Final Year Project

# An analysis of water purity using image analysis and infra-red imaging

Charles Kelly

Student ID: 16464276

A thesis submitted in part fulfilment of the degree of

**BSc. (Hons.) in Computer Science**

**Supervisor:** Assoc. Prof. Eleni Mangina

UCD School of Computer Science

University College Dublin

May 21, 2020

# Table of Contents

# Abstract

Clean drinking water is one of the top concerns in the modern world. On January 1st 2016 the United Nations created it's sustainable development goals including goal number six for clean water and sanitation[1]. This shows that clean water is still a key issue faced by many people internationally. The aim of this project is to make it simpler to identify if water is safe to consume.

This project will use image analysis and classification to test the purity of water samples by comparing them against a control sample of clean water that meets the European Union's and World Health Organisations water quality standards. Two sets of images will be collected: one retrieved from standard camera images and one retrieved from infra-red camera images. By using machine learning, deep learning and image classification a program will be created to determine if these images contain drinkable water. Similarly to the 2019 AquaSight project a convolutional neural network will be utilised in order to classify these images.[2]

The secondary goal of this project is to identify whether the classifier performing on infra-red or standard images is more successful at determining the purity of water. Finally, within this project the data set of the combined infra-red and standard images will be utilised (these images will be taken of the same samples of settled water) to determine if the combined image types will provide a higher accuracy than just a singular image type.

Depending on time constraints different image classification techniques will be used to create an image analysis algorithm that is capable of specifying the exact pollutant that is contaminating the water; if one does exist.

# Outline of Report Structure

**Interim Report Sections:**

- Abstract
- Project Specification
- Introduction
- Related Work and Ideas
- Data Considerations
- Outline of Approach
- Project Workplan
- Summary and Conclusions
- Bibliography

**Final Report Sections:**

- Data & Context
- Core Contribution
- Evaluation
- Final Conclusions

# Chapter 1: Project Specification

**Description** Infrared imaging is an imaging method that captures images of electromagnetic radiation which has a wavelength longer than visible light making it invisible to human sight. It has been used extensively in the medical and surveying fields of industry. By monitoring the differences in heat exuded by people or materials it is possible to determine if there issues with the subject.[3]

This principle could be relevant in testing the purity of water. So the idea of this project would be to create an image classifier to determine if a given water sample is pure. The base goal is to tell if the water is simply impure while comparing if the infrared imaging or standard imaging provides better accuracy. Our extended goal is to tell what is in the water and to detect harder to identify substances such as plastic residue.

**Core** The core of this project is to use infrared and standard cameras to compare samples of water to a control (pure water) in order to see if the water is contaminated. This will be carried out by using a convolutional neural network and image classifier to label images of water as either pure or impure. Then the amount of contaminant in the water will be continually lowered to establish the minimum amount of each material the classifier can detect. The contaminants that the system will be trained to identify in the later stages of the project include: oil, dirt, plastic residue, salt, algae, glass, alcohol. A set of couple infrared and standard images will also be compiled to determine if the coupling of these images will raise the accuracy of our classifier.

**Advanced** For the advanced features of the project it would be to identify the materials found within the water. This would creating a classifier that could identify all of the specified labels that could appear in our water samples. This would increase the usability of the detector so that it can now identify the materials rather than just letting the user know the water contains some foreign material.

# Chapter 2: Introduction

The basic premise of this project is to create an image classifier that can determine if a sample of water is fit for human consumption. The images will be gathered by using both standard and infra-red cameras to see if the varying types of images can increase the effectiveness of the classifier.

Compared to the chemical analysis method which requires laboratories and trained professionals to carry out, a pre-constructed image classifier would only require a laptop and camera. As you can see from the descriptions presented in APHA Method 4110: Standard Methods for the Examination of Water and Wastewater the current methods of testing water quality are intense and outside of the scope of knowledge for most people.[4] So the main aim of this project is to simplify the examination method because clean drinking water is a human right as stated by article 25 of the United Nations Universal Declaration of Human rights. By simplifying the method the hope is to increase the amount of people will be able to access clean drinking water.

Two points that should be addressed are; aren't there better solutions created by professionals and why use infrared cameras? For the first question please see a quote from Karel Horak, Jan Klecka, and Miloslav Richter, "Unfortunately there is a lack of completely automated vision-based systems in the literature using an image analysis".[5] So as you can see there is a sufficient gap in the field so that this projects contribution could make a difference and help progress the field for others. But as you can see from X.Zhang's turbidity analysis paper[6] and the Aquasight project[2] there does exist research into image analysis on water using a standard camera so this is why an infra-red camera has been incorporated into the project, in order to expand the scope of this project in comparison to those that only use standard camera images.

# Chapter 3: Related Work and Ideas

## 3.1 Standard Water Analysis

In order to understand the need for more streamlined and simpler methods of water analysis we must first examine the current scientific method of testing water quality. Many different techniques exist for the analysis of water including Ion chromatography[4], enumeration of colony count in water[7] and hydraulic fracturing[8]. What do all of these techniques have in common? They all require a vast amount of expertise in difficult fields such as chemistry and biology. They also require expensive equipment, laboratories and staff to perform. This shows that a cheaper more efficient solution is needed.

APHA's breakdown of ion chromatography with chemical suppression of eluent conductivity is a perfect example of the difficulty present in current water analysis methods. The apparatus needed to perform this experiment include "an injection valve, a sample loop, guard column, separator column, and fiber or membrane suppressors, a temperature-compensated smali-volume conductivity cell and detector and a strip-chart recorder capable of full-scale response of 2s or less". The sheer volume of equipment necessary to carry out this experiment proof the expense to operate these experiments is ludicrous especially regions of the world where this equipment would be harder to procure. This accompanied with the fact that trained professionals would be necessary to operate this equipment only add to the difficulty.

Another example of the difficulty present in the analysis of water quality can be seen in the schedule of accreditation presented by Advanced Laboratory Testing Ltd.[7] The amount of tests that need to be performed on water to ensure its purity is easily in the dozens. These include but are not limited to Determination of Total Nitrogen and Kjeldahl Nitrogen, Determination of Suspended Solids using Gravimetry and Determination of pH using Orion Star. The fact that so many tests exist in order to ensure the purity of water implies that industries producing clean water must have both the staff and equipment capable of handling all of these experiments. This would require a massive investment by the company for monitoring of components that may never be present in any of the water samples.

As you can see from Water Analysis: Emerging Contaminants and Current Issues the amount of contaminants that you have to monitor for in water samples is absurdly high.[8] To name a few you have Pesticide Degradation Products, Chiral Contaminants, Chemical Warfare Agents, Organotins and Arsenic. With just these components you would need an expert in both chemical and biological substances to ensure the proper monitoring and treatment of the water. This clearly shows that the minimum level of expertise needed for proper water quality analysis is extremely high.

Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring provides a different perspective on standard water analysis compared to the other works discussed so far in that it only discusses the analysis of natural water in the environment and the challenges that presents.[9]. While this doesn't provide much justification for this project it does pose the idea of an expansion into the project analysing open water source rather than just closed water samples. It discusses the types of analysis carried out on the water sources such as long term(monitoring), continuous(surveillance) and finite/intensive(survey). These methods of analysis seem like they could be completed most more efficiently with the aid of cameras and predesigned image classifiers. While this is out of the scope of this project it is relevant to show the possibilities that exist within this field of research.

## 3.2 Image Based Water Analysis

In order to prepare for this project a large amount of research into image based water analysis strategies and project had to be done in order to establish a firm understanding of the subject matter.

The best example of this is the AquaSight project.[2]. This project is extremely similar to our project in many aspects such as the use of standard camera images for classification, the use of two labels for pure and impure water, the convolutional neural network and the overall aim of the project to improve access to clean water. This is incredible for having a good comparison for how well the project is doing (we will see this later during the evaluation proportion of this report) but it also poses the problem that this project needs to be differentiated from the Aquasight project. A problem with the Aquasight project is that by just examining the water using a standard camera then unless you are incorporating object detection you are simply just examining the turbidity of the water which can be misleading e.g food colouring in water doesn't decrease the purity. That is one of the reasons infra-red cameras are used in this project in order to avoid this turbidity pitfall. However since the presented paper provides few details into the specifics of the project it is not helpful for much besides a comparison and evaluation.

A paper that discusses quite a similar approach to this project and the Aquasight project is "Water Quality Turbidity Detection Based on Image Recognition System Design and Implementation".[6]. Similarly to the Aquasight project X.Zhang's turbidity detection system using a neural network and standard camera to analyse water samples. While this is helpful from a comparison standpoint it has the same problems that the AquaSight project possesses. The simple monitoring of a waters turbidity is not all that informative. If your system only analysis's the turbidity of a water sample then it could easily be mislead by something as simple as food colouring. This further provides justification for the use of an infra-red camera in this project. While this project is quite similar to this one in scope and it has quite good accuracy it presents the same issue as the Aquasight project of not elaborating on the specifics of the neural network or anything technical within the project. This means the paper is only particularly relevant as a means of comparison and evaluation.

Before we begin discussing the papers pertaining to image classifiers that specifically target biological materials we will discuss two outliers that are neither extremely similar to this project nor analyse exclusively for biological material. Those being "Water quality analysis using an inexpensive device and a mobile phone" by Timo Toivanen, Sampsa Koponen, Ville Kotovirta, Matthieu Molinier and Peng Chengyuan[10] and "Image Processing Techniques Applicable in Wastewater Quality Detection: Towards a Hygienic Environment"[11] by Ibrahim Haruna Shanono, Mohd Razali Mohamad Sapiee, Khairul Azha Aziz, Nasir Hassan Zakaria Suleiman, Ashen Gomes1,, Chandima Gomes.

"Water quality analysis using an inexpensive device and a mobile phone" discusses the creation and implementation of their Secchi3000 device which is a is used along with an accompanying mobile app to determine the secchi depth of water. This accompanied by a turbidity analysis carried out by the same image classifier in the mobile app are used to determine the purity of the water. While this paper does take extra steps that the other two didn't by exploring portable equipment and analysing secchi depth it still has similar faults. Secci depth is similar to turbidity in that it is only a superficial examination of water and doesn't detect any clear substances. It is essentially an in-depth colour analysis of the water. Finally it also gives little specifics into the technical aspect of the project besides their self designed apparatus, which has no impact on this project.

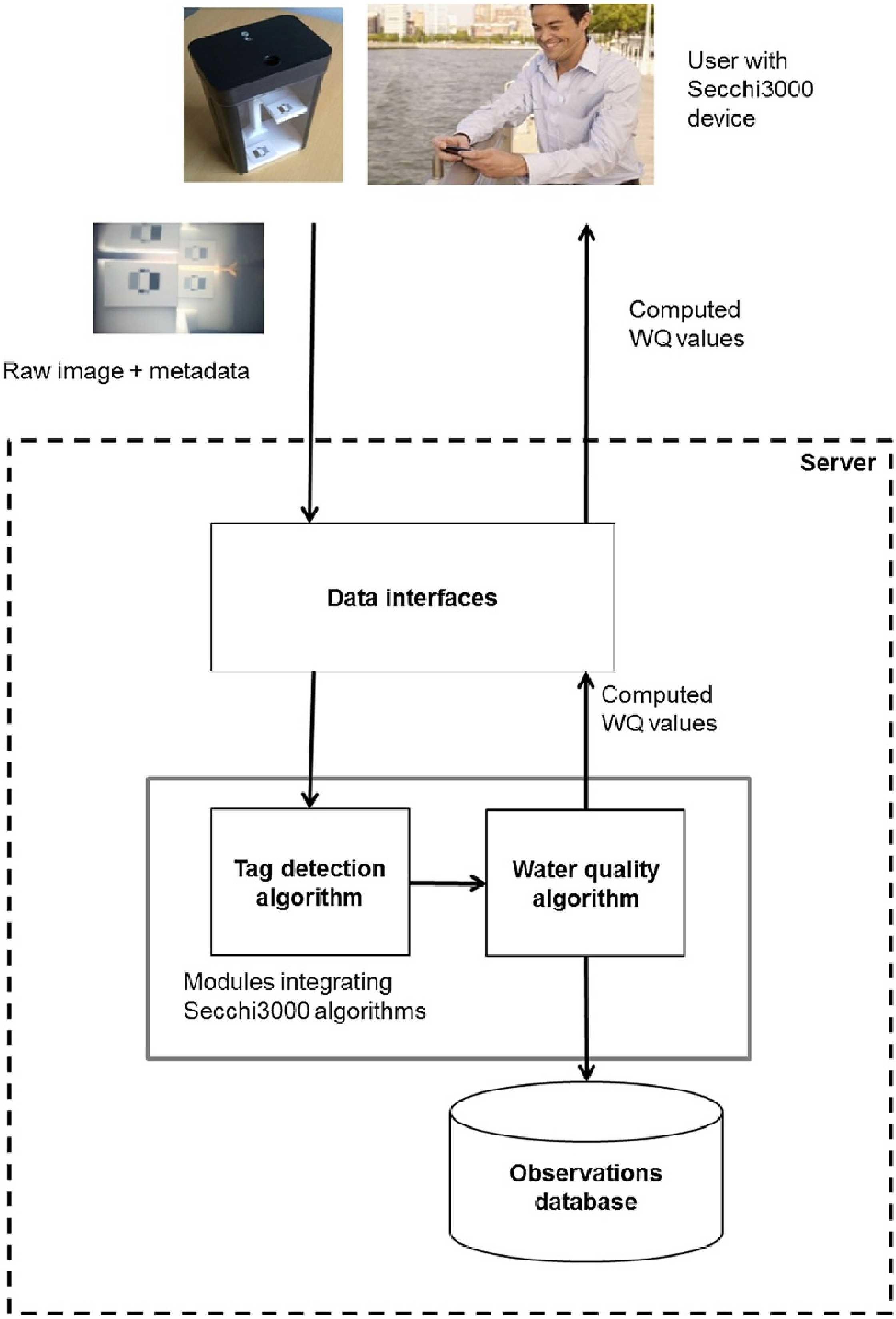



Figure 3.1: The proposed data flow diagram for the Water quality analysis using an inexpensive device and a mobile phone project

"Image Processing Techniques Applicable in Wastewater Quality Detection: Towards a Hygienic Environment" discusses water analysis from a different perspective to most of the other papers in that it talks about image analysis of wastewater in order to prevent environmental damage. It goes on to discuss the techniques used to briefly discuss some of the techniques used that they wish to replace with image analysis such as "Flow Injection Analysis (FIA), Sequential Injection Analysis (SIA), Multi-Commutated Flow Analysis (MCFIA) Multi-Syringe Flow Analysis (MSFIA), Lab-On-Valve (LOV) and Multi-Pump Flow Systems (MPFS)"[11]. This shows that in various industries that a simpler image analysis method is needed. It goes on to discuss the basics of image classification which at this point of the project are too simple to be of much help. While this paper does provide some in-depth discussion into its methodology it does so for its analysis of microscopic substances such as flocs. This isn't helpful because microscopic substances are not within the scope of this project.

The rest of the papers in this section will be discussing image classification based on the determining the presence of biological matter mainly micro organisms. Based on this fact, these papers won't be as helpful as those previously discussed because this project doesn't deal with microscopic material. But they do still provide some relevant information such as their classification algorithms, image preprocessing techniques and image collection methods.

