Muhammad Awais

Digitalization of Sustainable Energy Production Processes

Master's thesis in Electronic Systems Design
Supervisor: Dag Roar Hjelme
Co-supervisor: Jacob Joseph Lamb

June 2021

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Muhammad Awais

Digitalization of Sustainable Energy Production Processes

Master's thesis in Electronic Systems Design
Supervisor: Dag Roar Hjelme
Co-supervisor: Jacob Joseph Lamb
June 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Preface

In the master's program of Electronic Systems at the Norwegian University of Technology and Science, all students must complete the subjects Electronic System Design Thesis, giving 30ECTS. This project is lab work done for the master thesis, which ends in June 2021. The master thesis is supervised by Jacob J. Lamb, a researcher in the field of bioprocess sensor technology at IES, NTNU and Dag Roar Hjelme, a professor in the field of photonics at IES, NTNU.

Table of Contents

List of Figures

List of Tables

Abbreviations

MTOE	Millions of tonnes of equivalent oil
EU	European Union
UN	United Nations
AD	Anaerobic Digestion
VFAs	Volatile Fatty Acids
RGB	Red Green Blue
MDA	Multivariate Data Analysis
PCA	Principal Component Analysis
HCA	Hierarchical Cluster Analysis
LDA	Linear Discriminant Analysis
LCD	Liquid Crystal Display
ISO	International Standard Organization
GC	Gas Chromatography
IR	Infra-red
NIR	Near Infra-red
DIP	Digital Image Processing
mM	Milli Molar
mL	Milli Litres

Abstract

Biogas production is growing as a renewable energy source. In order to maximize energy production, it is essential to ensure that the production process is economical and efficient. This report discusses the anaerobic digestion (AD) process, which breaks down organic material with the help of microorganisms in the absence of oxygen, and the parameters involved in sensing of VFAs in the digester. It also covers the sensor characteristics required for the sensor to function in the biogas production process and how it works and is suitable for monitoring VFAs in the AD process. Volatile Fatty Acids (VFAs) are an intermediate compound made during the AD process, and their concentration influences the production process. In addition, their concentration can indicate the state of the biological process, with high VFA concentrations signifying biological deterioration within the process in terms of biological stability and feedstock overloading. Several traditional methods can measure and classify VFAs during the AD process to maintain stability and high efficiency. Despite this, these traditional methods are not fast, and usually, a trade-off has to be made between robustness and selectivity of the VFA detection. Here, an emerging technology for chemical sensing is suggested based on an optical sensor array to overcome these shortfalls. This sensor technology mimics the human olfactory system by having dyes instead of receptors in a nose to detect any smell in the environment and be sensitive to dilute concentrations.

By using dyes and image processing, it may be possible to detect and classify VFAs. This is achieved by assessing the interaction between the analyte (VFAs) and the dyes in terms of the observable color changes. The 21-dye colorimetric sensor array was evaluated experimentally to exhibit color change after interacting with VFAs. The color change of before and after image results in data in the form of a vector. However, this vector is high-dimensional data and is not readable. Therefore, data analysis and modeling tools must be used to interpret the optical sensor array to extract the information regarding the VFA presence. Several modeling methods can be used and are discussed with their pros and cons within this report. The main objective of this project is the development and testing of a sensor system that senses and classifies VFAs in AD to indicate process deterioration in a robust, low-cost, and automatic manner. The experiment shows that the sensor developed in this project provides the desired sensitivity and selectivity required to detect VFAs in the AD process. Sensitivity tests and mixed VFAs in different concentrations were used to examine the sensor system. The sensor was able to detect all types of mixed samples of all concentration levels low & high.

Chapter 1

1. Introduction

Energy demand is increasing at a rapid pace across the world. According to the latest figures from the U.S. Energy Information Administration, global energy consumption will grow by 28% between 2015 and 2040 [1]. The global energy supply is heavily reliant on non-renewable fossil fuels such as oil, lignite, coal, and natural gas, all of which contribute significantly to climate change [2]. Other significant contributors to greenhouse gas emissions are livestock, animal husbandry, wood combustion, and solid waste [3]. It's become apparent that a grave danger to the world and our way of life is looming on the horizon, with climate change escalating. Climate change is a serious threat to the world, and it is quickly becoming a reality. To avoid a climate disaster, an increased share of renewable resources is needed. Renewable and sustainable resources would have a crucial role in our society's energy future. By 2030, the Sustainable Development Targets of the United Nations seek to double the rate of energy efficiency improvements and boost the contribution of renewable energy sources [3]. Furthermore, European Union policies are striving to supply 32% of total energy from renewable energy by the year 2030. As part of an EU initiative, Norway has been a major contributor to reducing global carbon emissions [4]. Detailed theory about biogas and its parameters are discussed in the following sections, Table 1 includes names of some important journal papers, book chapters related to this theory, and discussion.

Table 1 - Overview of some important Papers and reason for their selection

Paper Title	Authors	Year Published	Criteria	Reference Number
"Perspectives of optical colorimetric sensors for anaerobic digestion"	Lamb et al.	2019	This article discusses the theoretical background, methods about colorimetric sensor in AD process and different measurement methods	18
"Volatile fatty acids as indicators of process imbalance in anaerobic digestors"	Ahring et al.	1995	Role of VFAs in AD process as indicator in investigated and relative changes caused by loading, temperature etc. which lays groundwork for this project	24
"Microbiology of AD"	J. J. Lamb	2020	Since the AD process is fundamental in biogas production these articles provide detailed explanation about parameters and steps occurring during AD process needed to do this project.	17
"Process Parameters Affecting Anaerobic Digestion"	P. Bajpai	2017		22

"Sensor systems for bioprocess monitoring"	Biechele et al.	2015	Discusses sensor features, problems in monitoring processes like AD and information about multiple sensors so, this sensor can be more efficient	30
"The optoelectronic nose: colorimetric and fluorometric sensor arrays",	Li et al.	2018	Detailed review of sensor arrays and dyes that are used in colorimetric sensors and provides the literature on which this sensor system is based upon "optoelectronic nose" concept.	33
"Optical sensor arrays for chemical sensing: the optoelectronic nose"	Askim et al.	2013		35

1.1. Biogas

There is a range of alternative routes to producing sustainable, clean energy. The one popular alternative route to renewable energy is biogas production: biogas is produced by breaking down organic matter into gases that result from organic waste [5]. It is expected that biomass-based renewable energy will meet the demand that continues to grow over the long term, and reduction in global carbon footprint as a result of the implementation of biomass-driven energy is very promising [6]. Developing the biogas production process further could significantly improve its potential and application as a clean energy source. Currently, only 3% of the electrical energy load is produced through the treatment of wastewater in developed countries. At the same time, it could also be a source of nitrogen and phosphorus for agriculture and energy if it is developed further [1]. Since biomass can be converted into different forms, such as solids, liquids, and gases, a wide-ranging set of energy applications, including heat, electricity [7], liquids, and as a transport fuel for vehicles, are well suited for this source of energy [8]. According to a report from the International Energy Agency, biogas' current production is 35 Mtoe (Million tonnes of oil equivalent), while the potential for biogas is around 570 Mtoe [9]. This shows that there is an extensive amount of "untapped" energy available through AD. Approximately 66% of biogas produced is used to displace fertilizers in agricultural processes. Although the combustion of biogas produces in 2018 was used for electricity and heating purposes. Until 2018 about 18 GW of power generation capacity based on biogas has been installed with a 4% increase every year between 2010 to 2018.

Even though, like fossil fuels, biogas releases carbon dioxide, and leaks methane, it is still considered a renewable energy source. It is because the production and use cycle is continuous, generating zero net carbon emissions [10]. Carbon absorbed in the growth of biomass source equals the released carbon when utilized as a source of energy hence zero carbon emissions [2]. This has resulted in a surge in biogas production in Norway and Europe in the last decade [11]. In

an anaerobic digester, the organic material is decomposed using the Anaerobic Digestion process, which produces biogas. The produced biogas contains 50-70% of CH_4, 30-50% of CO_2, and small amounts of H_2S, N_2, and H_2O [12]. Methane can then be used as fuel for generating electricity and other applications.

1.2. Anaerobic Digestion

AD is a natural process achieved through the digestion of biomass by microorganisms in the absence of oxygen. These microorganisms are used to break down the organic material resulting in biogas and digestate (a liquid-like wet mixture and can be used as fertilizer for crops) [13].
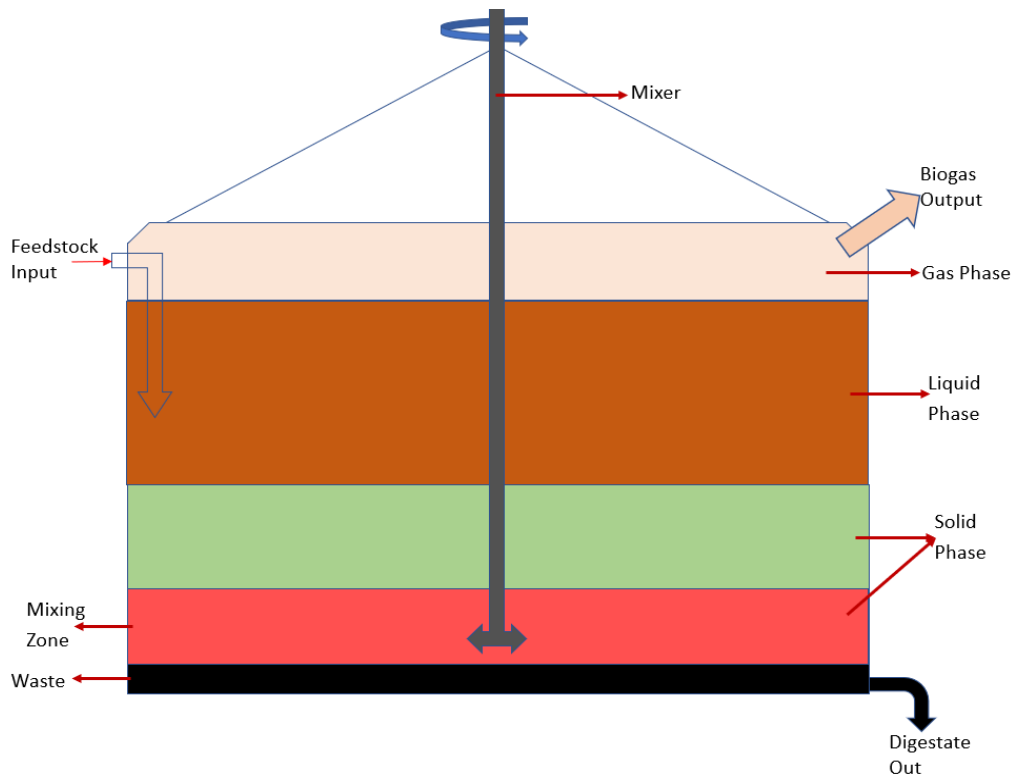
Figure 1.1 - Anaerobic digester design and feedstock phases.

