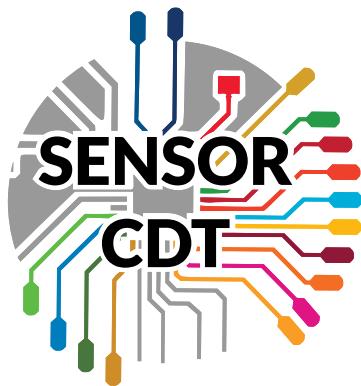


EARLY DETECTION OF CYANOBACTERIA CONCENTRATION IN FRESHWATER SAMPLES

Filip Ayazi, Justas Brazauskas, Stephen de Bank, Asher Dworkin, Terry Fawden, Hayley Gilbert, Panayiotis Ioannou, Sofia Kapsiani, Jake Stuchbury-Wass, Josephine Tumwesige, Sotirios Vavaroutas, Marie Zedler



September 2022

Final Report

EPSRC Centre for Doctoral Training in Sensor Technologies
for a Healthy and Sustainable Future

Department of Chemical Engineering & Biotechnology
University of Cambridge

Executive Summary

Cyanobacteria, also known as blue-green algae, are photosynthetic bacteria living on water surfaces and are usually present in relatively low concentrations. Climate change and eutrophication are two processes that promote rapid multiplication of cyanobacteria, which results in excessive amounts of cyanobacteria present in water, called blooms. These algal blooms often produce toxins which can have detrimental effects on humans and animals. Excessive amounts of cyanobacteria can create numerous issues, including human health and safety for recreational water activities, animals and other organisms living in freshwater systems.

In this project, we produced a sensing device for the early detection of cyanobacteria species. This device offers a low-cost solution for algae and cyanobacteria detection in water samples and includes an intuitive user interface which allows for usage by non-experts. We focused particularly on three cyanobacteria strains with different morphologies, namely *Synechocystis sp. PCC 6803*, *Anabaena sp. PCC 7120*, and *Synechococcus elongatus PCC 7942*. To facilitate this, a 3D-printed low-cost microscope and microfluidic channel is used to monitor cyanobacteria concentrations. A freshwater sample is flowed through a microfluidic channel and imaged with the microscope. Subsequently, a machine learning model is deployed for the automatic detection, classification, and quantification of cyanobacteria. The concentration over time can be used to warn about exponential growth.

With this project, we addressed the shortcomings of existing solutions. These include the relatively late detection through remote sensing, the need for expensive lab equipment, a lack of continuity of monitoring, and high costs of sensors. Our device achieves continuous monitoring in real time at any location of interest, with the option of wireless data transfer. Moreover, it enables automated use for non-specialists and allows for easy adaptation for different aims though the flexibility of the machine learning algorithm. The device was developed to align with the UN Sustainable Development Goals, with a particular focus on number 3 – Good Health and Wellbeing, number 6 – Clean Water and Sanitation, and number 14 – Life below Water. Our sensor is fully open source and we achieved a final device cost of less than £900, which is one magnitude lower than other existing solutions.

We envision our product to be used by citizen scientists to find early indicators of increased cyanobacteria presence, which can then serve as evidence to local governments to demand for further testing and action to remove the health hazards. Our device is also of interest to commercial outdoor swimming and water sports centres, who are dependent on healthy waters to ensure continuous use of their facilities. The system developed in this project will help the adoption of preventative measures early on to mitigate the negative effects of cyanobacteria on recreational water activities, drinking water sourced from freshwater, as well as other organisms in our ecosystems.

Acknowledgements

We would like to sincerely thank:

The Engineering and Physical Sciences Research Council (EPSRC) for funding the project;

Jenny Molloy, Pietro Cicuta, Tijmen Euser, Roisin Owen, Clemens Kaminski, Sensor CDT Program Manager and Administrators for continued support and advice throughout the project;

Stephen Rowden, Ana Fernandez-Villegas, Anna Bird, Yuqing Feng, Hassan Rahmoune, Danita Pearson, for their assistance in the project's laboratory work;

Nino Läubli for fabricating the microfluidic channels and advising on considerations for the microfluidic system;

Lara Allen for her continued support, advice and help with networking;

Laura Tome Maiz for her advice on outreach activities;

Charlie Wedd for assisting in the microfluidic pump selection;

Josh Easy and the MakerSpace personnel for printing the device;

Samuel McDermott for help getting set up with OpenFlexure;

Edward Ward and Jacob Lamb for aiding in the microscopy related elements to the project;

Gabi Kaminski for providing us lab access;

Valerian Sanga for advice on motor control board design;

To all the CDT and CEB directors and staff for their support and in particular Adriana Wolf, Latika Soogumbur and Kate Baker for all their work with the administration.

Contents

1 Project Outline	1
2 Science and Technology	4
2.1 Biology	4
2.1.1 Introduction to Cyanobacteria	4
2.1.2 Evolution of our Ideas	4
2.1.3 Microorganism Selection	5
2.1.4 Cell Culturing and Sample Preparation	6
2.1.5 Cell Counting Method	7
2.2 Microscopy	9
2.2.1 OpenFlexure	9
2.2.2 OpenFlexure Software	10
2.2.3 Custom Motor Control Board	11
2.2.4 Microscope Hardware	11
2.2.5 Brightfield Microscopy	13
2.2.6 Fluorescence Microscopy	15
2.3 Microfluidics	18
2.3.1 Microfluidic System Overview	18
2.3.2 Filtration	18
2.3.3 Microfluidic Pump	19
2.3.4 Flow Sensor	19
2.3.5 Microfluidic Channel	20
2.4 Systems and Design	21
2.4.1 System Overview	21
2.4.2 Hardware Design	22
2.4.3 Power System	23
2.4.4 Communications	25
2.4.5 Performance and Limitations	26
2.5 Machine Learning Solution	28
2.5.1 Machine Learning Models for Cyanobacteria Detection	28
2.5.2 Desktop Software for Training the Object Detection Model	31
2.5.3 Data Collection, Labelling & Augmentations	33
2.5.4 Running Inference On-Device	35
2.5.5 Conclusion	37
3 Inclusive Innovation, Market and Stakeholder Engagement	38
3.1 Market	38
3.2 Alternative Solutions	39
3.3 Stakeholder Engagement	41
3.4 Branding	42
3.4.1 Website	42
3.4.2 Dall-E 2	42
3.4.3 Logo and Naming	43
4 Future plans	44

1 Project Outline

The occurrence of toxic cyanobacteria blooms in freshwater ecosystems has been continuously reported over the last five decades worldwide [1]. Nevertheless, remote monitoring of cyanobacteria concentrations is not standard practice for drinking water suppliers or recreational freshwater spaces [2]. Although microscopic systems for the detection of cyanobacteria have been developed and commercialised, these are either laboratory instruments or have high costs. Alternative solutions such as sequencing are also not common practices among environmental agencies due the lack of funding and laboratory facilities [3]. Therefore, the aim of this project was to develop a low-cost and open-source microscopy system for remote detection of cyanobacteria concentrations in freshwater environments.

This was achieved through adapting the open-source OpenFlexure microscope to detect cyanobacteria species using brightfield microscopy. To enable continuous monitoring of freshwater samples a microfluidic system is included. This allows new samples to be collected and imaged autonomously over time. Cyanobacteria strains of different morphologies analogous to environmental strains were collected and cultured in the lab. A machine learning system is also included with a user-friendly user interface (UI) to allow non-experts to interact with the system.

The original objectives of the project included building a machine learning algorithm to distinguish toxic from non-toxic cyanobacterial strains. A few weeks into the project, we were challenged by expert opinions that deemed our approach infeasible. Several biology experts stated that a differentiation of toxic from non-toxic species, solely based on morphology, is extremely challenging due to cryptic diversity in cyanobacteria species as phenotypic plasticity that many strains exhibit [3]. After carefully evaluating these critiques, we decided to rephrase our objectives and instead develop a more generic tool for the detection of cyanobacteria, regardless of toxicity level. Not only was this more achievable in the given time frame, but also allows for broader research on cyanobacteria ecology.

Furthermore, as cyanobacteria autofluoresce when excited at a particular wavelength, we were interested in using this property for distinguishing them from other microorganisms that do not fluoresce in that part of the spectrum. Nevertheless, a second challenge arose at the beginning of August, when we made the decision to cease working with fluorescence microscopy. This was due to the time constraints of the project, we strongly believed that we would not be able to achieve good results with fluorescence imaging within a short amount of time. However, this is something that could be explored further in the future if time is not a limiting factor. Continuing with brightfield microscopy also gives us the opportunity of being able to expand our technique for detecting other pathogenic microorganisms that are often present in freshwater ecosystems that do not autofluoresce at strong intensity such as *Escherichia coli* (*E. coli*) or *Salmonellae*.

Our prototype has several important advantages over existing solutions, like the continuity of monitoring in real-time at any location of interest. A further advantage of our product is the ease of use and the adaptability for different imaging target species. This presents an exciting opportunity for ecology researchers to engage citizen scientists for wider data collection. In order to bring the project to its full fruition, sequencing is necessary to obtain ground-truth data to validate our results. Numerous opportunities for collaboration to test in local freshwaters arose. However, to ensure the accuracy of our device we would need to obtain ground-truth sequencing data, which was not possible

at this stage due to the time limitations of the project. We believe this is an exciting avenue to pursue in the future of this project.

During the project, we divided our team into smaller workgroups for six different research domains. We also assigned a project management team and a treasurer. These are listed in detail in Figure 1.

	Filip	Justas	Stephen	Asher	Terry	Hayley	Panayiotis	Sofia	Jake	Josephine	Sotirios	Marie
Project Management												
Project Managers												
Treasurer												
Biology												
Experts meetings												
Risk assesments												
Cell culture												
Cell counting												
Microscopy												
Microscope hardware												
Bright-field												
Fluorescence												
Motor control board												
Microfluidics												
Filtration												
Microfluidic pump												
Flow sensor												
Microfluidic channel												
Systems and Design												
Hardware design												
Power system												
Communications												
Machine Learning												
Labelling												
Model building												
Model training												
UI												
On device												
Marketing												
Logo design												
Website design												
DALLE images												
Stakeholder engagement												

Figure 1: Allocation of individual members to subgroups and tasks. Note different shades reflect levels of engagement.

Furthermore, each subgroup was allocated a group leader, to facilitate smooth communication and coordination across groups and to the management team. During the duration of the project, we found that this structure worked well for us and helped us to manage our time and resources effectively. To engage our stakeholders regularly, we organised weekly progress report meetings on Microsoft Teams, during which two researchers presented the team’s progress. These sessions were invaluable for gaining feedback from our expert stakeholders and ensuring continued interest in our project.

Aside from presenting in the weekly meetings, the two allocated team members also participated in a filming session where they provided a 2-minute summary of the project’s progress to be understood by a wider audience. Additionally, we held weekly meetings as a group to discuss the week’s progress, further steps, as well as complications, risks and any decisions that had to be made. Within subgroups, leaders organised meetings approximately once a week as well, at varying times, to discuss concrete steps to progress within their individual groups. This setting proved to be successful, as it enabled excellent communication between and within the subgroups with all members having a good understanding of what each subgroup is working on at all times.

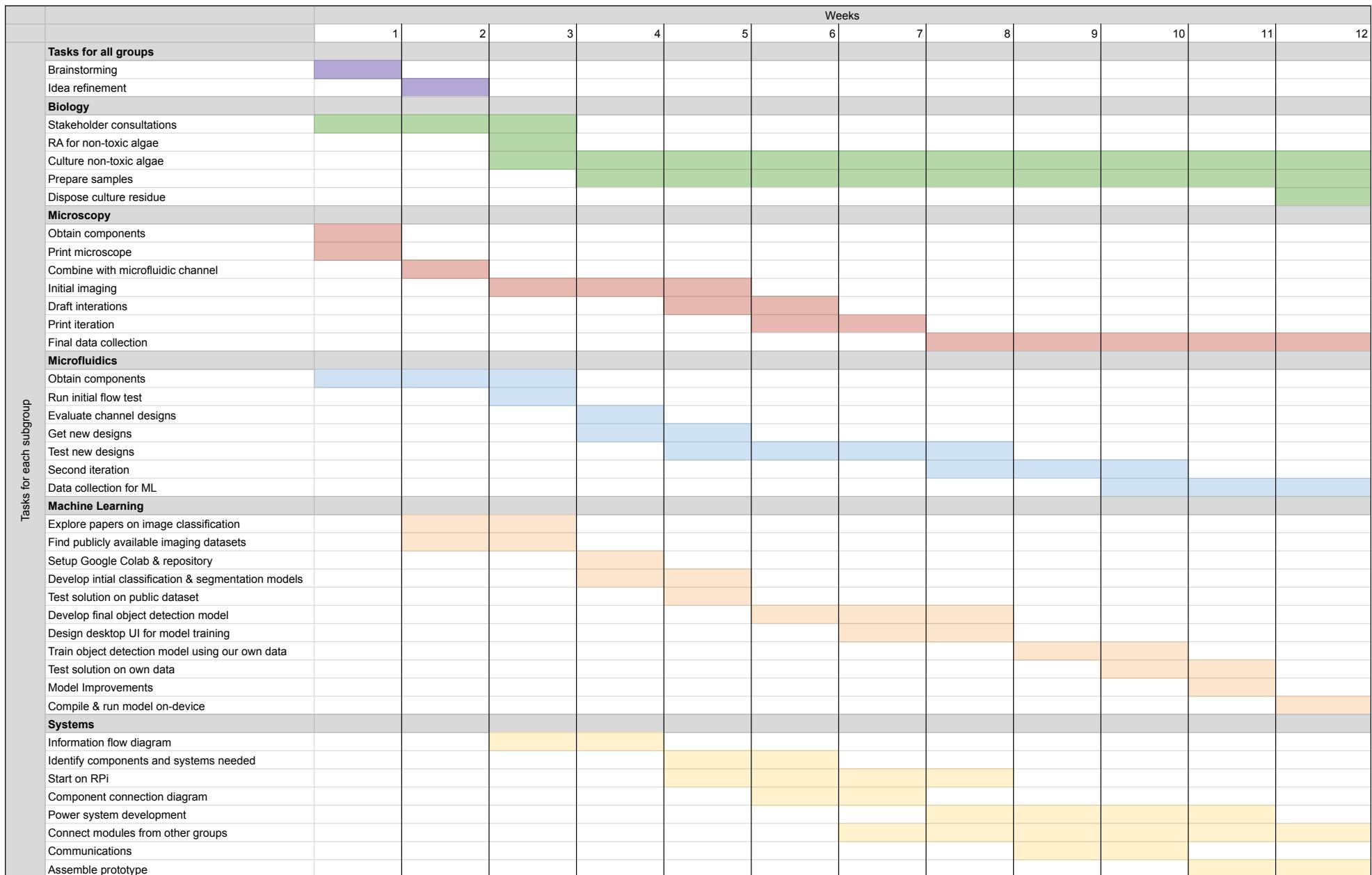


Figure 2: Tasks for each subgroup over time represented in a Gantt chart.

2 Science and Technology

2.1 Biology

2.1.1 Introduction to Cyanobacteria

Cyanobacteria, also known as blue-green algae, are oxygenic photosynthetic prokaryotes, that were responsible for the oxygenation of the Earth's atmosphere a billion years ago [4]. They are present in a wide range of habitats including freshwater, oceans and soil. Cyanobacteria are microscopic organisms with cell sizes ranging from 0.5 to 60 μm [5]. All cyanobacteria chlorophyll "a" pigment and many species additionally synthesise phycobilin or carotenoid pigments, which are responsible for the characteristic colour of cyanobacteria [6, 7]. The presence of these biopigments enables cyanobacteria to fluoresce at particular wavelengths [7].

Cyanobacteria are usually present in relatively low concentrations on water surfaces, but can also be found at deeper levels [8]. Excessive amounts of cyanobacteria can create numerous issues, including human health and safety for recreational water activities, animals and other organisms living in freshwater systems, as well as fishing. Climate change and eutrophication are two processes that promote rapid multiplication of cyanobacteria, which results in excessive amounts of cyanobacteria present in water, called blooms. These algal blooms often produce toxins which can have harmful effects humans and animals, with about 75% of surface bloom-forming cyanobacteria species producing toxins [9]. Potential health effects of cyanotoxins include gastroenteritis, fever, liver damage, neurotoxicity, and irritation of the skin, eyes, throat and respiratory tract [10].

2.1.2 Evolution of our Ideas

The original focus of this project was the development of a low-cost, open-source microscopy device suitable for real-time detection of toxic cyanobacteria species. As cyanobacteria blooms are often visible to the naked eye, we were interested in building an early detection system for the presence of toxic cyanobacteria strains in freshwater samples that could provide a warning before the formation of the toxic blooms. To evaluate the viability of this idea, we discussed it with several experts in the field, including Dr David Lea-Smith (Head, Lea-Smith lab at the University of East Anglia), Prof. Robert Field (Director, Manchester Institute of Biotechnology), Dr Paolo Bombelli and Dr Stephen Rowden (Department of Biochemistry, University of Cambridge). In all of our discussions, we were met with the same challenge; to the best of the experts' knowledge, toxic cyanobacteria strains could not be distinguished from non-toxic strains based on morphology characteristics alone. This claim is also supported by the literature, where for example Moreira *et al.* stated that morphological-based identification of cyanobacteria does not provide insight into the presence of toxicity [11]. Specifically, cyanobacteria display a vast diversity in their morphology, including unicellular, filamentous and colony-forming species [12]. Examples of cyanobacteria species of different morphologies are shown in Figure 3 which is taken from Dvořák *et al.* [13].

Distinguishing cyanobacteria species based on morphology requires the identification of structural characteristics that are easily recognisable even at lower magnifications [3]. However, this can be particularly challenging in species with small diameters, such as in the picocyanobacteria, as well as due to phenotypic plasticity [3]. For example, unicellular structures of diameters smaller than 2 μm are usually identified as *Synechococcus* or

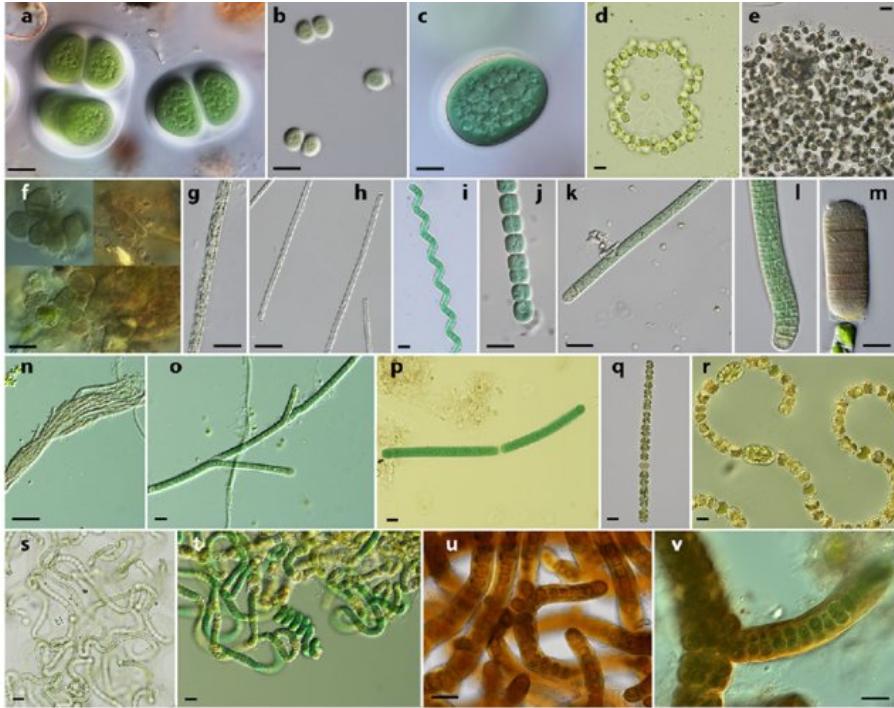


Figure 3: Diversity in cyanobacteria morphology. a) *Chroococcus subnudus*, b) *Ch. limneticus*, c) *Cyanothecae aeruginosa*, d) *Snowella litoralis*, e) *Microcystis aeruginosa*, f) *Pleurocapsa minor*, g) *Planktothrix agardhii*, h) *Limnothrix redekei*, i) *Arthrosphaera jenneri*, j) *Johanseninema constrictum*, k) *Phormidium* sp., l) and m) *Oscillatoria* sp., n) *Schizothrix* sp., o) *Tolypothrix* sp., p) *Katagnymene accurata*., q) *Dolichospermum plancticum*, r) *Dolichospermum* sp., s) *Nostoc* sp., t) *Nodularia moravica*, u) and v) *Stigonema* sp. Scale bars represent: au = 10 μ m, v = 20 μ m. Images and caption reproduced from Dvořák et al. [13]

Cyanobium, whereas, colony-forming species in the same range are classified as *Aphanothece*, *Aphanocapsa* or *Anathece* [3]. Nevertheless, *Cyanobium* strains under grazing pressure can also form colonies that closely resemble those of *Anathece*, potentially leading to misidentification [14, 15, 3]. Moreover, there are a few morphological characteristics that can be used to identify *Gloeobacterales* and *Synechococcales*, however, these can also appear in *Chroococcales* [3]. Therefore, the large cryptic diversity in cyanobacteria species makes it extremely challenging to distinguish between toxic and non-toxic species [3].