The first of these papers is "Water Quality Assessment by Image Processing" by Karel Horak, Ian Klečka and Miloslav Richter.[12] This paper details their examination of water purity through the introduction of specific micro organisms to water samples. These micro-organisms are Lemna Minor and Daphnia magan. They were used because Lemna Minor changes colour when it comes in contact with toxic material while Daphia magna begins to move slower in the presence of toxic substances. For the analysis of Lemna minor they used a classifier that detected when a sufficient colour change had occured between images. This technique could also be applied to comparing the turbidity of two different water samples which is helpful for the standard camera analysis of this experiment. To monitor the movement of Daphnia magna they used time delayed imaging and would calculate the difference between the images. If the difference between the two images with a given timeframe was below a certain threshold then this would imply the presence of toxic materials because the Daphnia has not moved enough. This however provides little to our project so it won't be discussed further. A point to note however is that the equipment setup used in their experiment is quite similar to the one proposed for this experiment with a top down view of the water. You can see this in figure 3.1

"Microorganisms Detection in Drinking Water Using Image Processing" by H L Fernandez-Canque, B. Beggs, E. Smith, T. Boutaleb, H. V. Smith and S. Hintea may seem similar to the previous paper but there are some key differences.[13] For one instead of using a micro-organism to tell if the water was pure the experiment was dedicated to detecting the presence of a specific micro-organism (Cryptosporidium). Another difference is that this experiment was geared towards drinking water specifically which makes it more relevant to our project. The most interesting aspect of this paper was the preprocessing of the images. The images were preprocessed with object detection to locate all objects not of an appropriate size for the specified organism. All of these objects were then filtered out of the image then a noise filter was run over the image. Finally the classifier was run on what was left in the image to see if it matched the appropriate organism. This preprocessing information is quite interesting and could possibly be incorporated into the project but not likely. Besides this the information of the classification techniques was almost absent from the paper so it didn't provide much information in that regard.

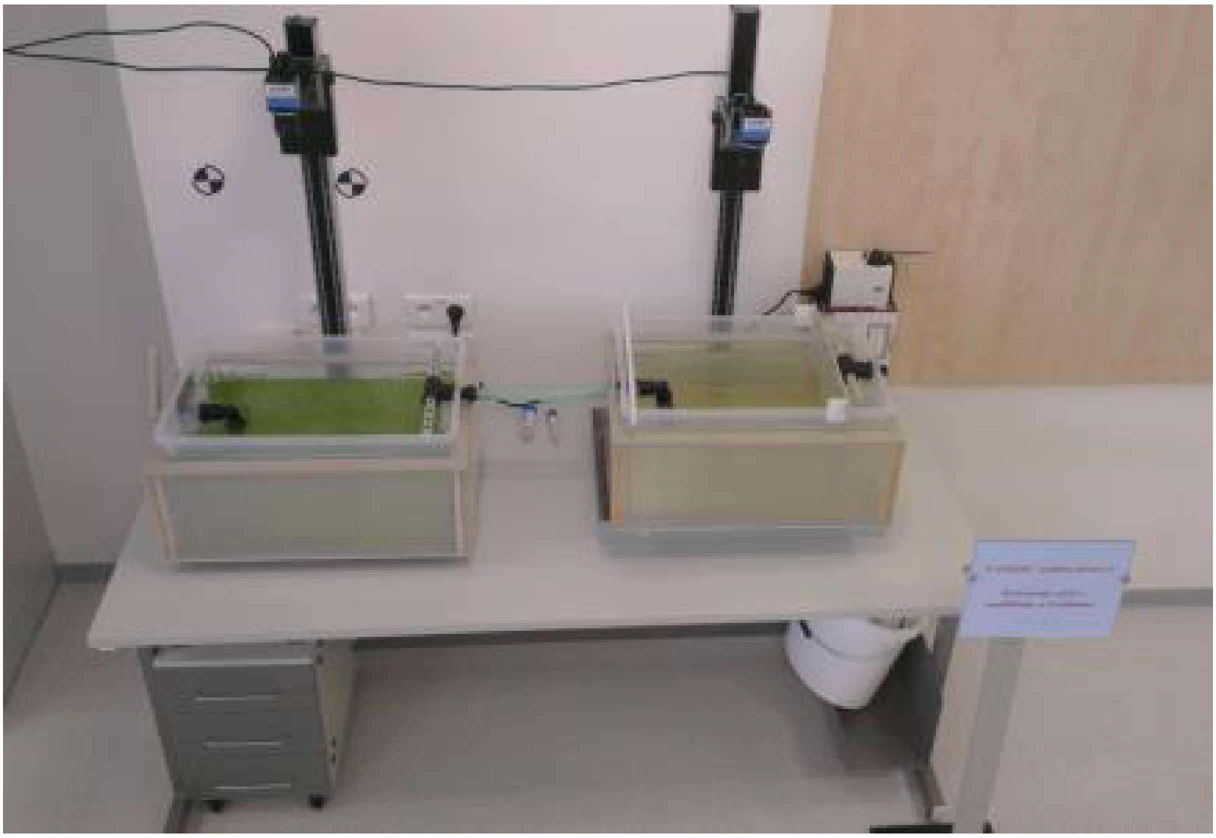


Figure 3.2: Water Quality Assessment by Image Processing, hardware setup

## 3.3 Image Classification and Deep Learning

There is currently only one source for this section being "Deep Learning For Computer Vision With Python"[14] by Dr Adrian Rosebrock. This is due to the fact that this document provides all the information necessary for the deep learning intended for this project. It provides in-depth examples of how to construct many different image classifiers. From the simple KNN strategy to the more advanced Convolutional Neural networks and everything in between. It provides coding examples, explanations on every aspect of the algorithms, advice on how to adjust them for better accuracy and comparisons between the algorithms so you can determine which is appropriate for you problem. As you will see in the outline of approach, most of the current methodology of this project has been based on this document. While this is unlikely to change other documents will more than likely be incorporated in to improve accuracy as the project continues.

## 3.4 Justification/Other

Three documents that were quite helpful for the project but didn't fall into the other categories include the "European union (drinking water) regulations"[15], "Drinking Water Quality Assessment" by J Aryal, Bina Gautam, Nawraj Sapkota [16] and "Parameters of water quality, interpretation and standards" by the world health organisation.[17]

The "European Union (Drinking Water) regulations" and the WHO'S "Parameters of water quality, interpretation and standards" provide a plethora of useful information for this project. More specifically they provide a list of chemicals/materials and how much of that said material can be found in a water sample before it is deemed unfit to drink. This is the standard by which the labels for the water samples in this project will be determined. So when determining if a given water sample is impure the amount of the substance present within it must be compared with the results found in this document. It also provides some explanations on terms used in the field of water analysis that could be helpful for the research being carried out in this project such as domestic distribution system.

Finally "Drinking Water Quality Assessment" by Aryal, J and Gautam, Bina and Sapkota, Nawraj is included here because it is a perfect justification piece for this project. The paper discusses the poor state of living in Nepal in terms of clean water. Throughout the paper the methods of collection and processing of the water are discussed but compared to other documents research for this project such as APHA the information provided here is trivial. They key point to take away from this paper is how poor the drinking water quality is. "85.71% of water samples showed higher Arsenic value (72) than WHO value" with the WHO value referring to the internationally recognized standard for drinking water.[17] This shows that in poorer countries clean drinking water is harder to come by and methods to check if any water found is clean is even scarcer. This provides perfect justification for the creation of this project, to help resolve problems like this.

# Chapter 4: **Data Considerations**

The only data to be used for this project is the collection of water sample images. There will be three sets of images; the standard camera images, infra-red images and the two previous image collections combined.

The cameras that will be used during the project have yet to be determined. Special considerations will have to be made into deciding the camera in order to maximise efficiency and prevent a large amount of preprocessing with the images. The cameras will be suspended above a tank filled with water in order to capture the images. This is done so that the glass from the tank doesn't interfere with the experiment and that as much of the image is filled with the water sample as possible.

At first the images will be given one of two labels, either pure or non-pure. Only water that falls under acceptable drinking requirements set by the EU[15] and WHO[17] will be assigned the pure label. The impure water samples will be manufactured by adding specific contaminants to a sample of pure water and tracking the amount of the contaminant added. The current contaminants currently are oil, dirt, plastic residue, salt, algae, glass and alcohol. More materials may be added later into the project but for now this list will be sufficient. The images will be marked with what material is in the water if one exists. This information may be used later in the project when the aim will be to determine what material is in the water. The water will not have more than one material in it at a time.

At least several dozen photos will be taken of each contaminant in the water sample. For images to be captured during the suspension of the materials the cameras would have to be set up on a time delayed capture system. This wouldn't be necessary if the infra red and standard cameras didn't have to capture the exact same image. But in order to ensure that any difference between their outputs is only from the camera types and not different data sets this is will be necessary. This will require the cameras be angled so that they are photographing the same section of the water to ensure minimal difference between the two datasets.

The images shouldn't need any cleaning up or noise removal. Since the camera models haven't been determined depending on the images produced by the cameras the images might have to be formatted to match the others images format. So the images may need to be cropped or have their shape adjusted in order to match. This will help eliminate most differences between the two data sets which will hopefully help in minimizing the bias between them to be exclusively the difference in camera type.

Their are two reasons for creating a personal database over trying to collect from external sources; it is much easier to create an extensive data set through taking the images rather than collecting them through online resources. Some positives for creating the database include the images can be easily formatted to the required standards and altering the specific materials in the water samples would be impossible from online resources. The second reason is that collecting an extensive data set with labelled images including infra-red images seemed near impossible especially in a standardized format that could compare the accuracy between the two image types.

By creating the data set it will limit the usability of the results in other settings because it eliminates a significant amount of the variance in the images meaning the product will be difficult to use outside of the previously established scenario simply due to the nature of image classification. But the scope of the experiment is more to prove the application is possible and easy to implement rather than to create a suits all purpose deliverable.

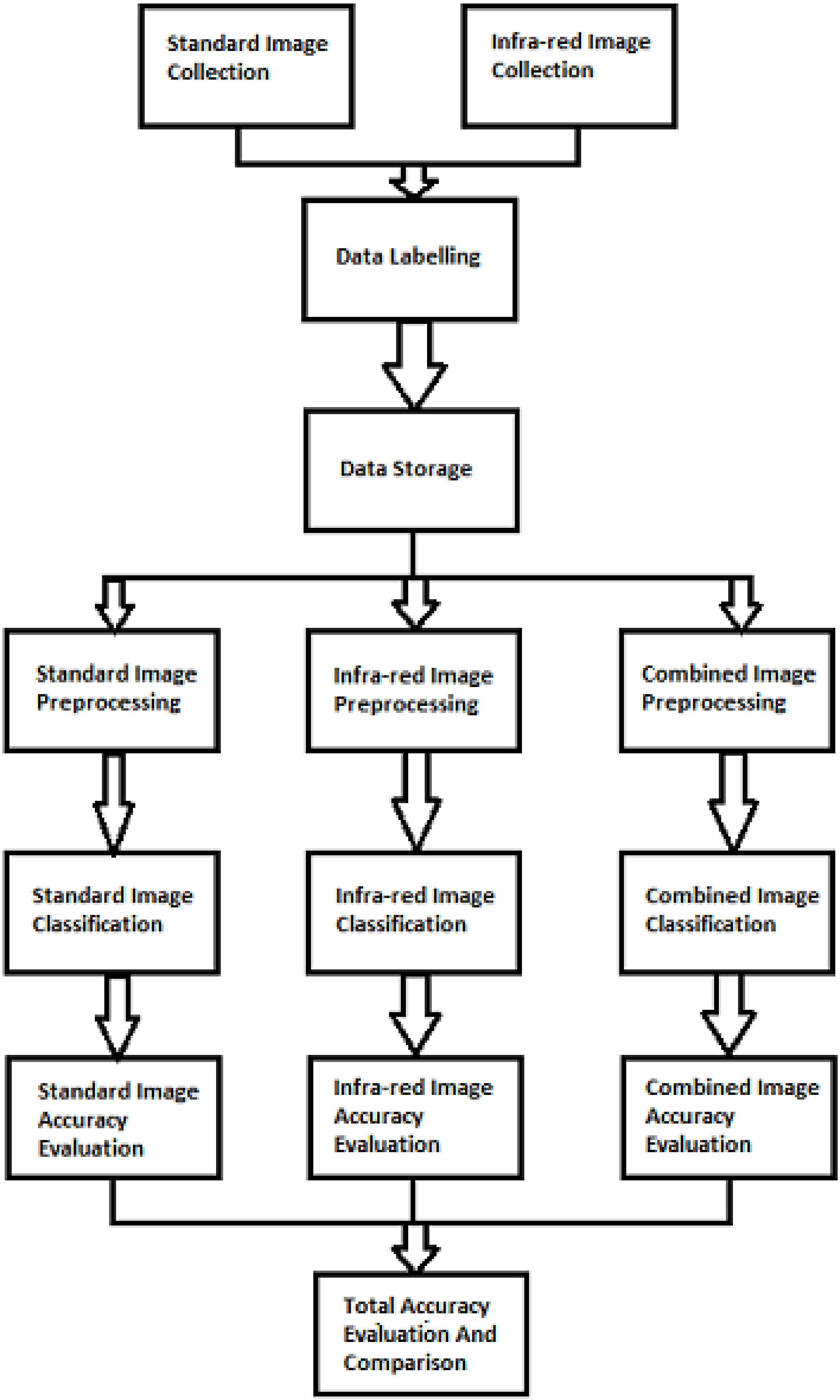



Figure 4.1: The proposed data flow diagram

# Chapter 5: **Outline of Approach**

For the initial stages of the project a simple four layer convolutional neural network will be used. This will be implemented using the python programming language and the functionality of the

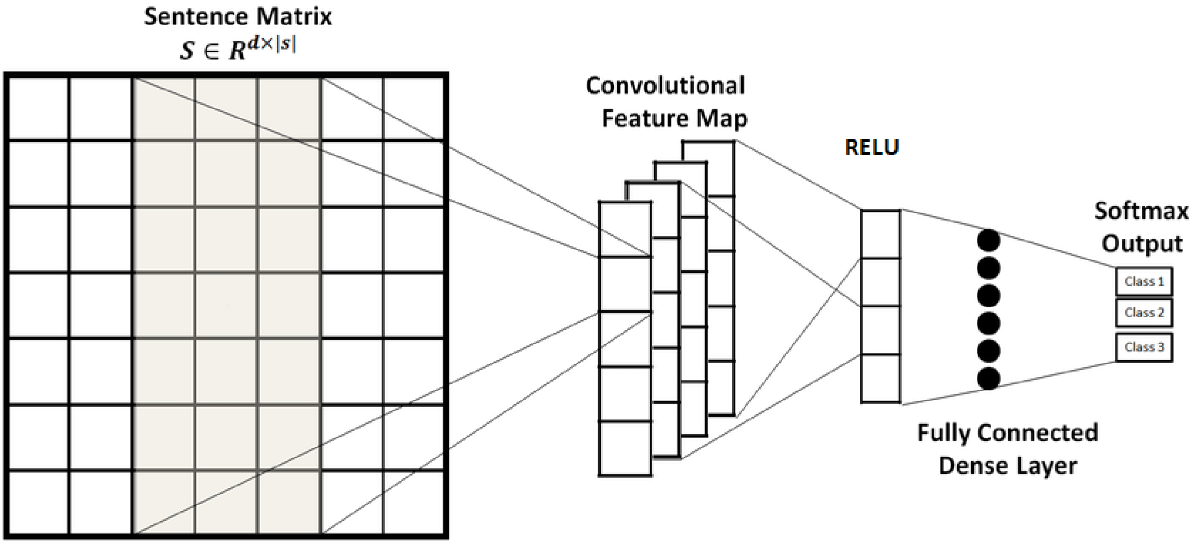



Figure 5.1: The proposed ShallowNet architecture

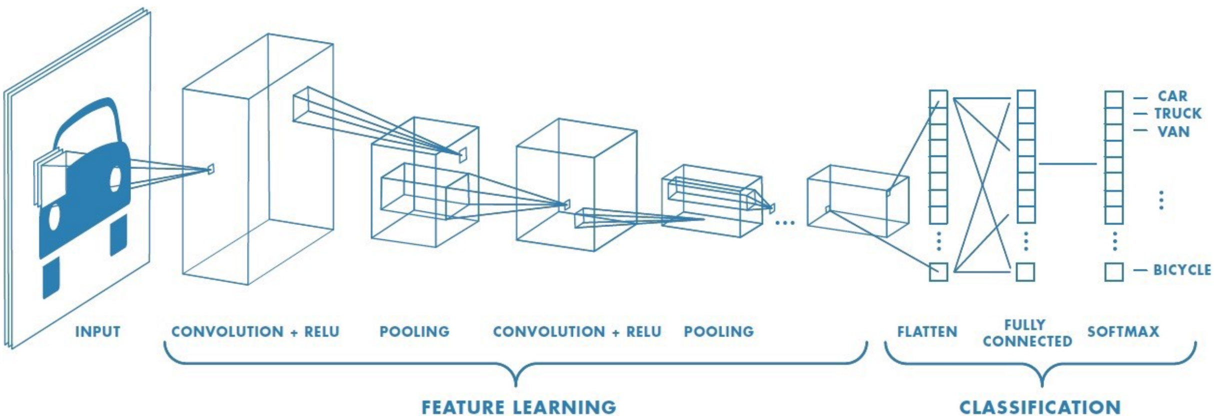



Figure 5.2: A more advanced version of the CNN used in this project

keras library, a open-source neural-network library written in Python. The reasoning for the simple structure will be discussed later in this chapter.