The AD process is series of biochemical reactions to break down biodegradable organics, which is described in four main steps; hydrolysis, acidogenesis or fermentation, acetogenesis, and methanogenesis [14], [12], [15] (Figure 1.2). During hydrolysis, macromolecules (e.g., carbohydrates, proteins, and lipids) are broken down into monomers (e.g., monosaccharides, amino acids, and long-chain fatty acids; LCFAs) [16]. Hydrolysis is achieved with extracellular enzymes created by specific types of facultative or obligatory anaerobes [17]. In acidogenesis, almost 70% of substances produced in hydrolysis are converted to H_2, CO_2, and acetate. During methanogenesis, these substances are directly used by methanogenic bacteria, but short-chain fatty acids still require further degradation [18], [19]. High partial H_2 hinders the acetogenic activity,

and the conversion of butyrate and propionate into hydrogen and acetate is possible at low hydrogen concentrations. By balancing H_2 through production from acetogens and consumption by methanogens, low concentrations of H2 can be achieved [20]. Methanogenesis is solely achieved with the help of methanogens such as acetotrophic methanogens and hydrogenotrophic methanogens producing CH_4 and CO_2 [17], [21]. Acetotrophic methanogens yield around 70% of the final methane, while 30% is produced by hydrogenotrophic methanogens through the conversion of H_2 to CH_4. Therefore, methanogenic organisms are a rate-limiting factor in the AD process because of their slow growth. Given below are the equations of chemical reactions that take place to produce methane.

$$CH_3COOH \longrightarrow CH_4 + CO_2 \qquad \text{(Acetotrophic Reaction)}$$

$$CO_2 + 4H_2 \longrightarrow CH_4 + 2H_2O \qquad \text{(Hydrogenotrophic Reaction)}$$

$$4CO + 2H_2O \longrightarrow CH_4 + 3CO_2 \qquad \text{(Hydrogenotrophic Reaction)}$$
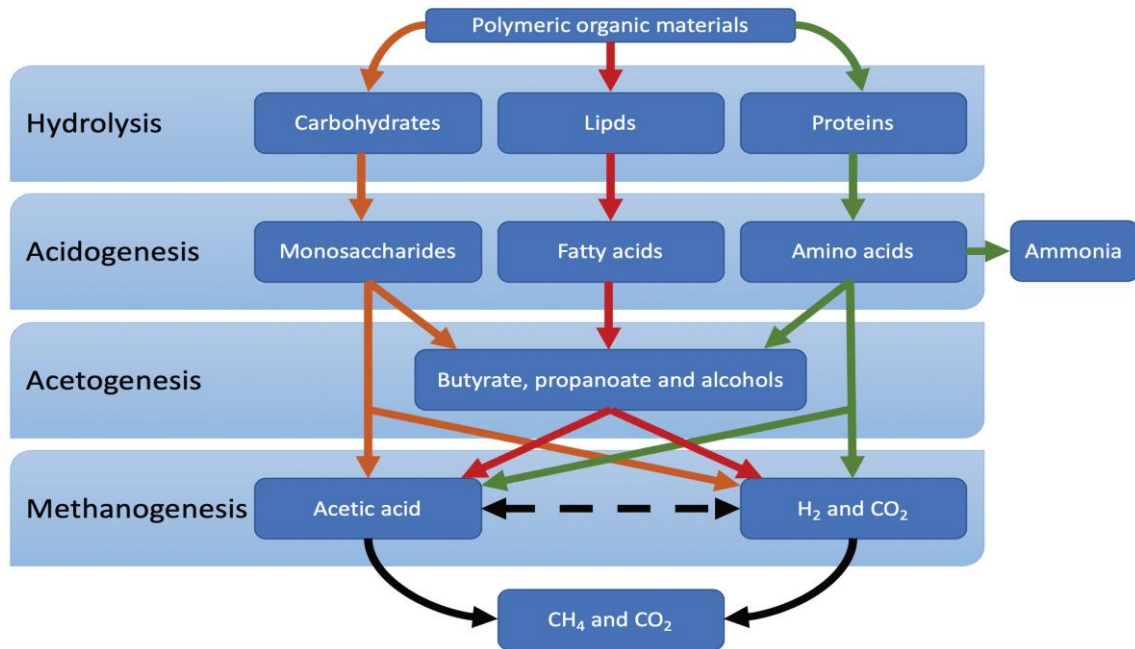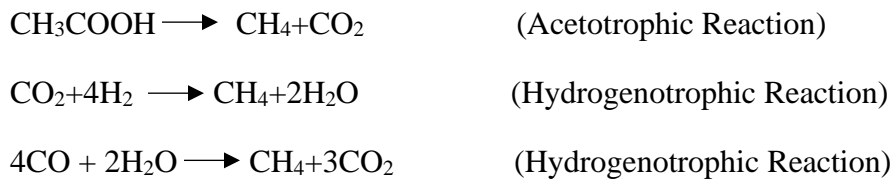
Figure 1.2 - Anaerobic digestion process stages. [14]

As described above, it is vital to ensure that methanogenic consortia have a healthy environment for optimal production. Critical parameters that affect the production are anaerobic conditions, temperature, system pH, volatile fatty acid mixing, toxicity, solid retention time, volatile solids loading rate, and hydraulic retention time [22].

Anaerobic conditions

The majority of the bacteria that are essential in the anaerobic system are obligate anaerobes (i.e., an organism that lives and grows in the absence of molecular oxygen). As a result, for optimal conditions, there must be no oxygen present.

Temperature

One of the most crucial factors in deciding the rate of reaction is the ambient temperature. The temperature has a significant effect on anaerobic processes, as it does on other biological processes. The optimal temperature ranges for bacteria are Psychrophilic (5–15 °C), Mesophilic (35–40 °C) Thermophilic (50–55 °C) [22].

System pH

There are two main classes of microorganisms, acidogens, and methanogens whose pH is significant for the AD process. For acidogens, a pH range between 5.5-6.5 is optimal, while for methanogens, 7.8-8.2. The activity of methanogens is inhibited by low pH, resulting in the accumulation of volatile fatty acids and hydrogen [22]. For both populations of microorganisms, the ideal operating pH range is 6.8 to 7.4. It should be noted that neutral pH is more desirable.

Mixing

Mixing is also a vital characteristic of the AD process. If not done correctly, it can lead to methane reduction, insufficient water stabilization, and non-uniform distributions of substrates and other microorganisms.

Toxicity

High levels of mineral ions can cause high toxicity, which results in the inhibition of microorganism growth. It should be noted that small amounts of mineral ions are essential for microorganisms (trace minerals), whereas higher levels inhibit growth.

Solid Retention Time

To allow an anaerobic process, it is essential that the microorganisms should not be washed out faster than their reproduction rate. The microorganisms are throughout the reactor but can become concentrated in the solid component. Therefore, it is important to have a long solid retention time (SRT) to maintain a high level of microorganisms. The average length of time the solids remain within the treatment system is termed the mean SRT [22]. Its formula is:

$$\text{SRT} = \frac{Mass\ of\ solids\ in\ digester\ (kg)}{Rate\ of\ removal\ of\ solids\ from\ digester(kg/day)}$$

Loading Rate

Feeding the digester with an appropriate quantity of the substances is essential as improper mixing resulting in pH drop and high levels of VFAs can cause process failure. Essentially it is to make sure that the microorganisms have the right amount of feed. Not so low that they become inefficient and slow, and not so high that the organisms are overloaded.

Hydraulic Retention Time

The average time wastewater is retained in a digester determining interaction time between microorganisms and pollutant is called Hydraulic Retention Time **(HRT)**.

$$\text{HRT} = \frac{Volume\ of\ digester}{Flow\ rate\ of\ wasterwater\ through\ digester}$$

The least value of HRT is desirable, but it should be sufficient enough for digestion of the substrate. It reduces reactor volume hence capital cost. HRT depends on the size and feedstock substance used in the AD process, but in a typical Continuously stirred tank reactor, the HRT for organic substance is around 2-5 days [23].
All the parameters described above are crucial for determining the state of the AD process, but they can change gradually. Hence, to detect a potential process failure at an earlier stage, VFAs are used as early indicators to determine the state of the process. Should there be an imbalance in the process, the accumulation of VFAs and alcohols can cause a pH drop.

1.3. Volatile Fatty Acids

Volatile Fatty Acids (VFAs) are essential elements in the AD process. They are a part of the AD process and are always present during the process, but their accumulation can cause deterioration. If they accumulate in the digester during the AD process, they can be indicators for process imbalances that lead to process failure. Acetic acid, butyric acid, and propionic acid are the focal VFAs of interest [24]. Even though these VFAs play an essential role in the production of methane for many organisms involved in the AD process, VFAs play an inhibitory role at high concentrations and can slow down or inhibit the AD process. VFAs are one of the first signs of degradation in the AD process, and therefore provide valuable data on the impending failures of the process [25]. It is vital to monitor and control the levels of VFAs produced in an anaerobic digester to avoid the collapse of the biological process. They are formed during the hydrolysis and acidogenesis stages of the process. Basically, water sludge is complex biopolymer (a substance having a large molecule with a repeating chain of molecules) such as protein and carbohydrates are broken down into monomers like fatty acids, sugars, amino acids by fermentative bacteria.

Monomers are then converted into short-chain fatty acids by further fermentation [26]. Acetic acid molecular formula $C_2H_4O_2$, butyric acid $C_3H_6O_2$, and propionic acid $C_4H_8O_2$ can be categorized as Short Chain Volatile Fatty Acids (SCVFAs).
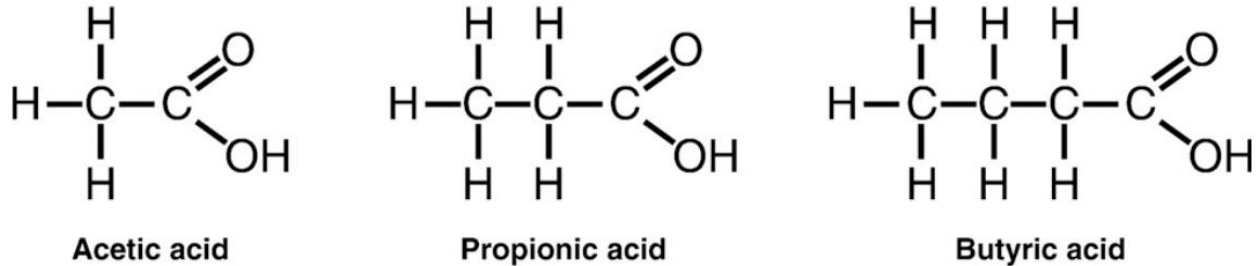
Figure 1.3 - Molecular structure of each VFA.

VFA concentrations are a necessary parameter for monitoring AD performance, and it is essential to monitor VFAs in order to understand the biological process. Furthermore, the process's stability can also be indicated by the ratio of VFAs to total alkalinity, with the target or recommended value remaining below 0.3 [27]. An ideally operating anaerobic digester gives concentrations of approximately 10, 5, and 0.3 mM (Milli Molar) for acetic, propionic, and butyric acid, respectively [18]. However, any degradation of the process within the reactor can result in spikes of up to 30, 20, and 3mM, respectively. Despite this, between various systems and feedstocks (raw material input to the digester such as sewage sludge, food waste), the precise concentrations of VFAs for indicating imbalances vary accordingly. It has also been suggested that the key VFA that indicates the stress status of the AD process is propionate [28]. Therefore, differentiation between individual VFAs is crucial, and there is a need for advances in sensor systems that can distinguish VFAs individually without considerable complexity and expense.

1.4. Process Monitoring

There are three phases in biogas production solid, liquid and gas, as shown in Figure 1.1, where a wide range of complex variables, including physical, chemical, and biological variables, must be observed [29]. It is preferable to measure all process variables in real-time since it leads to a more reliable and successful process, leading to a high-quality end product [30]. The anaerobic digester's liquid phase contains several types of substrates (solid and liquid), organisms, and dissolved gases. When taking measurements, it is beneficial to use the liquid phase to reduce the time lag in detecting changes and changes in the state of the process and to obtain a homogenous sample. However, traditional methodologies for detection require significant analytical analysis for some variables, resulting in a large time gap between sample acquisition and variable detection. Therefore, many current sensor systems rely on at-line or off-line measurements. As the metabolic

operation continues after the sample is taken, this raises some difficulties about sample treatment, resulting in a delayed reading of the AD state. Furthermore, current sensor technologies are vulnerable to biofouling and clogging due to the AD's environment. This reduces their sensitivity and lifetime and can result in inaccuracies in measurements.

In short, off-line sensors are laboratory tests requiring chemical analysis; thus, it takes a lot of time to check the process status at a certain time. But since the process is continuous, the output results will be from the process state in past time.

1.5.   Sensor Specifications

Improving monitoring technologies for AD is the cornerstone of achieving higher methane efficiency and implementing fully automated processes to minimize operating costs. Presently, for monitoring different parameters within a bioreactor, several techniques are available. These primarily involve the extraction of a sample from a bioreactor [18]. Improvements are required to make cheaper sensors with greater precision and specificity. Various methods of either in-line or at-line sensor systems are applied to AD monitoring [30]. In-line monitoring is when the sensor is directly interfaced with the bioreactor's internal environment. In-line sensors should be able to withstand digester pressure and function within the temperature range of 35-55˚C.