An alternative solution to microscopy is sequencing. However, sequencing is a more expensive method that requires the use of a laboratory and would therefore not be suitable for real-time monitoring. Taking the above into account we decided to switch the direction of our project to a microscopy device that can recognise any cyanobacteria species, whether toxic or non-toxic. The system will be able to detect an increase in cyanobacteria concentration or an existing high count of cyanobacteria cells in the water sample. This stresses the need for further testing, which could be performed through DNA sequencing to identify whether the cyanobacteria present are a toxic strain.

2.1.3 Microorganism Selection

The remarkable morphological diversity of cyanobacteria makes it important to train our microscopy system to recognise a variety of cyanobacteria morphologies that are

commonly present in freshwater samples. We therefore obtained from the Department of Biochemistry three cyanobacteria stains with different morphological features and sizes:

- (i) *Synechocystis* sp. PCC 6803 (morphology: circular)
- (ii) *Anabaena* sp. PCC 7120 (morphology: filamentous)
- (iii) *Synechococcus elongatus* PCC 7942 (morphology: rod-shaped)

Synechocystis sp. PCC 6803 and *Synechococcus elongatus* sp. PCC 7942 are unicellular freshwater cyanobacteria, that are commonly used as model organisms and are not known to produce any cytotoxins [16, 17]. In particular, *Synechocystis* sp. PCC 6803 is the most studied cyanobacteria strain due to its fast growth and amenability to genetic engineering [18, 16].

On the other hand, *Anabaena* is a genus of filamentous cyanobacteria that form blue-green algae blooms [19]. Several *Anabaena* strains are known to produce toxins such as microcystin as well as the neurotoxin anatoxin-a [20]. Toxic *Anabaena* strains have been reported worldwide, including *Anabaena bergii* in Australia, *Anabaena lapponica* in Finland and *Anabaena circinalis* and *Anabaena oscillarioides* in the Limpopo river basin in South Africa [21]. Herein, we obtained a laboratory strain *Anabaena* sp. strain PCC 7120 that does not produce any toxins [21].

Due to the morphological similarity between toxic and non-toxic strains we hypothesised that our algorithm may also be able to detect toxic strains with analogous structural features to the cyanobacteria we cultured. Working with non-toxic strains was preferred due to safety implications as well as faster approval of our risk assessments, which was particularly important due to the short length of this project.

Freshwater ecosystems are home to a myriad of microorganisms such as bacteria, algae, fungi and protozoa. To increase the robustness of our algorithm and ensure that it can distinguish cyanobacteria as opposed to other microorganisms and inorganic matter that are small enough to enter the microfluidic channel, we acquired two organisms to act as false positives. Specifically, common algae (*Chlorella Sorokiniana*) was obtained from the Department of Biochemistry and laboratory *E. coli* from the Department of Chemical Engineering and Biotechnology.

2.1.4 Cell Culturing and Sample Preparation

The cyanobacteria and algae were cultured using BG-11 medium. BG-11 medium consisted of compounds displayed in Table 1. The pH of BG-11 medium was 7.5 before autoclaving. After autoclaving and cooling the pH was about 7.1.

The live algal cultures were transferred and transported in single-use sterile 250ml Thermo Scientific Nunc cell culture treated flasks with filter caps. At the Department of Chemical Engineering and Biotechnology, the flasks were transferred into a culture cabinet at room temperature. After one to seven days, 10 ml of each strain was extracted and transferred to 100 ml glass flasks with 50 ml of fresh BG-11 medium. The delay in the transfers were due to waiting for approval of the risk assessment by the department's laboratory safety officer.

Two cultures of each cyanobacteria strain and algae were prepared and incubated at $30 \pm 2^\circ\text{C}$ and $21.5 \pm 6.5^\circ\text{C}$, with warm-white fluorescent tubes providing the light intensity of $8.6 \pm 2 \mu\text{E m}^{-2} \text{s}^{-1}$ and $35.6 \pm 2 \mu\text{E m}^{-2} \text{s}^{-1}$ respectively, and shaken at 200 to

Compound	Amount (grams per litre)
NaNO ₃	1.5
Na ₂ CO ₃	20 × 10 ⁻³
Na ₂ MG EDTA	100 × 10 ⁻³
ferric ammonium citrate	600 × 10 ⁻³
citric acid·H ₂ O	600 × 10 ⁻³
Ca ₂ Cl · 2 H ₂ O	3.6
MgSO ₄ · 3 H ₂ O	7.5
K ₂ HPO ₄ · 3 H ₂ O	4
H ₃ BO ₃	2.86
MnCl ₂ · 4 H ₂ O	1.81
ZnSO ₄ · 7 H ₂ O	222 × 10 ⁻³
CuSO ₄ · 5 H ₂ O	79 × 10 ⁻³
COCl ₂ · 6 H ₂ O	50 × 10 ⁻³
NaMoO ₄ · 2 H ₂ O	391 × 10 ⁻³

Table 1: Contents of BG-11 medium in grams per litre.

250 rpm. The culture growth was monitored using the culture colour by directly looking at the culture flasks and by adding a drop of the culture on a slide and observing the microorganisms using microscopy. The strains were transferred every 2-3 weeks with 1:50 dilution in new medium.

Cultures subjected to higher light intensity grew faster than those with low light intensity and were therefore used for sampling. The *E. coli* bacteria were cultured and provided to us by Anna Bird, a PhD student at the Department of Chemical Engineering and Biotechnology. All samples were prepared under a laminar flow hood to minimise the risk of contamination of the laboratory environment and the sample, and a 70% ethanol solution was used to clean surfaces. Furthermore, all culture residues were also sterilised by treatment using Virkon disinfectant and autoclaving. The samples were pipetted onto a slide or extracted using a syringe and pumped into a microfluidic channel for observation and imaging.

2.1.5 Cell Counting Method

Determining cell counts of unicellular, colonial and filamentous cyanobacteria was critical for our goal of monitoring potentially toxic cyanobacteria. In order to easily measure the biovolume of the cyanobacteria in our flow, a hemocytometer was used. A hemocytometer consists of a thick glass microscope slide with a grid of perpendicular lines etched in the middle. The grid has specified dimensions so that the area covered by the lines is known, which allows us to count the number of cells in a specific volume of solution.

The most common type of hemocytometer was used, and has an “H” shape engraved in the middle that encloses two mirror-like counting chambers, and provides the cover slip mounting area. The depth of each chamber with the cover slip is known to be 0.1 mm. Therefore, the volume and cell density can be determined using the equations:

$$\text{Volume} = \text{Area of the counting chamber} \times \text{Depth} \quad (1)$$

$$\text{Cell density} = \frac{\text{Number of cells}}{\text{Volume}} \quad (2)$$

2.1.5.1 Sample Preparation

The culture flask was swirled under a laminar flow hood to evenly suspend the cells before an aliquot was extracted for sampling. The cell suspension was further diluted to a 1:30 dilution in new medium to avoid having no fewer than ten and no more than 50 cells over each square millimeter in the counting chambers. Different dilution ratios were tried and we found that 1:30 was an appropriate concentration.

After cell suspension dilution, 10 μL of the suspension cells was removed from the sample flask and mixed with 10 μL Trypan Blue solution in an Eppendorf tube (dilution factor = 2). Trypan Blue is a stain that selectively stains dead cells.

Before loading the sample, the hemocytometer and the cover slip were cleaned with 70% ethanol, and dried with lens paper. The clean and dry slide was placed on a cleaned work surface and the cover slip mounted on the hemocytometer to cover the reflective counting chambers. Using a pipette, the cells were resuspended by drawing the Trypan Blue treated cell suspension from the Eppendorf tube. The suspension was then gently expelled back into the tube several times each time expelling the cells and media while moving the pipette across the bottom of the tube to gently discharge cells settling at the bottom.

The micropipette was then used to draw 10 μL of the cell suspension, and gently fill both chambers under the coverslip, allowing the cell suspension to be drawn out by capillary action. The loaded hemocytometer was then placed on a microscope stage and the counting grid brought into focus using a 40X magnification.

2.1.5.2 Cell Counting

A hand tally counter was used to count live and stained cells in all four sets of 16 corner squares (see Figure 4). Before we started counting, it was decided that we will deploy the hybrid of the quick and logical count methods to choose squares that give a good overall representation of the cells on the slide. The logical count representative approach counts the cells in the four corner squares and the middle square of the hemocytometer's grid, while the quick count method only counts the cells in the two squares that are diagonally opposite each other [22]. The hybrid method we used counted cells in all four 16 square sets because some of our cells were too small to consistently count the cells in the middle set of squares. This resulted in the quick count approach alone not being as representative as we hoped.

Furthermore, for consistency and to avoid counting the same cell twice, cells in all four sets of 16 corner squares were counted (illustrated with orange borders in Figure 4) except those on the far right edge and lower bottom edge. Consequently, the cell/mL and total cells in the sample were calculated using the formulae below:

$$\text{Cells/mL} = \frac{\text{Total cells counted}}{\text{Number of squares counted}} \times \text{Dilution factor} \times 10,000 \quad (3)$$

where 10^4 is the correction factor that converts 0.1 mm³ to 1mL,

$$\text{Total Cells} = \text{Cells/mL} \times \text{Total sample volume} \quad (4)$$

Cell counting with the hemocytometer is error-prone, with errors often as high as 20–30% [23]. We experienced issues including: the presence of air bubbles; debris in the chamber; overfilling of the chamber and incomplete filling of the chamber that resulted

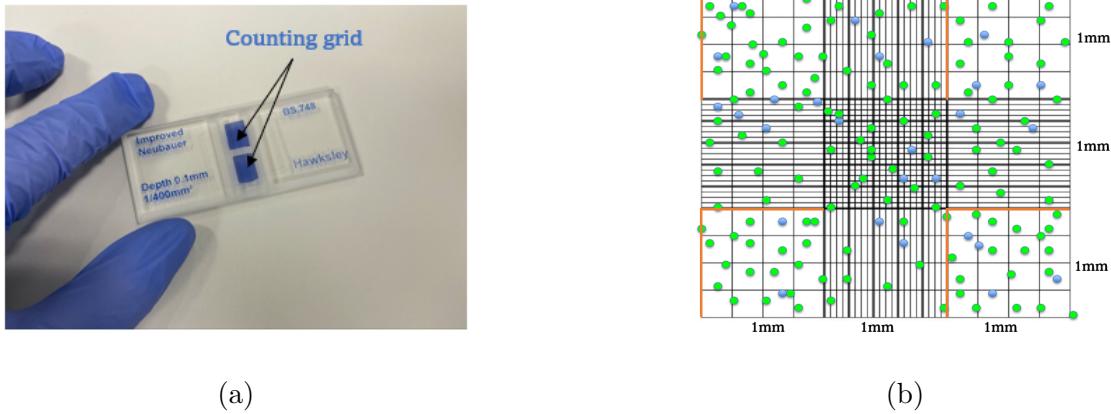


Figure 4: Hemocytometer overview: a) The counting grids are located on the reflective surface of the hemacytometer and are visible using a microscope, b) Representation of the view of one counting chamber of the hemacytometer through a microscope. The circles represent cells that had previously been cultured (green for live cell and blue for dead cells). Live and dead cells in each of the four quadrants denoted by orange borders in the counting chamber of the hemacytometer were counted.

in cells not being evenly distributed. Additionally, clustered cells caused an inability to distinguish between the cells and other sample contaminants. Issues also arose if there were too few cells to count or too many cells to count.

To overcome the above challenges and minimise errors, we sought help from personnel in the department with expertise in cell counting. With them we developed counting guidelines and took care to prepare dilutions correctly and pipette carefully. We repeated the count until the appropriate cell concentrations were determined for each strain. Typically, for each strain, cell counts were repeated at least three times.

2.2 Microscopy

This section discusses the microscopy techniques, equipment, and software used to image the cyanobacteria samples.

2.2.1 OpenFlexure

The journey of OpenFlexure can be traced back to the encounter between Dr Richard Bowman and Professor Alexandre Kabla at the University of Cambridge [24]. Professor Alexandre Kabla led the development of a low-cost 3D-printed professional microscope based on open-source technologies as part of the OpenLabTools initiative. The project aimed to build a modular, easy to replicate microscope to improve and adapt to different research purposes [25].

However, Dr Bowman observed that most microscope parts are not printed in practice. He developed an automated microscope produced using 3D printing, capable of sub-micron positioning, which we know as OpenFlexure [26]. The OpenFlexure Microscope (OFM) is an open-source, 3D-printed and fully automated laboratory microscope with motorised sample positioning and control. Initially, the microscope was co-developed between the University of Bath and the Tanzanian engineering company STICLab [27]. The OFM project aims to make automated microscopy accessible to all by enabling local

production.

The microscope is highly customisable, and the design has been refined for precise motion control and configurable high-performance imaging [28]. The microscope boasts an interchangeable optics module, enabling different imaging modalities, including trans- and epi-illumination, polarisation contrast imaging, and epi-florescence imaging.

The server software is distributed as a pre-built SD card image for a Raspberry Pi microcomputer. It includes a user-interface for more plugins for functionality beyond the basics, making the OFM appropriate for resource-limited settings whilst allowing adaptation for a wide range of applications [29, 30, 31].

The features mentioned above, as well as the open-source design, ease of production, and customisation and maintenance of the equipment without external service engineers, have contributed to the OFM's use as a prototyping tool in developing novel microscopes. In Tanzania and Kenya, locally produced OFM design microscopes are being used for educational, scientific and clinical applications [27]. Collins *et al.* also notes that the designs have been replicated in maker spaces and academic labs in numerous countries, including Peru, Germany, Ghana, the USA, and the UK [29].

Despite the several advantages, some limitations must be addressed when adapting the OFM design for a specific application [32, 29, 33]. These include:

1. Limited load capacity due to a smaller range of mechanical motion compared to mechanical stages;
2. Decreased resolution due to field curvature and aberration towards the edges of the field of view;
3. Increased noise at the edges of the image due to the Raspberry Pi camera design (the sensor is designed for use with a short focal length lens);
4. The use of the Laplacian filter for edge detection is slow on the Raspberry Pi, and;
5. Image acquisition time is significantly affected by write-speed fluctuations.

By adapting the OFM design for our project, our goal is to exploit the automation of imaging samples over long periods using time-lapse imaging. This is to monitor the long-term behaviour of biological systems in a data-efficient manner, while also contributing to the project by extending the functionality beyond the basics.

2.2.2 OpenFlexure Software

The OFM project uses a flexible and extensible software with a web of things based API and a user-friendly web based UI [34].

To control the pump and illumination system with the newly developed motor control board, two new extensions were developed, called illumination and stepstick. These bypass the regular *pysangaboard* python module and directly communicate with the motor control board. These extensions can be found in Reference [35]. The illumination extension can become part of the standard software distribution when the new motor control board becomes the default option for OpenFlexure (see Section 2.2.3).

The microscope software exposes both the standard microscope features and extension features through its API and the top level script uses this to control the microscope and the pump.

2.2.3 Custom Motor Control Board

To control the motors, the OFM uses a custom unipolar stepper control board called Sangaboard. The existing board has several drawbacks, including lack of availability and does not have all the features needed for this project.

To address these issues, we designed a custom board, taking into consideration input from many OFM users at OpenFlexureCon and advice from the original Sangaboard author. It integrates features needed for this project and makes sure this board can be adopted as the latest version of the Sangaboard for the OpenFlexure project. The board was designed using open source software (KiCAD) and is open source with design files available in Reference [36].

The new board is based on the Raspberry Pi RP2040 micro-controller which is widely available and significantly more capable than the older Atmel designs, making the new board more flexible. This choice is supported by a Sangaboard firmware rewrite [37] (mostly completed before this project), which also makes it much easier to extend the firmware with new features.

To give extra flexibility in motor control, the new design supports four motors instead of three. It features three illumination control channels (2 PWM based channels capable of running at over 100 kHz and 1 programmable constant current source for precise control). The board is a Raspberry Pi hat and uses hardware UART on the Pi for communication, eliminating the need for a USB cable between the Pi and the Sangaboard. Both devices can be powered by a single 3A USB-C connector. The board also exposes all unused pins on a header connector, making it easy to extend the system further. As the pump chosen for this project used a bipolar Nema17 stepper motor, a custom extension board was designed and manufactured to control this motor, confirming the utility of the extension option.

Testing also revealed some minor issues: a capacitor was incorrect and had to be desoldered to enable communication with onboard flash; and motor wires interfered with the shield connection, which will be solved by adding suitable cutouts to the shield design. A stackable 24 pin connector is hard to find, so the next iteration will use a 26 pin variant which is more available. A sample board has been sent to the original Sangaboard author for further feedback which will be integrated to the final version before it becomes the new OpenFlexure microscope motor control board.

The board and the pump control shield are shown in Figure 5.

2.2.4 Microscope Hardware

The microscope (displayed in Figure 6) was printed using black tough polylactic acid (PLA) with a Ultimaker 3D printer. Tough PLA was chosen for its affordability and strength, aligning with our goal for an affordable and durable device. The device was printed in black to absorb stray light, which prevents internal reflection in the optics module. The other colours of PLA can be partially transparent; printing in black ensures no external light enters the optics module.

The design was based off the OpenFlexure Delta Stage [38]. In this project we modified components and added novel features, such as a sample holder which slides the microfluidic channel in place. We focused on developing the Delta Stage for brightfield and fluorescence imaging. The differences between the microscope setup and imaging techniques will be discussed in Sections 2.2.5 and 2.2.6 for brightfield and fluorescence microscopy respectively. The CAD designs for the microscope can be found in the Open-



(a) Sangaboard v0.5.

(b) The stack including the Raspberry Pi, the new Sangaboard v0.5, and a shield to control the pump.

Figure 5: New motor control board and the full control stack.

