Ruíz-Santaquiteria, J.; Espinosa-Aranda, J.L.; Deniz, O.; Sanchez, C.; Borrego-Ramos, M.; Blanco, S.; Cristobal, G.; Bueno, G. Low-cost oblique illumination: An image quality assessment. *J. Biomed. Opt.* **2018**, *23*, 1–14. [CrossRef] [PubMed]
26. Piper, J.; Piper, T. Microscopic Modalities and Illumination Techniques. In *Modern Trends in Diatom Identification: Fundamentals and Applications*; Cristóbal, G., Blanco, S., Bueno, G., Eds.; Springer Nature: Cham, Switzerland, 2020; pp. 53–93. [CrossRef]
27. Brenner, J.F.; Dew, B.S.; Horton, J.B.; King, T.; Neurath, P.W.; Selles, W.D. An automated microscope for cytologic research a preliminary evaluation. *J. Histochem. Cytochem.* **1976**, *24*, 100–111. [CrossRef]
28. Brown, M.; Lowe, D.G. Automatic Panoramic Image Stitching using Invariant Features. *Int. J. Comput. Vis.* **2007**, *74*, 59–73. [CrossRef]
29. Preibisch, S.; Saalfeld, S.; Tomancak, P. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics* **2009**, *25*, 1463–1465. [CrossRef] [PubMed]
30. Valdecasas, A.G.; Marshall, D.; Becerra, J.M.; Terrero, J.J. On the extended depth of focus algorithms for bright field microscopy. *Micron* **2001**, *32*, 559–569. [CrossRef]
31. Sun, Y.; Duthaler, S.; Nelson, B.J. Autofocusing in computer microscopy: Selecting the optimal focus algorithm. *Microsc. Res. Tech.* **2004**, *65*, 139–149. [CrossRef] [PubMed]
32. Mir, H.; Xu, P.; Chen, R.; van Beek, P. An autofocus heuristic for digital cameras based on supervised machine learning. *J. Heuristics* **2015**, *21*, 599–616. [CrossRef]
33. Yazdanfar, S.; Kenny, K.B.; Tasimi, K.; Corwin, A.D.; Dixon, E.L.; Filkins, R.J. Simple and robust image-based autofocusing for digital microscopy. *Opt. Express* **2008**, *16*, 8670–8677. [CrossRef]
34. Vaquero, D.; Gelfand, N.; Tico, M.; Pulli, K.; Turk, M. Generalized autofocus. In Proceedings of the 2011 IEEE Workshop on Applications of Computer Vision (WACV), Kona, HI, USA, 5–7 January 2011; pp. 511–518. [CrossRef]
35. Choi, D.; Pazylbekova, A.; Zhou, W.; van Beek, P. Improved image selection for focus stacking in digital photography. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 2761–2765. [CrossRef]
36. Li, W.; Wang, G.; Hu, X.; Yang, H. Scene-Adaptive Image Acquisition for Focus Stacking. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1887–1891. [CrossRef]
37. Papini, A. A new algorithm to reduce noise in microscopy images implemented with a simple program in python. *Microsc. Res. Tech.* **2011**, *75*, 334–342. [CrossRef] [PubMed]
38. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 4th ed.; Pearson: London, UK, 2017.
39. Meiniel, W.; Olivo-Marin, J.C.; Angelini, E.D. Denoising of microscopy images: A review of the state-of-the-art, and a new sparsity-based method. *IEEE Trans. Image Process.* **2018**, *27*, 3842–3856. [CrossRef] [PubMed]
40. Pappas, J.L.; Stoermer, E.F. Legendre shape descriptors and shape group determination of specimens in the Cymbella cistula species complex. *Phycologia* **2003**, *42*, 90–97. [CrossRef]
41. Pappas, J.; Kociolek, P.; Stoermer, E.F. *Quantitative Morphometric Methods in Diatom Research*; Springer Nature: Cham, Switzerland, 2014; Volume 143, pp. 281–306. [CrossRef]
42. Gelzinis, A.; Verikas, A.; Vaiciukynas, E.; Bacauskiene, M. A novel technique to extract accurate cell contours applied for segmentation of phytoplankton images. *Mach. Vis. Appl.* **2014**, *26*, 305–315. [CrossRef]
43. Rojas Camacho, O.; Forero, M.G.; Menéndez, J.M. A tuning method for diatom segmentation techniques. *Appl. Sci.* **2017**, *7*, 762. [CrossRef]
44. Ruíz-Santaquiteria, J.; Bueno, G.; Deniz, O.; Vallez, N.; Cristobal, G. Semantic versus instance segmentation in microscopic algae detection. *Eng. Appl. Artif. Intell.* **2020**, *87*, 1–15. [CrossRef]
45. Dimitrovski, I.; Kocev, D.; Loskovska, S.; Džeroski, S. Hierarchical classification of diatom images using ensembles of predictive clustering trees. *Ecol. Inform.* **2012**, *7*, 19–29. [CrossRef]
46. Kuang, Y. Deep Neural Network for Deep Sea Plankton Classification. Technical Report 2015. Available online: https://pdfs.semanticscholar.org/40fd/606b61e15c28a509a5335b8cf6ffdefc51bc.pdf (accessed on 28 July 2020).