In contrast, at-line monitoring is when a sensor is outside of the digester in a sampling loop or similar. Continuous monitoring can be accomplished with in-line measurements, providing measurements with a minimum time lag, ensuring a more reliable and effective process that results in a high-quality product. However, due to the composition of the digester and the digestate, the positioning of the sensor is important. The process would need multiple sensors in the case of an inhomogeneous material to portray the entire process. To achieve the best process control, it is crucial that sensor technology can provide precise measurements of the desired variable in an environment consisting of multiple components. Therefore, the ideal sensor must have necessary analytical characteristics such as precision, sensitivity, and selectivity.
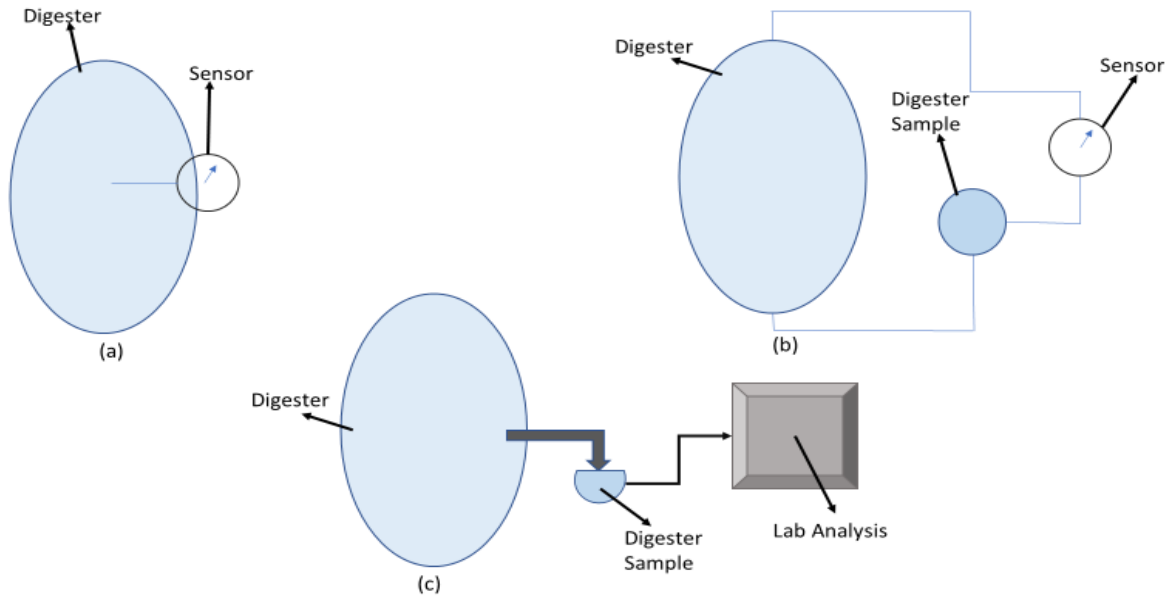
Figure 1.4 – Overview of different arrangement methods of sensing (a) Shows in-line measuring directly interfaced with the environment (b) Represents at-line sensing and (c) Depicts off-line sensing method.

High dimensional data is preferable to perform multivariate data analysis since it can use a diverse set of sensors and simultaneously target multiple variables, reducing the maintenance cost of sensors. Therefore, apart from the sensor features stated above and keeping the harsh environment of the digester in mind, a non-invasive sensor is more suitable for reliability, less prone to biofouling, and low maintenance cost [30].

In essence, live monitoring of the AD process depends on the sensor's response time, i.e., the time it takes to get the result. Response time depends on analysis duration, solid retention time, and filtration volume. So, robustness, accuracy, selectivity, sensitivity, low response time, and consistent, clear, and stable values are all desired and expected features of a good in-line sensor.

1.6. Conventional VFA Sensing Techniques

Several different VFA detection methods exist, as shown in Table 2, but they are off-line detection methods (i.e., a sample is taken from the digester and tested in the lab). These off-line detection methods consume a lot of time to analyze the sample, leading to delays in process control and, eventually, a bad AD process. Also, some methods do not provide high specificity and selectivity, which makes them less desirable. Despite this, gas chromatography (GC) and titration are commonly used as detection methods as well as spectroscopy. Table 3 gives an overview of detection methods and compares their properties and costs.

Table 2 - VFA detection methods [15]

Spectroscopy	Electrochemical	Chromatography	Other
Fluorescence	pH	Gas Chromatography	Acoustics Chemometrics
Infra-red	Redox Potential	Liquid	Mass Spectrometry
Near Infra-red	Electronic Tongue	Headspace gas chromatography	Microwaves
Raman	Electronic Nose	Paper chromatography	Titration

Table 3 - VFA detection method characteristics [18]

Technology	Accuracy	VFA distinction	Sample processing	Expertise	Computation	Duration (mins)	Cost (kUSD)	TRL
GC	5	5	5	5	2	60	30	6
Titration	3	1	3	3	2	30	15	9
IR-spec	4	5	5	5	5	60	50	4
Colorimetric	3	5	2	3	3	2	2	2

1.6.1. Titration

Titration is a method to determine the concentration of a known analyte using quantitative chemical analysis by adding it to a standard reactant of known concentration with which the analyte reacts in a definitely known proportion. Titrimetry is widely used for chemical analysis in laboratories as it is a straightforward method to determine the concentration of a known chemical analyte based on pKa and pH.

It has been used for research and development for VFA detection resulting in reliable and robust measurements. By using the titrimetric method, there is no need for large sample filtration while quantifying the VFAs (depending on the consistency of the digestate). By simply taking a sample from the anaerobic digester, a few simple sample pre-treatments are performed (typically filtering or centrifugation), and then the acidic concentration of the sample can be determined. To monitor titration analysis in practice, a computer-based system is used to automatically quantify VFAs and adjust the protocol automatically based on VFA concentrations.

Automatic titration is easily applicable to any AD process to perform VFA analysis because it is a simple and robust procedure. One example of automatic titration is "Anasense", an on-line analyzer for AD monitoring. It is an automated machine based on titration that can measure VFAs, pH, and alkalinity to enabling the monitoring of critical parameters in the AD process. Its quick result time of 10-15 min allows easy and efficient monitoring.
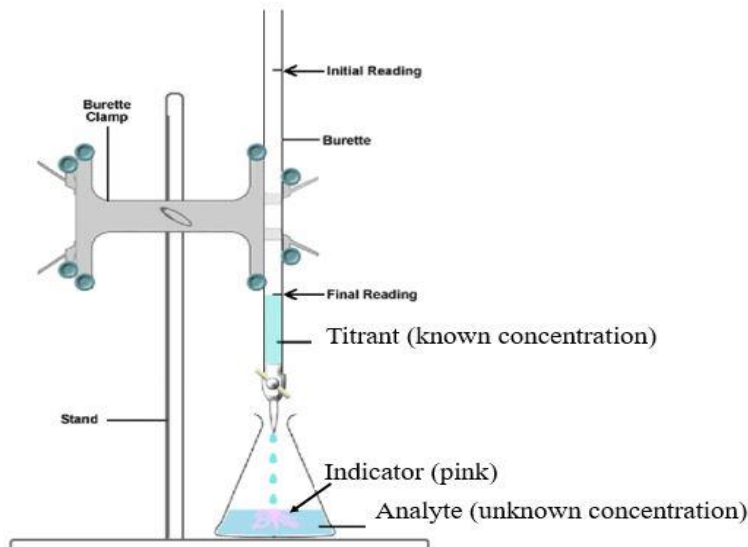
Figure 1.5 - Titration method overview & apparatus. [31]

The shortcoming of this method is that it does not detect individual VFAs but rather gives the total acidity of the mixture. Many interfering elements, including acids or bases (e.g., lactate and phosphate), are present in large amounts in AD, which is one of the reasons for the decreased accuracy of titrimetric methods.

1.6.2. Chromatography

Chromatography is a method to separate components in a mixture. It separates solutes that are dissolved in the same solvent. Chromatography has two phases mobile phase and a stationary phase. There are two main types of chromatography Gas Chromatography (GC), in which the mobile phase is a gas, Liquid Chromatography (LC), in which the mobile phase is a liquid. The sample is transported in a mobile phase that carries it through a stationary phase material. The different components of the sample mixture travel at different speeds through the solid phase, causing them to separate (e.g., separated due to size or interactions). The shortcoming of titration methods can be filled by the high selectivity and sensitivity of LC. LC methods provide highly accurate results for analytical measurements and should always be used to calibrate any VFA detection method. In general, these sensors are off-line and in a controlled laboratory; however, there have been reports of their incorporation into VFA research in AD reactors. This type of measurement involves on-site digestate filtration prior to analysis, and although the process is automated, regular human intervention is required to maintain the analysis system. By replacing the filtration stage with a pre-treatment sample cell, attempts have been made to eliminate the need for filtration methods that are prone to biofouling [32]. This pre-treatment changes the temperature, pH, and ionic strength of a sample to decrease the solubility of VFA, yielding a gaseous sample of the VFAs for analysis using gas chromatography (GC; similar to LC, but the sample is in a gas

phase as opposed to a liquid phase). Despite this, chromatographic methods are not feasible for smaller-scale biogas plants because of the need for a chromatography machine and the high cost associated with system procurement, maintenance, and expertise. Figure 1.6 illustrates a basic chromatography method working in the separation of components.
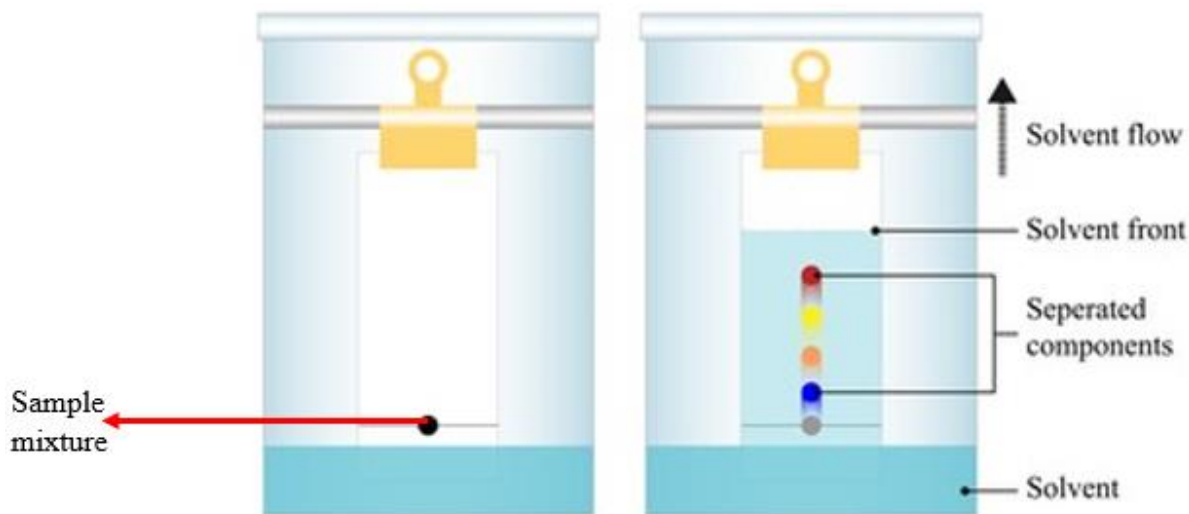
Figure 1.6 - Working of chromatography method.

1.6.3. Spectrometry

Electromagnetic Radiation refers to the propagation of the waves through space carrying energy. Spectroscopy is the study of the interaction between substances or chemicals with electromagnetic radiation as a function of frequency or wavelength. The electromagnetic (EM) spectrum ranges from 1 Hz to over 1020 Hz. Figure 1.7 shows different wavelengths of the EM spectrum and their classification. It is classified into waves based on the frequency of the photons.
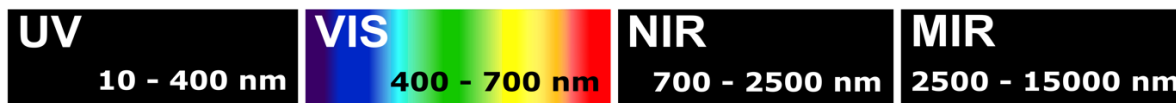
Figure 1.7 - Electromagnetic spectrum. [33]

Energy in EM waves is carried in the form of packets called photons. The energy carried by photons interacts with electrons excites them, and they change their energy level or depict certain vibration patterns, which in turn changes the characteristics of the electromagnetic waves. There are several types of spectroscopy (e.g., absorption, Raman, dielectric, and fluorescence). As an example, absorption spectroscopy, which can be done as a measure of transmission or reflection of EM radiation through the material. The material absorbs the energy in electromagnetic radiation as a function of wavelength. The light transmits through the material at certain wavelengths without absorption, as depicted in Figure 1.8 absorption spectra, and the rest of the energy is

absorbed by the substance. The transmitted light can be calculated using Beer's Law. This results in a light transmission profile that is the light source minus the light absorbed in the material.