The intended architecture that will be used is referred to as ShallowNet by the Deep learning for computer vision with python starter bundle[14]. It starts by taking in the address of the dataset from the command line. The dataset will either be the standard camera images and/or the infra-red images along with their respective labels (clean or unclean). These images will then be put through a number of preprocesses.

The first preprocess will scale the images to a predetermined size. It is currently unknown what size the images will be set to, this will be determined after testing is carried out with the cameras.

The next proprocess the image to array process will order the channels of a given input image based on our image data format setting. For the initial stages of this experiment the default keras settings of channels last will be used. The raw pixel intensities are then scaled to a 0-1 range.

The data set will then be split into a training and testing set. More than likely this will be a 75% testing 25% training split but this could change as the project develops. Since the labels are binary (either positive for contaminant or negative) the labels will not need to be converted into vectors. Until the contaminant classifier is implemented.

Then a Stochastic gradient descent will be created with a specified learning rate. This learning rate is so far not decided, currently it will be set to 0.005 but this will more than likely be changed as experimentation continues.

Next the shallow net model will be constructed. To construct this model it must be supplied with the height, width, depth and class of the images. In this case width and height haven't been determined, depth refers to the RGB values so it is three and class refers to the class labels your images could be assigned so two for clean and unclean. The model will order depth, width and height differently depending on the specified data format. Our model will have it ordered as (height, width and depth).

The convolutional layer is then added to the model. As of right now the layer will have 32 filters each of size 3x3. As with all our parameters these are all flexible and will more than likely be changed later on in the experiment.

For our activation layer we will use Relu (rectified linear unit) simply because it is the most commonly used activation strategy in deep learning. Later on this might be changed to leaky relu or parametric relu if that improves the accuracy.

Then our model will be flattened into a 1D array, a dense layer will then be created using the same number of nodes as class labels (2). Finally we will apply the softmax classification to the model which will return the class label probabilities for each class in our data set.

After the model has been instantiated with the shallow net architecture, cross entropy will be set as the loss function and our previously defined SGD is set as our optimizer. Then our model will be trained using a standard backpropogation neural network algorithm. Both the predictions and loss will be calculated using the standard NN methods.

The reason why a simplistic approach was chosen to start this project is because it will make the project easier to redesign and develop. If the project started with the most complex and thought to be accurate solution to the problem then there is a large chance a more accurate method could have been missed during the development process. For this reason it is possible that the first iteration of this project could be a basic neural network implementation in order to test all possible solutions.

For the hardware in this experiment a fish tank will be filled with water. This water will have different materials added to it throughout the experiment in order to diversify the images for the data set.

The two cameras will be suspended above the water tank pointing downwards towards the water. They will be angled so that they are taking extremely similar images in order to eliminate diversity between the data sets.

If getting the cameras to take extremely similar images proofs too difficult to accomplish then the images will simply have to be cropped in order for them to contain the same area captured.

See figure 5.1 for an illustration of the projected hardware setup.

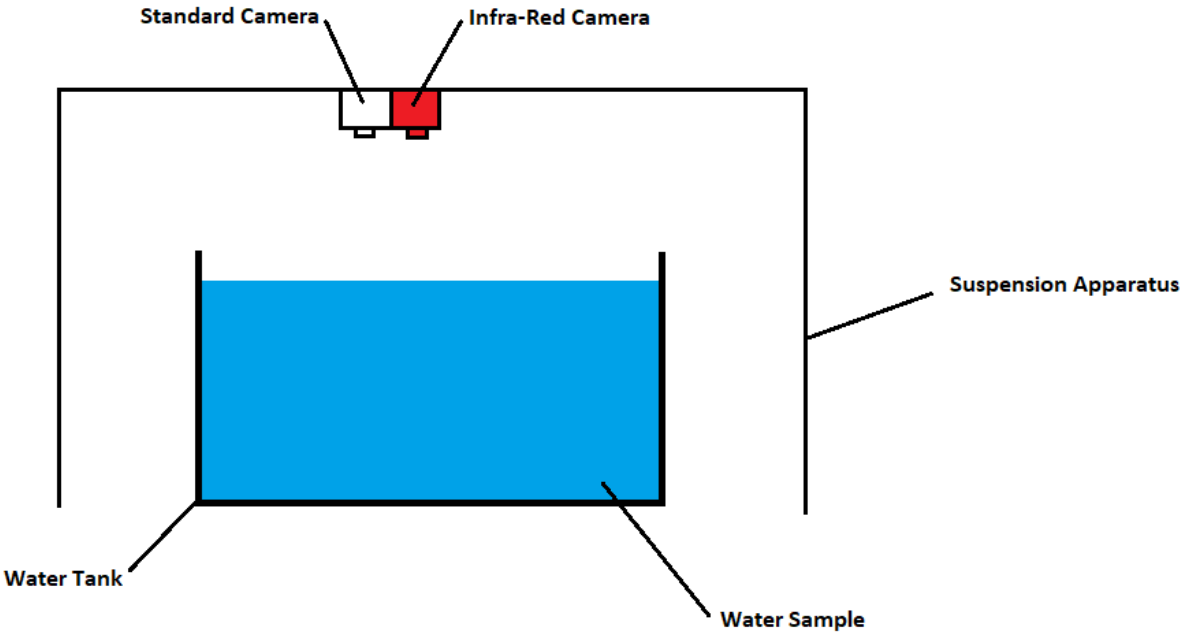



Figure 5.3: The structure of the hardware

# Chapter 6: Project Workplan

The current plan for the project is to spend the time between the submission of this report and the interview preparing for the interview by constructing my presentation and further researching various aspects of the project such as better CNN formulas.

From the completion of the presentation until the first week of the new year will be spent further researching improvements to the CNN formulas. The cameras will hopefully be examined and experimented with while ensuring their specifications and produced image quality are noted. This needs to be done to know if any adjustments need to be made and so we know how much preprocessing must be carried out on the image before they can be classified by the algorithm.

Next during second week of the new years will be spent collecting the images for the data set. This will more than likely take numerous days of changing the set ups and ensuring the standard of the images are good.

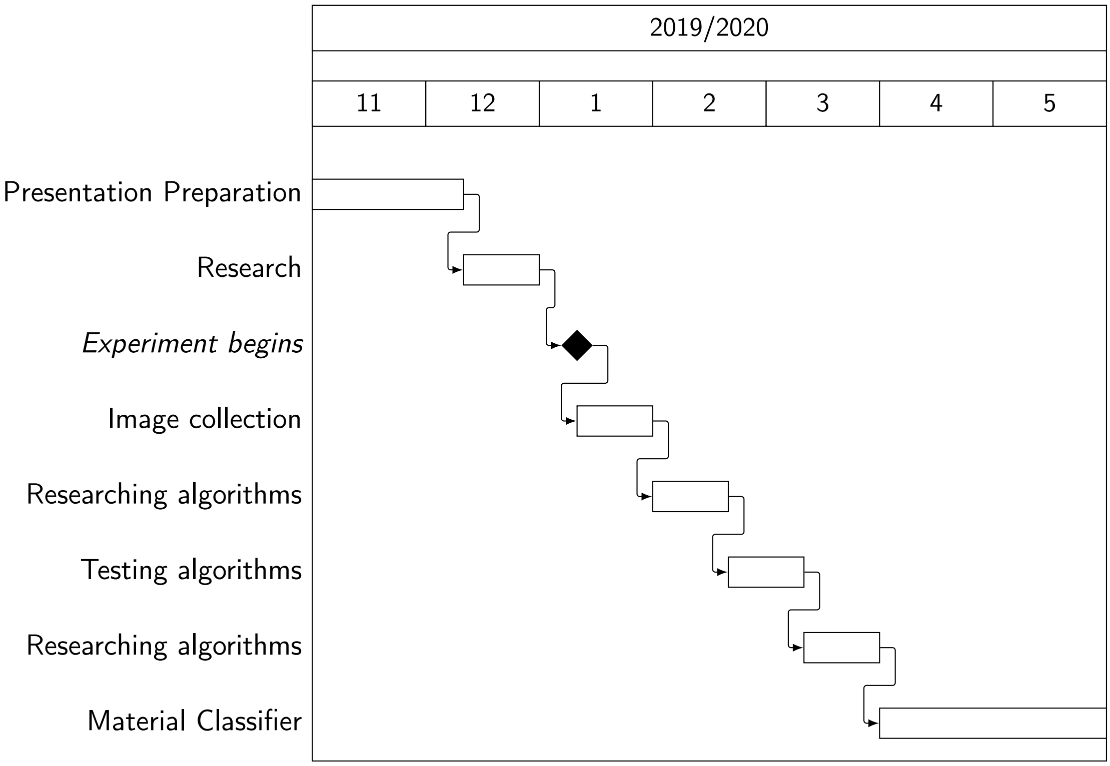
After this the remainder of the time will be spent testing and adjusting the CNN algorithms to maximise efficiency. Unfortunately not much more information can be given about these activities because the amount of change needed is currently unknown.

If a acceptable accuracy can be reached before the sixth week of the second semester of college then progress onto the contaminant identifier will begin. This process of development will be quite similar to the previous experiment. It will simply require changing the CNN in order to classify for a larger label set. This will more than likely lead to a drop in accuracy so the rest of the time will be spent attempting to improve the accuracy of this classifier assuming the original classifier has reached an acceptable standard.

The evaluation of this project will be have two layers. First the accuracy of the classifier itself. The higher the accuracy of the classifier the more successful the project will have been. A good benchmark to compare to is the AquaSight project and its 96% accuracy.[2]

The second layer of evaluation will be the completion and accuracy of the contaminant classifier. If the project is done to a well enough standard and is on schedule then we will move on to the classification of the contaminants in the water. Similarly to the first evaluation this can be judged based on the accuracy that the classifier returns.

This is the best method of evaluation of a classifier because it is not affected by personal bias and relies solely on the numbers presented to you by the program.



2019/2020
11
12
1
2
3
4
5
Presentation Preparation
Research
Experiment begins
Image collection
Researching algorithms
Testing algorithms
Researching algorithms
Material Classifier

# Chapter 7: Final Report Additions

## 7.1 Final Report Sections:

- Data & Context
- Core Contribution
- Evaluation
- Final Conclusions - Societal Contribution
- Final Conclusions - Final Thoughts

## 7.2 Project Link

Github Link or go to the next url: `https://github.com/charliekelly13/Final-Year-Project`

# Chapter 8: Data And Context

## 8.1 Camera Types

For this project the original idea was to use one camera with different lens in order to create different image types for an image classifier. However as the project progressed it seemed much more reasonable to incorporate infrared and thermal cameras instead of just lens. This was partially due to the difficulty of getting the appropriate lens and the availability of the other cameras. So the project had five camera types: standard, microscopic, thermal, long wave infrared and shortwave infrared.

**Standard:** The default/standard camera was a Blackfly S USB3 using a C-mount Navitar lens. This produced images on the visible light spectrum. These images were essentially used as the control of the experiment. Based off of several experiments carried out during this project images of visible light have limited usability in water image classification. This is due to the reliance on the turbidity of the water for gathering information on the contaminants. This is a problem because several of the contaminants can produce images with similar turbidity. These images were quite useful but on their own they produced quite a poor classifier. The cameras specific software "Spinview SDK" was used to capture the images.

**Microscopic**: Similar to the standard images these images were produced with the Blackfly S USB3 except this camera used a computar 1.0x, 5 Megapixel Telecentric Lens in C-Mount for 1" Sensors. This produced images on the visible light spectrum that resolved down to six microns. Similar to the standard images these images also suffered from a dependency on the turbidity of the water. However these images presented their own set of problems: they were very dependent on the position of the camera such that if the camera wasn't directed at an area the contaminant was present it couldn't detect the contaminant due to the small visual range of the camera, it also hindered the classifier when it came to edge detection. Since the field of view for the camera was so small it struggled to view enough of the contaminant for feature and edge detection to be useful. I used the same software for these images as I did with the standard images.

**Near Infrared:** For the near infrared images a NIR USB Mobius camera was used and it was operated using the built in windows camera functionality. This camera captures near infrared light between the wavelengths of 1,100 and 3,000 nanometers. Since this camera does not work off the off the visible light spectrum the classifier based off of it is not as dependent on the turbidity of the water in order to make informed decisions compared to the standard image classifier. This indicates these images being very useful for the classifier and even independently they produced good results.

**Long-Wave Infrared:** This camera captured images between the wavelengths of 8,000 to 12,000 nanometers also known as long-wave infrared. This camera was built into a Raspberry PI computer and worked off computers built in camera functionality. Similar to the near infrared camera since these images were not caught on the visible light spectrum it is less susceptible to the turbidity of the water. The different wavelength allows for this camera to cover a different category than the other cameras which allows for a deeper collection of information by the classifier therefore increasing it's accuracy.