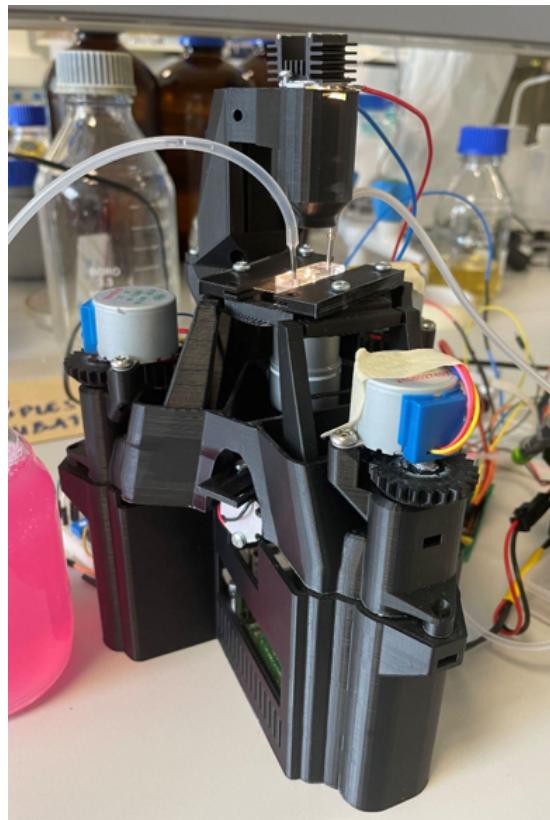


Figure 6: The modified OpenFlexure Delta Stage.

Flexure GitLab repository [39]. A list of the components used can be found in Reference [40]. The microscope was controlled using a Raspberry Pi 4B model and the images were captured using a Raspberry Pi V2 camera. The stage was moved using three 28BYJ-48 micro geared stepper motors [41], which were controlled by the custom motor control board (Section 2.2.3).

The lens used did not change for brightfield and fluorescence imaging: PMMA condenser lens with a focal length of 5 mm, LeeTun 100X oil objective (with a numerical aperture of 1.25), and ThorLabs AC127-050-A tube lens with a focal length of 50mm. These lenses were chosen to be as cost effective as possible, while obtaining images of the required magnification and resolution. The Raspberry Pi V2 camera has a pixel size of $1.12\text{ }\mu\text{m} \times 1.12\text{ }\mu\text{m}$; with the 2 by 2 colour grid [42], this creates a net pixel size of $2.24\text{ }\mu\text{m} \times 2.24\text{ }\mu\text{m}$. The smallest cyanobacteria investigated is of the order of $1\text{ }\mu\text{m}$, giving a required Nyquist frequency of 500 nm.

The objective lens has a back focal length of 150mm. Due to the size constraints of the optics module, the camera is placed in front of the image plane. To correct this distance and reduce the magnification, a tube lens is added so an image with a larger field of view is formed on the camera [43]. This is calculated via,

$$M_f = \frac{f}{f - p} \quad (5)$$

where f is the focal length of the tube lens and p is the distance to the virtual image plane (-141.5mm, as the tube lens is 8.5mm from the objective [43]), giving a magnification of the tube lens, M_f , of 0.261. The total magnification (M) of the system is given by,

$$M = M_o \times M_f \quad (6)$$

where M_o is the magnification stated by the objective. Therefore, for our system we have a total magnification of 26.1.

Despite this reduction in the expected magnification, the system was still able to resolve to the required Nyquist frequency with the pixel size of the camera (as $\frac{2.24\text{ }\mu\text{m}}{26.1} = 85.8\text{ nm} < 500\text{ nm}$). Therefore, the main resolution limit is from the diffraction limited optics and not the camera.

2.2.5 Brightfield Microscopy

In brightfield microscopy the sample of interest is illuminated with white light. The sample scatters this light, which is collected by the objective lens. The scattered light interferes and creates contrast (from the differing phases), which creates an image of the sample when projected onto the camera by the tube lens. For this imaging technique, the transmission illumination setup was used as shown in Figure 7.

A white light LED was used with a heat sink to aid in heat dissipation. As brightfield microscopy is a coherent imaging method, the resolution was calculated using Abbe theory,

$$r_{\text{brightfield}} = \frac{\lambda}{NA} \quad (7)$$

where $r_{\text{brightfield}}$ is the resolution, λ is the wavelength of light, and NA is the numerical aperture of the objective. Due to the absence of data from the LED manufacturer, the resolution was obtained for the range for visible light (400 nm – 700 nm), which was calculated to be 320 nm – 560 nm.

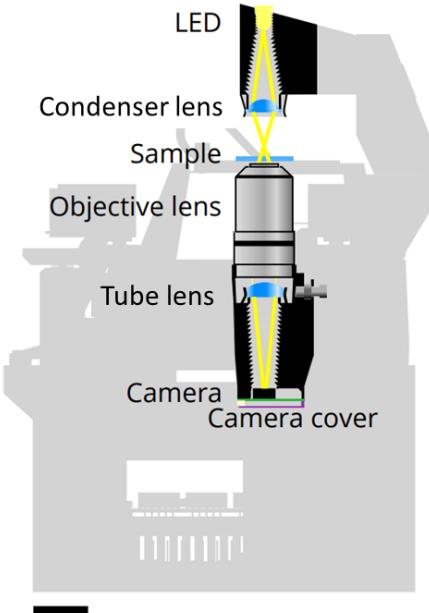


Figure 7: Schematic to show the transmission setup for the OpenFlexure Delta Stage. Adapted from Figure 2a from [41]. The scale bar represents 2cm.

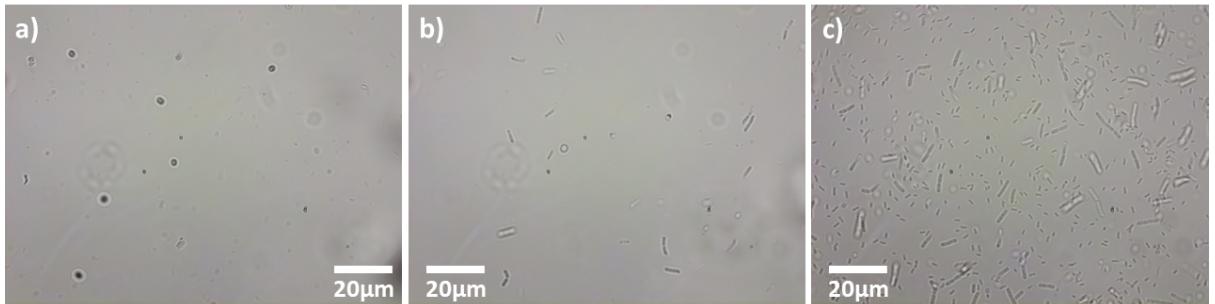


Figure 8: Brightfield images of a) *Synechocystis* sp. PCC 6803, b) *Synechococcus* elongatus PCC 7942, and c) *E. coli*.

The depth of field (d) in the image is dependent on the objective lens' NA and the wavelength of the illumination light (λ). To calculate the depth of field for the brightfield setup,

$$d = \frac{\lambda \cdot n}{NA^2} + \frac{n \cdot e}{M \cdot NA} \quad (8)$$

where n is the refractive index of medium on the objective lens, e is the smallest distance resolved by the camera, and M is the magnification of the objective. Similar to the resolution calculation, the depth of field was calculated for the range of visible light. Given that the refractive index of the objective oil was 1.5, the effective magnification for the 100X objective in this system was 26.1, and the smallest resolvable distance is 2.24 µm, the depth of field for the brightfield setup was calculated to be 384 nm – 672 nm.

The images obtained with this setup are shown in Figure 8. The camera was set to have a high digital gain and short exposure time, to reduce image blur of the cyanobacteria moving through the channel.

An advantage to brightfield imaging is the ease of finding the channel and obtaining optimum focus on the moving cyanobacteria. However, low contrast arises from the illu-

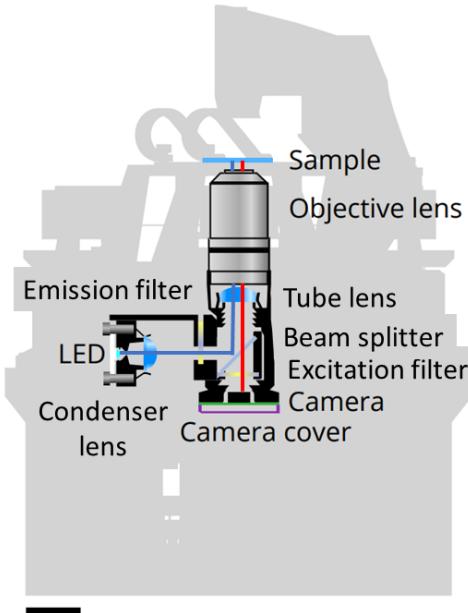


Figure 9: Schematic to show the reflection setup for the OpenFlexure Delta Stage. Adapted from Figure 2b from [41]. The scale bar represents 2cm.

mimation light also being transmitted onto the camera. Additionally, dirt and dust on the optics and imperfections in the microfluidic channel are imaged, creating a background. Though the background is easy to identify (as the background does not move in between images), the background could still be counted as cyanobacteria by the machine learning algorithm. To ensure that this did not happen, background images were taken with only growth medium in the channel.

2.2.6 Fluorescence Microscopy

To improve contrast and eliminate background, we used the reflection illumination setup (Figure 9) to image the cyanobacteria using fluorescence microscopy. Cyanobacteria contain chlorophyll-like pigments which autofluoresce [44]. In fluorescence microscopy these autofluorescent pigments fluoresce at a certain wavelength. These can be imaged by exciting the pigment with a shorter wavelength and measuring the emitted red shifted wavelength. By knowing the wavelength of the fluorescence and using the correct filters, the cyanobacteria can be selectively imaged.

To determine the filters and LED required, a Tecan Spark plate reader was used to obtain the fluorescence absorption and emission spectra of the cyanobacteria samples. Figure 10 shows that each strain absorbs at around 450 nm and Figure 11 displays that each strain emits at 690 nm. Therefore, a royal blue LED with a peak wavelength of 450 nm, a short pass 550 nm excitation filter, a long pass 600 nm emission filter, and a 560 nm cut-on dichroic mirror were added to the modified optics module. All the filters were from Comar Optics, as they can be easily cut to size and were affordable compared to other filters on the market.

As fluorescence microscopy is an incoherent imaging method, the Rayleigh criterion was used to calculate the resolution,

$$r_{\text{fluorescence}} = \frac{\lambda_{\text{ex}}}{2NA} \quad (9)$$

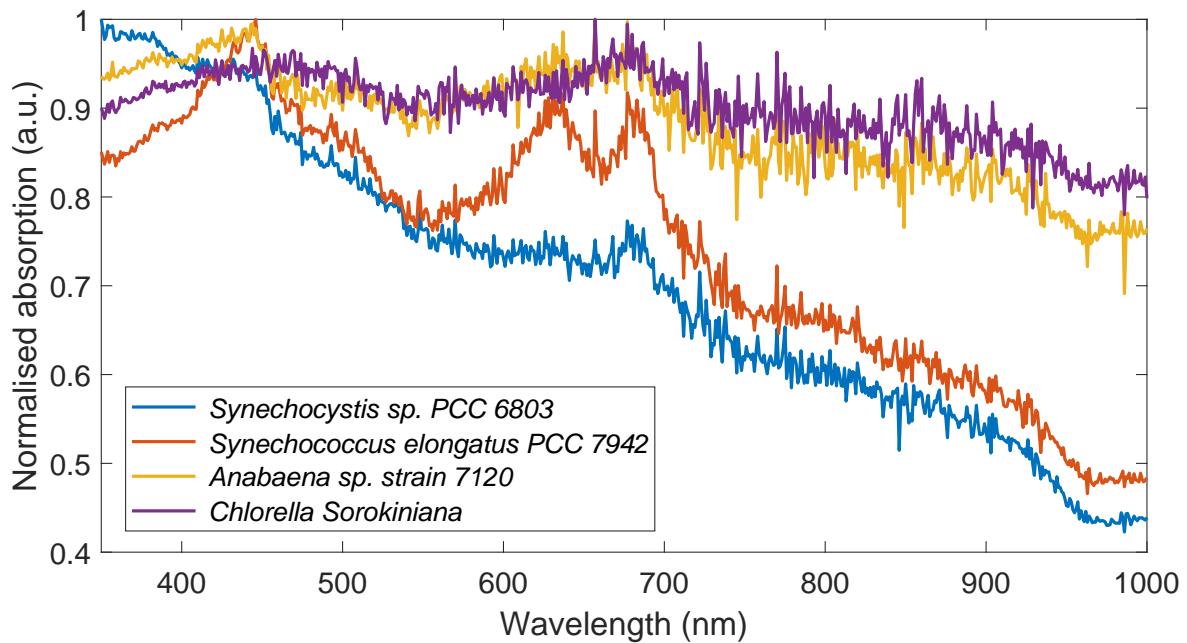


Figure 10: The fluorescence absorption spectra of the cyanobacteria samples *Synechocystis* sp. PCC 6803 (blue), *Synechococcus* elongatus PCC 7942 (red), and *Anabaena* sp. strain PCC 7120 (yellow), as well as the false positive *Chlorella* Sorokiniana (purple).

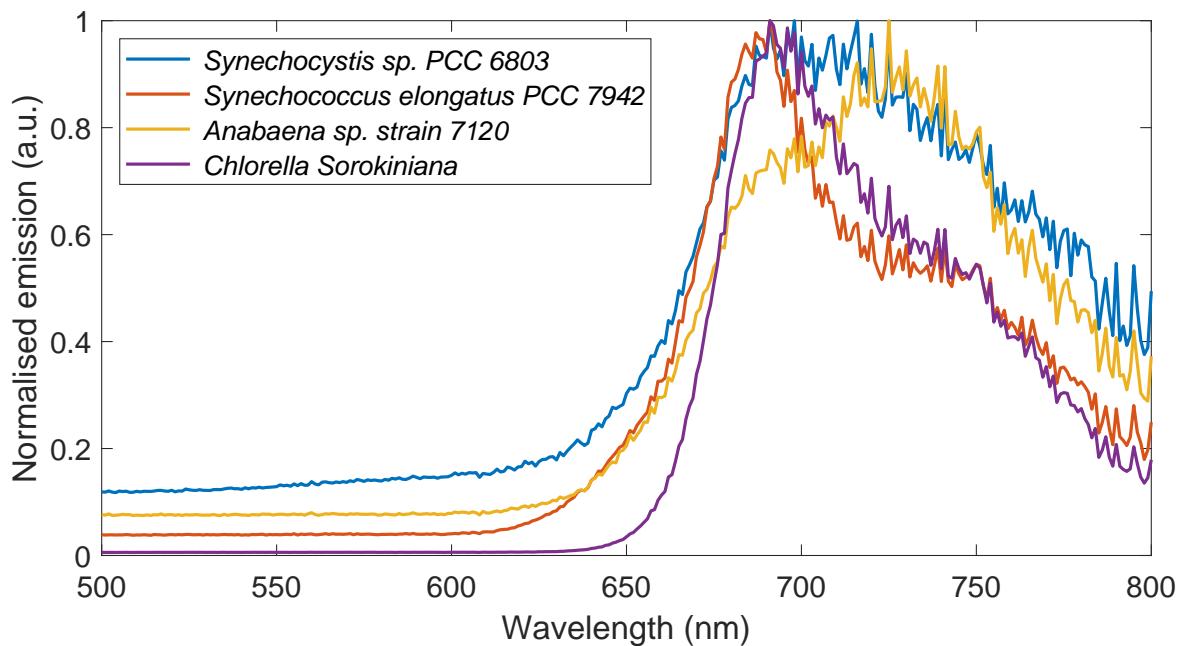


Figure 11: The fluorescence emission spectra of the cyanobacteria samples *Synechocystis* sp. PCC 6803 (blue), *Synechococcus* elongatus PCC 7942 (red), and *Anabaena* sp. strain PCC 7120 (yellow), as well as the false positive *Chlorella* Sorokiniana (purple).

where $r_{\text{fluorescence}}$ is the resolution and λ_{ex} is the wavelength of the excitation source. The blue LED has a peak wavelength of 450 nm; therefore, the resolution of the fluorescence setup was 180 nm. Using Equation 8 and substituting the excitation wavelength of 450 nm, the depth of the field for the fluorescence setup was calculated to be 432 nm.

The initial fluorescence images are shown in Figure 12, where the sample was pipetted onto a slide to remain stationary.

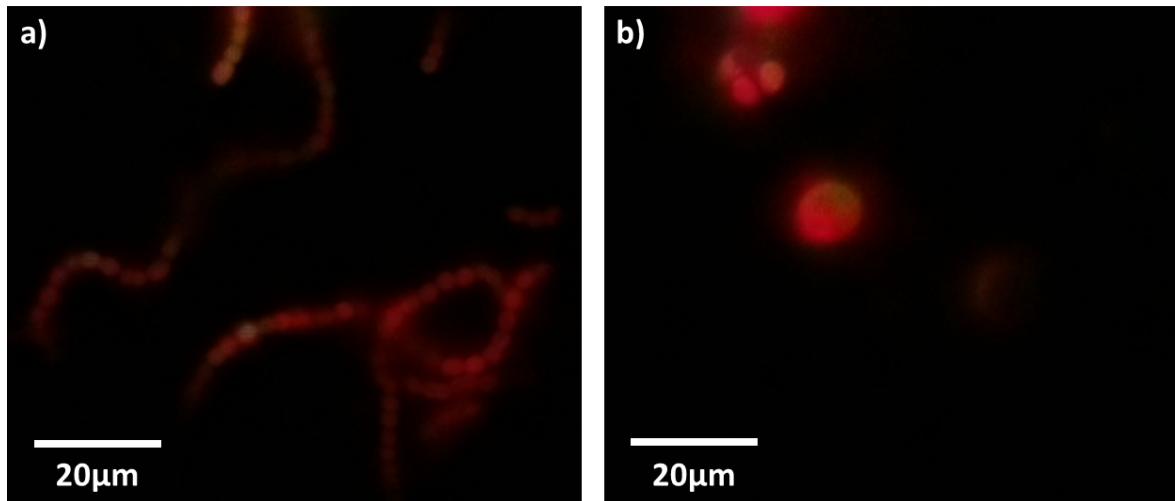


Figure 12: Fluorescence images of a) *Anabaena sp. strain PCC 7120*, and b) *Chlorella Sorokiniana*.

Images of the samples in the microfluidic channel were unable to be obtained due to the long exposure times required. Longer exposure times are required to obtain a stronger fluorescent signal, as more light is able to hit the camera. As the sample moves, the intensity of the signal is spread over multiple pixels, effectively reducing the fluorescence intensity and contributing to poor image quality due to motion blur.

In addition to this, obtaining a fluorescence signal from the samples was inconsistent. It was hypothesised that the cyanobacteria were dying due to not being split regularly enough and the cells subsequently not being supplied with sufficient growth medium. For future work, better working protocols for handling the cyanobacteria need to be put in place.

During initial testing, it was confirmed that the samples did fluoresce using a more advanced lab-based fluorescence microscope with a 60X oil objective (Figure 13). The images shown in Figure 12 were obtained shortly after. Future project and microscope development could focus on obtaining extended access to the lab-based microscope to ensure the samples autofluoresce, before running further fluorescence imaging using the Delta Stage. Initial iterations were explored, but were proven to be unsuccessful, including increasing the current through LED to increase the excitation power.