47. Lai, Q.T.K.; Lee, K.C.M.; Tang, A.H.L.; Wong, K.K.Y.; So, H.K.H.; Tsia, K.K. High-throughput time-stretch imaging flow cytometry for multi-class classification of phytoplankton. *Opt. Express* **2016**, *24*, 28170–28184. [CrossRef]

48. Dai, J.; Yu, Z.; Zheng, H.; Zheng, B.; Wang, N. A hybrid convolutional neural network for plankton classification. In *Asian Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 102–114. [CrossRef]
49. Zheng, H.; Wang, R.; Yu, Z.; Wang, N.; Gu, Z.; Zheng, B. Automatic plankton image classification combining multiple view features via multiple kernel learning. *BMC Bioinform.* **2017**, *18*, 570. [CrossRef]
50. Zheng, H.; Wang, N.; Yu, Z.; Gu, Z.; Zheng, B. Robust and automatic cell detection and segmentation from microscopic images of non-setae phytoplankton species. *IET Image Process.* **2017**, *11*, 1077–1085. [CrossRef]
51. Pedraza, A.; Bueno, G.; Deniz, O.; Cristóbal, G.; Blanco, S.; Borrego-Ramos, M. Automated Diatom Classification (Part B): A Deep Learning Approach. *Appl. Sci.* **2017**, *7*, 460. [CrossRef]
52. Bueno, G.; Deniz, O.; Pedraza, A.; Ruiz-Santaquiteria, J.; Salido, J.; Cristóbal, G.; Borrego-Ramos, M.; Blanco, S. Automated Diatom Classification (Part A): Handcrafted Feature Approaches. *Appl. Sci.* **2017**, *7*, 753. [CrossRef]
53. Tang, N.; Zhou, F.; Gu, Z.; Zheng, H.; Yu, Z.; Zheng, B. Unsupervised pixel-wise classification for Chaetoceros image segmentation. *Neurocomputing* **2018**, *318*, 261–270. [CrossRef]
54. Pech-Pacheco, J.L.; Cristóbal, G.; Chamorro-Martinez, J.; Fernández-Valdivia, J. Diatom autofocusing in brightfield microscopy: A comparative study. In Proceedings of the 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain, 3–7 September 2000; Volume 3, pp. 314–317.
55. Li, J. Autofocus searching algorithm considering human visual system limitations. *Opt. Eng.* **2005**, *44*, 113201. [CrossRef]
56. Pertuz, S.; Puig, D.; Garcia, M.A. Analysis of focus measure operators for shape-from-focus. *Pattern Recognit.* **2013**, *46*, 1415–1432. [CrossRef]
57. Yang, S.J.; Berndl, M.; Ando, D.M.; Barch, M.; Narayanaswamy, A.; Christiansen, E.; Hoyer, S.; Roat, C.; Hung, J.; Rueden, C.T.; et al. Assessing microscope image focus quality with deep learning. *BMC Bioinform.* **2018**, *19*, 1–9. [CrossRef]
58. Mann, D.G. The species concept in diatoms. *Phycologia* **1999**, *38*, 437–495. [CrossRef]
59. Mann, D.G.; Vanormelingen, P. An inordinate fondness? The number, distributions, and origins of diatom species. *J. Eukaryot. Microbiol.* **2013**, *60*, 414–420. [CrossRef]
60. Chan, T.F.; Vese, L.A. Active contours without edges. *IEEE Trans. Image Process.* **2001**, *10*, 266–277. [CrossRef]
61. Li, Y.; Belkasim, S.; Chen, X.; Fu, X. Contour-based object segmentation using phase congruency. In *Int Congress of Imaging Science ICIS*; Society for Imaging Science and Technology: Rochester, NY, USA, 2006; Volume 6, pp. 661–664.
62. Verikas, A.; Gelzinis, A.; Bacauskiene, M.; Olenina, I.; Olenin, S.; Vaiciukynas, E. Phase congruency-based detection of circular objects applied to analysis of phytoplankton images. *Pattern Recognit.* **2012**, *45*, 1659–1670. [CrossRef]
63. Libreros, J.; Bueno, G.; Trujillo, M.; Ospina, M. Automated identification and classification of diatoms from water resources. In *Iberoamerican Congress on Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 496–503. [CrossRef]
64. Dutta, A.; Zisserman, A. The VIA Annotation Software for Images, Audio and Video. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; ACM: New York, NY, USA, 2019. [CrossRef]
65. Redmon, J.; Farhadi, A. Yolo 9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271. [CrossRef]
66. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:abs/2004.10934.
67. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *arXiv* **2015**, arXiv:abs/1511.00561.
68. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:abs/1409.1556.
69. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. In *Computer Vision—ECCV 2014*; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 740–755. [CrossRef]
70. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988. [CrossRef]

71. Viola, P.; Jones, M. Robust Real-time Object Detection. In Proceedings of the Second International Workshop on Statistical and Computational Theories of Vision–Modeling, Learning, Computing, and Sampling, Vancouver, BC, Canada, 13 July 2001.
72. Annunziata, R.; Trucco, E. Accelerating convolutional sparse coding for curvilinear structures segmentation by refining SCIRD-TS filter banks. *IEEE Trans. Med Imaging* **2016**, *35*, 2381–2392. [CrossRef]
73. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation, Inc.: La Jolla, CA, USA, 2012; pp. 1097–1105. [CrossRef]
74. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009. [CrossRef]
75. Bochkovskiy, A. YOLOv4—Neural Networks for Object Detection (Windows and Linux Version of Darknet). 2020. Available online: https://github.com/AlexeyAB/darknet (accessed on 14 August 2020).