**Thermal:** For the thermal images a thermal Seek usb camera was used. The camera was plugged

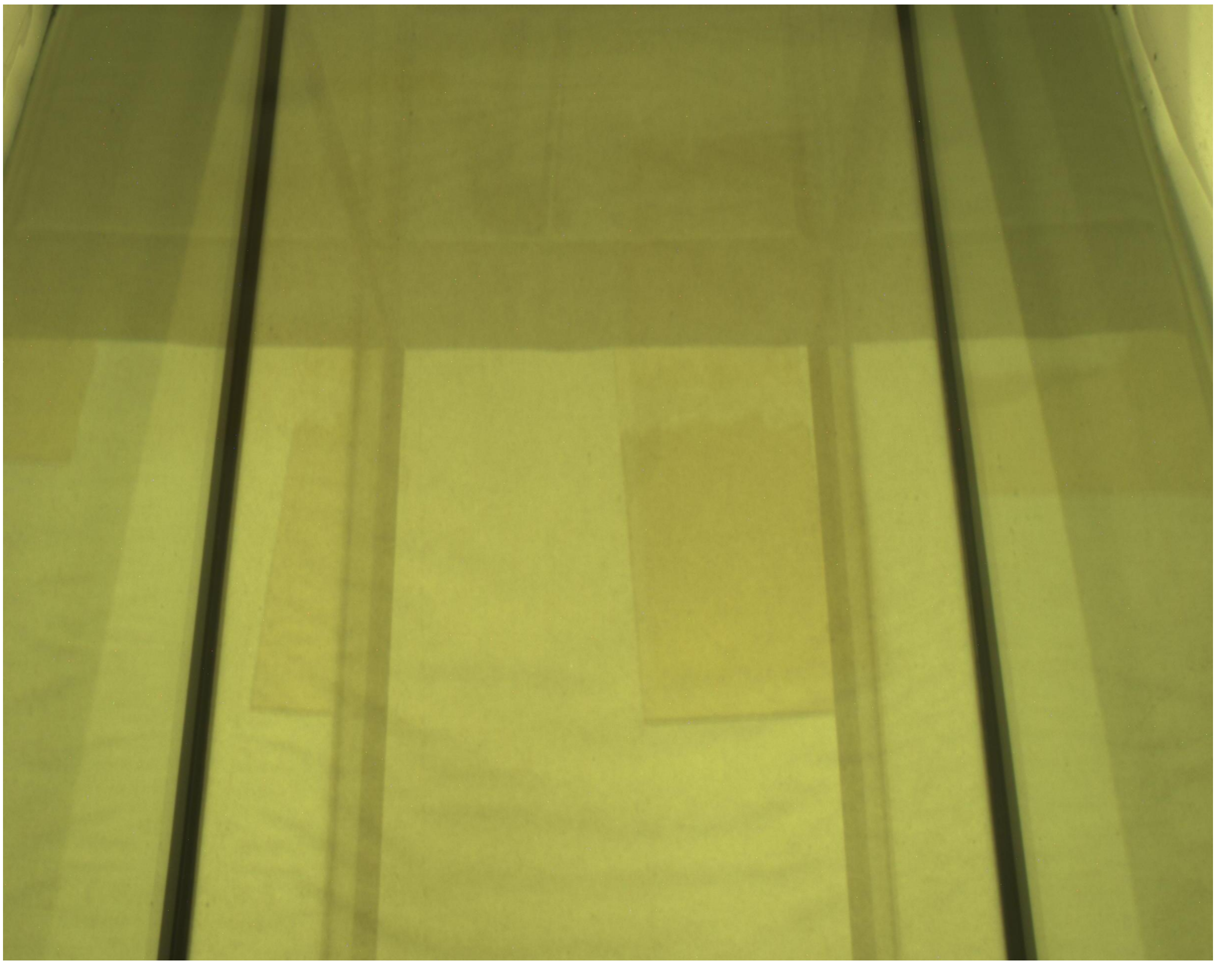


Figure 8.1: Sample Standard Image

Figure 8.2: Sample Algae Image

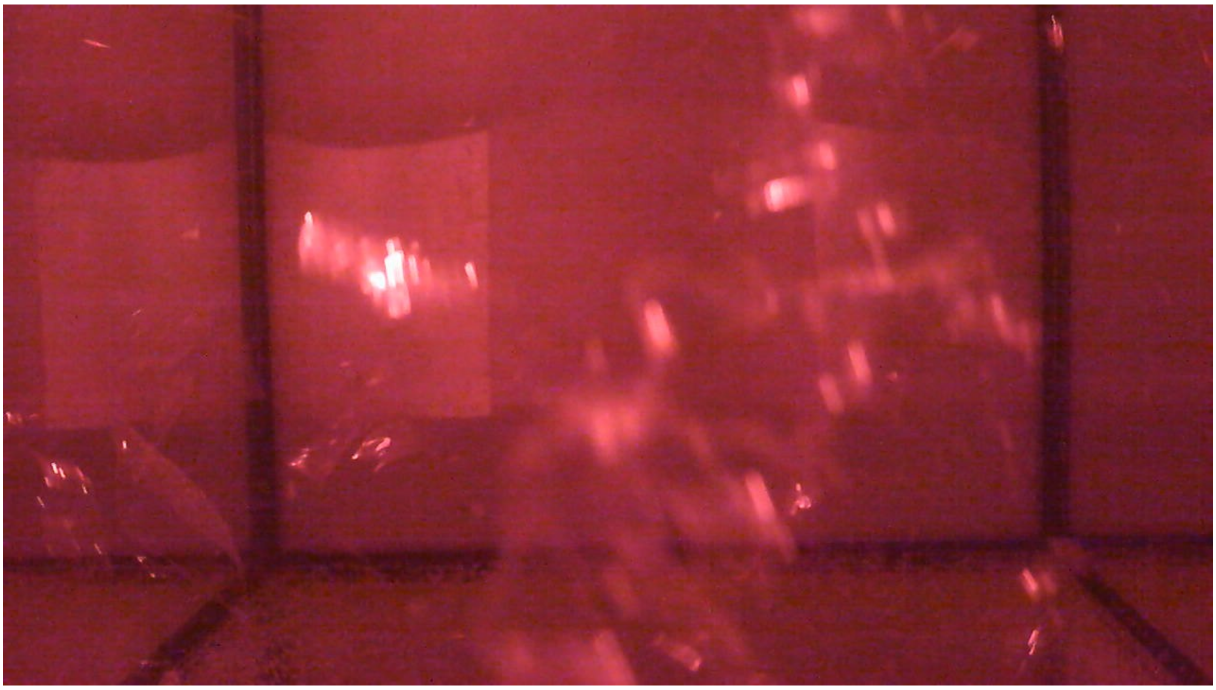


Figure 8.3: Sample Near Infrared Image

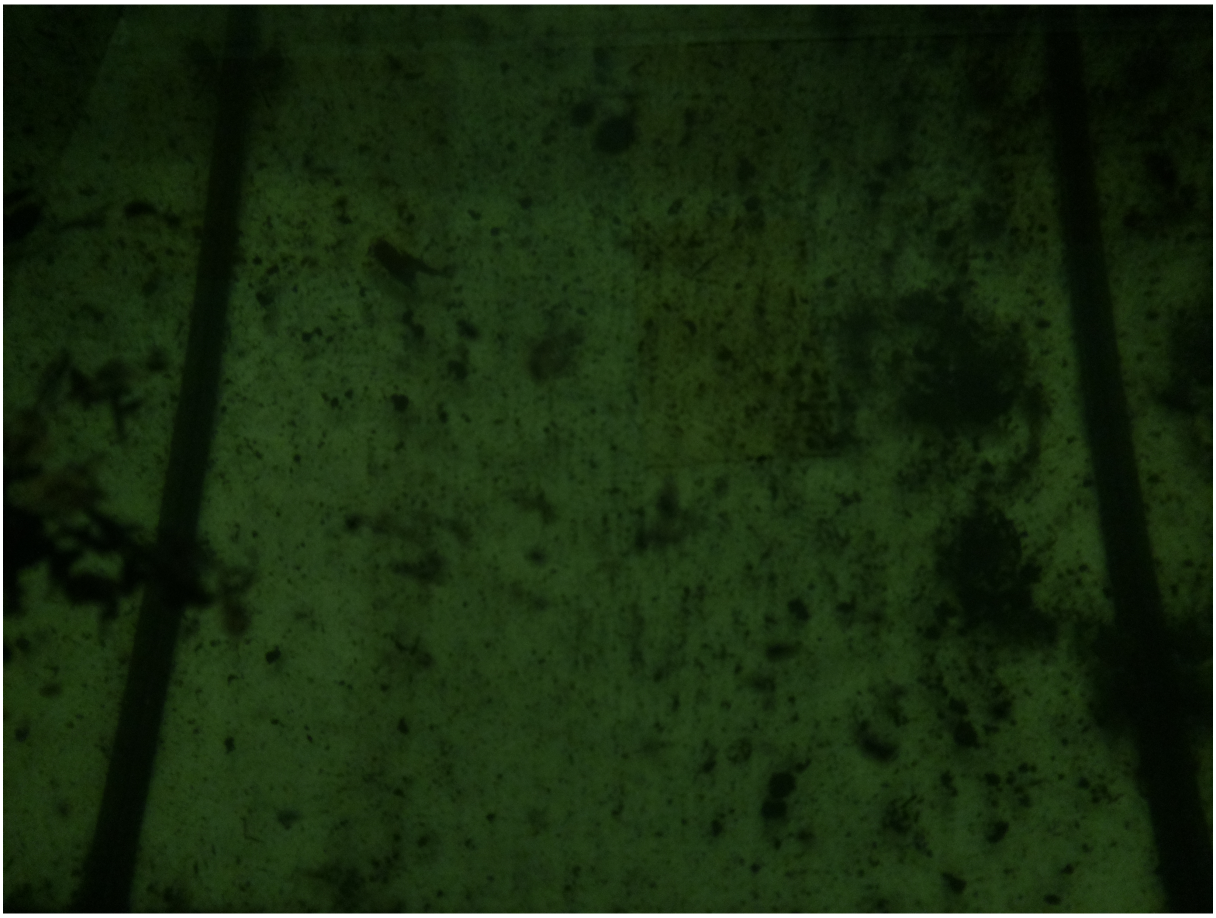


Figure 8.4: Sample Long Wave Infrared Image

into the usb port of a phone and using the seek thermal phone app the thermal images were taken. This is another form of long-wave infrared image but instead the wavelengths captured are from 9,000 to 14,000. With rises in temperature more wavelengths of this range are produced which allows these images to capture differences in temperature values. Initially this was expected to be very helpful to the project and that the differences in heat would be able to detect the different heat levels of the substances in the water. This whoever turned out not to be the case and instead the cold water and the contaminants produced quite similar heat signatures in the images. This might have been rectifiable by putting heat through the water in some form then the differences in heat patterns caused by the insulation capabilities of each contaminant could have been viewed. But due to the setup of the experiment and the time limit this was not feasible so the images produced by this camera were not of great benefit to the project.

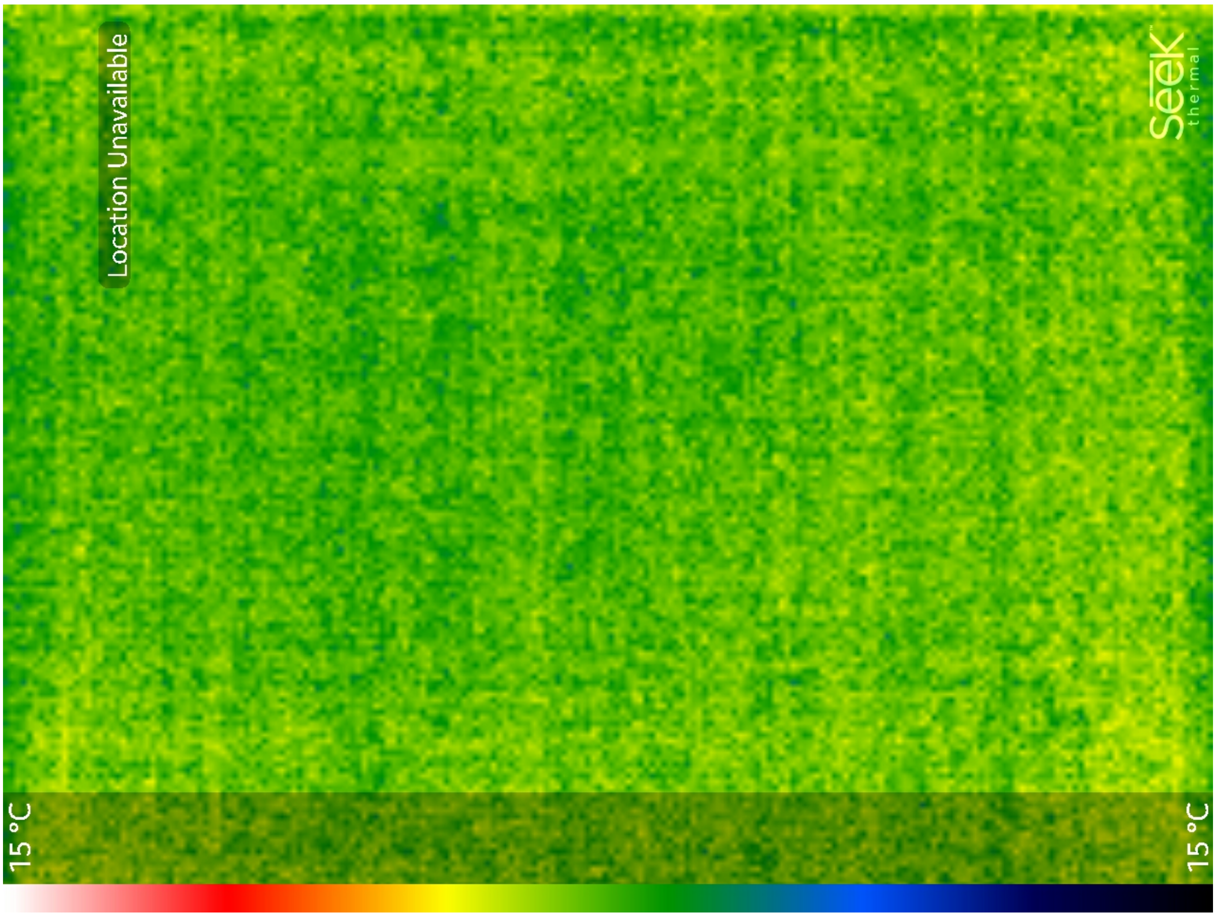



Figure 8.5: Sample Thermal Image

## 8.2 Image Alterations

Alterations to the individual images were kept to a minimum for this project. Since all the cameras were positioned in the same location with the same angle and the lighting was kept the same throughout there was very few alterations that had to be made to the images.

The only alteration of note was done on the standard camera images. The specific lens that was on the camera caused a large black rim to appear around the circular image. So ImageMagick was used to crop all of these images to remove the black border and make the images square to match the rest of the images. The cropping of these images led to a one percent increase in the classifier containing the near infrared, standard and long wave infrared images.