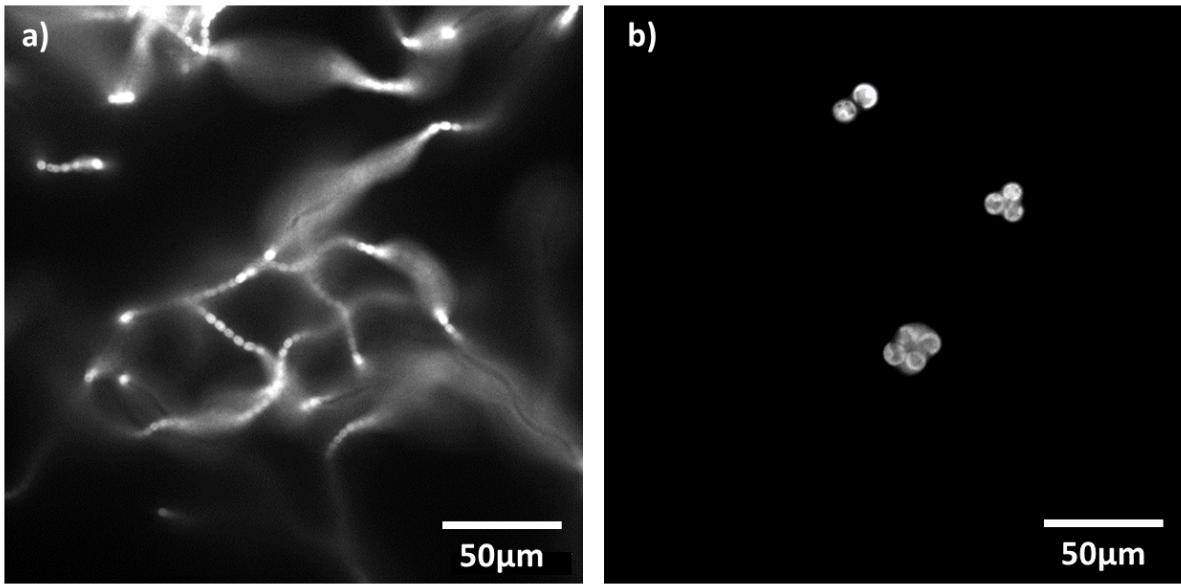


Figure 13: Fluorescence images of a) *Anabaena* sp. strain PCC 7120, and b) *Chlorella Sorokiniana* obtained using an advanced lab-based microscope.

2.3 Microfluidics

In this section, the development and final implementation of the microfluidic system is discussed.

2.3.1 Microfluidic System Overview

Figure 14 shows the overview of the microfluidic system. The main implementation is to allow for continuous flow which facilitates the continuous monitoring of cyanobacteria in freshwater systems. This is done by having a single stream, where a sample is pumped using a peristaltic pump though a filter and flow sensor before reaching a microfluidic channel, where the sample can be imaged using the OFM. The sample can then be returned unaltered to the environment. In the following sections, each part of the system are expanded on.

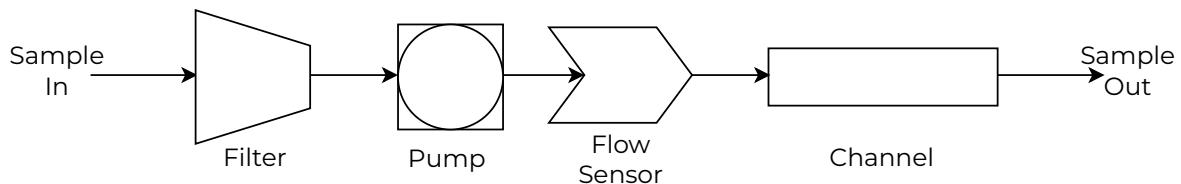


Figure 14: Microfluidic System Structure.

2.3.2 Filtration

In the prototype development, a Bartels mp-f filter with a porosity of 20–60 μm was used in conjunction with a surgical mask to filter environmental samples. This regime is referred to as normal flow filtration (NFF). Filtering environmental samples is necessary to ensure that fouling is minimised downstream in the system, particularly in the

microfluidic channel. This is due to the very small (20 µm by 200 µm) dimensions of the channel, making them highly susceptible to blockage from grit and other large particles that can be suspended in freshwater systems. As the system is intended to be somewhat autonomous with little operator interaction, great care needs to be taken to prolong the lifespan of the microfluidics. By having this filter setup, we can ensure that the lifespan of the channel is extended. However, as an NFF system was employed, this increase is unlikely to be long. Due to time constraints and prioritisation of more relevant parts of the system such as the pump, we were unable to test the longevity of the filters. However, it is expected that the filters will likely be fouled in less than a month, due to the small size of the filters utilised as well as the NFF setup. To improve this in later iterations, it would be worth implementing tangential flow filtration (TFF). This is used in many settings from microfluidics to pharmaceuticals to both prolong filter lifespans and increase material throughput [45]. It works by flowing a sample over a membrane instead of through it, allowing a more gentle filtration. This means that the membrane has an increased lifespan, as it takes longer for it to be fouled. A system like this would be very suitable for Cyanovision given that only small volumes of liquid are under test.

2.3.3 Microfluidic Pump

During development, two types of pump were tested. First, a Bartels mp-6 was used. This is a piezoelectric pump that has a low power consumption of typically 50 mW [46] and is developed specifically for use in microfluidic systems. It is controllable through I²C which made it very attractive for an autonomous battery powered system. However, in the lab it was found that the back pressure caused by the microfluidic channel meant the pump did not have the capability to flow liquids at a high enough rate for our application. Additionally, any air present in the liquid flowing through the pump would cause it to stall, requiring manual interaction to rectify. Though a bubble trap was investigated to prevent this, it was found that it was not effective in reducing the amount of air present. Tighter fittings and other models will need to be investigated in any future development.

The other type of pump considered was a peristaltic pump. These work by using a rotor to press a flexible tube containing a fluid to induce flow [47]. This has the benefit of isolating the pump hardware from the samples, ensuring no fouling of the system. Initially, 3D printed pump designs were investigated. This was to fit in with the affordability and open access nature of the project, allowing the technology to be as accessible as possible. However, this was problematic as the tolerances required to reliably print a functioning peristaltic pump were too great for a regular 3D printer.

This motivated the final pump which was a Fdit 40 stepper peristaltic pump. This had the benefit of being able to support higher pressure flow to allow the effective pumping of liquid through the system, whilst being easily controllable in software. This is due to the stepper motor driving the pump being compatible with the motor driver routines of the OpenFlexure software. This was very effective when used in conjunction with the flow sensor to give a closed loop feedback on the system. Typical flow rates achievable with this regime when changing samples were 25 µL/min to 30 µL/min which was adequate for the long term monitoring envisaged.

2.3.4 Flow Sensor

It is key for the effective use of the peristaltic pump that there is feedback from a flow sensor to ensure when images are taken there is minimal motion blur. This is to ensure

the machine learning image analysis can correctly analyse and distinguish particles in the water. As the flow rates achieved by the pump are so large, the camera is not capable of tracking particles in motion due to their velocities. Therefore the flow sensor feedback is required to stabilise the samples for measurements.

The Sensirion SLF3S-0600F flow sensor was used in this project. It uses I²C communications so easy to interface with wider system software. It has a resolution of 1 $\mu\text{L}/\text{min}$ with a range of $\pm 2000 \mu\text{L}/\text{min}$ with a response time of 20 ms making it suitable for the low flow rate of the system [48].

2.3.5 Microfluidic Channel

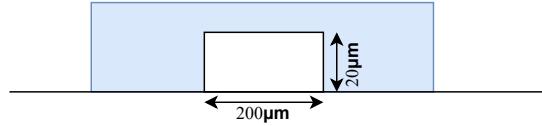
The key component of the microfluidics system is the microfluidic channel. This is the device that will constrain the sample in space for imaging while allowing continuous flow for environmental monitoring. As shown in Figure 15, the channel used in the project is made of polydimethylsiloxane (PDMS) which is bonded to a glass slide and is mounted to the microscope using the dedicated sample holder printed for our application.

When developing the system, several channel dimensions were tested including those of 80 μm depth. As shown in Figure 15a the final dimensions chosen were 200 μm by 20 μm . Firstly, as discussed in Section 2.2.5 and 2.2.6, the depth of field is very shallow, being of the order of hundreds of nanometres. As the intention is for continuous monitoring, it is not feasible to fix samples to a glass slide. Hence, to increase the probability of sample flowing across the field of view in focus, the height of the channel should be as shallow as possible. Through testing different dimensions, it was determined that a height of 20 μm is the optimal configuration. This is due to both fabrication limitations and back-pressure caused by the channel dimensions. To reliably fabricate PDMS channels, dimensions below 20 μm are challenging due to the increased risk of channel collapse [49]. As the dimensions of the fluidic path decrease to meet the channel the back pressure increases, requiring the pump to have much greater pressure to support flow [50]. With these two limitations, the furthest we could reliably push the dimensions in the lab were as already discussed.

The width of 200 μm was chosen as it ensured that the side walls of the channel are not in the field of view of the microscope. This is necessary to ensure that the machine learning algorithm does not erroneously detect cyanobacteria because of textures viewed in the side walls of the channel. Ideally, there would be scope to have a wider channel, but the molds used to produce microfluidic channels in this project did not support widths greater than 200 μm . It is thought that greater width would lead to a better system as the depth will remain constrained but the back pressure can be reduced. Due to time restrictions, we were unable to create a new mask to produce and test wider channels.



(a) Microfluidic Channel on OpenFlexure Stage.



(b) Microfluidic channel cross section. The blue region indicates PDMS and the central blank region is the channel.

Figure 15: Microfluidic Channel.

2.4 Systems and Design

This section details the work that went into designing and constructing the system and why certain design choices were made. This section concludes by discussing the performance of the system.

2.4.1 System Overview

The system can be split up into roughly three parts: the microscope and microfluidic system, power system, and compute and communication system. The individual components of these systems and their connections are shown in Figure 16. The Figure shows what parts of the system belong to each subsystem. The whole device is controlled by a Raspberry Pi which communicates with the other elements of the system. The Pi runs a master program that makes calls to the different services needed such as the Coral, OpenFlexure server and communications device. A second key component is the motor controller board. This is a custom board designed for this project based on the original OpenFlexure Sangaboard. The board allows the system to power and control the pump and illumination whilst taking commands from the OpenFlexure server for ease of use. This board is discussed in detail in Section 2.2.3.

The microscope and microfluidics section represents the microscope hardware including: OpenFlexure Delta Stage; illumination; pump, motors; flow sensor; microfluidic channels; inlet and outlet tubing. The setup is further discussed in the Microscopy (2.2) and Microfluidics (2.3) sections. The power system includes: voltage down converter (Buck); charging circuit; solar panel; and LiPo batteries. This system is further discussed in the Power Systems (2.4.3) section, but the goal was to design a system that can sustain itself in reasonable sunlight. The compute system includes: Raspberry Pi; 4G and GPS dongle; and the Coral Edge tensor processing unit (TPU). This system controls the device, communicates internally and externally as well being able to quickly run ML models. The Raspberry Pi used is the 4B 8GB model, this is the most powerful and most expensive model available. It was chosen because more computing power was desirable for testing the system, but a user could recreate the system with a less powerful, cheaper

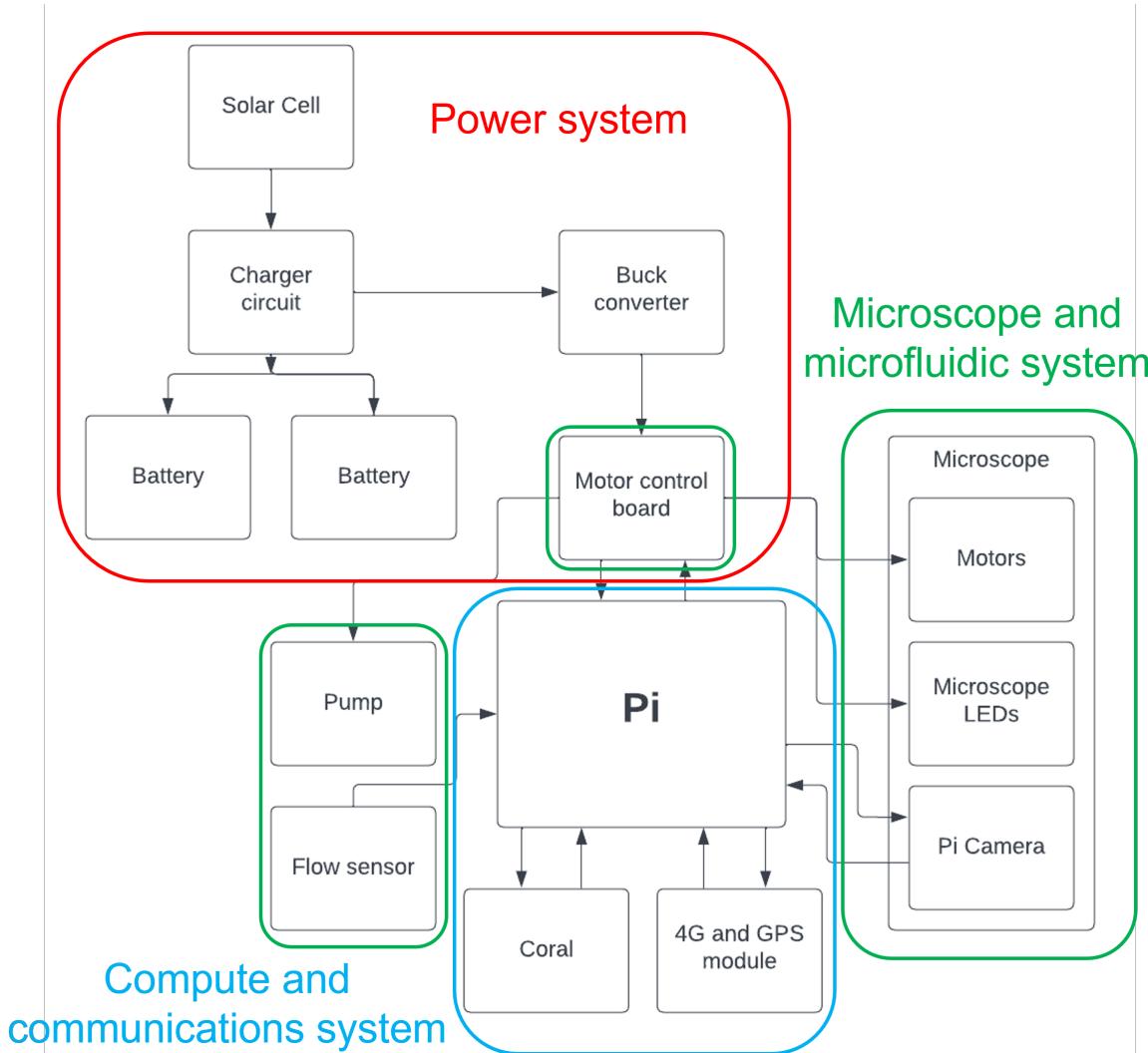


Figure 16: System flowchart.

Pi.

2.4.1.1 Costing

The cost for implementing the system comes to around £900. A break down of the cost into different parts of the system is shown in Figure 17. However, some of the cost reported was used for testing parts and as stated a high end Raspberry Pi was used. An end user may find the pricing is lower than this, especially if building multiple devices as certain parts cannot be ordered as one item.

2.4.2 Hardware Design

The hardware design of the device had the specification of being: easy to move; resilient to being continuously deployed in the wild; and contain all the different developed systems. To realise this we started finding a watertight casing that could fit all the components inside.

To design the device we used a CAD software. Figure 18 shows a screenshot of the different layers designed in CAD. It can also be seen how the layout was made by adding

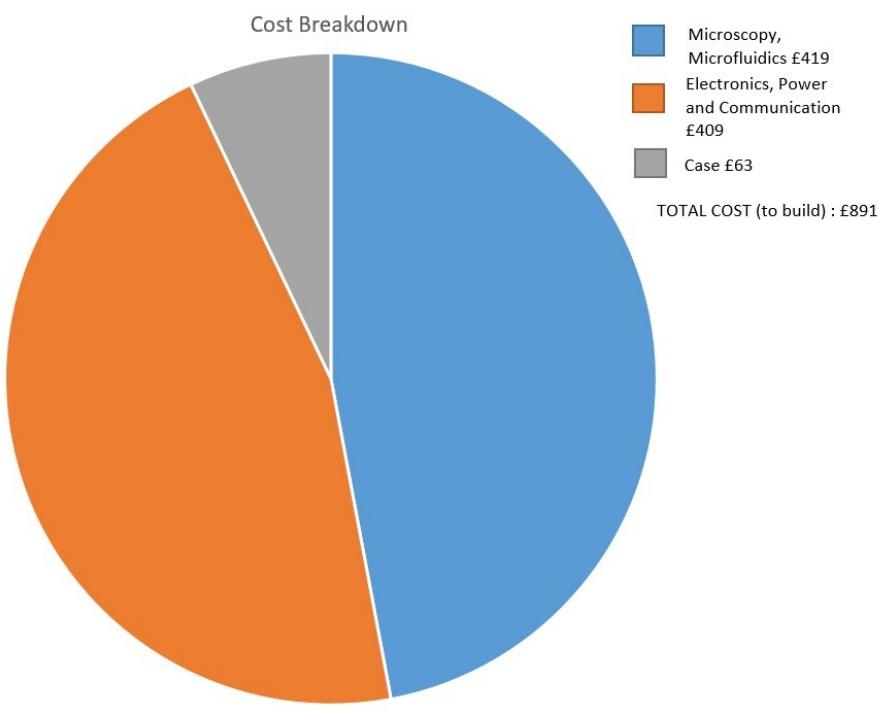


Figure 17: Pie chart of cost for different parts of the system.

proxy shapes for the different parts of the system to correctly position the mounting holes. With this done, the CAD files can be sent to a laser cutter device to create the cutouts for each layer from acrylic. The CAD is also flexible should a user wish to modify the design. For example, they may wish to add more batteries in which case the holes can be added to the CAD and printed. Four holes, one at each corner, of the layers can be seen that are used to hold the layers in place with an M5 threaded bar running through. Space has been left between the sides of the layers and the casing. This gives space for the battery cables to fold around the side to meet the power system.

The prototype created for the project is shown in Figure 19, with the casing and layered build shown. It can be seen on the casing where the charging port and cutout for the power button are. On the other side of the case two ports for the sample inlet and outlet are cut. The M5 threaded bars used to support the layers with a nut and washer on each side are shown in this Figure, along with the 3D printed feet that sit into the cutouts at the bottom of the case. The device is made so that 1 m of bar, the easiest to purchase, cuts into four sections to use as the supports. If a different case were chosen then the feet could be redesigned if needed with the M5 thread hole for the bar to screw into.