Similarly, when measuring the transmission spectrum in absorption spectroscopy, the sample will not absorb the energy. The electromagnetic radiation will pass through it except certain wavelengths where incident light will be absorbed, as depicted in Figure 1.8. Fluorescence is the emission of electromagnetic radiation by a material that has absorbed energy or electromagnetic radiation. This radiation emission is captured by the sensor and measured as shown in Figure 1.9. The emitted energy wavelength is longer compared to the absorbed energy wavelength (i.e., Stokes shift).
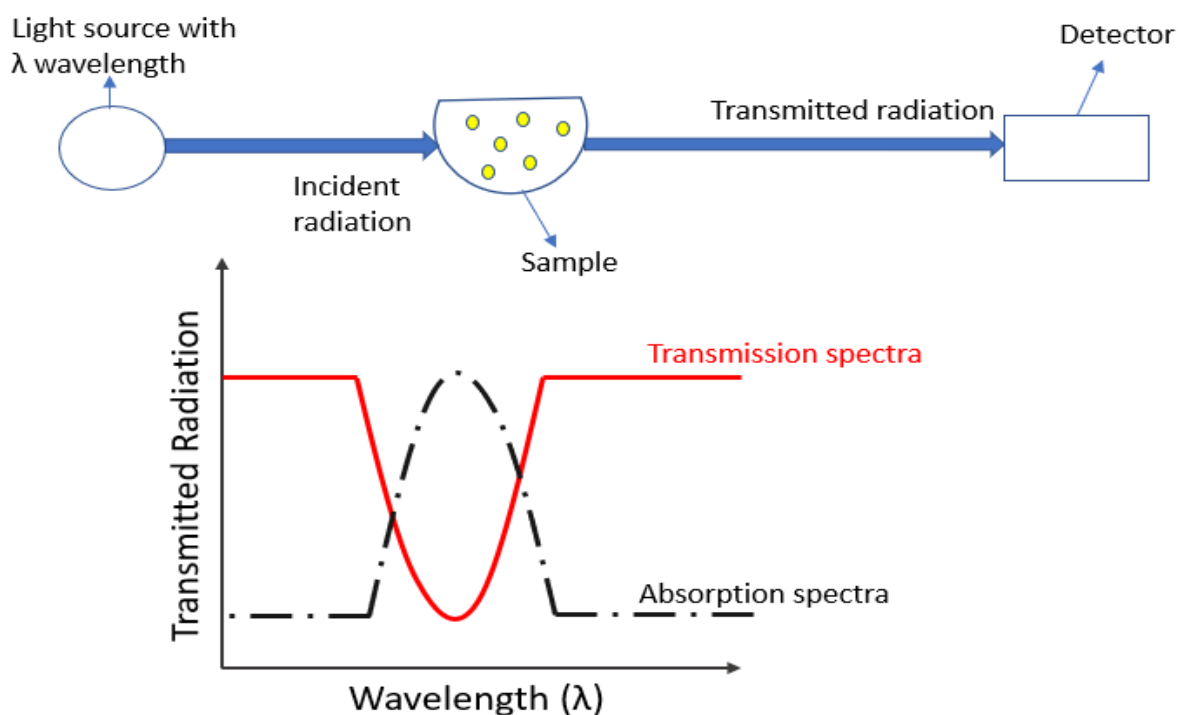
Figure 1.8 – Setup layout for measurement of transmission & absorption spectrometry.

Near-infrared (NIR) and infrared (IR) spectroscopy can be used to classify and investigate several chemicals present in a sample. On the IR spectrum, all organic and inorganic compounds have particular spectral vibrational signatures, and the more these vibrational modes are excited, the greater the specificity of analyte determination. Therefore, IR spectroscopy can provide fast, sensitive, and reliable anaerobic digester multi-analyte sampling. In addition, IR sensors are easy to integrate into an on-line in-line measurement device, using mostly non-invasive technologies via direct beam or optical fiber methods in a recirculation loop. However, sample treatment to obtain a particle and gas-bubble free sample exposes the system to biofouling and clogging. In addition, there is a need for extensive calibration measures to achieve appropriate results.

Alterations in the digester's feedstock require new calibration to be done since the calibration curve only applies to a particular feedstock.
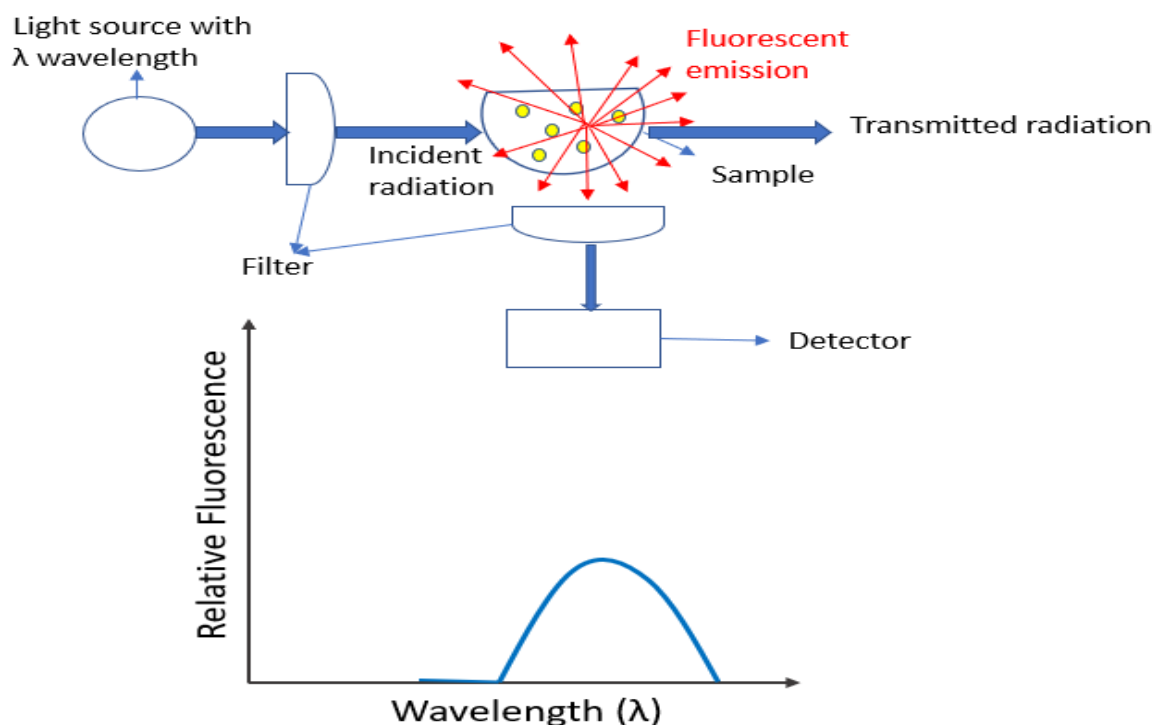
Figure 1.9 – Generic setup for measuring fluorescence.

1.7. Sensor System Theory

The aim of this project is to develop an optical sensor to detect VFAs using chemical sensing as opposed to traditional electrochemical sensors that involve current and voltage to measure something. Chemical sensing detects small or significant changes in the chemical environment to give an overview of the state of a process. It is based on the concept of optical nose and tongue, i.e., the sensor detects changes that mimic olfactory and gustatory (sense of smell and taste) of animals or humans. Humans have around 400 hundred receptors that are able to differentiate over 10,000 scents [34]. This type of sensing based on the olfactory system allows us to have high selectivity and sensitivity, thus discriminate many chemicals over different levels of concentrations. The sensor uses optical properties, hence infra-red, visible, or ultraviolet light, to determine changes in so-called receptors for sensing.

The sensor used in this project utilizes many dyes that act as receptors to detect changes in a chemical environment. The optical nose sensor is based on a dyes array, where the dyes can interact with an analyte causing the dye to change optical properties. These array-based sensors use different individual cross-reactive dyes that react with the analytes through physical interaction

14

[18]. A combination of array-based techniques that use a chemically diverse collection of cross-reactive sensor elements with modern digital imaging techniques will create a composite response pattern for any analyte as a specific optical fingerprint. The most popular optical sensors of this type are based on colorimetric changes resulting from intermolecular chromophore interactions (molecules used to capture or sense light energy, where the chromophore is the part that allows the molecule to change conformationally when struck by light) or fluorophores (fluorophore absorbs light energy of a specific wavelength and re-emits light at a longer wavelength) with the analytes. The data collected from imaging is then processed and analyzed to get a useful result. Different approaches can be taken for analysis few of them are discussed in the forthcoming section, namely Principal Component Analysis, Hierarchical Cluster Analysis, Linear Discriminate Analysis [34].

The chemical sensors based on the concept of optical tongues & noses have been tested in the field for different applications in industries like food, clinical, security screening, etc., and have provided good results. But the application of chemical sensors to detect VFAs for biogas production has not been made before; therefore, it is a new perspective to VFAs detection.

1.8. Analysis Tools

Chemical sensing based on optical noses & tongues involves optical characteristics like imaging, but the sensor itself cannot give any picture about the state of the process or measurement we want to take. In order to get the result, we have to analyze the data from digital imaging. Data obtained from images is usually raw, i.e., difficult to interpret because it is usually a high dimensional vector or matrix including random values, which provides no information about the state of the process. The dimension of output data depends on how many chemical properties our sensor is sensing. Using the sensor for the differentiation of several chemical substances and determining their concentrations gives high-dimensional data. It is an advantage because higher dimensions mean the sensor is able to differentiate among many types of chemicals.

For the specific application in biogas reactors, the data from the sensor array has very high dimensionality and gives no straightforward interpreted knowledge about the AD process. Using data analysis, the high dimensionality of the data can be reduced to fewer independent components, improving the interpretation of the data [34]. Euclidian distance (ED) approaches can be used to reduce the dimensionality significantly for comparing a known sample with a reference solution. This is a useful and straightforward approach for determining the sensitivity and cyclability of the array sensor in the laboratory.

Despite using the ED approach, when using complex mixtures (as in the industry), robust approaches are required. There are different approaches to perform qualitative and quantitative

computation using Multivariate Data Analysis (MDA). Three popular approaches are Principal Component Analysis (PCA), Hierarchical Cluster Analysis (HCA), and Linear Discriminant Analysis (LDA) [18]. Generally, data analysis is divided into multiple types like clustering, regression, and classification. Clustering analysis combines data into a set of groups (clusters), while classification tries to predict information of a sample based on old data [34].

The analysis methods mentioned above can be biased or unbiased. As the name suggests, a biased method is one perspective in which some information from earlier results or measurements is incorporated regarding samples and might deviate from real value and repeating procedure multiple times will lead to different mean value than a real value, while unbiased has no external information and when repeated multiple times can converge to the real value. Biased methods may provide predictive analysis making the process more complicated using a training set. In contrast, non-biased approaches provide a semi-quantitative analysis of the results [34].

A critical step in the analysis is the choice of a distance measure, RGB color space, in this case, because it helps to improve the performance of the computational tool in classification and clustering [34]. Several types of distance measurement methods are available (e.g., Euclidean, Manhattan, and City Block) [35]. Euclidean distance is commonly used in clustering. Euclidean distance is the shortest distance between two points. Data analysis and machine learning algorithms use this distance to measure similarity or dissimilarity between two vectors or observations to classify them into groups/classes. High dimensional output requires a more sophisticated analysis approach in order to obtain some meaningful information. This is the reason we need computational tools like PCA, HCA, and LDA.

1.8.1. Euclidean Distance

Euclidean Distance (ED)is simply a measure of length between two points in plane or space in mathematics. It is calculated using the distance formula. For 2-dimensional points with coordinates for point P (p1,p2) and for Q (q1,q2):

$$ED = \sqrt{((q1\text{-}p1)^2 + (q2\text{-}p2)^2)}$$

Similarly, for 3-dimensions or higher subtracting same axis coordinates, e.g., P (p1, p2, p3) and Q (q1, q2, q3) applying the same formula, we get q3-p3, q2-p2, q1-p1, and so on.