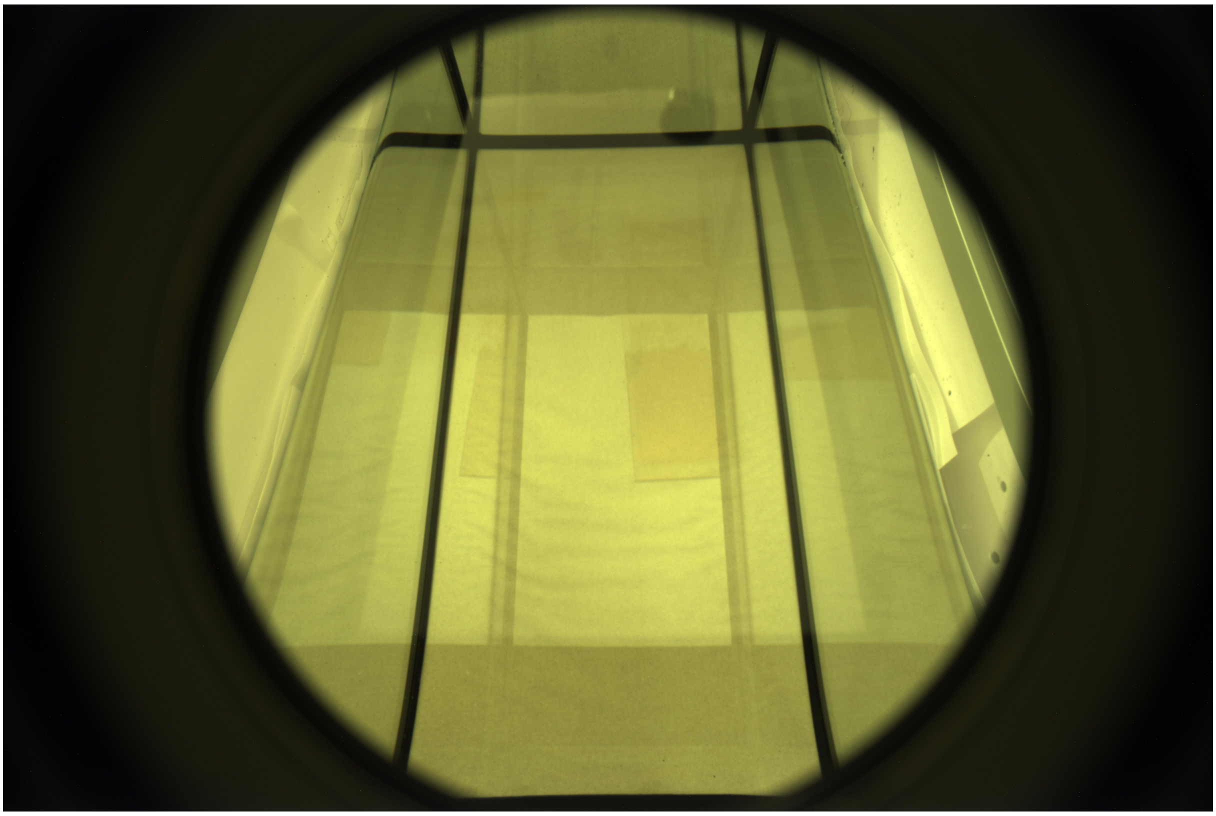


Figure 8.6: Non cropped Standard Image

## 8.3 Data Considerations

All of the data-sets were split into two different types, top down and front facing images. For the top down images the cameras were suspended above the water tank and the front facing images were taken by cameras propped in front of the tank.

The tank sides were covered in white paper in order for the images not to be effected by the background of the image collection environment because that information would be useless for this image classifier. For lighting a lamp was positioned on the opposite side of the tank to where the camera was. The lamp was positioned above the tank pointed down at the tank to fully illuminate the tank. The lamp was the only source of light in the room. From testing different types of lighting it was discovered that anymore light or lighting at a different position would cause a glare on the top down images that lowered the image quality for this experiment especially in the microscopic and standard images.

Each one of the subsets of data had their own breakdown of how many images were collected for each of the contaminants. Excluding soil and clean every contaminant had three hundred and twenty images per camera. Soil and clean had less images because it was believed that more images would not have provided more visual information if their image amount was increased. Soil

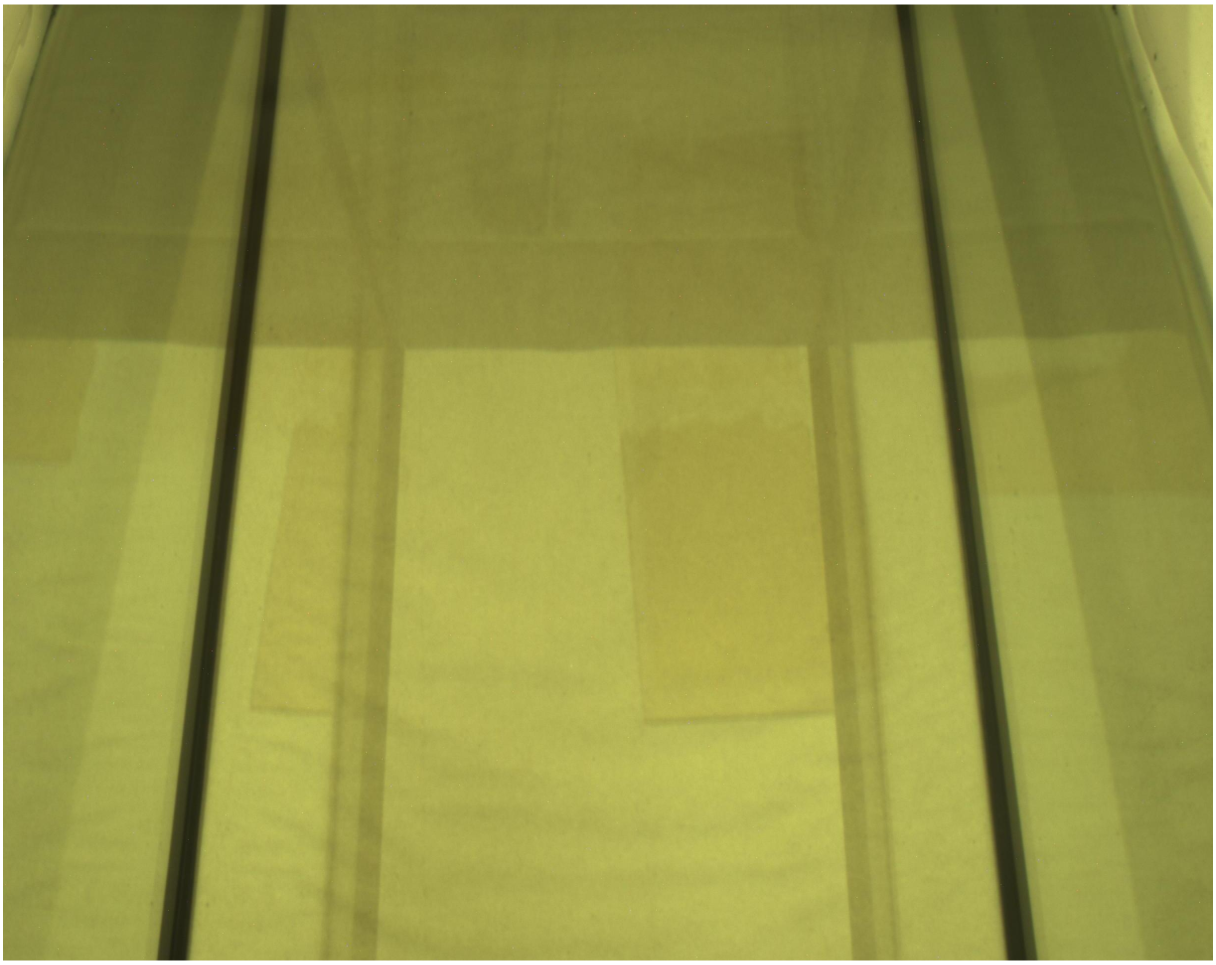


Figure 8.7: Cropped Standard Image

in higher concentrations made the water appear wholly black in the lighting so adding more soil seemed irrelevant because it wasn't producing any visual differences in the water. Similarly for clean water no additional visual information could be gathered from additional images because with the lighting being the same all of the images of clean water would present the same visual information. So the amount of images were left smaller than the other data sets. More images of soil at lower concentrations were not taken due to the saturation issue becoming known at a later point of the project and a time restraint issue didn't allow for the earlier part of the experiment to be repeated. Also adding more images to the lower quantity could cause a further imbalance within the database.

Salt and sand had the exact same breakdown of contaminant quantity. They both had nine sets of images of their contaminant in different quantities. These quantities were one gram, five grams, ten grams, twenty grams, fifty grams, one hundred grams, five hundred grams and one thousand grams. These sizes were chosen to give a good diversity of contaminant sizes while remaining realistic to what could appear in a body of water that is fifty liters in size. Also some research was put into the amount of salt that is considered dangerous or unhealthy to have in water, which is proposed to be one gram per fifty litres of water.[18] Since salt is considered the least dangerous of these substances because it is regularly consumed it was used as the basis for the quantities of the other contaminants. This may be a naive approach since the other contaminants such as soil and sand could be dangerous at higher or lower quantities but there is not much information available on how much of these substances are allowed in water before it is considered unfit to drink however it seems likely they are less fit for consumption that salt so presenting them in the same quantities should ensure they are also unfit at these quantities.

The point at which images of soil were no longer taken was after the one hundred gram mark due to as mentioned previously the over saturation of the water caused the images to become too dark and after testing adding more soil did not change the images on the camera because it had already become over saturated at that point. So instead of the three hundred and twenty images that salt and sand had, soil instead had two hundred and forty images per camera.

The algae and plastic images worked off of their own system. They were more limited because acquiring plastic and algae was more difficult compared to the other contaminants. For both of these contaminant types the quantities were kept smaller, there was less variation in the quantity amounts but the amount of images per quantity was increased.

For plastic the quantities was lower because the goal was for the classifier to be able to detect plastic in low amounts. Also it was very difficult to get large quantities of plastic that could appropriately be in water. So for plastic there was three hundred and twenty images split between five groups each with thirty two images per camera position. The quantities started at two grams and went up in increments of two for each group with a maximum of ten grams.

Algae posed the same problems a plastic in that larger quantities were harder to acquire and that detecting lower quantities was more important. The key difference is that in larger quantities algae poses the same problem as soil in that it over saturates the colour of the water and adding more algae wouldn't effect the overall image information produced. In algae the was three hundred and twenty images split between four groups of different quantities each with forty images per camera position. The quantities one gram, two grams, five grams and ten grams.

## 8.4 Image Comparisons

From the experiments carried out through this project the standard, near infrared and long wave infrared images were the most helpful for the experiment. Using these three image types produced the best results overall, this will be discussed in the results section but by using these three images types the accuracy was roughly ninety two percent versus using all five of the image types which produced an accuracy of roughly seventy eight percent. Next will be a breakdown of why these images produced the results they did along with an individual breakdown of the results they give when classified individually. There will also be a discussion on the results based on the individual contaminants.

The standard camera got quite poor results at fifty eight percent average accuracy. The reasons could be that salt and plastic were quite poor because with a standard camera turbidity has a major impact on the information gain. Since plastic and salt do not produce obvious colours that affect the turbidity it will be difficult for any information to be gained from these clear images. Another prevalent factor could have been that the turbidity of soil and sand can become quite similar at higher quantities. Both producing colours a dark brown colour at higher quantities. This can make it difficult for a classifier to gain information on sample images if multiple image categories produce similar looking images. This clearly backs up the previous point that the reliance on turbidity makes relying on exclusively standard images ineffective for a water contaminant classifier.

The microscopic images produced very bad results at fifty two percent average accuracy which is the worst performing of the classifiers in this experiment. The reasons for this are quite similar to the ones affecting the standard images such as being very dependent on turbidity while turbidity is not a decisive factor when it comes to these images with several contaminants producing the same turbidity. Another key issue with the microscopic images is that it had a startlingly bad accuracy when it came to the salt contaminant. To be precise it had an accuracy of twelve percent. This was more than likely caused by classifier identifying the salt images as clean images because on the normal light spectrum it would be quite difficult to tell the difference at most levels of contamination. They would only be distinguishable when the water becomes highly saturated with salt. Also the microscopic camera has a much smaller view span so it was very dependent on the contaminant being positioned directly in front of the camera which didn't always occur especially in smaller quantities.

The final of the poorly performing image types was thermal which had an accuracy of fifty three and a half percent average accuracy. This is mainly due to the previously discussed camera types section but there are some additional points of interest. The two contaminants that thermal does not perform well on are algae and plastic. This could be because these substances appeared in the lowest quantities. The maximum they appear at is ten grams while the other contaminants appear in up to one thousand grams. So it could be the quantity of the substances which is affecting the heat of the water and therefore providing more information to the classifier. It also could be that the insulation values for algae and plastic are the more similar to water than the other contaminants meaning there is little change in the heat patterns in the water. This is one theory but it is not relevant to this experiment so no further research was carried out on it.

The long wave infrared images performed exceptionally well compared to the other image types with an average accuracy of ninety percent. There is little to discuss with this image type since all of the image types performed very well. The only image type that had a large difference from the average score was plastic which had an accuracy of eighty one percent. Since the next lowest accuracy rating was eighty four percent this is that large of a difference. More than likely the reason for this is that the specific kind of plastic doesn't register as clearly as the other contaminants in long wave infrared images.

Not much can be said about the near wave infrared images. They achieved the best accuracy of all of the image types with a staggering ninety four percent. This result is good enough to be used as its own classifier. While this is a small sample size this shows that near wave infrared cameras may be the best for water contaminant classification, at least with regards to these specific contaminants. The margin of difference is minimal between all of the contaminants, the only contaminant below ninety percent is salt which is probably being mistaken for clean water in small quantities like with other classifiers. Clean water however almost has a perfect accuracy at ninety nine percent which shows this classifier has learned how to detect the presence of any contaminant to the degree it can almost perfectly detect an image with the absence of any contaminants.

A discussion on certain contaminants has already been given throughout the previous section so this will be a brief overall breakdown of the specific components. Salt was consistently one of if not the most poorly identified contaminant amongst the classifiers. As previously discussed this is probably due to the extreme similarities to clean water and plastic at lower quantities. Plastic has the same problem but to a lesser degree. This is probably due to the distinct edges create by the plastic pieces in the water being recognized through edge detection in the CNN. Clean water typically got the highest accuracy when classified. This is more than likely due to the classifier eliminating the option upon the detection of any material. Algae consistently received high scores (excluding in thermal images) due to its unique shape and colour. Finally sand and soil typically presented a high accuracy excluding standard and microscopic.

# Chapter 9: Core Contribution

This section will be dedicated to the development and implementation of the project.

## 9.1 Data Collection

Over the period of two months the images were collected in a controlled environment. A sixty litre tank was filled with fifty litres of water to allow room for the contaminants to be added. For salt, sand, soil and algae they were weighed out on a weighing scale. Unfortunately the weighing scale only moved in increments of one gram so their may be a discrepancy at the micro-gram level in the measurements.

As previously discussed salt, soil and sand were measured in the same increments. All three of these contaminants were store bought (food salt, plant soil, construction sand). Plastic was collected from clear food bags each weighing two grams. These clear bags seemed best for this experiment because it helped the classifier learn to detect see through plastic. These pages were cut into small pieces so they could be dispersed around the water. Algae was collected from the wild on a beach. It was washed to remove the other contaminants that could have present on the algae such as salt, soil and sand.

Contaminants were put into the tank one at a time. So for one gram of salt each camera would take it's set of images then the camera would be changed. This would be taken from one of the two capture angles then changed to the other camera angle then this process would be repeated for all the camera types again. This would then be repeated for all the quantities for a specific contaminant. Usually it would take two days to complete a contaminant, this was mainly because the images could only be taken during the night because the lighting needed to be kept consistent throughout the images to ensure that it wasn't effecting the overall classification results and the capture environment did not have a good method for blocking out sunlight.

Different lighting's were tested but the many of them would cause shadows, glare on the water also some would work for specific image types but not others. It also lead to some of the image types becoming too dark while the lighting would appear perfectly for other image types so it required a lot of testing to figure out the appropriate lighting to use for the images and the one that was used achieved the best quality images.

## 9.2 Neural Networks

For this project many different kinds of neural networks had to be tested to see which would achieve the best results. Five different networks were tested through this project. A basic convolutional neural network (shallownet), a single convolutional layer CNN with L1 and L2 regularisation, a two convolotional layer CNN with L1 and L2 regularisation and dropout, a four convolutional layer CNN with L1 and L2 regularisation, dropout and max pooling with each convolutional layer and a four convolutional layer CNN with L1 and L2 regularisation, dropout and average pooling neural network with difference in dropout.