2.4.3 Power System

The power system has been designed to support long term off-grid deployments of the device. We estimated that to achieve this goal, the system should be fully functional on battery power only for at least 30 hours, with a solar panel with enough power to recharge the system in approximately 4 hours of direct sunlight. With the panel producing some power even in less than ideal conditions, this combination of batteries and solar should allow it to run continuously in most situations. Furthermore, sampling frequency could be reduced or the system could be kept idle if there is an extended period of no sunlight

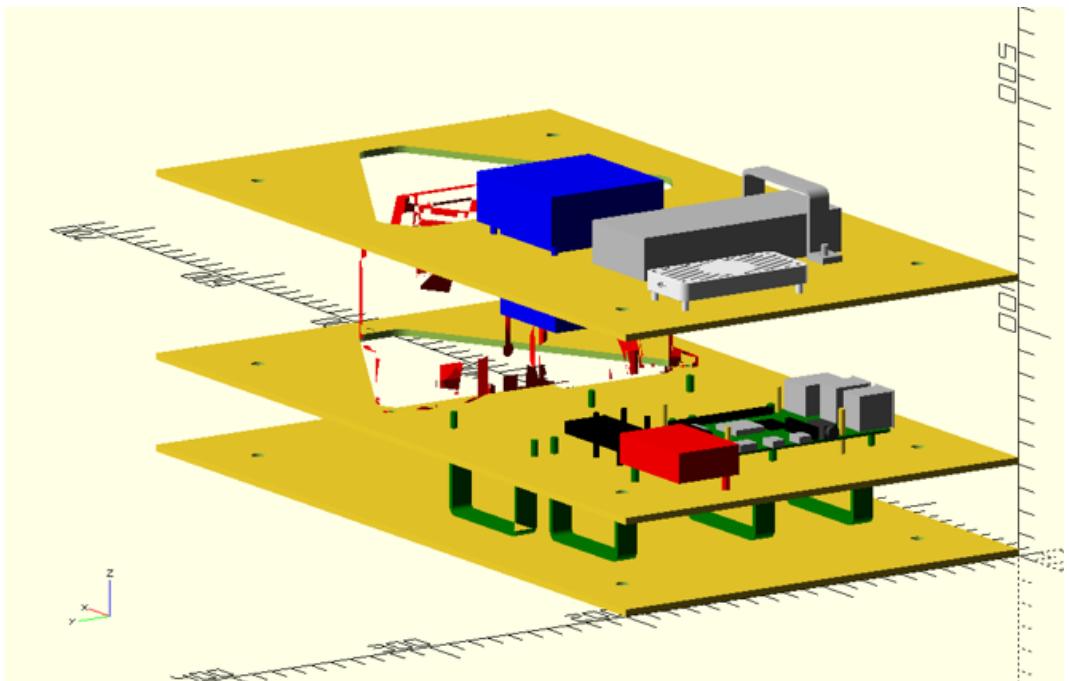


Figure 18: Image of the CAD design showing the layers used.



Figure 19: System in its final casing (left) and a built version of the layered design (right).

draining the batteries.

Based on the estimated average power draw of approximately 6.5 W we selected 2 9000 mAh 3S LiPo batteries, giving a total of 200 Wh of nominal capacity. LiPo batteries offer a good combination of power density and safety. To achieve lower costs in low resource areas a standard 12 V car battery could be used instead.

The battery pack is charged using an off the shelf solar charging circuit with maximum power point tracking with input from a 19V solar panel through a barrel jack connector. This appears to be the most common nominal voltage for portable solar panels, so this allows most flexibility in solar panel choice. The solar panel chosen for our build has 60W of nominal output to satisfy the 4 hour per full charge requirement.

The 11.1 V of the battery pack is stepped down to 5 V using an off the shelf buck converter module. This then runs everything except the peristaltic pump, which requires higher voltage and is driven off the raw battery output.

2.4.4 Communications

The systems team implemented a communications system to transfer information from the device to the user. This is centred around the SIM7600G-H 4G dongle, which creates a 4G network for sending data. The dongle also includes a GPS module and supports SMS messaging, which may be used where there is no data available, data network, or an urgent message needs to be sent.



Figure 20: The SIM7600G-H 4G dongle.

Compared to alternative networking solutions such as WiFi modules, HAT-based 4G modules and simple ethernet connections, the dongle is relatively affordable. Also it is flexible enough to enable communications wherever there is a mobile network while interfacing conveniently with the Raspberry Pi via USB. Initialisation of the device is simple by sending AT commands through the Python serial communications library, meaning it can be integrated directly into the master program.

The communications system makes use of two main features. Firstly, it connects with a Google Sheets document using the `gsheets` Python library and the Sheets API of the corresponding Google account, enabled by saving the OAuth Client details on-device. This allows the document to be updated automatically in real time with the GPS location, timestamp and current state of the system. Such a straightforward structure

Device	Flexible?	Low Cost?	Easy to Interface?
WiFi Dongle	X	✓	✓
Ethernet	X	✓	✓
4G HAT	✓	X	X
4G Dongle	✓	✓	✓

Table 2: Comparison of different communications modules.

allows for a network of Cyanovision devices to be deployed and monitored simultaneously, allowing for a wide geographical area to be covered in more detail.

Secondly, the system connects with Google Drive using the PyDrive Python library, again by carrying out the necessary authorisation steps with the Google Drive API. In the event of a positive system reading, the images are automatically uploaded to the drive for a third party to inspect and distribute them to any other stakeholders. This provides a platform for a human-in-the-loop system, giving scope to improve model accuracy over time. Data bandwidth and power is saved by only sending positive images for human inspection.

2.4.5 Performance and Limitations

The power consumption of the entire system was measured using an INA219 current sensor connected between the battery and the device, sampling every 200 ms. The power consumption during different phases of operation and a worst case scenario when every part of the system is being used at the same time, including a benchmark running on the Raspberry Pi, is shown in Figure 21. Average powers obtained during this test are listed in Table 3. The average power is lower than our earlier estimates. This means that only using a single larger battery, or a pair of smaller batteries would likely be suitable. With the current power system, the device should be able to continue working for approximately 48 hours with no external power.

Activity	Total power / W	Net power / W	Estimated on time per cycle
Idle	3.18 ± 0.36	6 h	
Pump active	8.18 ± 0.39	5.00 ± 0.53	1 h (assuming 2 m tube)
Pump holding	4.54 ± 0.36	1.36 ± 0.52	0
Stage moving	7.69 ± 0.34	4.51 ± 0.50	5 min
GPS	4.05 ± 0.54	0.870 ± 0.64	30 min
Illumination	9.54 ± 0.32	6.36 ± 4.83	5 min
Weighted total	4.2 W		

Table 3: Power consumption with its standard deviation during different phases of the imaging cycle and estimated times spend in these phases during an assumed 6 hour cycle. The expected average power of 4.2 W is lower than the estimate used when designing the power system and a smaller battery would likely be able to satisfy the capacity requirements.

Performance of the pump system was measured using the flow sensor discussed earlier in Section 2.3.4 to determine flow rate response at different pump motor speeds. The

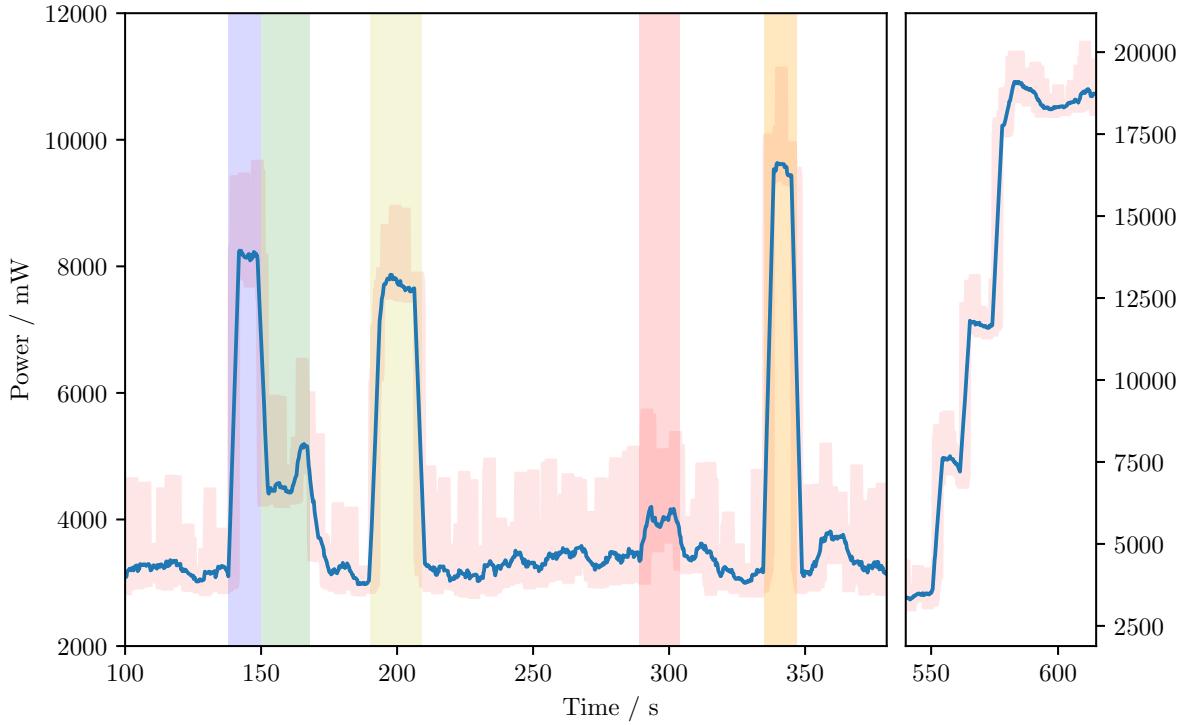


Figure 21: Power consumption at different phases of operation. The highlighted sections represent in order: pump on; pump holding; stage moving; GPS attempting a fix; LED on. The right side of the Figure shows power consumption with all of these on at the same time and verifies that the power system can handle this worst case scenario. During this test, battery voltage dropped from 11.31 V to 11.00 V representing only a minimal drop below the packs nominal voltage of 11.1 V with no effect on the rest of the system.

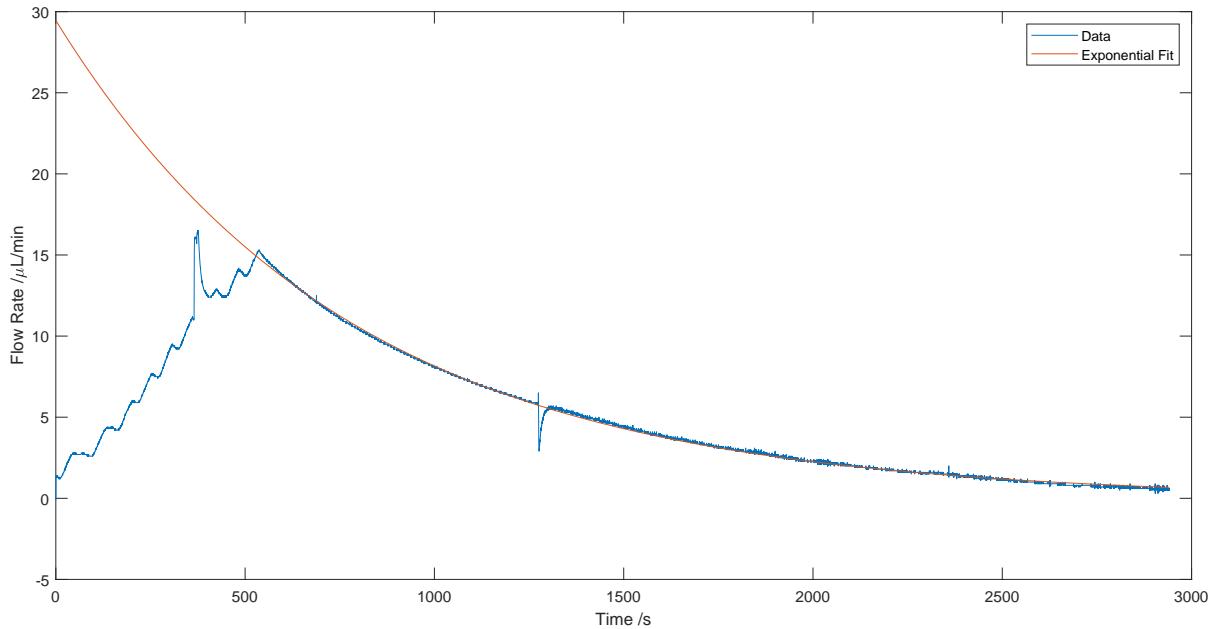


Figure 22: Flow rate decay when starting at $15 \mu\text{l}/\text{min}$ in a $150 \mu\text{m}$ by $20 \mu\text{m}$ channel. The exponential fit shows in this configuration the flow reduces by a factor of 2 in nine minutes.

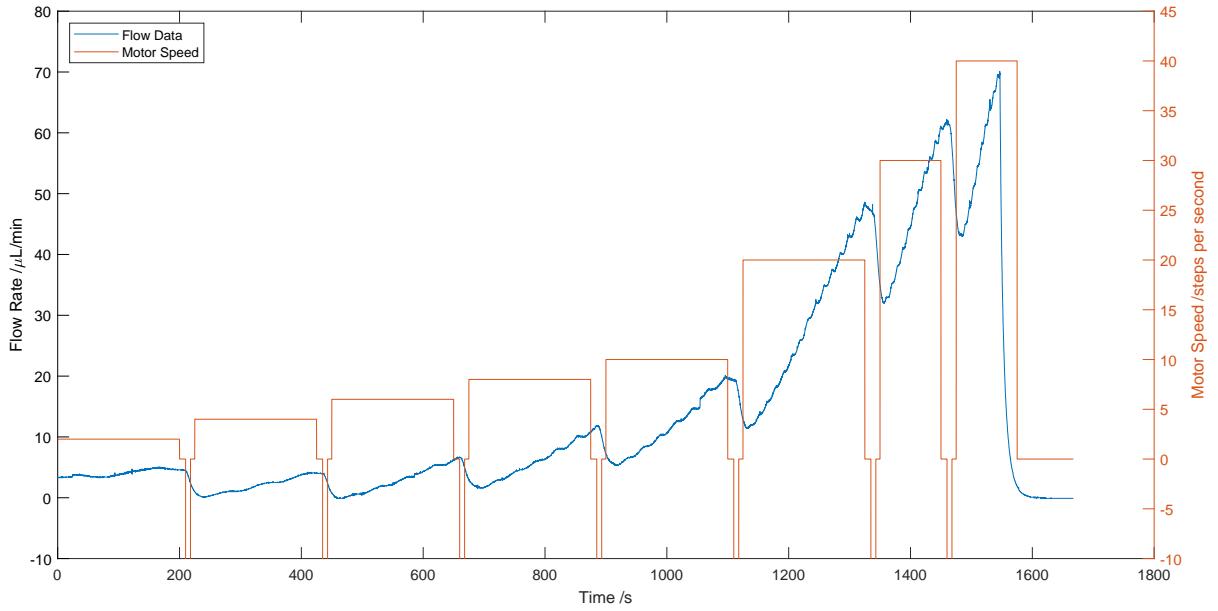


Figure 23: Flow rates in a $150\text{ }\mu\text{m}$ by $20\text{ }\mu\text{m}$ channel when the pump is ramped. At 1550 seconds, the system failed due to one of the tubing joints detaching, hence the drop to $0\text{ }\mu\text{l}/\text{min}$. In orange are the pump speeds over time. When the speed is negative, it is running at -100 steps per second. This is cropped for clarity.

results are shown in Figure 23. The resistance of the microfluidic channel and the flexibility of the tubing causes significant capacitance in the system before the channel. The pressure inside the tube increases and the tube expands slightly, so even when the pump is switched off pressure is maintained. This very slowly decreases by flowing the liquid through the channel. As shown in Figure 22 it takes nine minutes for the flow rate to halve. This effect can be counteracted by reversing the pump direction to equalise the pressure on both sides of the channel. At high pump speeds, the pressure increase causes failure of the channel or the tubing joint,s so pump speeds need to be kept slow when there is liquid flowing through the channel. Initial system loading can be performed at high flow rate, but this must be stopped before the liquid hits the channel. As shown in Figure 23, the point of failure is $70\text{ }\mu\text{L}/\text{min}$ which corresponds to a pump speed of 40 steps per second.

To deliver a fresh sample to the channel, assuming a tubing length of 2 m and 1 mm inner diameter, requires the pumping of 1.57 ml of liquid, at $30\text{ }\mu\text{l}/\text{min}$. This requires 53 min of pumping whenever a new sample is loaded.

2.5 Machine Learning Solution

2.5.1 Machine Learning Models for Cyanobacteria Detection

Apart from designing the hardware-based part of this project, completing it required developing significant parts of software as well. Our microscope system is able to capture images of the sample flowing through the microfluidic channel. However, there needs to be a machine learning solution analysing these pictures to detect whether they contain large concentrations of cyanobacteria or algae. As such, we have experimented with various machine learning solutions, ranging from Convolutional Neural Networks (CNNs) for classification and UNet models for segmentation to end-to-end models for object

detection.

Our final machine learning solution consists of a TensorFlow-based object detection model that is capable of finding algae and cyanobacteria of various morphologies. It can detect filamentous, rod-shaped and unicellular cyanobacteria in the samples fed to it. Additionally, we have designed a desktop-based software with a user-friendly interface that can be used to train this model, making our system more accessible. Throughout the development of the machine learning aspects of this project, the ability to run inference on-device directly on the Raspberry Pi embedded in our microscope hardware was also of utmost importance. Therefore, we always tried to make our models as lightweight as possible for running them at an Edge TPU accelerator.

2.5.1.1 First Steps: Classification CNN Model & Segmentation UNet Model

As an initial step towards the development of the machine learning part of this project, we explored publicly and freely available microscopy imaging datasets. This was necessary as the images from our team's microscope could not be captured until all the hardware components of the system were in place. The completion of the hardware-based parts of the system relied on the respective components getting selected, ordered, delivered and assembled into one functional unit. Also we were constrained by the various data collection experiments getting completed. This was a time-consuming process, so to overcome a potential bottleneck on the development of the machine learning solution, we used a labelled dataset of environmental microorganisms [51] and started exploring and building various models using it.

An initial model that we developed was a sequential CNN [52] for image classification with eight layers, including three 2D convolutional blocks, a block to flatten the input, and a dropout layer to avoid overfitting the images amongst others. After carefully optimising the hyperparameters of this model, we trained it using data from the environmental microorganisms dataset for 30 epochs both in a binary and in a three-class classification manner. Through these experiments and by following a 70%/30% training/validation split, we were able to achieve an $\approx 80\%$ and $\approx 70\%$ prediction accuracy, respectively, for images on this dataset.

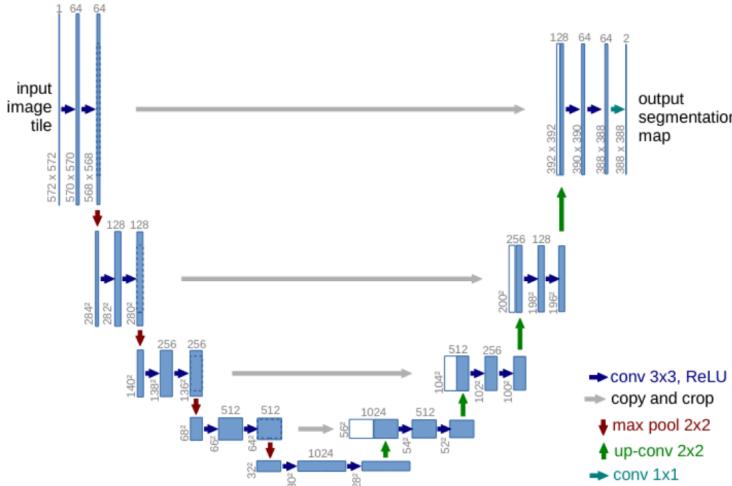


Figure 24: U-Net Architecture [53].

A further model we developed during the early phases of the machine learning progress

was based on the U-Net architecture [53] for image segmentation. As the images in the microorganisms dataset all featured a significant amount of background, we also expected this to be the case with ours. Since this is the case with almost all microscopy datasets, we identified image segmentation as the optimal approach to isolate only the objects of interest in the image. Despite some pre-trained generalist algorithms for cellular segmentation like Cellpose [54] existing, we chose to build and optimise a custom-made U-Net model. This is more flexible and allows us to compile it in a TensorFlow Lite format for deployment on the Edge TPU accelerator connected to the Raspberry Pi of our microscope system. A graph depicting the U-Net architecture can be found in Figure 24.