1.8.2. Regression Analysis

Regression analysis is a statistical method used to estimate the relationship between one dependent variable and one or multiple independent variables. It is utilized to determine the strength of the relationship between variables, and based on this relationship, the prediction for a future relationship can also be assessed. There are several types of regression analysis, such as linear,

nonlinear, polynomial, etc. Basically, regression estimates how the dependent variable will change with respect to change in independent variables and how they affect the value of the dependent variable. In this report, linear regression analysis is done to find a relation between color change and the concentration of the VFAs. Mathematically it is represented with the following equation:

$$y = bx + c$$

'y' is the dependent variable; 'b' is the slope of the regression line, rate of change in y with respect to change in 'x'; 'c' is the y-intercept.

1.8.3. Hierarchical Cluster Analysis

HCA is an algorithm that combines similar objects or the ones that are closer to each other into groups (clusters). The data to be clustered is based on the ED between them, and therefore can be seen as a developed ED method, so a short distance between objects means they will combine in a cluster. Hence, HCA does what ED can do, and in addition, it is used to combine objects in clusters. The two main types of HCA are as follows:

1. Agglomerative: A bottom-up method where each object is initially single and at each step merges with other objects to form larger clusters. This continues until we have one cluster remaining at the end.

2. Divisive: It is a top-down approach that starts with one large cluster, and at each iteration, the heterogeneous cluster is divided. Iterations are done until all objects are in their own clusters.

The agglomerative approach is usually used for HCA, where clusters are merged by linkage criterion (e.g., averages and variance). The most common linkage criteria is Ward's minimum variance, where the cluster variance at each iteration is minimized [34]. A resulting tree-dendrogram is obtained where all parent and child clusters are connected to each other Figure 1.10. The figure depicts that firstly, two closest data points by distance together form a cluster. Then two closer clusters join to form one big group. This is repeated until one cluster remains. The clusters are formed based on the Euclidean distance between the data points [36]. In terms of chemical analysis, connectivity displays relative similarity (i.e., which samples are like each other), and distance displays the magnitude of this similarity [34].

Usually, HCA provides measurements about samples or organisms that are identical to each other, thereby showing how closely related they are with each other [37]. It is a popular method for presenting genetic data to show how closely related certain biological species are to each other.

There are three limitations of HCA. Firstly, being an unbiased method, where it cannot perform predictive analysis since it does not incorporate any previous information or earlier results. Secondly, dendrograms need to be recreated completely whenever new objects (or, in this case, analytes) are used, increasing the computational burden and restraining the use for qualitative analysis only.
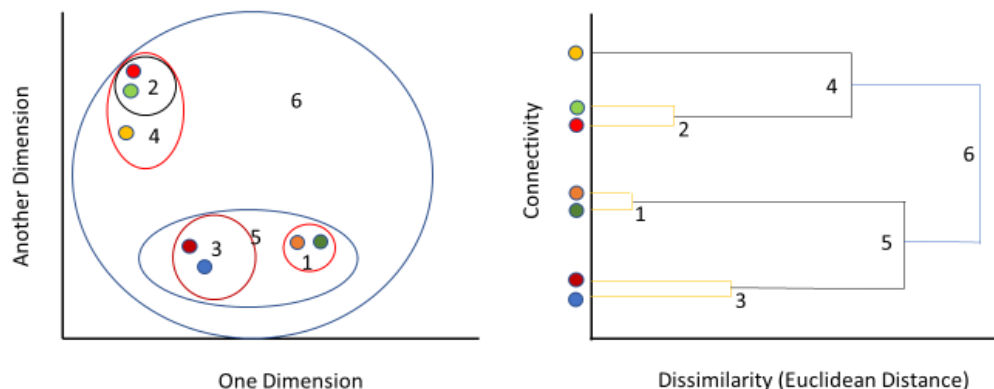
Figure 1.10 - A figure is showing hierarchical cluster analysis of multidimensional data (two dimensions only) (left) and how this analysis tool based on clustering forms a dendrogram linking the clusters(right). [36]

The third limitation is that noisy data may lead to misinterpretation of dendrograms because of one quantitative axis (ED).

1.8.4. Principal Component Analysis

PCA is an unbiased method that reduces high dimensional data to fewer linear independent components while maintaining maximum possible variance. It creates a new orthogonal set of dimensions of independent components in the data set. The first dimension has the highest variance, while the second has the second-highest variance, and so forth. Therefore, data is described by its variance with high variance, meaning that the sample varies a lot across the dimension, and for small variance, the sample is almost similar or closer to the average value across the dimension. The goal is achieved by reducing the dimensions while keeping 95% of the variance.

This method is widely used as it is readily available and included in most mathematical software packages. Even for a large number of dyes in an optical array, usually, two or three principal components are significant to fully describe the variance [34]. When many different types of analyte classes are being investigated, it is desirable to have a sensor array with a large chemical reactivity space (having high dimensional output) in order to gain from the array's high specificity and reduce the chance of overlap between classes of analytes [34] Figure 1.11.

This approach can then be used to test the principal components by measuring the variation that is generated by changes occurring within the bioreactor [25]. To show the number of principal components of the data set and of eigenvalues associated with it, a scree plot is generally made, which is the measure of the influence of a component on the data. Finally, a score plot is used to visualize various samples based on principal components with their cumulative contribution.
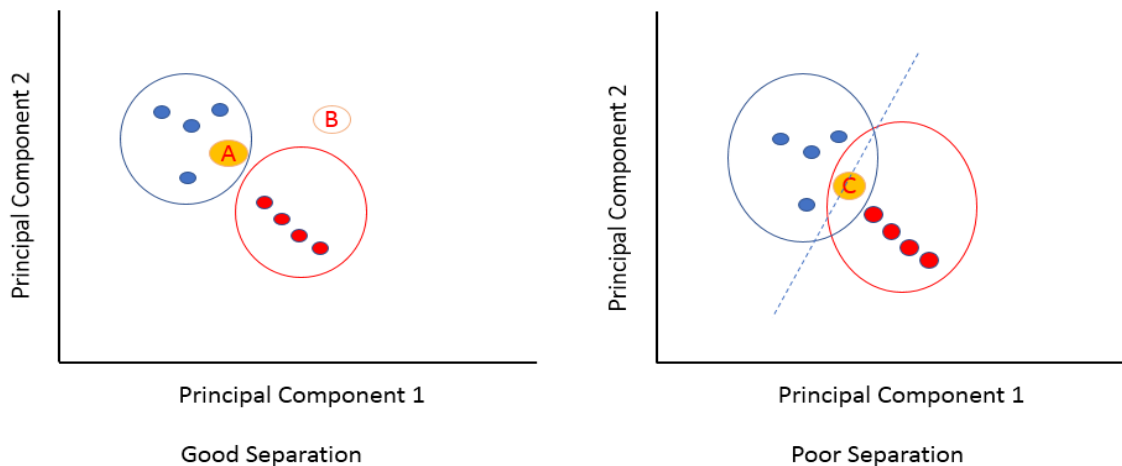
Figure 1.11 - The score plot with red and blue classes is shown. Three points A, B, and C are unknown, and circles represent confidence intervals. A good separation is achieved with high dimensional data set, and it can be concluded that A belongs to the blue class while B does not belong to any class. While data set with poor separation (right) is ambiguous, and nothing can be said for sure. A dashed line separates the two classes. [36]

Being an unbiased method like HCA can perform evaluation but not accurate predictive analysis [18]. However, it can perform rudimentary predictions by projecting the sensor response onto the dimensional space of PCA, but this is done based on the assumption that sample variation is lower than the variation of PCA dimension classes, i.e., with high dimensionality to low sample size.

1.8.5. Linear Discriminant Analysis

LCA shares some similarities with PCA because it also creates an orthogonal dimension consisting of linear components and also reduces the dimensions of the given data set. LDA tries to optimize the data by selecting dimensions to maximize the ratio of between-class variance to with-in-class variance for better discrimination ability [34]. Since it gives the maximum discrimination (separability), the threshold value is normally a mean of two classes. This optimization gives better differentiation of the sample classes compared to PCA Figure 1.12.

However, unlike PCA, the new set of dimensions are separated based on already known classes and not the total variance in the data. There are two major features of LDA. Firstly, LDA works better for high-dimensional data, like PCA. Secondly, the criteria for maximization used for selecting discriminant dimensions and the mathematical method for implementing this criterion

are analytically solvable and known, making use of LDA even more secure. In addition, to get dependable results in LDA, a high sample size compared to sample classes is required. In order to obtain an accurate prediction matrix may require a sample size that is 100 times greater or even more. Due to the requirement of a large sample size relative to sample classes, the prediction matrix is unreliable when the sample number is not considerably larger than sensor dimensions. This poses a problem for high dimension sensors with smaller sample sizes.
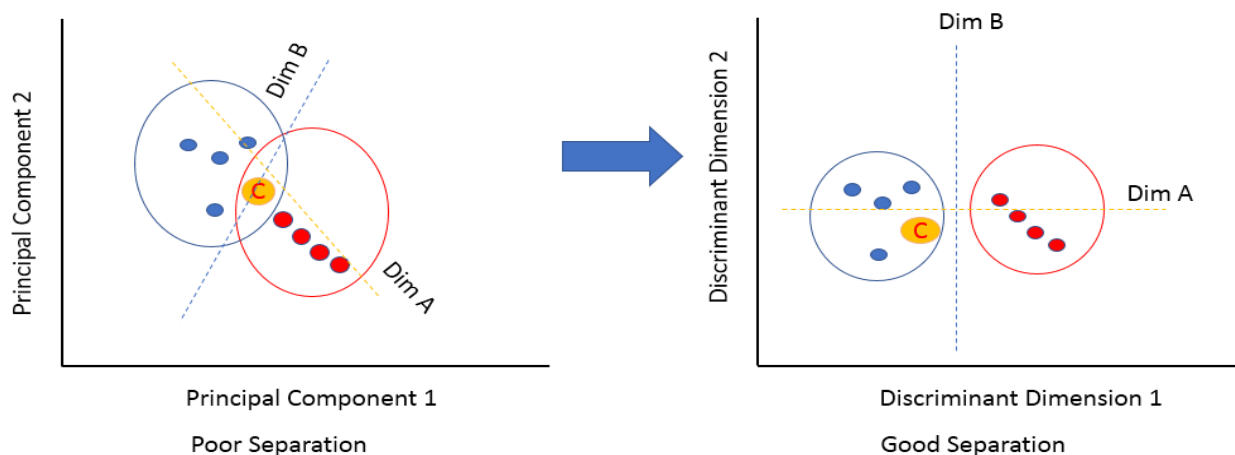
Figure 1.12 - Score plot of PCA (left) and LDA (right). From figure 1.11, where C was unclear in PCA, by using LDA, it can be concluded that C belongs to the blue class where dimensions A, B correspond to the discriminant dimension in LDA. [36]

1.9. Objective of the Project

This project aims to develop a chemical sensor that can detect Volatile Fatty Acids accumulation to determine the condition of the process (i.e., whether it is stable or deteriorating) to take necessary measures that will result in an efficient process. Dyes will be used to assess chemical samples, and data analysis of the interaction will be performed using software to extract the results. This work is a part of a larger project called "Digitalization in operation, monitoring and control of large-scale biogas plants" (DIGI), which aims to develop new technology to improve the operation, monitoring, supervision, and control of Biokraft's large biogas plant at Skogn, Norway.

Firstly, an array of dyes is set up in a small experiment testing setup, and a high-resolution camera is mounted directly in front of the dye sensor array. The image of the array is taken before any chemical interacts with dyes, and the other one is taken after the reaction. Both pictures are then analyzed. The data analysis discussed in section 1.8 is performed on output obtained from the images to sense and identify the chemical or substance under investigation.

Chapter 2

2. Methodology

This chapter discusses details about the experimental setup, the components, and the substances used in the experiment. It also includes information about how the experimental chemicals were prepared and how the experiments were performed.