### 9.2.1 Basic structure

The goal of this experiment was to create a system that upon entering an image of water the system could identify what the contaminant if any was. This is done by creating a training model using our own original dataset. Images are then entered into the training model and evaluated to see what class of image it falls under. To replicate this process of entering unseen images into the model we used a training testing split. Seventy five percent of the data was used to train the model to recognise specific class while the other twenty five percent was treated as the unseen data that needed to be classified.

This program is currently run off Google Colab but could be run locally or on a different online service. The images are stored on a personal google drive and accessed by the program using colabs specific functionality. These images are processed to be the same size (32, 32) then are converted to arrays of their rgb pixel values so that the images can be treated mathmetically in the neural network section for multiple with the weights and bias.

All of the layers in this program were implemented using the keras and tensorflow libraries.

### 9.2.2 Convolutional Neural Network

The purpose of a convolutional neural network is for the system to be able to learn filters that allow it to detect features within the image such as edges or "blobs". These low level features can be used as building blocks that form more complex structures. In this specific experiment those complex structures could be the shapes of the plastic or algae in the water.

This is performed by having different types of layers performing specific actions on the images. These actions are performed by the layers of the convolutional neural network. The specific layers that will be discussing are the convoltional layer, RELU activation layer, pooling layer, dropout layer and fully connected layer.

The convolutional layer contains K learnable filters that will be applied to the image. Each filter has a set height and width and is smaller than the input image. The filters convolve the image by moving across the image then altering the image with features such as sharpen etc. After applying all of the K filters we now have K, 2-dimensional activation maps. Then the maps are stacked along the dimension of our array to form out input. These filters move left to right on our image, they move in increments known as stride. In this experiment the stride is set to one because retaining image information is critical for this experiment and reducing the computational size is not needed for this experiment. The activation layer is applying the nonlinear RELU activation function to the image. The activation layer must directly follow the convolution layer. While activation layers are technically not layers because there are no parameters or weights learned here but they are included into the architecture because they are important parts of the network architecture. The RELU (Rectified Linear Unit) is used because it is the most common activation function and from the experimentation it seemed to do its job perfectly. This function takes an input and returns the maximum between the input and zero. This activation function is used to check if the "neuron fires" for a positive reaction.

The pooling layer is typically used to reduce the spatial size of the input volume. It is known to help reduce computation in the network and reduce overfitting. There are two kinds of pooling max

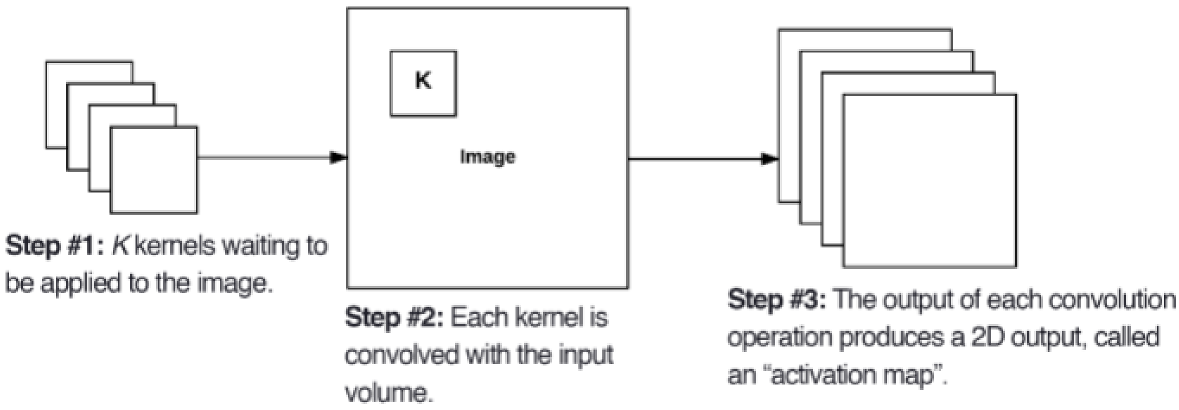



Figure 9.1: Example of Convolutional layer

and average pooling. These benefits are achieved similar to convolutions. A small matrix is moved across the image matrix and in the case of max pooling the largest number of in the pooling matrix is kept and the others are dropped, in the case of average pooling the average of the numbers in the matrix are taken and the rest are dropped. The size of the pooling matrix is 2x2 in these experiments and the stride is set 2,2. Regularizers are a method of reducing overfitting in deep

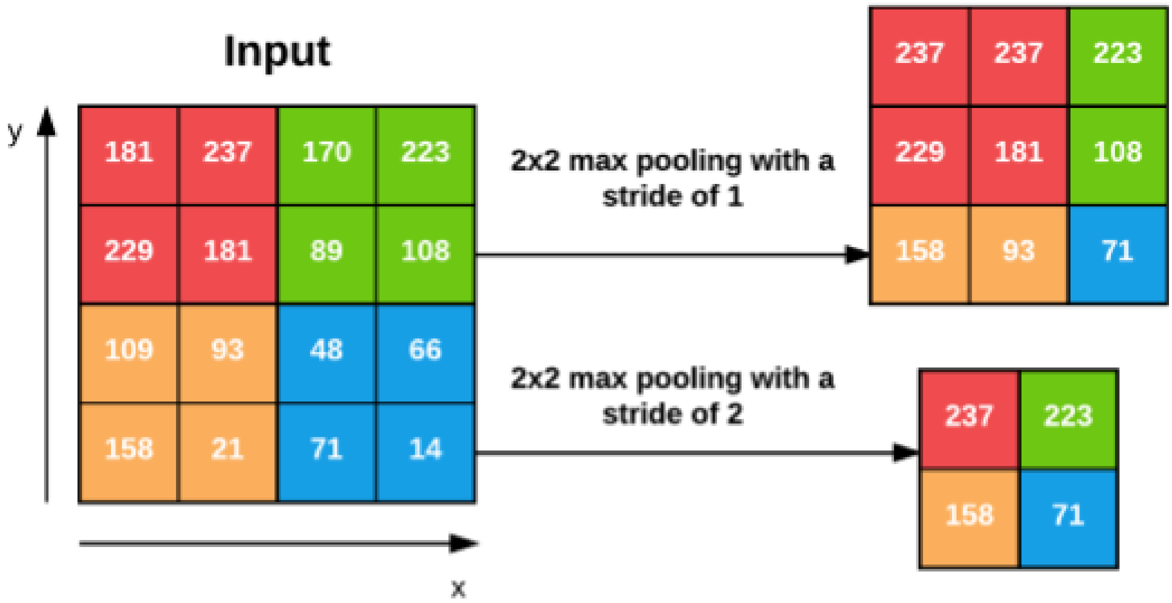



Figure 9.2: Example of Pooling layer

learning projects. This promotes generalization and makes the classifier more accurate for a larger amount of data. This is done by updating our loss and weight values based off of a regularization algorithm. In this project when regularization is used we use both L1 and L2 regularization because in some renditions of the project overfitting was a small issue.

The dropout layer randomly removes inputs from your training set based on a probability inputted into the dropout function. The two different dropout values used in this experiment were 0.1 and 0.2. Dropout is form of regularisation that aids with over-fitting. It typically leads to an increase in testing accuracy but a decrease in training accuracy. It aids with over-fitting because we are explicitly altering the network architecture at run time. By randomly dropping connections in the network this ensures that no single node is responsible for "activating" when presented with a specific pattern. This ensures that there are multiple redundant nodes that activate on similar inputs which helps our network generalise.

The fully connected layer is simply the implementation of a simple back propagation neural network structure as explained earlier in this report. This layer is the one that makes predictions based off of the model and it uses categorical cross entropy as it's loss function. The batch size is set to fifty and the amount of epochs is set to one hundred.

### 9.2.3 Tested networks

**Network 1:** INPUT => CONV => RELU => FC

**Network 2:** INPUT => CONV => RELU => FC including L1 and L2 regularization

**Network 3:** INPUT => [CONV => RELU]*2 => DO => FC including L1 and L2 regularization

**Network 4:** INPUT => [CONV => RELU => MAX PO]*4 => DO => FC including L1 and L2 regularization

**Network 5:** INPUT => [CONV => RELU]*4 => AVG PO => FC including L1 and L2 regularization

# Chapter 10: Evaluation

## 10.1 Results and Conclusions

Each one of the networks is tested ten times to get the average accuracy. The recorded values for the experiments are accuracy and F1-Measure. Accuracy is an essential statistic to track and F1 Measure is good to track in order to ensure the classifier is working properly with true positives, true negatives, false positives and false negatives. These results were evaluated on the premise that the training and validation curves for loss and accuracy determine whether the classifier is overfitting. Specifically the amount the curves match each other implies the lack of overfitting. E.g if the curves match eachother perfectly then there is no overfitting present.[14]

### 10.1.1 Network 1

This network has an input layer, one convolutional layer and a fully connected layer.

| Standard Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 70.3% | 75% |
| Salt | 45.6% | 45.9% |
| Soil | 51.1% | 53.3% |
| Sand | 52.4% | 52.9% |
| Algae | 48.5% | 48.1% |
| Plastic | 55.6% | 60.3% |
| Weighted Average | 54% | 55.8% |

The standard images produced a poor accuracy and the believed reasons why has been explained previously in the report. The F1-measure is slightly ahead of the standard accuracy which shows that the accuracy is accurate to the actual performance. But as we can see from the graph the differences in the curves indicates there is an issue with the classifier. Since the results are poor this more than likely indicates underfitting with this classifier. The direction of the curves indicates that the classifier is learning over epochs.

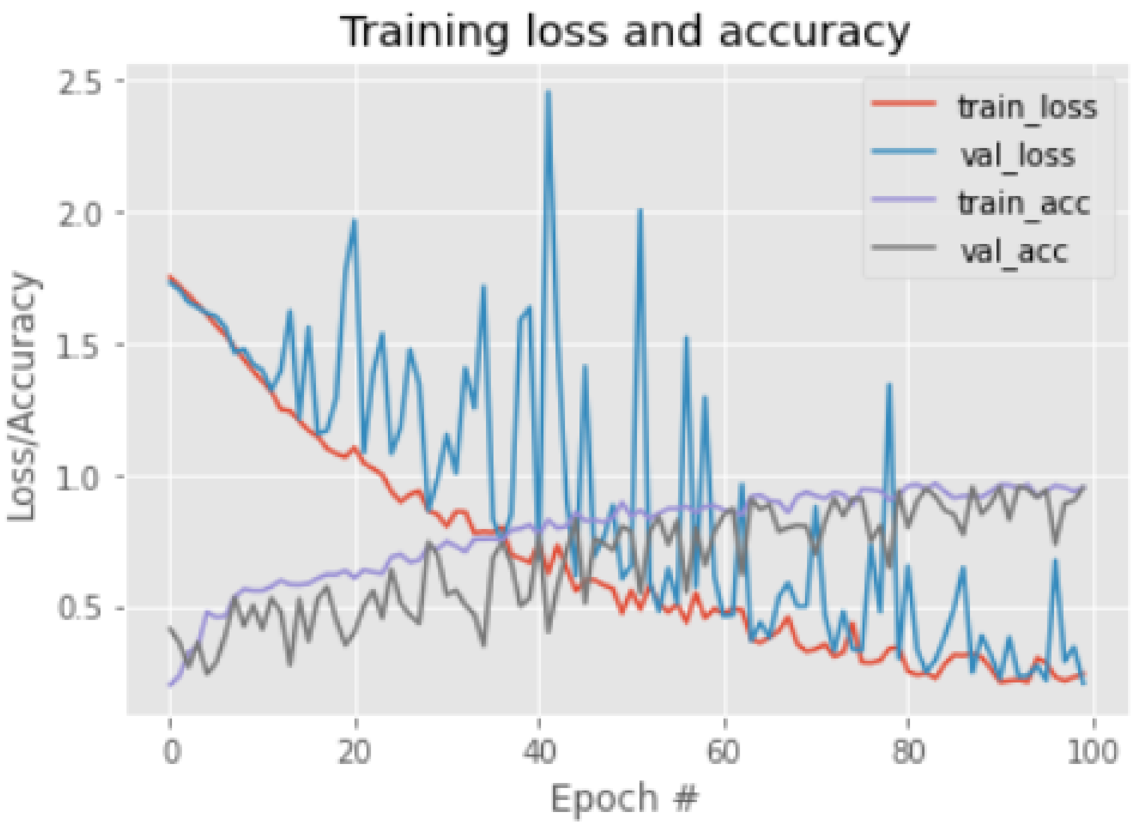



Figure 10.1: Network 1 Standard training and validation graph

| Near Infrared Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 99.3% | 98.6% |
| Salt | 89.3% | 94% |
| Soil | 93.2% | 87.7% |
| Sand | 92.8% | 86.4% |
| Algae | 95.6% | 94.3% |
| Plastic | 93.5% | 95.7% |
| Weighted Average | 94% | 92.4% |

Near wave infrared images produced the best accuracy out of the single image type classifiers. Similar to the other classifiers salt was its worst performing contaminant in terms of accuracy. The F1 measure is slightly lower than the standard accuracy but its still incredibly high so it's showing that the classifier is performing well with all versions of true and false positives. There is some variation in the curves but that becomes a bit less prevalent as the epochs progress. This does show that there is some overfitting with the classifier but not to any impactful degree. The direction of the curves indicates that the classifier is learning over epochs.

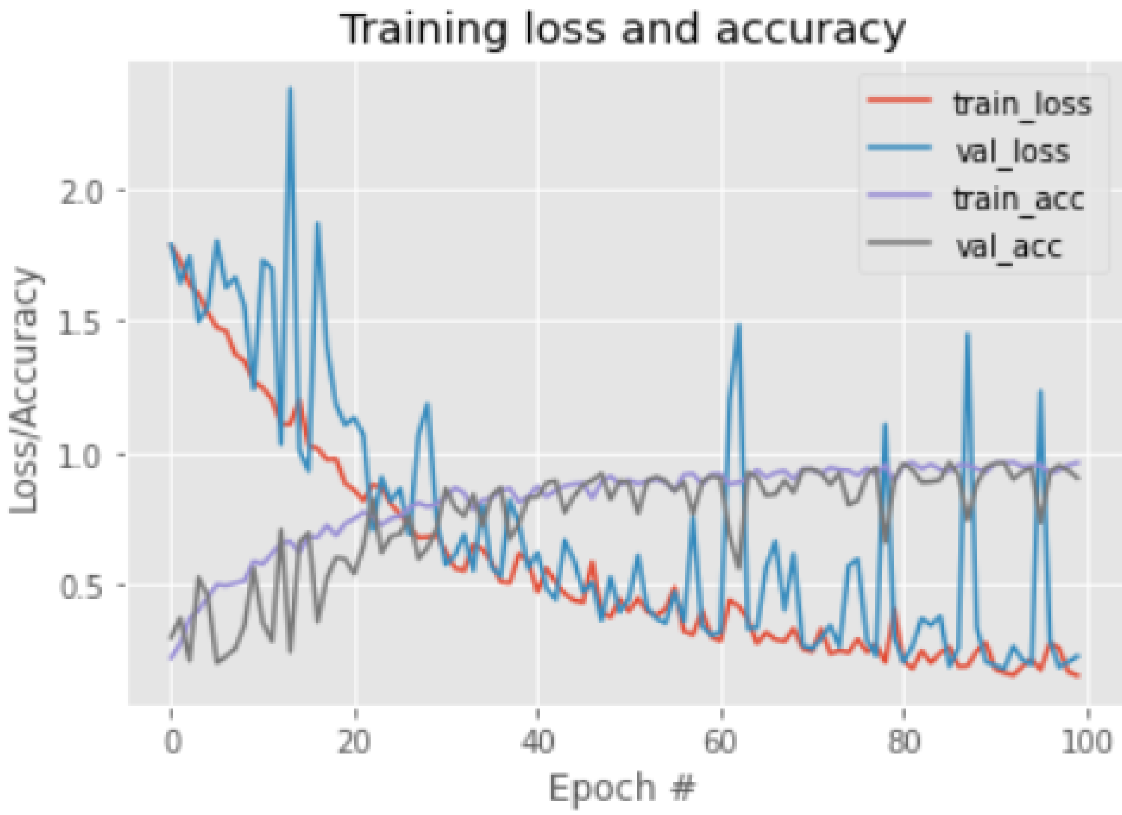



Figure 10.2: Network 1 Near Infrared training and validation graph

| Microscopic Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 62.8% | 55.2% |
| Salt | 12% | 24.2% |
| Soil | 48.9% | 41.5% |
| Sand | 59.8% | 58.1% |
| Algae | 65.3% | 44.9% |
| Plastic | 40.3% | 29.1% |
| Weighted Average | 51.7% | 42.3% |