Although the results of both early-stage models were encouraging, throughout carrying out the relevant experiments we identified that all the samples we were examining from the publicly available dataset featured one or only a few microorganisms per image. As such, once the images from our team’s microscope system started becoming available and we identified that they included a large number of bacteria per image, we realised that shifting our focus to object detection would be the optimal way forward for the task in hand.

2.5.1.2 Final Approach: Object Detection Model

As the first samples from our microscope were captured and after thorough consultation to refine the project objectives, we shifted our focus to building a machine learning model for object detection. Object detection aims to automatically locate one or more objects of the same or various types in an image. This is similar to how humans would be able to recognize a given object, but in an enhanced and more precise manner using computer vision models. Object detection involves drawing a bounding box around each of the objects identified in an image, while combining with the task of classifying each of the identified objects and assigning a label to them.

To create our object detection solution, we explored numerous frameworks and options. Given the fact that we wanted our model to be as lightweight as possible for running on the Coral accelerator, we identified the TensorFlow Lite Model Maker library [55] as being the optimal toolkit to develop it. This allows importing our own pre-processed dataset for training and uses transfer learning to reduce the amount of training samples needed. This was of utmost importance for our use case, as the training dataset that was captured during the course of this project was relatively small, with approximately 110 samples. This approach was the optimal solution for the specific requirements of this project and the size of the dataset that we had available.

After feeding our processed dataset in a CSV format to the model and running the training on a GPU using Google’s Colab platform [56] for 40 epochs, we were able to get some encouraging results. Our object detection model achieved an AP50 average precision that was over 60%, which is a good metric given the number of objects present in the images. The samples captured all featured a significant amount of background, some of which consists of dust and noise. The model had to learn to distinguish which of the items in the image are cyanobacteria or algae purely based on their shape and whether they were in focus or not. However, as we wanted to enhance its precision further, we explored additional options like augmenting our dataset, as discussed in Section 2.5.3. Luckily, we had already gained experience with data augmentation from our early stage experiments, and when feeding the augmented training dataset to our model which now

consisted of over 200 samples we were able to increase its AP50 average precision to 70%. At this stage, we used 204 samples for training, 15 for validation and 25 for testing purposes.



Figure 25: Examples of the Object Detection model output.

Some pictures of the objects detected in the part of our dataset reserved for testing can be found in Figure 25. This depicts samples containing both cyanobacteria and algae. Despite the fact that the cyanobacteria concentration is particularly high in these cases, it is evident that the model can appropriately and accurately detect them.

2.5.2 Desktop Software for Training the Object Detection Model

Creating a machine learning model capable of detecting algae and cyanobacteria in images was our top priority. However, once we completed its development we identified that users that are not very familiar with coding would not be able to train a new model using their own image samples. This is because the entirety of our training implementation relied on writing Python code and running it through a Google Colab-based notebook. As such, to make this procedure more user-friendly, we designed a desktop-based software that can run on one’s personal computer. This prompts the user to choose if they wish to start training their object detection model, select their microscopy dataset and make some parameter decisions before invoking the model training.

To develop the desktop-based UI we used the Python programming language. Although other programming languages like Java are often better suited for creating Graphical User Interfaces (GUIs), we opted to develop ours using Python. This is because the code to train our object detection model was also coded in Python and this made the interoperability of the two flawless. The desktop software had to function as one piece, guiding the user through the process and running the model training all in one go. More specifically, to develop the GUI we used the Tkinter Python library [57], as it was open-source, widely-used, and available on both macOS and Windows systems. It also supported window graphics that resemble those of one’s operating system, so it was the optimal approach for the user-friendly aspects that we were trying to achieve. With the help of this library, we were able to place a set of graphical elements on a window that pops up when the user invokes our desktop software, with the option to invoke the respective function that trains the model with the appropriate parameters.

As depicted in Figure 26, our desktop software greets the user with an intuitive screen. Once they select the “*Train Object Detection Model*” button it requests them to choose the directory containing their processed microscopy dataset via a friendly file dialogue. At the next screen, depicted in Figure 27, they are then requested to choose the number

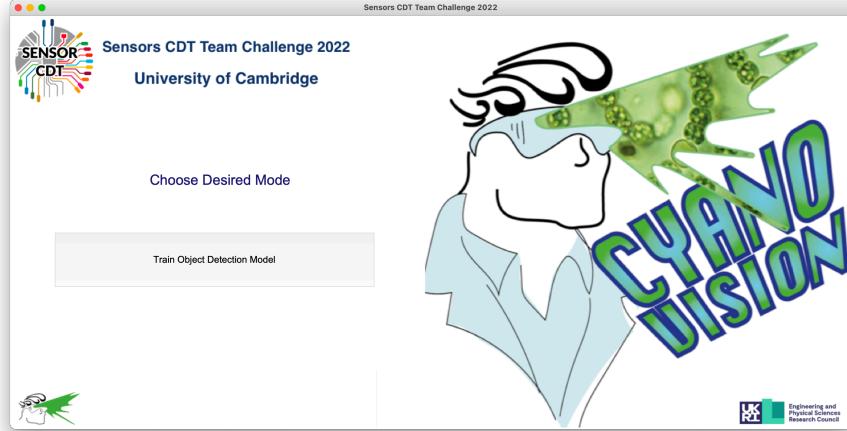


Figure 26: Welcoming screen of desktop software for training the Object Detection model.

of epochs they wish to use during the training, with clear text indicating which is the slower but more precise option and which is the faster but less precise one. After making their selection, they are then able to run the object detection model training, by simply clicking the button on the bottom half of this screen and to check its progress on the terminal. Finally, the trained model is saved in the same directory as their dataset. This makes it easier to find and to potentially transfer it to the Raspberry Pi of their microscope allowing inference to run for new images on-device, as discussed in Section 2.5.4. We have already trained a model using the dataset collected during the course of this project and experts in the field can still use Google’s Colab service to do so. But, this desktop-based solution should also make training the object detection model on new data easier for everyone.

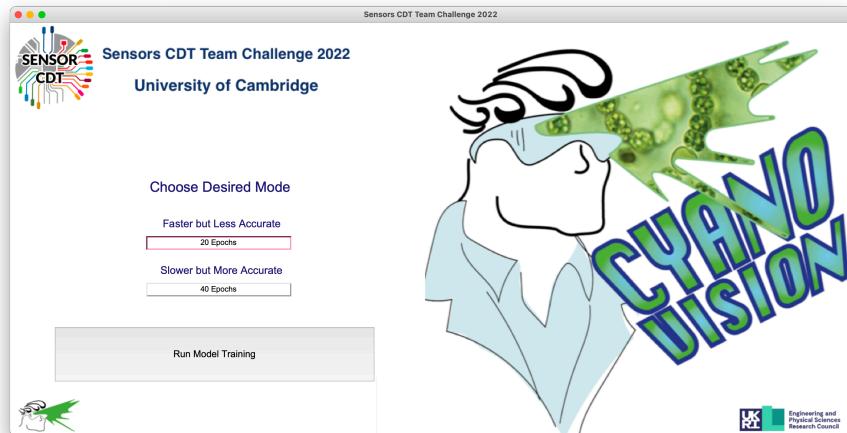


Figure 27: Main screen of desktop software for training the Object Detection model.

2.5.3 Data Collection, Labelling & Augmentations

2.5.3.1 Dataset Generation

As described in the previous section, in the early phases of the project we expected to find a sample dataset of microscopy images with bacteria online. However, it soon became apparent that no suitable dataset exists, and the ones that did seem promising like the environmental microorganisms [51] dataset, were not fit for our needs. After doing further research, a decision was made to rely on our own generated dataset in order to create our own computer vision model for algae and cyanobacteria detection. Hence, our own dataset was generated from 112 original images captured with our microscope, and we decided to call this the *Genesis* dataset. The collected images were labelled manually using the LabelBox software [58] and, because of this, our dataset comes with both image files as well as labels for any further machine learning related projects for algae and cyanobacteria detection. In order to obtain better model accuracy, we later augmented the original training dataset of 112 images to 204 using data augmentation software to have a larger sample size.

2.5.3.2 Data Labelling

We prepared our *Genesis* dataset to train our machine learning model so that it can detect and distinguish cyanobacteria and algae. The training of the model works by learning the features (patterns) that are associated with each of the objects, cyanobacteria and algae in our case, in the labelled images. Since we are using an object detection method, as discussed in Section 2.5.1, labelling our samples involved drawing a box around each individual instance of the cyanobacteria and algae cells in the microscope images manually. Since we have two classes, cyanobacteria and algae, a different box was used for each, as shown in Figure 30).

For this method to work well, many labelled images were required and in our case we labelled the 112 images captured during the course of this project. Labelling images is generally very time consuming, but with good project management we managed to label them successfully.

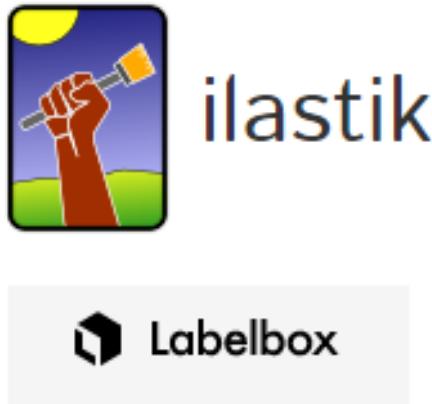


Figure 28: Logos. Top: ilastik [59]. Bottom: LabelBox [58].

There are a lot of open-source programs for labelling images and our initial approach was to use ilastik [59] (Figure 28), an image segmentation program that can also be used

as an automatic labelling tool. The plan was to label a few images and then use the automatic tool to label all of our remaining images based on our manual labels. This way, we could then manually modify some of the mistakes in the labels and use the automatic tool again. In order to get the images labelled correctly, a few iterations of this were going to be needed. This way, all images could be labelled without the need to manually label each individual cell in all of them.

While using ilastik to label our images sounded ideal, several problems arose. The main problem was its inability to handle more than 10 images at once. The program would make the computer very slow to run, thus making it more time consuming than manually labelling all of the images. Additionally, due to the time-consuming nature of image labelling, a lot of our members had to get involved. The fact that the labelling process in ilastik was not straightforward meant that it would be harder and more time-consuming for everyone to learn how to do it.

In the end, to label the images we decided to use the LabelBox framework. LabelBox [58] is an online software platform that contains labelling tools where multiple people can log in and collaborate on the same project. This and the fact that labelling with LabelBox is very easy, made it ideal for our use case. Furthermore, we had the ability to monitor our progress, as depicted in Figure 29), and review each other’s work, helping us to be consistent with our labelling.

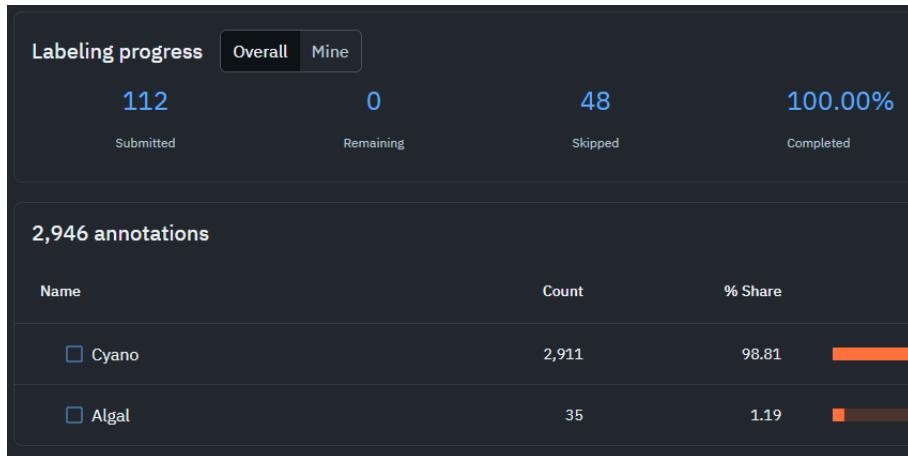
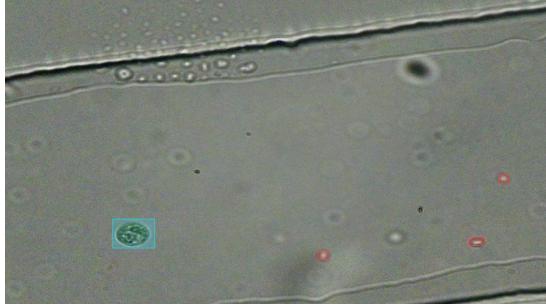


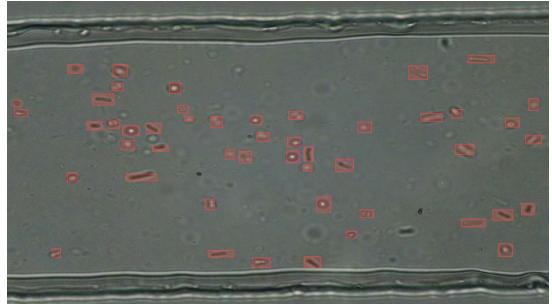
Figure 29: LabelBox [58] interface (snapshot). We can see the total number of images labelled (112) and the overall progress. 48 images were skipped because they were taken before some improvements to the set-up, resulting in blurry shots.

Figure 30 shows two examples of our labelled images. Algae cells are within the cyan-colored bounding boxes, while cyanobacteria are within the light red boxes. In Figure 30a, we can see a labelled image corresponding to low concentrations of algae and cyanobacteria, while Figure 30b captures cyanobacteria at a much higher concentration.

Since our aim is the early detection of algal blooms, images that are more representative of what we would get when our device is deployed in the wild would look more like Figure 30a that shows a low concentration example. However, images like Figure 30b were useful to include to increase the number of cyanobacteria instances in the labelled data, hence helping our model learn how to detect and distinguish them better. Nevertheless, bacteria being out of focus and background spots, due to dust and imperfections in the optics or the microfluidic channel, created ambiguity in labelling. As such, we labelled all the cells (instances) in the images that were moving and that were visible with



(a) Low concentration example.



(b) High concentration example.

Figure 30: Examples of our labelled images. Cyanobacteria are within the light red bounding boxes, while algae cells are within cyan-colored boxes. Anything not labelled is either background features (due to dust or imperfections in the optics and the microfluidic channel) or bacteria that are way out of focus.

the naked eye. One of the most important parts of the labelling was to be consistent as to which cells we label. LabelBox allowed all of the members that took part in the labelling process to work seamlessly together, helping us to make sure that our labels were consistent across all the images.

2.5.3.3 Preparing & Augmenting the Data for Training

For the machine learning model to train, the labelled data needs to be converted to the correct format. Since we are using the TensorFlow Lite Model Maker framework, as discussed in Section 2.5.1, the training data needed to be converted to a TFRecord file format. In order to achieve this and also augment our data, we used RoboFlow [60], a computer vision platform that helps users prepare data for machine learning applications.

Data augmentation is performed to improve the robustness of machine learning methods. In our case, we increased the amount of training data and hence the model performance by adding modified copies of our existing images. After some experimentation with different augmentation settings, we decided to apply the following modifications, as they gave us the best results: brightness variations from 0% to 8%; clockwise rotations of 90 and 180 degrees; as well as horizontal or vertical flips. This increased our total number of training images to 204, and even allowed room to have 15 images left for validation and 25 images left for testing purposes. Since we want our model to be able to detect cyanobacteria and algae cells regardless of the brightness of the microscope images and the position or orientation of the cells within the images, all of the above modifications were the optimal augmentations to include. With this step completed, everything that we required to train our model, as discussed in Section 2.5.1, was ready.

2.5.4 Running Inference On-Device

2.5.4.1 Introduction to the Coral USB Accelerator

In order to run inference using our model on-device we used the Edge TPU, a small application-specific integrated circuit (ASIC) designed by Google that provides high performance machine learning co-processing for low-power devices [61]. These devices are produced with Coral, Google’s platform for developing accelerated AI on embedded devices. The device we used was the Coral USB Accelerator, which provides a simple

interface with the Raspberry Pi via USB [62].

The Edge TPU is optimised for running neural networks at a high speed. The compatible models use only 8-bit signed integers as opposed to standard 32-bit floating points, reducing the required memory by 75% and greatly reducing the number of operations, saving time and energy. Additionally, its architecture relies on the structure of state-of-the-art neural networks; while the operation at each individual node is simply a multiplication and addition of the data in the form of a tensor, the latency comes from the large number of calculations required. The Edge TPU is composed of a large pipeline of specialised multiply accumulate gates arranged so as to parallelise the calculations and, therefore, greatly speed up the execution of the network.

2.5.4.2 Compiling our Model

To make the model compatible with the Coral USB Accelerator, we needed to compile the model to a quantised *.tflite* file. This process involved the aforementioned 32-bit to 8-bit quantisation and restructuring the model to use the Edge TPU runtime libraries. This was performed in another Google Colab notebook. This allows the process to be easily replicated by a user who has trained their own model, independent of the available operating system.

2.5.4.3 Running Inference On-Device

In practice, setting up the Edge TPU accelerator to integrate with our system involved the following steps:

1. Select model architecture;
2. Train model with desired dataset;
3. Quantise and compile model to 8-bit *.tflite* file format;
4. Transfer model to the Raspberry Pi via SSH;
5. Install PyCoral and associated dependencies on the Raspberry Pi;
6. Call inference on model in master code.

Alongside implementing the co-processor into our system, we carried out some of our own benchmark testing to compare with that stated on the Coral AI website [61]. This involved uploading a representative training image (shown in Figure 31) taken from the microscope and collecting the time to execute 100 inferences using both the Edge TPU processor and the Raspberry Pi CPU. The results are illustrated in Figure 32.

From inspecting Figure 32, it is apparent that the Edge TPU greatly reduces the inference time, particularly once the tensors have been loaded onto the device (after trial 1). However, the mean speedup is only $2.4\times$, which is lower than the $10\times$ advertised on the Coral AI website for the MobileNetV2 model [61]. This may be due to the relatively low computational demands of classifying a single image compared to a large dataset, alongside overhead from the Python libraries when compared to Coral's C++ benchmarking code. Despite this, such a speedup would be useful in a real-time application or where a large volume of data is being processed - perhaps in the case where an imminent cyanobacterial bloom has been detected and the user wishes to check this urgently. In



Figure 31: Training image used to benchmark the Edge TPU accelerator.

any case, if we wanted to make the Cyanovision product cheaper it could include the Coral USB Accelerator as an optional component which can be selected according to the needs of the user.

2.5.5 Conclusion

The machine learning deliverables for this project are numerous and they can be found in Reference [63]. Over the span of the project we were able to experiment with several machine learning approaches related to algae and cyanobacteria detection, finally settling on a novel object detection computer vision model. In order to aid those with less experience with machine learning models, we have created a user-friendly UI so that the general public and scientists from a broad range of disciplines can train the model using their own data. Furthermore, we are also releasing world-first algae and cyanobacteria-focused dataset for future machine learning projects and research at large. The future users of our Cyanovison toolkit are expected to expand this dataset by submitting their own collected algae and cyanobacteria images, thus making the dataset even more valuable. Consequently, the *Genesis* dataset becomes a further novel open-source deliverable that fills a niche for algae and cyanobacteria images online that other researchers can use in the future.