2.1. Dyes

The sensor array consists of a variety of dyes for different analytes, each with different specificities. The choice of dyes in a colorimetric array is made depending on the application, the sensitivity to several analytes, or to a more specific group of analytes. The matrix holding these dyes may also play a role in enhancing their chemical selectivity, either by altering the dye's local environment or by immobilizing the dye molecules in a sterically confined environment [34]. Furthermore, the addition of co-chemicals to modify the reactivity of the dyes is also possible [34]. The types of chemical sensors and their applications have increased considerably during the last decade due to the addition of a wider variety of dyes. The different dye classes available include; Lewis acid-base dyes, Brønsted acid-base dyes (i.e., pH indicator dyes), large permanent dipole dyes for local polarity detection, and hydrogen bonding (i.e., solvatochromic, vapochromic, or zwitterionic dyes), redox-responsive dyes and chromogenic aggregative colorants [34]. The array of colorimetric sensors is based on the intermolecular interactions between the indicator and the analyte. Since the option of chemo-responsive dyes or fluorophores is the primary factor in the optical sensor array's functionality, the results will be affected by the array content. Related properties for the material of dyes are inertness to gasses and liquids, high surface area, and resistance to a wide pH range. The dyes are prepared in gel-type material to hold them in wells of a bowl-shaped container with transparent glass to capture images. Small wells are 2mm in diameter drilled at equal distance in order 6x7to hold the dyes, which will change their color with respect to the chemical or substance added to it. Table 4 shows a complete list of the tested dyes and the dyes which were at the end used in the sensor array and discarded dyes.

Table 4 - List of tested dyes (Dyes highlighted in colors were used in sensor array & discarded dyes in red)

Type	Dye	Water Solubility
Redox Dyes	N-phenyl-p-phenyl diamine	Insoluble
	O-dianisidine	Insoluble
	N-N'-diphenyl-p-phenylene diamine	Insoluble
	O-tolidine	Insoluble
	Diphenyl amine	Insoluble
Lewis Acid Dyes	Palladium (II) acetate	Insoluble
	Octaethyl-porphine cobalt (II)	Insoluble
	5,10,15,20-Tetraphenyl-21H,23H porphine	Insoluble
	5,10,15,20-Tetraphenyl-21H,23H porphine zinc	Insoluble
	5,10,15,20-Tetraphenyl-21H,23H porphine copper (II)	Insoluble
	Copper (II) phthalocyanine	Insoluble
Solvatochromic	Nile Red	Insoluble
	Dansyl chloride	Insoluble
	Reichardt's Dye	Insoluble
	Merocyanine 540	Insoluble
Brønsted acid-base dyes	Phenol red	Insoluble
	Methyl red	Insoluble
	Cresol red	Insoluble
	Acridine orange base	Insoluble
	1-naphthyl red hydrochloride	Insoluble
	Thymol blue	Insoluble
Discarded Dyes (Unused)	M-cresol purple	Soluble
	Bromophenol blue	Soluble
	Ethyl viologen dibromide	Soluble
	Nitra zine yellow	Soluble
	Tetrakis-porphyrin iron (III)	Soluble
	Aluminum phthalocyanine	Soluble
	N-N-dimethyl-6-propionyl-2- naphthylamine	Soluble
	Methyl viologen dichloride hydrate	Soluble
	Indigo Carmine	Soluble
	Bromocresol green	Soluble
	Thionin acetate salt	Soluble
	Malachite green chloride	Soluble
	1-ethyl-4-(2-hydroxy styryl) pyridinium ion	Soluble
	Chlorophenol red	Soluble
	Lissamine green B	Soluble
	Bromocresol purple	Soluble
	Pyrocatechol Violet	Soluble

2.2. Sensor System Design

To detect VFAs accumulation, we need a sensor system that can detect and check process deterioration through chemical changes in dyes after interacting with samples. It should be able to operate automatically without any human interference and perform in the reactor's environment and test multiple times before need to change dyes. It must be able to take data for computation and apply tools to give results. Lastly, connecting this sensor system via the internet to monitor remotely will be advantageous, as a physical presence to check sensor output is not needed, but remote monitoring is not in the scope of this project.

2.2.1. Theory and Components

To test the theory in the lab, a sensor system to allow automatic control of sensor components and liquid samples has been developed. It consists of a camera, raspberry pi, pumps, valves, relays, and power supplies. The camera is the cA4024-8gc (Basler, Germany), which is 12 MP with 4024 px X 3036 px resolution.

Three main physical features of a camera that determine an image's quality are International Standard Organization (ISO) sensitivity, shutter speed, and aperture. Maintaining the balance between the three is vital for a quality image. ISO in a digital camera is the signal amplification coming from the sensor of the camera. If the ISO is set too high, it will increase the noise and reduce the resolution, contrast, and color. Shutter speed is a measure of how long the light is captured. The time allowed for the light to enter should not be too long, as the camera is not completely stable during image capture., This will cause blurry images. The aperture is a measure of the amount of light hitting the sensor affecting the depth of field of an image.

A raspberry pi is being used as a controller that gives ON/OFF signals to the relays, takes input from the camera, and controls the image data's computation in Figure 2.1(a). There are 4 relays used to control two valves and two pumps. The relays turn the pumps on and off when the input signal to these relays changes. The pumps and valves help in controlling the quantity of sample that goes through the dye array. After interacting with the dyes, the sample goes to a waste container Figure 2.1(b). An ethernet cable is used to connect the camera to the Raspberry pi, as this allows the camera to act as a unit on a local area network, so images can easily be transferred to the Raspberry pi. The system has a touch-screen LCD used to interact with the controller.
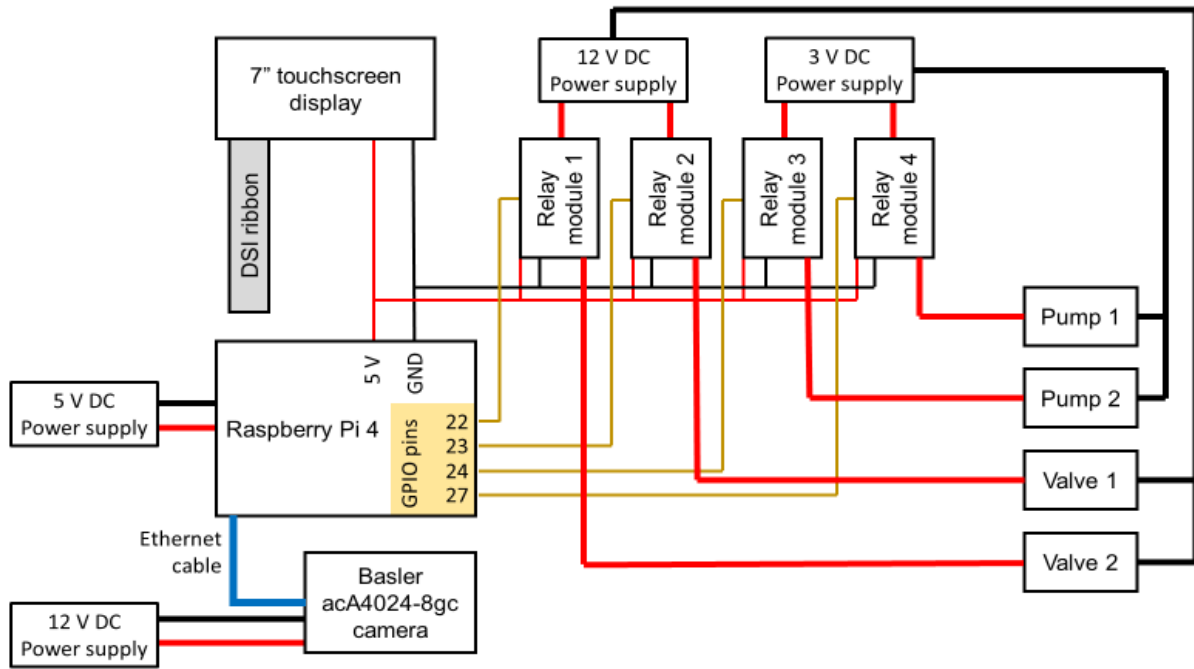
(a)
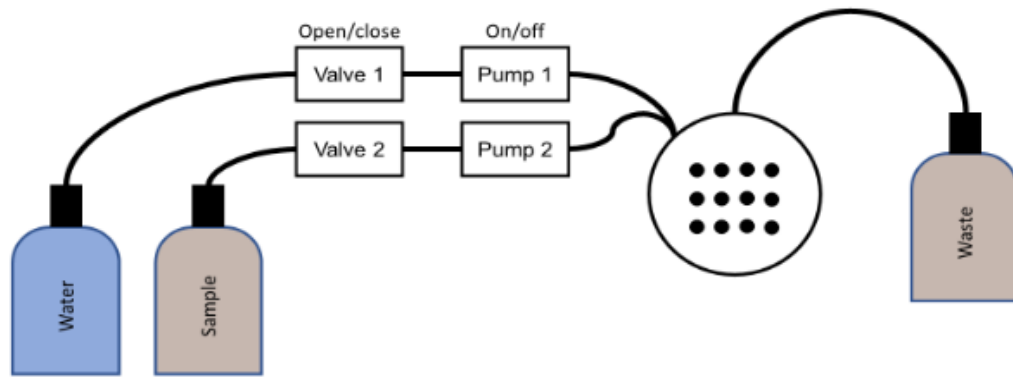
(b)

Figure 2.1 - Connections and layout of the experimental setup. (a) shows how the connection between components is made. (b) shows the sample and water intake of the dyes.

Table 5 - Summary of components in the sensor

Sensor System	Equipment
Raw Data	Microcontroller
	Camera
	Light
	Pumps
	Dyes Array
Data Analysis	Raw data computation using Minitab
	LDA using python's scikit library

2.2.1. Construction of Sensor

The setup consists of a wooden box with all components attached to it, as shown in Figure 2.2. Four relays are wall-mounted on the inside of this box. The camera is also fixed and aligned with the array to capture the dye array's maximum detail. The dye samples are inside a small glass chamber with 2mm wide wells, as shown in Figure 2.1(b) above. The array has 7 x 6 wells, so 42 different dyes can be used. This glass chamber is constructed with guiding pins at 3 different points to keep the position of holes the same to get the same array orientation upon changing the array dyes. Two pumps along with two valves are mounted to circulate water and sample through the dyes array container. The pumps circulate according to the value fixed by the user so, it is certain that the chemical in the container is able to interact with dyes.
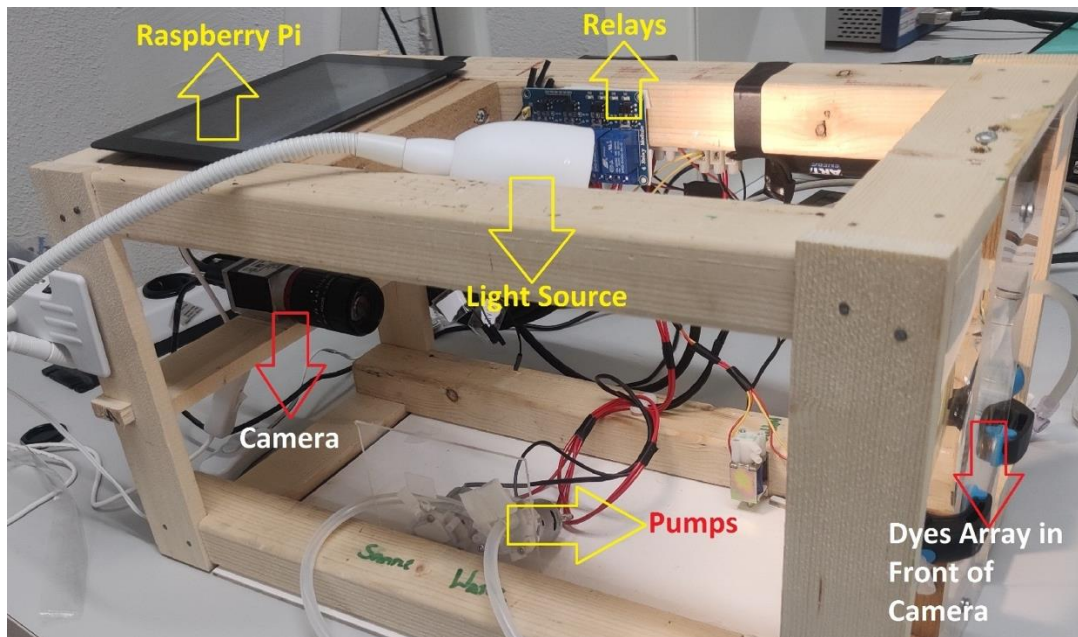
Figure 2.2 - Experiment rig.