Microscopic images were the worst performing of the individual classifiers in terms on pure accuracy. A breakdown of the individual contaminants can be found in the data and context section. As is clear from the graph the curves are not behaving as they did in the other graphs, this is because the loss rate isn't decreasing as quickly as the other classifiers and the accuracy isn't increasing as quickly. This shows clearly the classifier isn't performing well in any aspect. Also the differences in the curves shows there is more than likely some underfitting going on with the classifier

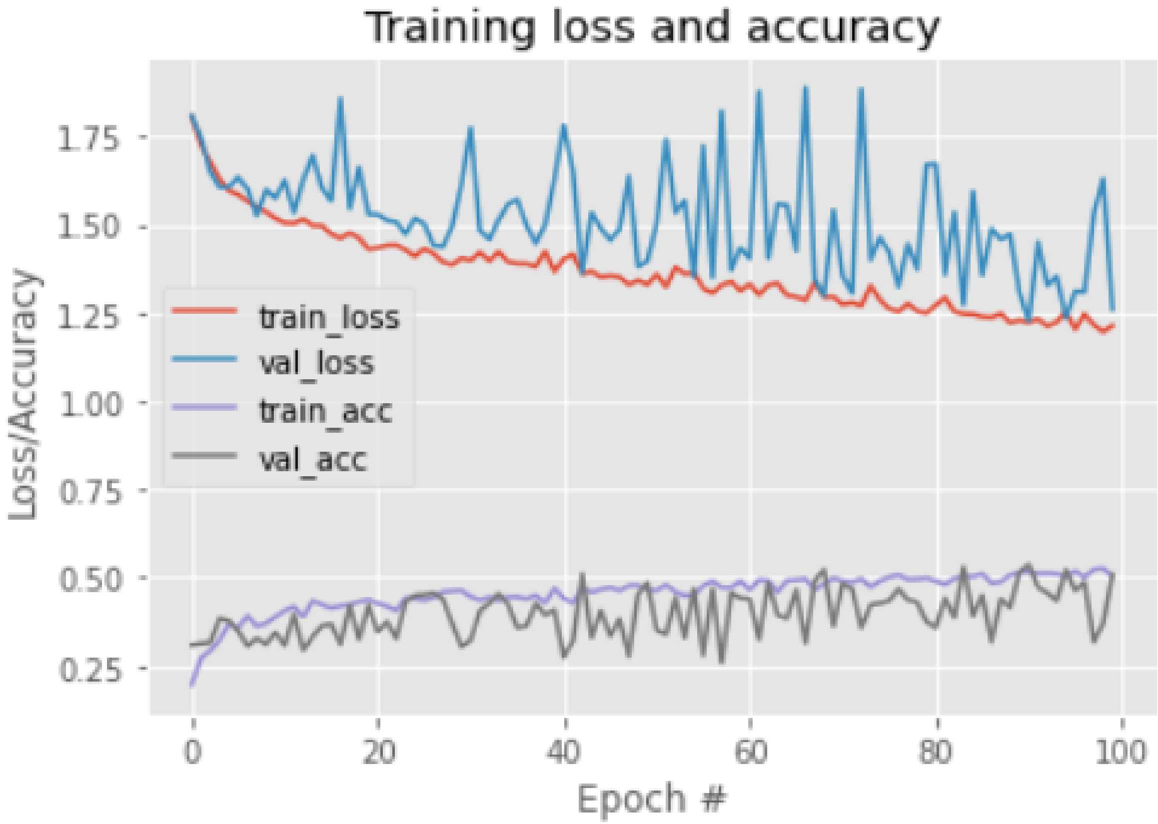



Figure 10.3: Network 1 Microscopic training and validation graph

| Thermal Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 70.5% | 30.5% |
| Salt | 80.7% | 39.2% |
| Soil | 62.8% | 39.9% |
| Sand | 59.2% | 70.1% |
| Algae | 34.9% | 37.1% |
| Plastic | 33% | 26.2% |
| Weighted Average | 53.5% | 41.3% |

Thermal images were the worst performing in terms of the F1-Measure quality which in some peoples opinions is more important that accuracy. So it's difficult to determine whether this or the microscopic image classifier is the worst of the two. Similarly to the microscopic graph this graph shows that the classifier is performing poorly based on the direction of the curves. There seems to be little evidence of over fitting or under fitting present in this classifier based on how closely the curves match eachother.


Figure 10.4: Network 1 Thermal training and validation graph

| Long Wave Infrared Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 94.6% | 96.5% |
| Salt | 95.9% | 80.2% |
| Soil | 90.1% | 88.8% |
| Sand | 85.2% | 86.8% |
| Algae | 86.1% | 82.8% |
| Plastic | 81.8% | 88.7% |
| Weighted Average | 89.8% | 87.5% |

Long wave infrared images are the second most accurate in terms of accuracy and F1-Measure. There is not much to discuss in regards to this classifier. The curves here are very similar indicating there is no overfitting or underfitting in this classifier. The direction of the curves indicates a well performing and learning classifier.

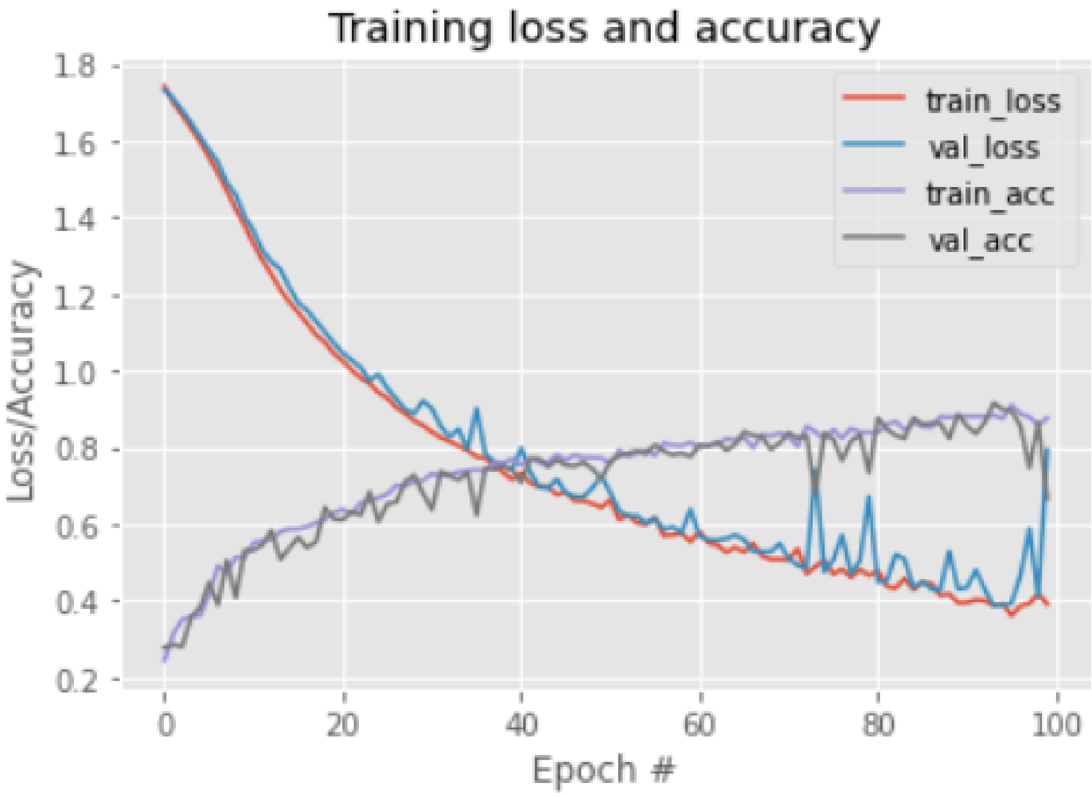



Figure 10.5: Network 1 Long wave infrared training and validation graph

| Three-Fold Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 98% | 98.4% |
| Salt | 85.5% | 84.8% |
| Soil | 90.6% | 91.1% |
| Sand | 94.1% | 87.1% |
| Algae | 95.6% | 94.2% |
| Plastic | 95% | 94.7% |
| Weighted Average | 93.5% | 92.2% |

Three fold refers to a combination of standard, near wave and long wave infrared images. It did not perform as well as solely the Near infrared image classification, this may have just been because of the limited number of experiments. This experiment was carried through to the additional networks because it was the most accurate of the classifiers that isn't based on a single image type. Since the goal of this experiment is a combined classifier it was the obvious decision to pursue this network throughout the rest of the experiments. The reason all of the image types weren't tested on each network was because of time restraints limiting the amount of experiments that could be run. The graph presented is very similar to the one associated with the near infrared images which shows the classifier is learning and only slightly over fitting.

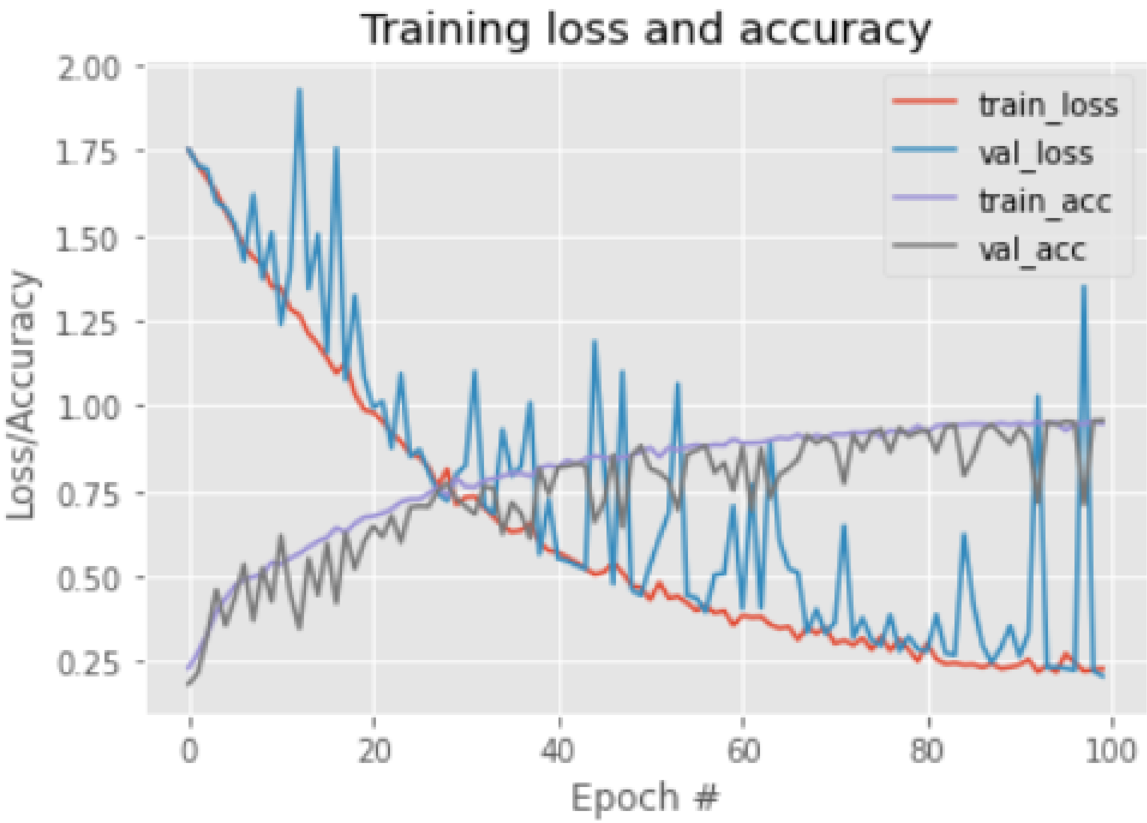



Figure 10.6: Network 1 Standard, Near wave infrared, Long wave infrared training and validation graph

| Complete Image Set | Accuracy | F1-Measure |
|---|---|---|
| Clean | 82.6% | 80% |
| Salt | 82.6% | 65.5% |
| Soil | 80.2% | 71.3% |
| Sand | 75.3% | 77.6% |
| Algae | 78.1% | 73.4% |
| Plastic | 72.2% | 65.6% |
| Weighted Average | 78.3% | 73.2% |

The combined classifier for all of the image types produced an adequate accuracy which was between the lower and higher performing classification accuracy's. Based off of this the microscopic and thermal images were eliminated in order to improve accuracy and F1-measure for future experiments. The graph does not present any signs of over or under fitting to any major degree however the curves show that they are learning slower than normal which is probably the reason for the mediocre accuracy.

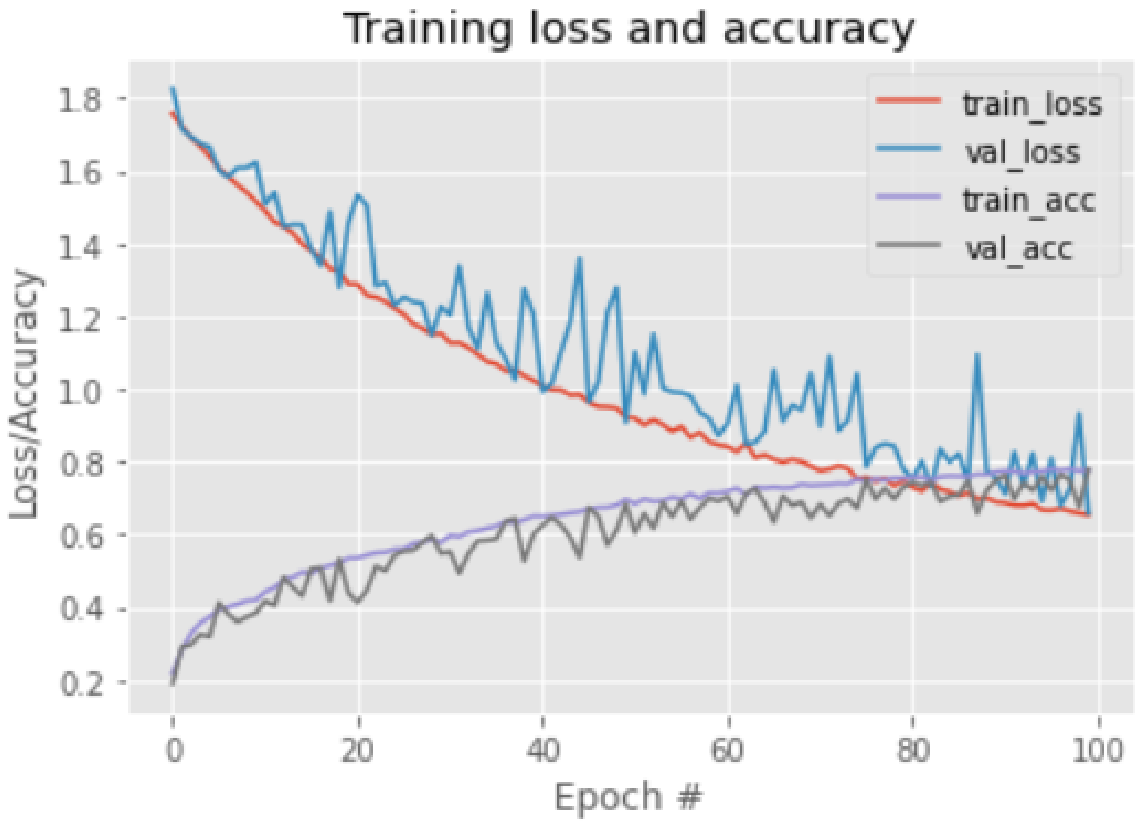



Figure 10.7: Network 1 All Image Type training and validation graph

### 10.1.2 Network 2

This network has a convolutional layer, RELU activation layer, fully connected layer and L1 and L2 regularizers.

| Three-Fold Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 96.1% | 97.1% |
| Salt | 90.1% | 81.7% |
| Soil | 85.7% | 86.4% |
| Sand | 91.9% | 85.1% |
| Algae | 91.5% | 89.1% |
| Plastic | 91.9% | 93.8% |
| Weighted Average | 91.1% | 89.3% |

The addition of regularisation techniques leads to a small decrease in accuracy. The difference between the accuracy's could just be based on the individual experiments and if a larger sample size was carried out then the results may have been more similar. The difference in accuracy's is not large enough over ten experiments to decide on which network performed better. The graph is very similar to the network 1 three image graph in that it shows mild overfitting but not as much as the network one graph and a good training and validation curve.