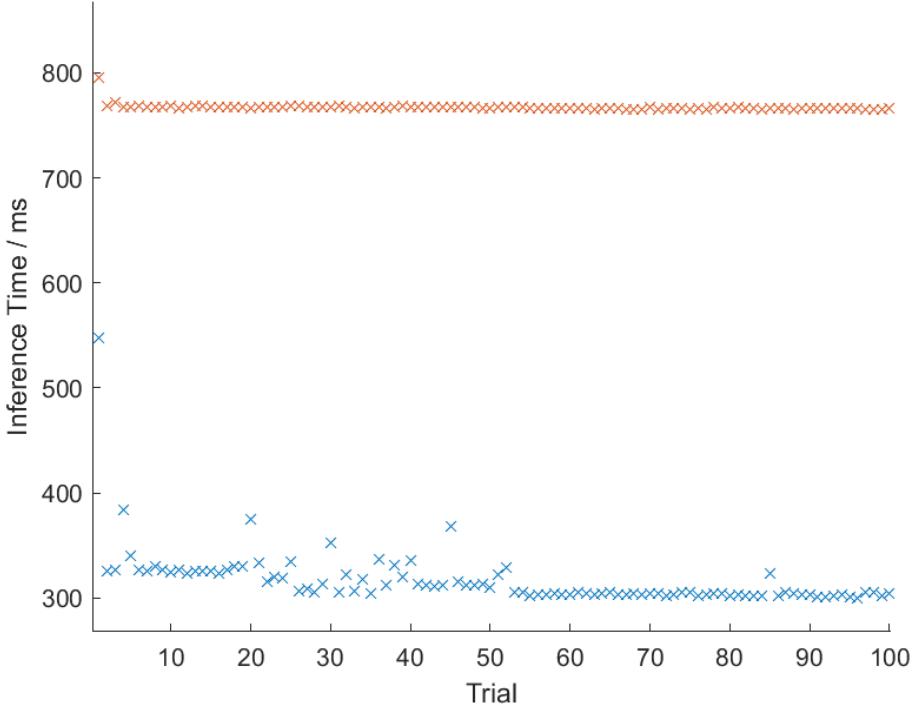


Figure 32: Comparison of the model inference times with and without the Edge TPU accelerator for 100 trials.

3 Inclusive Innovation, Market and Stakeholder Engagement

3.1 Market

In 2007, the city of Wuxi in China was faced with a water crisis, when the drinking water delivered to approximately 2 million residents turned coloured and odorous [64]. For a few weeks, the residents did not have access to tap water and needed to rely solely on bottled water [65]. The water crisis was attributed to a toxin-producing cyanobacteria *Microcystis aeruginosa* bloom reported in Lake Taihu, the drinking water source of the city. This was caused by a combination of nutrient enrichment and industrial pollution [64, 65].

In 2014, a “do not drink” policy was imposed in the city of Toledo in Ohio, USA. The drinking water supplied by Lake Erie had the growth of a toxic cyanobacteria bloom affecting 400,000 people [1]. Cyanobacteria blooms are not new in Lake Erie, which have been increasingly reported in the last two decades [1]. Toxic cyanobacteria blooms have also been reported in the UK during hot weather. Most recently, in August 2022 a blue-green algae bloom was detected in Milton Country Park lake near Cambridge [66].

During the past century animal and human poisoning associated with cyanobacteria and cyanotoxins has been well documented [67]. Death of farm animals, pets and wildlife due to drinking water infected with toxic cyanobacteria has also been reported in literature [67]. In fact, toxic and non-toxic cyanobacteria cases are increasing worldwide, both in terms of the number of species and magnitude [68]. Even so, real-time monitoring and detection of cyanobacteria and cytotoxins are not a common practice for water bodies and freshwater recreational spaces [2].

Due to these harmful effects of cyanotoxins on humans and animals, numerous commercial centres for water sports and recreation are severely affected in the case of toxic cyanobacteria blooms. Examples include outdoor swimming ponds and lakes, water sports venues and competition events, as well as fishing activities. An early detection of increased cyanobacteria presence would benefit these businesses vastly in tackling blooms before they lead to venue closures and economic losses. Furthermore, cyanobacteria could be used as an indicator species for eutrophication [69]. This could be of great economic value, because freshwater eutrophication in England and Wales is estimated to cause damage costs of 75.0114.3 million GBP, as well as leading to policy measures against these damages at a cost of 54.8 million GBP [70].

A low-cost, remote and reliable device would also greatly benefit water suppliers. Cyanobacteria occurrence can severely affect the reputation and trustworthiness of the water suppliers resulting in a large financial burden. This can also lead to temporary closure of the water intake plants, as for example in the case of Lake Taihu, where the estimated cost related to the closure of the lake is more than 100 million RMB [71].

Additionally, costs arise from the water treatment required after the occurrence of toxic cyanobacteria blooms [71]. However, the growth of toxic cyanobacteria does not only affect the water suppliers, but also the people who rely on the suppliers for the drinking water. Changing water suppliers is not always possible for some areas, as only a single supplier may be available, and relying on bottled water can be expensive and unsustainable.

Finally, the discovery of toxic cyanobacteria blooms can also result in a decrease of tourism in freshwater spaces, as many tourist activities benefit from high water quality and visibility [71]. Due to the open-source nature of the project, Cyanovision can be 3D printed and assembled anywhere in the world. For example, it could be of particular interest for the MakerSpace in South Africa to monitor cyanobacteria in the environment, as they have previously been reported in Limpopo river basin and other freshwater ecosystems [21]. Additionally, it opens up opportunity for citizen science to engage the local community and children with science. Also, our device can be used by the swimmers and people enjoying recreational activities in lakes and rivers to gather evidence of increasing cyanobacteria population, in order to request further testing from the government.

3.2 Alternative Solutions

The general results for interest in cyanobacteria and algae detection are that there is a clear demand for it, especially around the time when seasonal toxic algae blooms start appearing in lakes and rivers. A Google trends analysis illustrates this well in the English-speaking world. The trends closely correlate with the peak algae bloom cycle at around mid-to-late August every year, with the exception of 2020, as seen in Figure 33.

When investigating the market, we found half a dozen companies specialising in cyanobacteria detection using purpose-built field-deployable sensors, and several other companies specialising in algae bloom detection using satellite data (e.g. Reference [72]). For the purposes of this report, we will mostly be focusing on physical algae detection sensors, since satellite-based technologies are only able to detect toxic algae blooms after they happened. This is directly in contrast with our product, which is better suited for early detection purposes in local lakes and rivers.

All of the products found on the market use fluorometer-based techniques to measure toxic cyanobacteria and algae concentrations. A fluorometer works based on the principle

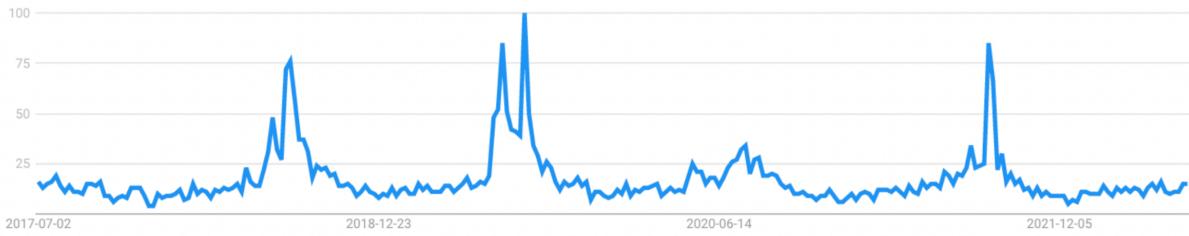


Figure 33: Overview of Google Trends data for Algae bloom shows clear seasonal patterns. The y-scale represents search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means that there was not enough data for this term. The data was collected worldwide over the last 5 years.

where samples are measured on their visible spectrum fluorescence after excitation by a particular wavelength of light [73].

The majority of products on the market do not come as standalone sensors and require add-ons to function properly, despite their relatively steep prices [74, 75, 76]. Because of this, some manufacturers had to be contacted separately to receive a more accurate quote for their products.

Overall, we found that the alternative solutions are much more expensive (starting from 2000 GBP) and often require additional tools to set up properly. Additionally, we found no existing solution with similar technology that we used, especially lacking in the software department where no field-deployable product offered machine-learning-based data analysis tools for the collected data. Below you will find a rundown of the most suitable alternatives to our product we found available on the market.

The AlgaeROW uses optics specific to the excitation fluorescence of chlorophyll “a” [76]. The derived concentration of chlorophyll “a” provides an indication of the occurrence of algal blooms [76]. This product comes with an IP68 enclosure and offers flexible customisation options on how the user may choose to deploy the device and transfer the data over several telemetry options. This is one of the few standalone devices that come fully prepared to be deployed upon arrival. Some additional features include a depth sensor that is able to readjust the sensor readings based on how deep the algae sensor is submerged in the water. The asking price for the device is around 8200 GBP.

YSI ProDSS Total Algae Sensor is a digital smart sensor for helping monitor, mitigate, and manage the impacts of harmful algal blooms [77]. Unlike the AlageROW device, YSI ProDSS is a dual channel sensor, measuring both chlorophyll and blue-green algae. The sensor is not a standalone device and thus requires a multiprobe sonde device to be able to operate. The price for the sensor is 2800 GBP, and the required multiprobe sonde (e.g. Aqua TROLL 500 Multiparameter Sonde [78]) is 1050 GBP, with the total ready-to-use price for the standalone set being over 4000 GBP.

When it comes to lab-based algae detection systems, FlowCam proposes cyanobacteria and algae detection from water samples [79]. The standalone device is fluorometer-based and is capable of differentiating cyanobacteria from other algae, using a combination of excitation wavelength, phycocyanin fluorescence measurement, and image recognition software. FlowCam was the only device we found on the market that also used camera and traditional microscopy techniques in addition to a fluorometer to provide algae

concentration readings.

Finally, in academic research, we found that Miller *et al.* describe a self-contained, field-deployable, fluorometer-based algae sensing solution called Panther Buoy [80]. The buoy is similar to our proposed device, featuring real-time algae monitoring capabilities, with attached solar panels for energy harvesting and 4G+GPS communication backhaul. The device is open-source and has seen extensive field testing. The total unit cost is around 8600-10400 GBP.

3.3 Stakeholder Engagement

We engaged with several stakeholders throughout the project, both internal and external. Internal consultations with the project advisory members and any interested stakeholders took place in the form of weekly progress meetings organised through Microsoft Teams. The Teams application was also used to publish the weekly update presentations, allowing the broader department body to gain access and keep abreast with the project progress.

Beyond the Department of Chemical Engineering and Biotechnology, we reached out to several stakeholders with expert knowledge about the critical themes of the project, such as microbial organisms, freshwater systems, monitoring, branding and marketing. The initial communication with these stakeholders was often via email and proceeded with a scheduled phone or online video call. Some of these consultations resulted in chain referrals extending our reach through our stakeholders' networks. For example, Lara Allen connected us to several potential end-users with critical real-world perspectives on how to make the best use of our device. Furthermore, we spread the word about Cyanovision to our acquaintances. Our efforts roused interest from local swimming groups, a watersports park, and green officers of several college within the larger Cambridge community.

During expert consultations on microbial and monitoring of freshwater systems, we realised that our initial goal of monitoring and identifying toxic cyanobacteria in the freshwater system was unrealistic. Experts informed us that identifying toxic cyanobacteria based on morphology would be difficult, if not impossible. This understanding allowed us to revise our goal to ensure that it is simple, meaningful, achievable, and could be delivered within the project's time frame.

Our marketing consultations identified wild swimmers, freshwater health and advocate officials and users of rivers and lakes as recreational spaces as the optimal end-users. Alternatively, some suggestions backed Cyanovision to be used for biology and ecology research, and educational science. For example, enabling citizens to engage in science and be part of the decision-making process for local policies that ultimately ensure their health and safety. An overview of the potential end-users we identified can be found in Figure 34.

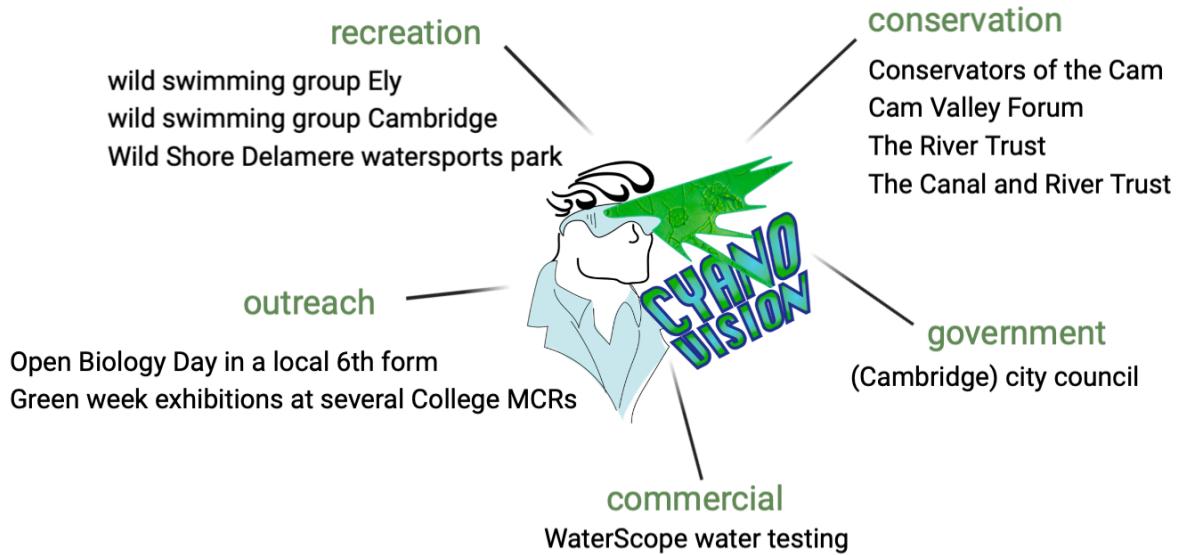


Figure 34: Overview of the stakeholders who could be potential end-users of our device.

3.4 Branding

This section covers the project’s branding, including the title, website and visual communication tools, like the logo.

3.4.1 Website

In order for our product to be accessible to the wider public, we have created a website available at www.cyantist.xyz. The frontend of the website is based on a free, readily available bootstrap design that we further modified to suit our needs. The website is running a Flask-based backend written in Python, with hopes to connect the machine learning APIs in a similar manner to the proprietary UI stand-alone app. This would allow users to submit and train their datasets only using the website in the future.

The website’s role is two-fold. First, it acts as the main advertising hub of the project: presenting what it is; how it came to be; and who created it. Secondly, it acts as the main page for the documentation, describing all of the project pieces including our main deliverables in detail.

Overall, the website contains the following parts: the main landing with the introduction; problem statement; our solution; and device features. We introduce the project, as well as the main problems behind harmful algae blooms and the need for their timely detection. Next we provide links to the projects’ documentation including hardware, software and assembly instructions, as well as a short FAQ section covering the essentials of the project. Finally, we have a rotating carousel of images showcasing the team and a form to contact us.

3.4.2 Dall-E 2

Given the high-tech nature of the project, instead of taking our own promotional images, we decided to generate them using the newly released Dall-E 2 AI image generator [81]. We have generated a several images with algae-related prompts for Dall-E 2 to process, and selected the ones currently on the website. We believed it fitted the website well,

as it added an artistic and a home-made element to the website that matched our DIY, open-source nature of the project better. The generated images will also become part of the official project merchandise in the form of laptop stickers.

3.4.3 Logo and Naming

The project's title - Cyanovision - is a portmanteau of cyanobacteria and computer vision, the two key aspects of the project. The title was selected because it conveyed the essence of the project clearly. The next part was coming up with the logo for the project. The motif of vision came back when designing the logo in form of a superhero who can see algae with their own eyes. Except in our case the superhero also comes sporting a lab coat and safety goggles for protection against harmful algae blooms. It is noteworthy that the green image of algae on the full-scale logo is our own collected image from the *genesis* dataset. The cyanovision superhero, also known as the cyantist, has also become the domain for our website, since cyanovision was unfortunately unavailable. Finally, a smaller, simplified logo exists, where the use of full-scale logo would not be appropriate, e.g. as a watermark on presentation slides.

4 Future plans

We have demonstrated that the designed device is fit for purpose. However, a significant body of work remains to fine-tune the design for deployment and application. The critical steps will include device testing and verification in different contexts, improving the robustness and efficiency of the device both in terms of hardware and software, and expanding and finalising the steps for collaboration and further development.

In terms of verification, we need sizeable data to confirm that the cyanobacteria approximations determined by the system are accurate. Furthermore, it will require the design of stress-test experiments for the device to determine its limits in ability in differing deployments. The initial project timeline restricted testing to the lab bench and examining images of only a few strains of cultured cyanobacteria. Additionally, the system uses an approximation of calculating the concentration of cyanobacteria from a small sample of water and extrapolating this to the whole body of water.

There is a need to build our data set on the accuracy of the measurements when compared to “gold” standard testing. In practice, many factors may be a barrier to successful deployment. These include the location of sample extraction (near the surface of the water or at a certain depth), filtration system, heterogeneity of water body ecosystems, and the device’s reliability over time. This data must be collected from consecutive studies examining each determining factor to identify the scope of application and potential optimisations. The Milton Country park will be our initial field testing site.

Additionally, we intend to expand our data with samples containing a variety of microorganisms for a rich data set and to verify that the machine learning model is sufficient for meaningful output. The additional microorganisms will ideally be common bacterial species found in freshwater in the locations of interest. Our project aims to provide a solution for monitoring microbial organisms that may harm humans, animals or other organisms in the freshwater ecosystem. Therefore the inclusion of a diverse range of microorganisms in our database aligns with our cause.

Furthermore, we would like to explore the concept of absolute and relative measurements. However, determining the precise and accurate concentration of cyanobacteria in lake water from a relatively small sample size is complex. The current system determines whether there is an increase in concentration compared to the previous measurement (which is what makes it an early warning system). However, a more applicable and useful system would be able to quantify the absolute concentration and link this to whether concentrations are at a dangerous level to health. Achievement of this capability could expound the application of the solution, elucidating the nature of the algal bloom growth and its progression.

Moving forward, we would like to re-explore the use of fluorescence as an alternative imaging modality for higher sensitivity and specificity in detection. The initial testing was promising, but was too complicated to idealise in the time frame. We believe it is achievable and within reach to add this functionality to the existing setup to provide more robust identification. Immediate steps will be obtaining fluorescent reference materials, such as microbeads, and testing the channels with standard commercial microscopes available in the Department of Chemical Engineering and Biotechnology before testing and optimising our device.

In addition, there is a need to overcome initial challenges, such as the integration of the flow sensor, which can make the system more automatic. Flow steadiness and regulation were issues, and it was challenging to collect accurate data without a closed

loop control system. By integrating software control to the system, it could automatically run to settle the sample when images are taken. Further functionality could be added to assist in diagnostics if any problems occur in the microfluidics.

We would also want to run longitudinal studies to test the device's continued reliability over time, particularly the durability of the moving components and the filtration system. Similarly, security issues were raised and there is room to address them regarding locking and tracking the device to protect the hardware. Also, we considered different deployment systems; a later model may include floating or submerged system features which raise some interesting possibilities.

Beyond the hardware and deployment, we would like to start testing whether the data from the microscope is good enough for the machine learning algorithm to be extended to the detection of other microbial organisms of interest. As mentioned above, many interested groups with different applications need to establish if the device is customisable as we expect. We anticipate this to be a relatively direct reprogramming step, provided with a sufficient sample data set. The design and system may be helpful to OpenFlexure Project in general, as some stakeholders have expressed interest in using some software and hardware ideas for their projects.

Finally, the bulk of the work is putting the device forward and advertising. As an open-source system, we would want to communicate the merits and potential of the system to a wide variety of interested stakeholders and pursue the collaborations that have already been discussed. Each collaboration will give the device greater scope and data for modifications and benefit those who use it. We will seek more public engagement opportunities, in addition, to several calendar dates that have already been marked out for showing the device applications. The website is running, and many advertisement materials are ready to showcase the device's functionality. The more excitement and interest we can drum up coupled with our vision for the future, the greater the system's benefit to society.