2.3. Working of Sensor

Colorimetric sensor arrays are based on the use of an embedded indicator (i.e., a dye) in a matrix. There is either a fluorophore or a chromophore present in the indicator. A light source (e.g., a light-emitting diode) is aimed at illuminating the sensor array's dyes, while the imaging device is located in front of the matrix, allowing the imaging for digital imaging. Water or the sample is then pumped through the dye array. The interaction of the indicator with the analyte may result in a change in the dye's color. A difference map is created by determining the color change between images prior to and after the chemical of interest is present to detect colorimetric changes Figure 2.3. To evaluate the difference between the color of the dye before and after, the RGB values of the indicator dyes are subtracted pixel by pixel, generating RGB difference data that can be used for further quantitative and statistical analysis to determine the analyte that caused the color change.
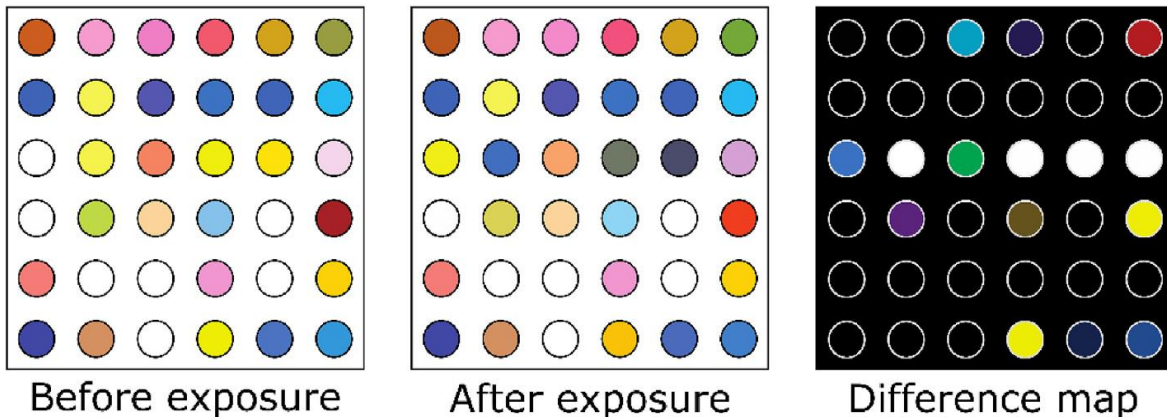
Figure 2.3 - Representation of colorimetric array. A 6x6 array is shown before subjecting it to an analyte (left), the middle is after exposure, and the right one represents the RGB difference map of the first two figures. [18]

2.3.1. Sensor Control

To control the sensor, a programming script is uploaded to the microcontroller, i.e., Raspberry Pi in this case. For the purpose of automation and repeated sample testing, a program that can operate the sensor system in order to achieve automatic testing is needed. This program should be able to use the components in specific order to make the sensor work and perform sample testing. A code can be written for this function in any programming language. In this project, the script is written in Python. The filename extension of Python is '.py' in this case, while there are others as well.

Pypylon is a camera software package including drivers and tools for straightforward interfacing of Basler cameras. It is easy to use package that can be integrated into Python. Additional libraries used in the Python script are:

- **Matplotlib**—Library for plotting, creating static and animations for visualizations.
- **Numpy**—Library that supports mathematical functions and large dimensional matrices.
- **GPIO**—Library that helps control input/output devices and also interface with raspberry pi.
- **Skimage**—Library designed for image processing.
- **Scipy**—Scipy contains optimization, linear algebra, integration, etc.
- **Python Image Library**—Used for opening, saving, and manipulation of images.
- **Scikit**—Library containing machine learning and statistical modeling tools for data analysis.

The first script to run before an experiment with the sensor array is the Reference.py (see appendix). This pumps water into the glass chamber and takes the sensor array's initial image to use as a reference. Then the main script called Control.py (see appendix) can be used to undertake an experiment. When executed, it calls for many functions that have been written to carry out different tasks discussed below. Overall, the script execution and experimental process are as follows:

1. Water is pumped for cleaning the array, then a reference RAW format image of the sensor array is taken with the camera, date, and time marked.
2. Then this image is read to determine the RGB values of the dyes and is converted into a JPG image to reduce the memory requirement. The RGB is averaged across a 10x10 pixel area, so it is the average of 100 pixels in the dye well.
3. The RGB difference is then calculated with reference to the new image's RGB values, and the image was taken through the Reference.py script, giving a calculated difference between the RGB values.
4. The sample is pumped into the glass chamber and interacts with the dyes.
5. Then this sample is imaged and read to determine the RGB values of the dyes and is converted into a JPG image to reduce the memory requirement.
6. The RGB difference is then calculated with reference to the new image's RGB values and the image taken through the Reference.py script, giving a calculated difference between the RGB values.

This flow can be run automatically indefinitely to see how the sensor array reacts to the analyte through multiple cycles (Figure 2.4).
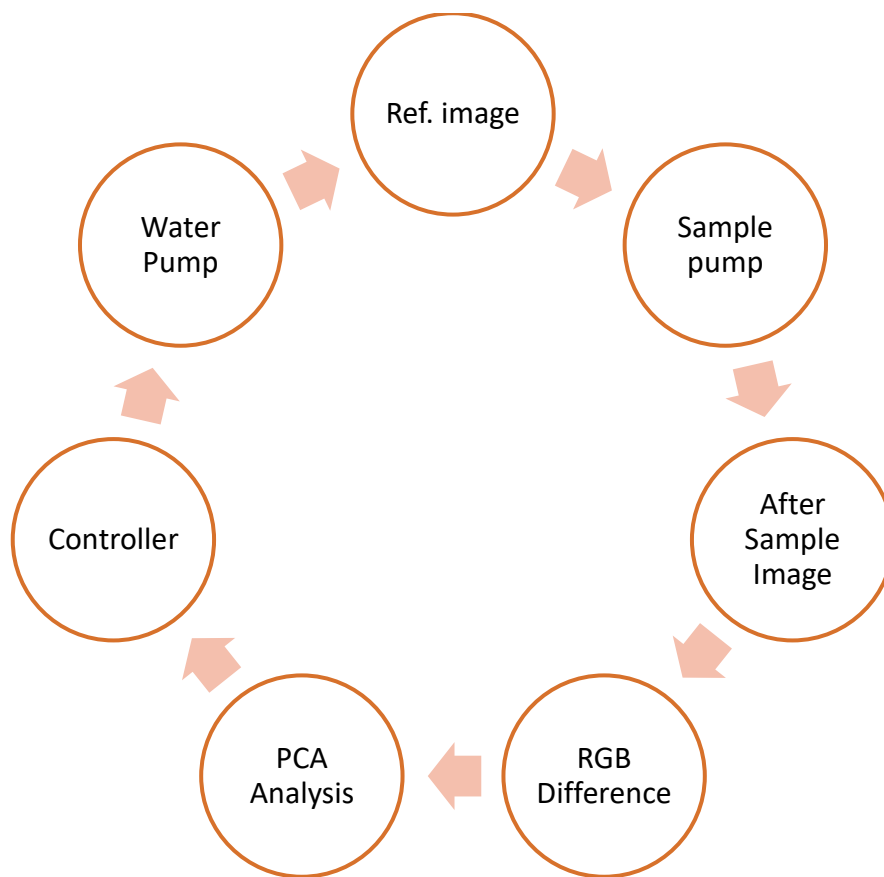
Figure 2.4 - Stepwise representation of script execution in Raspberry pi.

2.4. Experimental Method

This section discusses in detail how the experiment was conducted and the substances, components that were prepared and used. The experiment includes various substances in different concentrations to analyze our chemical sensor's accuracy, sensitivity, and selectivity. The first step is to choose the correct dyes from different types of dyes. We aim to detect VFAs, i.e., acids, so multiple dyes from Brønsted acid-base dyes and Redox dyes change color with pH change, thus good indicators for detecting acids. Dyes are receptors in the optoelectronic nose and thus are a crucial part of the experiment. After selecting dyes, a proper container for our sensor array is required, which can hold dyes and the digestate/water to interact with the array.

The next important part is to take high-quality pictures from which color changes in dyes, i.e., the RGB difference discussed above, will be calculated. For good pictures and calculating RGB difference, there are few things to be kept in mind. Firstly, camera settings and the camera position should also remain the same, i.e., it should be fixed as depicted in Figure 2.2 to avoid blurry images due to any movement or shaking. Secondly, the sensor array container must have the same physical position, and lightning should be the same for all pictures. It is so to ensure that before and after

28

images have the same dye spot and the comparison of the same dyes is made for both pictures. The dye samples are inside a small glass chamber with 2mm wide wells, as shown in Figure 2.1(b) above.

Lightning the sensor array is also important part appropriate lighting on the sensor matrix for imaging is important. A generic powerful white LED white was used in the experiment. The sensor array's appearance is also dependent on the material and the angle of the light incident upon it. Various characteristics (e.g., shape, color, surface defects, and transparency) of the material may be illustrated by altering the source's angle, which influences how the light is reflected. The colorimetric sensor array aims to detect the colors of the dye spots while minimizing surface detail. Another factor regarding how the light probes the array is the use of diffusing material. If the light has a direct path to the array, there will be an unequal distribution of lighter and darker spots in the picture. Utilization of a diffusing material helps soften and disperse the light resulting in an even distribution of the light over the object.

2.4.1. *Dyes Preparation*

Out of 38 mentioned in Table 4, 21 different dyes, the ones that were selected were water-insoluble. We want water-insoluble dyes; otherwise, they will simply dissolve in water and not give any output. Most of the dyes in this experiment come in powder, which means that we cannot put them directly into wells of the glass container because they will not stick inside well holes and simply pour down. So, to make the dyes in gel form, firstly sol-gel recipe was tried. Sol-gel is a process in which the formation of inorganic colloidal suspension (sol) and gelation of the sol in a continuous liquid phase (gel) [38]. The process involves the fabrication of metal oxides. Sol-gel is widely used in biosensors and optical sensors because it is transparent chemically, thermally, and mechanically stable, therefore suitable for harsh environments [39]. To make sol-gel, Tetramethylorthosilicate (TMOS), Methyltrimethoxysilane (MTMS), methanol, and pure water were combined in the molar ratio 1:1:11:5 [40]. The mixture was stirred for 2 hours at room temperature. But this method was ineffective because dye concentration in sol-gel was high, and dyes leaked into the water resulting in less accurate readings.

The second method was to mix dyes with Agar powder. Agar is a naturally gelatinous powder and makes an excellent thickener for different products like soup, other liquids. To make this solution, 1 gram of Agar is mixed with 100 milliliters (ml) of water and stirred for about 15 minutes at 80˚C temperature. After the mixture is made, a small dye quantity was mixed with agar solution in a small lab sample tube while the mixture was hot. Each dye was then mixed, and after the solution cooled down, dyes were in thick gel form liquid, which was able to stick in wells of the container without leaking into water or other substance. Figure 2.5 shows the dyes after mixed with Agar and hot water to be used in dye wells.
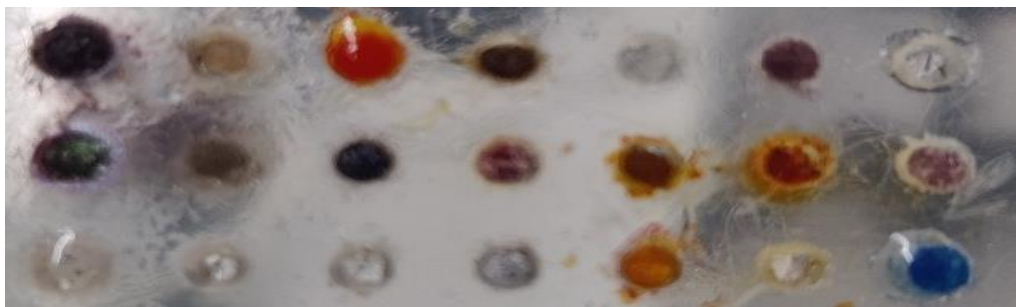
Figure 2.5 - New dyes sensor array before experiment (Top 3 rows of well are filled with dyes).