Figure 10.8: Network 2 training and validation graph

### 10.1.3 Network 3

This network has two convolutional layers, two RELU activation layers, a dropout layer, a fully connected layer and L1 and L2 regularizers.

| Three-Fold Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 99.8% | 99.5% |
| Salt | 98.9% | 97.9% |
| Soil | 98.4% | 98.5% |
| Sand | 99% | 97.9% |
| Algae | 96.2% | 96.5% |
| Plastic | 96.7% | 96.9% |
| Weighted Average | 98.4% | 98.2% |

The inclusion of additional convolutional and activation layers along with the dropout layers lead to a great improvement on the classifier accuracy and there is no distinct presence of over or underfitting. We can see this because the validation and training curves match which implies an appropriate classification process. This network produces the highest accuracy and with a lack of overfitting this network seems the best out of the ones tested. So if one network was chosen to be the final deployment it would be this one. There seems to be mild overfitting which is evident from the spikes in the validation loss but this seems to stop by the end of the experiment.

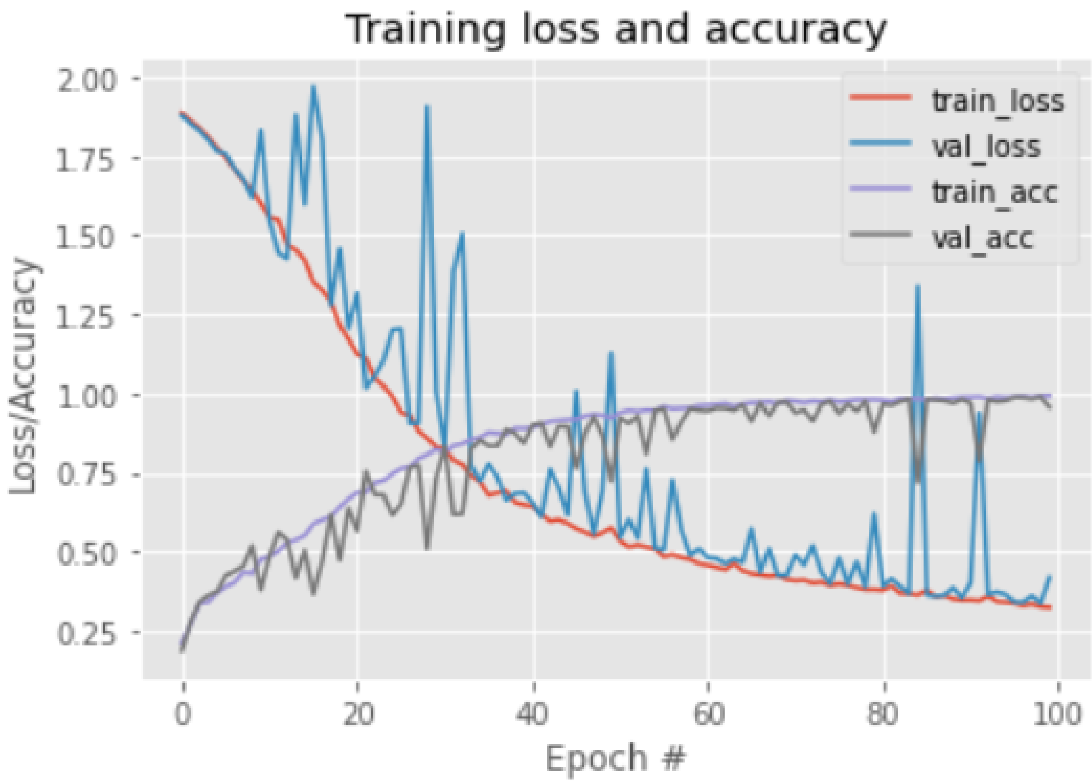



Figure 10.9: Network 3 training and validation graph

### 10.1.4 Network 4

This network has four convolutional layers, four RELU activation layers, four max pooling layers, a dropout layer, a fully connected layer, L1 and L2 regularizers.

| Three-Fold Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 77.2% | % 82.2 |
| Salt | 76.3% | 33.6% |
| Soil | 74.7% | 63.6% |
| Sand | 61.6% | 62.9% |
| Algae | 85.3% | 81.3% |
| Plastic | 65.9% | 57.8% |
| Weighted Average | 73.9% | 66.5% |

The addition of pooling layers causes a significant drop in accuracy. This seems to be because the classifier isn't learning as quickly as the others. From how similar the curves are in the graph there doesn't appear to be any over or under fitting going on. Based on the trajectory the curves are progressing they seem to begin learning at a much faster rate near the end of the experiment. It could be that with a larger number of epochs the classifier would have learned sufficiently to produce better accuracy. If this experiment was carried out again this would be a key aspect to test.

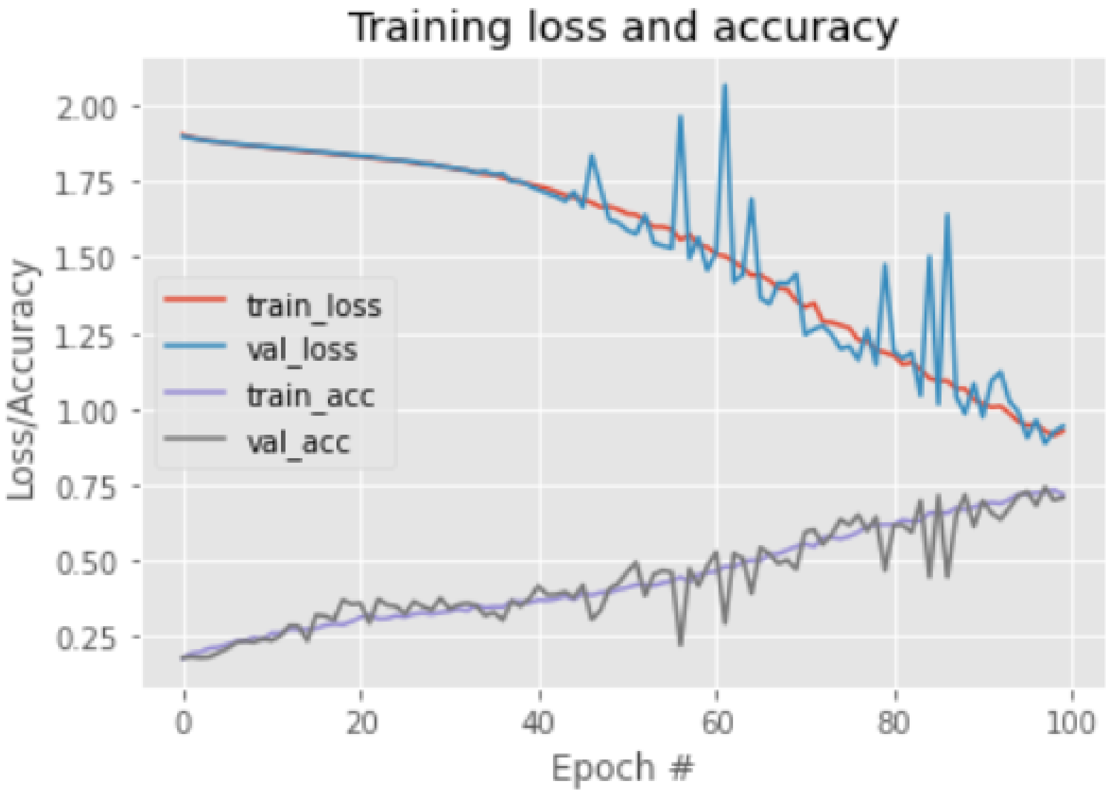



Figure 10.10: Network 4 training and validation graph

### 10.1.5 Network 5

| Three-Fold Images | Accuracy | F1-Measure |
|---|---|---|
| Clean | 99.2% | 98.8% |
| Salt | 93.4% | 93% |
| Soil | 97.2% | 96.8% |
| Sand | 95.7% | 96.3% |
| Algae | 98.2% | 97% |
| Plastic | 96% | 97.3% |
| Weighted Average | 97.1% | 96.9% |

This network has an exceptionally high accuracy at ninety seven point one percent which is the second highest accuracy of all of the multi image classifiers. While the accuracy of classifier is less than network three it is evident from the graph that there is less overfitting occurring in the classifier. This can be seen because there are less drastic spikes in the validation loss than when compared to the graph from network three.


Figure 10.11: Network 5 training and validation graph

# Chapter 11: Final Conclusions

## 11.1 Interim Report Conclusion

In summary this project is quite fleshed out and will soon be ready to start. For the remainder of the time most of it will be spent simply refining the classifier in order to improve accuracy. Based off of the AquaSight project this project should be able to achieve similar results so hopefully in around the 90 percentile.
Based off of the current trajectory the basic classification of the project should be completed within a reasonable amount of time so sometime may be spent on the contaminant identifier.
As further research has been performed the uniqueness of this project has diminished since projects such as AquaSight and X.Zhangs turbidity classifier exist. Thankfully the implementation of infra red sets this project apart from similar ones in the field and hopefully puts this project ahead of the competitors.

## 11.2 Societal Contribution

This project produced an excellent image classifier for water contaminants. While it may not be a general classifier due to the specific environment that the images were taken in, it stills proves the concepts that a multi image type water contaminant classifier is a reasonable goal for a project with more resources.

It is quite possible that a new project could be constructed that achieves more generality in this field. All that would be necessary is for their to a larger and more general dataset, more extensive experimentation into how changing the networks in more minute ways would effect the overall performance and more time and a great classifier could be devised.

Such a classifier would be a great contribution to society as previously discussed in this report. It would make checking and maintaining the quality of drinking water much simpler. So this project should be considered a societal success if it can aid in the development of such a classifier.

## 11.3 Final Thoughts

| Networks | Accuracy | F1-Measure |
|---|---|---|
| 1 | 93.5% | 92.2% |
| 2 | 91.1% | 89.3% |
| 3 | 98.4% | 98.2% |
| 4 | 73.9% | 66.5% |
| 5 | 97.1% | 96.9% |
| Weighted Average | 90.8% | 88.62 % |

As you can see the two best performing classifier were network three and five. Network three had better accuracy but as was previously discussed overfitting was less present in network five.

Overall network five was chosen to be the superior network because while network three did provide better accuracy it more than likely will not perform as well on newly inputted data due to the mild overfitting present. Since network five has less overfitting present it will more than likely perform better on new images tested in the network. Since identifying newly entered data is the key objective of an image classifier this was the key goal of the project. Therefore network five is the final chosen network but all five will be uploaded to the github repository.

Overall this project has been considered a massive success. The initial goal of achieving an above ninety percent accuracy was achieved with some classifiers coming close to one hundred percent in their accuracy. The final product is different to the original design in that it focuses on specific contaminants rather than if the water is clean or not. This was a good direction to take the project because it made the project much more informative. Overall more experiments could have been performed and the data set could have been larger but the project was an overwhelming success.

# Acknowledgements

First off I would like to confirm that all work done in this report is my own. Secondly I would like to thank my supervisor Eleni Mangina for continually aiding me in this project and my parents for being a constant guiding presence and extra set of hand for this whole project.

# Bibliography


1. *Water and Sanitation - United Nations Sustainable Development* https://www.un.org/sustainabledevelopment/water-and-sanitation/.
2. Gupta, A. & Ruebush, E. AquaSight: Automatic Water Impurity Detection Utilizing Convolutional Neural Networks. *CoRR* **abs/1907.07573.** arXiv: 1907.07573. http://arxiv.org/abs/1907.07573 (2019).
3. Ayaz, H., Shewokis, P. A., Bunce, S. & Onaral, B. *An optical brain computer interface for environmental control* in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2011), 6327–6330.
4. American Public Health Association American Water Works Association, W. E. F. *APHA Method 4110: Standard Methods for the Examination of Water and Wastewater* (American Public Health Association, 1917 - 1992).
5. Horak, K., Klečka, J. & Richter, M. *Water Quality Assessment by Image Processing* in (July 2015).
6. Zhang, X. *Water quality turbidity detection based on image recognition system design and implementation* in *First International Conference on Real Time Intelligent Systems* (2016), 63–70.
7. Ltd, A. L. T. *Schedule of Accreditation* (Advanced Laboratory Testing Ltd., Sept. 2013).
8. Richardson, S. D. & Ternes, T. A. Water Analysis: Emerging Contaminants and Current Issues. *Analytical Chemistry* **90.** PMID: 29112806, 398–428. eprint: https://doi.org/10.1021/acs.analchem.7b04577. https://doi.org/10.1021/acs.analchem.7b04577 (2018).
9. Chapman, D. V. *Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring - Second Edition* (EFN Spon, 1996).
10. Toivanen, T., Koponen, S., Kotovirta, V., Molinier, M. & Chengyuan, P. Water quality analysis using an inexpensive device and a mobile phone. *Environmental Systems Research* **2,** 9. ISSN: 2193-2697. https://doi.org/10.1186/2193-2697-2-9 (Oct. 2013).
11. Shanono, I. H. *et al.* Image Processing Techniques Applicable in Wastewater Quality Detection: Towards a Hygienic Environment. *American Journal of Remote Sensing* **6,** 29–42 (2018).
12. Horak, K., Klečka, J. & Richter, M. *Water Quality Assessment by Image Processing* in (July 2015).
13. Fernandez-Canque, H. *et al.* Micro-organisms detection in drinking water using image processing. *cell* **15,** 4–2 (2006).
14. Rosebrock, D. A. *Deep Learning For Computer Vision With Python* (Nov. 2018).
15. Union, E. EUROPEAN UNION (DRINKING WATER) REGULATIONS (2014).
16. Aryal, J., Gautam, B. & Sapkota, N. Drinking water quality assessment. *Journal of Nepal Health Research Council* **10,** 192–6 (Sept. 2012).
17. Organisation, T. W. H. Parameters of water quality, interpretation and standards (2001).
18. New York State, D. o. H. Salt and Drinking Water (February 2020).