References

- [1] G. S. Bullerjahn, R. M. McKay, T. W. Davis, D. B. Baker, G. L. Boyer, L. V. D'Anglada, G. J. Doucette, J. C. Ho, E. G. Irwin, C. L. Kling *et al.*, "Global solutions to regional problems: Collecting global expertise to address the problem of harmful cyanobacterial blooms. a lake erie case study," *Harmful Algae*, vol. 54, pp. 223–238, 2016.
- [2] H. Duan, M. Tao, S. A. Loiselle, W. Zhao, Z. Cao, R. Ma, and X. Tang, "Modis observations of cyanobacterial risks in a eutrophic lake: Implications for long-term safety evaluation in drinking-water source," *Water Research*, vol. 122, pp. 455–470, 2017.
- [3] M. Albrecht, T. Pröschold, and R. Schumann, "Identification of cyanobacteria in a eutrophic coastal lagoon on the southern baltic coast," *Frontiers in microbiology*, vol. 8, p. 923, 2017.
- [4] Z. Zahra, D. H. Choo, H. Lee, and A. Parveen, "Cyanobacteria: Review of current potentials and applications," *Environments*, vol. 7, no. 2, p. 13, 2020.
- [5] A. K. Srivastava, A. N. Rai, and B. A. Neilan, *Stress biology of cyanobacteria: molecular mechanisms to cellular responses*. CRC Press, 2013.
- [6] D. K. Saini, S. Pabbi, and P. Shukla, "Cyanobacterial pigments: Perspectives and biotechnological approaches," *Food and chemical toxicology*, vol. 120, pp. 616–624, 2018.
- [7] B. A. Whitton and M. Potts, "Introduction to the cyanobacteria," in *Ecology of cyanobacteria II*. Springer, 2012, pp. 1–13.
- [8] R. Mur, O. M. Skulberg, and H. Utkilen, "Cyanobacteria in the environment," 1999.
- [9] R. W. Zurawell, H. Chen, J. M. Burke, and E. E. Prepas, "Hepatotoxic cyanobacteria: a review of the biological importance of microcystins in freshwater environments," *Journal of Toxicology and Environmental Health, Part B*, vol. 8, no. 1, pp. 1–37, 2005.
- [10] W. W. Carmichael, "Health effects of toxin-producing cyanobacteria: "The CyanoHABs"," *Human and ecological risk assessment: An International Journal*, vol. 7, no. 5, pp. 1393–1407, 2001.
- [11] C. Moreira, V. Ramos, J. Azevedo, and V. Vasconcelos, "Methods to detect cyanobacteria and their toxins in the environment," *Applied microbiology and biotechnology*, vol. 98, no. 19, pp. 8073–8082, 2014.
- [12] T. Enikő, V. Lerf, I. Tóth, T. Kisantal, G. Várbíró, G. Vasas, B. Viktória, J. Görgényi, Á. Lukács, Z. Kókai *et al.*, "Uncertainties of cell number estimation in cyanobacterial colonies and the potential use of sphere packing," *Harmful Algae*, vol. 117, p. 102290, 2022.

- [13] P. Dvořák, A. Pouličková, P. Hašler, M. Belli, D. A. Casamatta, and A. Papini, “Species concepts and speciation factors in cyanobacteria, with connection to the problems of diversity and classification,” *Biodiversity and Conservation*, vol. 24, no. 4, pp. 739–757, 2015.
- [14] J. Jezberová and J. Komárková, “Morphometry and growth of three *synechococcus*-like picoplanktic cyanobacteria at different culture conditions,” *Hydrobiologia*, vol. 578, no. 1, pp. 17–27, 2007.
- [15] P. Huber, N. Diovisalvi, M. Ferraro, S. Metz, L. Lagomarsino, M. E. Llames, M. Royo-Llonch, J. Bustingorry, R. Escaray, S. G. Acinas *et al.*, “Phenotypic plasticity in freshwater picocyanobacteria,” *Environmental Microbiology*, vol. 19, no. 3, pp. 1120–1133, 2017.
- [16] Y. Yu, L. You, D. Liu, W. Hollinshead, Y. J. Tang, and F. Zhang, “Development of *synechocystis* sp. pcc 6803 as a phototrophic cell factory,” *Marine drugs*, vol. 11, no. 8, pp. 2894–2916, 2013.
- [17] Y. Yang, V. Lam, M. Adomako, R. Simkovsky, A. Jakob, N. C. Rockwell, S. E. Cohen, A. Taton, J. Wang, J. C. Lagarias *et al.*, “Phototaxis in a wild isolate of the cyanobacterium *synechococcus elongatus*,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 52, pp. E12378–E12387, 2018.
- [18] H. Iijima, Y. Nakaya, A. Kuwahara, M. Y. Hirai, and T. Osanai, “Seawater cultivation of freshwater cyanobacterium *synechocystis* sp. pcc 6803 drastically alters amino acid composition and glycogen metabolism,” *Frontiers in microbiology*, vol. 6, p. 326, 2015.
- [19] K. Gilbride, “Chapter 8 - molecular methods for the detection of waterborne pathogens,” in *Waterborne Pathogens (Second Edition)*, second edition ed., H. Bridle, Ed. Academic Press, 2021, pp. 237–292. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780444643193000083>
- [20] A. J. Ouellette and S. W. Wilhelm, “Toxic cyanobacteria: the evolving molecular toolbox,” *Frontiers in Ecology and the Environment*, vol. 1, no. 7, pp. 359–366, 2003.
- [21] M. Magonono, P. J. Oberholster, A. Shonhai, S. Makumire, and J. R. Gumbo, “The presence of toxic and non-toxic cyanobacteria in the sediments of the limpopo river basin: Implications for human health,” *Toxins*, vol. 10, no. 7, p. 269, 2018.
- [22] ChemoMetec, “How to count cells with a hemocytometer,” *ChemoMetec, Cell Counting Blog*, 2022. [Online]. Available: <https://chemometec.com/how-to-count-cells-with-a-hemocytometer/>
- [23] EMS, “Electron microscopy sciences: Neubauer haemocytometry,” *Electron Microscopy Sciences, Technical Data Sheets*, 2021. [Online]. Available: <https://www.emsdiasum.com/microscopy/technical/datasheet/68052-14.aspx>
- [24] C. Thornett, “The quest for open science,” *TechRadar*, Nov 2017. [Online]. Available: <https://www.techradar.com/news/the-quest-for-open-science>

- [25] O. Initiative, “Open-source through the lens of a microscope,” *Cambridge Department of Engineering*, Dec 2013. [Online]. Available: <http://www.eng.cam.ac.uk/news/open-source-through-lens-microscope>
- [26] D. R. Bowman, “Dr richard bowman,” *NanoPhotonics Centre*, Jan 2016. [Online]. Available: <https://www.np.phy.cam.ac.uk/people/rwb27>
- [27] J. Stirling, V. L. Sanga, P. T. Nyakyi, G. A. Mwakajinga, J. T. Collins, K. Bumke, J. Knapper, Q. Meng, S. McDermott, and R. Bowman, “The OpenFlexure project. the technical challenges of co-developing a microscope in the UK and Tanzania,” *2020 IEEE Global Humanitarian Technology Conference (GHTC)*, 2020. [Online]. Available: <https://doi.org/10.1109/ghtc46280.2020.9342860>
- [28] “High precision autofocus on a low cost microscope: Automating blood sample imaging on the openflexure microscope,” *Proceedings of the Microscopy Congress 2021 incorporating EMAG 2021*, 2021. [Online]. Available: <https://doi.org/10.22443/rms.mmc2021.127>
- [29] J. T. Collins, J. Knapper, J. Stirling, J. Mduda, C. Mkindi, V. Mayagaya, G. A. Mwakajinga, P. T. Nyakyi, V. L. Sanga, D. Carbery, and et al., “Robotic microscopy for everyone: The openflexure microscope,” 2019. [Online]. Available: <https://doi.org/10.1101/861856>
- [30] R. Bowman, “Robotic microscopy for everyone with the web of things: The openflexure microscope,” *Proceedings of the Microscopy Congress 2021 incorporating EMAG 2021*, 2021. [Online]. Available: <https://doi.org/10.22443/rms.mmc2021.133>
- [31] T. Matsui and D. Fujiwara, “Optical sectioning robotic microscopy for everyone: The structured illumination microscope with the openflexure stages,” *Optics Express*, vol. 30, no. 13, p. 23208, 2022. [Online]. Available: <https://doi.org/10.1364/oe.461910>
- [32] S. D. Grant, G. S. Cairns, J. Wistuba, and B. R. Patton, “Adapting the 3D-printed Openflexure microscope enables computational super-resolution imaging,” *F1000Research*, vol. 8, 2019. [Online]. Available: <https://doi.org/10.12688/f1000research.21294.1>
- [33] S. McDermott, “Automation of imaging and control of the openflexure delta stage using the openflexure matlab client.” *Proceedings of the Microscopy Congress 2021 incorporating EMAG 2021*, 2021. [Online]. Available: <https://doi.org/10.22443/rms.mmc2021.219>
- [34] J. T. Collins, J. Knapper, S. J. McDermott, F. Ayazi, K. E. Bumke, J. Stirling, and R. W. Bowman, “Simplifying the openflexure microscope software with the web of things,” *Royal Society Open Science*, vol. 8, no. 11, p. 211158, 2021. [Online]. Available: <https://royalsocietypublishing.org/doi/abs/10.1098/rsos.211158>
- [35] “OpenFlexure / microscope-extensions / staging · GitLab.” [Online]. Available: <https://gitlab.com/openflexure/microscope-extensions/staging>

- [36] F. Ayazi, “Filip ayazi / sangaboard-rp2040 · GitLab.” [Online]. Available: <https://gitlab.com/filipayazi/sangaboard-rp2040>
- [37] ——, “Files · sangaboard-v5 · filip ayazi / sangaboard-firmware · GitLab.” [Online]. Available: <https://gitlab.com/filipayazi/sangaboard-firmware/-/tree/sangaboard-v5>
- [38] openflexure.org, “Openflexure delta stage,” 2022. [Online]. Available: <https://openflexure.org/projects/deltastage/>
- [39] GitLab, “Openflexure,” 2022. [Online]. Available: <https://gitlab.com/openflexure>
- [40] S. McDermott and R. Bowman, “Openflexure delta stage,” 2022. [Online]. Available: <https://build.openflexure.org/openflexure-delta-stage/v1.2.1/>
- [41] S. McDermott, F. Ayazi, J. C. anb Joe Knapper, J. Stirling, R. Bowman, and P. Cicuta, “Multi-modal microscopy imaging with the openflexure delta stage,” *Optics Express*, vol. 30, no. 15, pp. 26 377–26 395, 2022.
- [42] Sony, “Diagonal 4.60 mm (type 1/4.0) 8 mega-pixel cmos image sensor with square pixel for color cameras, imx219pqh5-c,” 2022. [Online]. Available: https://www.opensourceinstruments.com/Electronics/Data/IMX219PQ.pdf?fbclid=IwAR3TC85gDy66RecdtCkYbq_-BPqsqHQOVZuB5vWdfKLXsRRuHZtXUGOdJQ8
- [43] J. Stirling, “Ofep2: Improved rms optics configurations,” 2021. [Online]. Available: <https://ofep.openflexure.org/OFEP2/?fbclid=IwAR22ntPhlY2hAaCanGuXl21in1Ax9TocUgvDfT9XM130ezmrlLs89k24bVQ>
- [44] B. Z. Rousso, E. Bertone, R. A. Stewart, P. Hobson, and D. P. Hamilton, “Automation of species-specific cyanobacteria phycocyanin fluorescence compensation using machine learning classification,” *Ecological Informatics*, vol. 69, p. 101669, 2022.
- [45] T. Musumeci, A. Leonardi, A. Bonaccorso, R. Pignatello, and G. Puglisi, “Tangential flow filtration technique: An overview on nanomedicine applications,” *Pharmaceutical Nanotechnology*, vol. 6, no. 1, pp. 48–60, 2018. [Online]. Available: <https://www.ingentaconnect.com/content/ben/pnt/2018/00000006/00000001/art00008>
- [46] *Bartels Micropumps*, Bartels Mikrotechnik, 3 2021, rev. 4.4.
- [47] J. Klespitz and L. Kovács, *Peristaltic pumps - A review on working and control possibilities*, 01 2014, pp. 191–194.
- [48] *SLF3S-0600F Liquid Flow Sensor*, Sensirion, 8 2020, version 0.7.
- [49] Y. Xue, D. Kang, Y. Ma, X. Feng, J. A. Rogers, and Y. Huang, “Collapse of microfluidic channels/reservoirs in thin, soft epidermal devices,” *Extreme Mechanics Letters*, vol. 11, pp. 18–23, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352431616301286>

- [50] N. Dalir and H. Saffari, “Influence of variation of pipe diameter on pressure drop predictions of the new modified three-fluid model inside condensing vertical pipes,” *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering*, vol. 230, no. 1, pp. 36–44, 2016. [Online]. Available: <https://doi.org/10.1177/0954408914538784>
- [51] Z. Li, C. Li, Y. Yao, J. Zhang, M. M. Rahaman, H. Xu, F. Kulwa, B. Lu, X. Zhu, and T. Jiang, “Emds-5: Environmental microorganism image dataset fifth version for multiple image analysis tasks,” *Plos one*, vol. 16, no. 5, p. e0250631, 2021.
- [52] R. Pramoditha, “Coding a sequential Convolutional Neural Network (CNN),” *Towards Data Science*, Jun 2022. [Online]. Available: <https://towardsdatascience.com/coding-a-convolutional-neural-network-cnn-using-keras-sequential-api-ec5211126875>
- [53] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for biomedical image segmentation,” *CoRR*, vol. abs/1505.04597, 2015. [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [54] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, “Cellpose: A generalist algorithm for cellular segmentation,” *Nature Methods*, vol. 18, no. 1, pp. 100–106, Jan 2021. [Online]. Available: <https://doi.org/10.1038/s41592-020-01018-x>
- [55] TensorFlow, “Tensorflow Lite model maker,” *TensorFlow Documentation*, May 2022. [Online]. Available: https://www.tensorflow.org/lite/models/modify/model_maker
- [56] Google, “Google Colaboratory,” *Colab*, 2022. [Online]. Available: <https://colab.research.google.com/notebooks/intro.ipynb>
- [57] D. Amos, “Python GUI programming with Tkinter,” *Real Python*, Mar 2022. [Online]. Available: <https://realpython.com/python-gui-tkinter/>
- [58] Labelbox Inc, “The data engine for AI,” *Labelbox*, 2022. [Online]. Available: <https://labelbox.com>
- [59] S. Berg, D. Kutra, T. Kroeger, C. N. Straehle, B. X. Kausler, C. Haubold, M. Schiegg, J. Ales, T. Beier, M. Rudy, K. Eren, J. I. Cervantes, B. Xu, F. Beuttenmueller, A. Wolny, C. Zhang, U. Koethe, F. A. Hamprecht, and A. Kreshuk, “iLastik: interactive machine learning for (bio)image analysis,” *Nature Methods*, Sep. 2019. [Online]. Available: <https://doi.org/10.1038/s41592-019-0582-9>
- [60] Roboflow Inc, “A Computer Vision developer framework,” *Roboflow*, 2022. [Online]. Available: <https://roboflow.com>
- [61] Google Inc, “Edge TPU.” [Online]. Available: <https://cloud.google.com/edge-tpu>
- [62] ——, “USB Accelerator Datasheet,” 2019. [Online]. Available: <https://coral.ai/docs/accelerator/datasheet/>
- [63] The CDT Team Challenge Authors, “Cyanovision - final ML repository,” *Google Drive*, 2022. [Online]. Available: <https://drive.google.com/drive/folders/1SC4DjTgI5ylN5W1q7h3sAdxPPt9OMgb?usp=sharing>

- [64] X.-j. Zhang, C. Chen, J.-q. Ding, A. Hou, Y. Li, Z.-b. Niu, X.-y. Su, Y.-j. Xu, and E. A. Laws, “The 2007 water crisis in wuxi, china: analysis of the origin,” *Journal of Hazardous Materials*, vol. 182, no. 1-3, pp. 130–135, 2010.
- [65] Y. Liu, W. Chen, D. Li, Z. Huang, Y. Shen, and Y. Liu, “Cyanobacteria-/cyanotoxin-contaminations and eutrophication status before wuxi drinking water crisis in lake taihu, china,” *Journal of Environmental Sciences*, vol. 23, no. 4, pp. 575–581, 2011.
- [66] “Toxic blue-green algae discovered in Milton Country Park lake,” *BBC News*, Aug. 2022. [Online]. Available: <https://www.bbc.com/news/uk-england-cambridgeshire-62497000>
- [67] L. Giannuzzi, “Cyanobacteria growth kinetics,” 2019.
- [68] B. Z. Rousso, E. Bertone, R. Stewart, and D. P. Hamilton, “A systematic literature review of forecasting and predictive models for cyanobacteria blooms in freshwater lakes,” *Water Research*, vol. 182, p. 115959, 2020.
- [69] B. Van Geel, B. Odgaard, and M. Ralska-Jasiewiczowa, “Cyanobacteria as indicators of phosphate-eutrophication of lakes and pools in the past,” *PACT (Rixensart)*, no. 50, pp. 399–415, 1996.
- [70] J. N. Pretty, C. F. Mason, D. B. Nedwell, R. E. Hine, S. Leaf, and R. Dils, “Environmental costs of freshwater eutrophication in england and wales,” 2003.
- [71] D. P. Hamilton, S. A. Wood, D. R. Dietrich, and J. Puddick, “Costs of harmful blooms of freshwater cyanobacteria,” *Cyanobacteria: an economic perspective*. John Wiley & Sons, pp. 247–256, 2014.
- [72] US EPA, “Cyanobacteria Assessment Network (CyAN).” [Online]. Available: <https://www.epa.gov/water-research/cyanobacteria-assessment-network-cyan>
- [73] “Fluorometer - biology,” University of York. [Online]. Available: <https://www.york.ac.uk/biology/technology-facility/molecular-interactions/mi-equipment/mi-fluorometer/>
- [74] “TriLux Fluorometers: 3 in 1 algae sensor by chelsea...” [Online]. Available: <https://www.environmental-expert.com/products/trilux-3-in-1-algae-sensor-120691>
- [75] “Blue green algae sensor,” OTT Hydromet. [Online]. Available: <https://www.ott.com/products/sensors-108/blue-green-algae-sensor-162/>
- [76] “ALGAL: bloom sensor,” Laser Diagnostic Instruments. [Online]. Available: <https://ldi.ee/products/algal-bloom-detector/>
- [77] “Ysi prodss total algae sensor.” [Online]. Available: <https://www.fondriest.com/ysi-prodss-total-algae-sensor.htm>
- [78] “Aqua troll 500 multiparameter sonde - in-situ.” [Online]. Available: <https://insitu.com/en/aqua-troll-500-multiparameter-sonde>
- [79] “FlowCam Cyano | cyanobacteria and algae flow imaging cytometer,” FluidImaging. [Online]. Available: <https://www.fluidimaging.com/products/flowcam-cyano>

- [80] T. R. Miller, W. Tarpey, J. Nuese, and M. Smith, “Real-time monitoring of cyanobacterial harmful algal blooms with the panther buoy,” *ACS ES&T Water*, 2022. [Online]. Available: <https://pubs.acs.org/doi/10.1021/acsestwater.2c00072>
- [81] OpenAI, “DALL-E 2,” *OpenAI*. [Online]. Available: <https://openai.com/dall-e-2/>