2.4.2. Method

After dyes preparation, the array container is fixed into its place, and a picture is taken. The captured image is not the reference picture; instead, it is used to determine each dye's pixel location. This pixel location of 7 columns and 6 rows is then written in our Python code, and our code will use these pixels to generate RGB difference values from images captured later on after pixels entry water is filled inside the array container and rested for 10 minutes. A "reference picture" is then taken, and following that container is filled with a VFA solution with a specific concentration and rests for 5 minutes. The image is captured and compared with the reference image to calculate the RGB difference. Again, water is filled up, rests for 10 minutes, and a different image is generated. This process is repeated multiple times with several concentrations of VFA and various combinations of VFAs, which will be discussed in the next section. Figure 2.6 shows an overview of the experimental method.
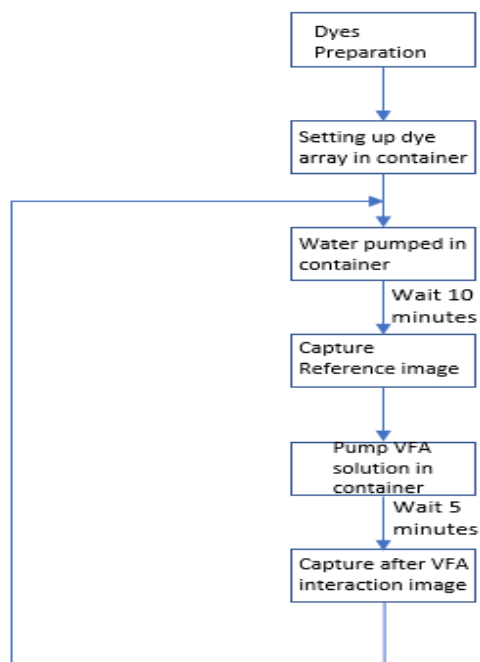
Figure 2.6 - Flow diagram of the method followed in experiments.

The code's output is a ".csv" file with 1 and 63 data values that include random numbers. Since we are using 21 dyes with RGB calculation of each dye, it generates 21x3 (63) data values. The python script has a counter that increases every time code is executed to save a new file for each different reading, including different VFAs. Each file contains 63 data points which give no meaningful result but will be used for Data Computation.

2.5.   Experiments

For checking the sensor's sensitivity and selectivity, experiments were done with different combinations of substances and different concentration levels. All experiments consist of multiple individual readings giving raw output. It is important to note that all the experiments give output in the form of raw data, which can be interpreted using the software. All of these experiments will be discussed in this section and results in the following chapter.

2.5.1.   Sensitivity Tests

Firstly, to check sensor sensitivity over multiple measurements, 30mM (of acetic acid solution in water is prepared. 1.72 mL of 99% pure acetic acid was mixed with 1 L of water. Two experiments were done to check the sensitivity over different periods.

In the first experiment, all 20 readings were taken in a repeated cycle between water and 30mM acetic acid. Once with water in the container after resting for 5 minutes and once with 30mM acetic acid solution in the container, it rested for 5 minutes. A new sensor array was prepared, and all of the readings were taken in one sitting and output of 20 RGB difference files.

In the second experiment, all 50 readings were taken at different times using the same sensor array. All alterations in this experiment were only between water and 30mM acetic acid. A new sensor array was made, and 10 cycles of water and 30mM acetic acid were performed with 5 minutes gap. Then subsequent 30 cycles were done after 24 hours with the same sensor array; during this time sensor array was resting filled with tap water. The last 10 alteration cycles were done after 5 days with the same 5 minutes gap on the previously made dye array.

2.5.2.   Concentration Tests

In order to check the concentration detection by the sensor, different acid concentration solutions were tested. Firstly, acetic acid from low to high concentration was tested. The concentration values of the VFAs were according to the expected values during the AD process. In order to examine whether low-level and high-level VFAs are detectable by our sensor, each acid was prepared for several concentrations that can be expected during the AD process.

For acetic acid, being the main VFA in the AD process, 30mM, has been considered a high level, resulting in process failure or deterioration. 0.1mM, 0.5mM, 1mM, 2mM and 5mM acetic acid

solutions were prepared to perform concentration test. 30mM solution was already tested in the sensitivity test. A new array was set up and filled up with water, and rested 10 minutes in the reference image container. Once the reference image was taken, 0.1mM acetic acid solution was interacted with the dyes, followed by water. Again, after water 0.3mM, the acetic acid solution was filled in the array. The same process was done for higher concentration levels, with lower concentration coming first.

Propionic acid was the next VFA that was investigated. For propionic acid, accumulation of 5mM in the AD process is normal, but if this level goes up to 20mM, then it results in process failure. A similar test to acetic acid was conducted with a newly made dyes array starting from 0.1mM, 0.5mM, 1mM, 2mM, 5mM propionic acid solution with testing from lower to higher concentrations. Propionic acid has the most increase among VFAs in accumulation when the process starts to fail almost 4 times the normal level.

Similarly, the last VFA butyric acid's the desired concentration is 0.3mM, but any degradation in AD can raise the level up to 3mM around 10 times the normal level. Sensitivity was examined with the concentration of 0.1, 0.5, 1, 2, 5 mM, following the same method as with the other VFAs.

2.5.3. Mixed VFAs Tests in Water

After conducting sensitivity and concentration tests, to obtain data for data computation to see if our sensor is able to differentiate between individual VFAs when they are mixed together and exposed to the dyes array.

The first solution tested was a mixture of Acetic & Propionic Acids in water. A complete list of chemical combinations and respective concentrations is given in Table 6. The solution was prepared using 70mM acetic acid by mixing 4 mL of 99% pure acetic acid in 996 mL of water and 100mM propionic acid by mixing 99.3 mL water with 0.748 mL of 99% pure propionic acid. After making these two solutions different 100 mL concentration solutions of 70mM acetic & 2mM propionic, 70mM & 6mM, 70mM & 10mM, 70mM & 20mM were made. Following this, a new sensor array was made to test the sensor. Water is allowed to rest in the array for 10 minutes to capture the reference image and then replaced with a mixture of 70mM & 2mM acids resting for 5 minutes to take the difference image. The same procedure was done for remaining solutions from lower to higher concentrations of propionic acid. The output was RGB difference was raw data to be computed with Minitab software.

The second test was done with Acetic & Butyric Acids by preparing a 70mM acetic acid solution as done in the previous experiment and 100mM Butyric acid solution by mixing 0.918 mL of 99% pure butyric mL acid. Following this 100 mL solutions of concentrations of 70mM acetic & 1mM

butyric, 70mM & 3mM, 70mM & 5mM, 70mM & 10mM were prepared. A new sensor array is made, and the same procedure is repeated for the solution of acetic & propionic acids.

The final test of VFAs was done by mixing different concentration solutions of all three acids. 70mM acetic, 100mM propionic and 100mM butyric acids prepared in previous experiments were combined together in concentrations of 70mM acetic & 1mM propionic & 20mM butyric acid, 70mM & 10mM & 20mM, 70mM & 20mM & 20mM, 70mM & 20mM & 10mM, 70mM & 20mM & 1mM to make 100 mL solution for each combination. Test was done on newly made sensor array. Starting with mixture of 70mM acetic & 1mM propionic & 20mM butyric acid, the next readings were in same order as written in text above.

2.5.4. Mixed VFAs Tests in Digestate

Testing the sensor in thick black liquid obtained from inside the digester to simulate sensor performance in actual condition. To do this experiment, VFAs were mixed together in the same concentration and manner as in the previous experiment.

The first pair of acids, i.e., acetic & butyric, then acetic & propionic, and at last solution, all three VFAs were made in the same concentrations as when tested in water. 70mM acetic acid, by mixing 4 mL pure acetic acid and 996 mL of digestate mixed firmly for 10 minutes. 100mM propionic & butyric acid in digestate were prepared by combining 0.748 mL & 0.918 mL of respective acids in 99 mL digestate.
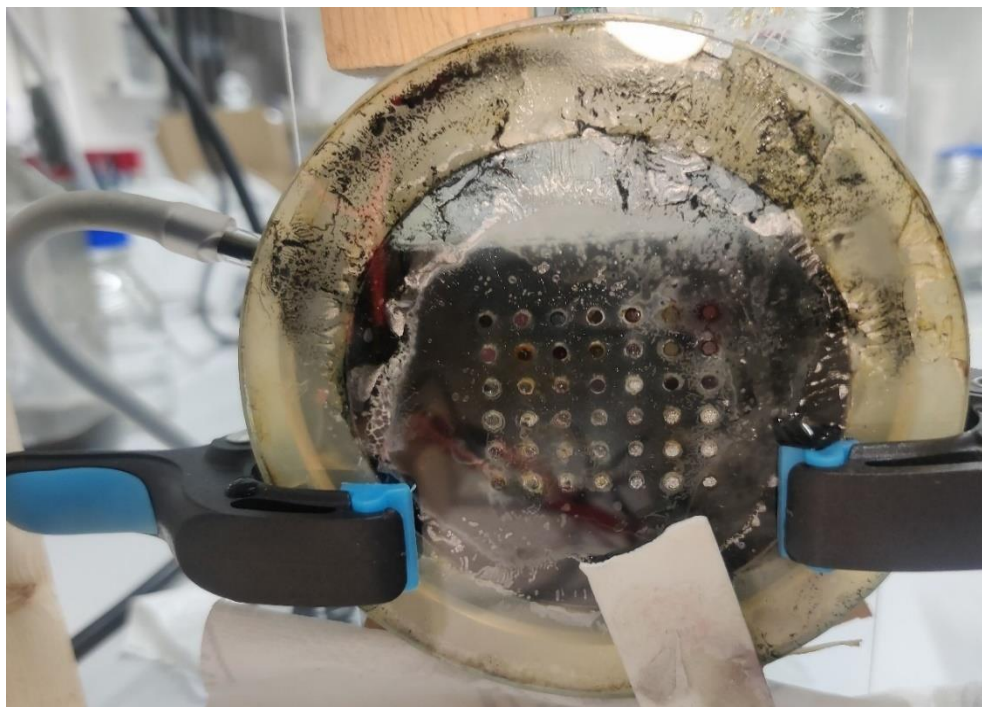
Figure 2.7 - Resting digestate mixed with VFAs in dyes sensor array.

Firstly, the solution of acetic & propionic acid in concentrations same as water experiment was done in ascending order of acid concentration. After every acid solution, water was pumped into the array and allowed to rest for 10 minutes. Similarly, by setting up a new array solution of acetic & butyric acid is tested from weaker to stronger concentrations with water pumped in after every acid reading. At last, all three VFAs mixed in digestate, having different acid concentrations, were tested in order from weaker to stronger acids with a similar procedure.

Table 6 - Summary of chemical combinations & concentrations used in experiments

Test	Chemical Combination	Individual Concentration of Chemicals mM (millimolar)
Concentration	Acetic Acid	0.1
		0.5
		1
		2
		5
	Propionic Acid	0.1
		0.5
		1
		2
		5
	Butyric Acid	0.1
		0.5
		1
		2
		5
Mixed Volatile Fatty Acids (Both in water and digestate)	Acetic & Butyric Acids	70 Acetic & 1 Butyric
		70 Acetic & 3 Butyric
		70 Acetic & 5 Butyric
		70 Acetic & 10 Butyric
	Acetic & Propionic Acids	70 Acetic & 2 Propionic
		70 Acetic & 6 Propionic
		70 Acetic & 10 Propionic
		70 Acetic & 20 Propionic
	Acetic, Propionic & Butyric	70 Acetic, 1 Propionic & 20 Butyric
		70 Acetic, 10 Propionic & 20 Butyric
		70 Acetic, 20 Propionic & 20 Butyric
		70 Acetic, 20 Propionic & 10 Butyric
		70 Acetic, 20 Propionic & 1 Butyric

2.6. Data Analysis Software

The experiments performed yield raw data from RGB difference images. Since the sensor consists of 21 dyes, with each dye giving 3 values, we get 21x3 (63) values to be computed. Firstly, PCA analysis is performed on the data from experiments. PCA is a simple method that gives good visualization of the correlation of our chemicals. It can easily classify/identify different concentrations used in our experiment based on our data. HCA is used to display similarity between different substances in this project. A dendrogram in tree-like form shows the similarity between clusters, in this case, the concentrations of mixed VFAs.

Even though LDA and PCA are somewhat similar for LDA, the sample size must be much larger than the number of classes analyzed (e.g., a large number of different concentration tests of acetic & propionic or acetic & butyric). Even though we get 63 rows of data for each reading, it is not high in terms of statistical calculations, so data from multiple experiments was combined to perform LDA in Minitab.

2.6.1. ED Calculation

The sensitivity results are based on ED calculations using "Microsoft Excel" and difference vector values. First, all 63 values from the difference vector are squared and summed up. Since the difference vector is already the RGB value of the after image minus the before image value, subtraction between two points is already done in the difference vector. Then the square root of the sum is taken. Figure 2.8 shows 63 values in the same column and their ED calculated at last. Excel formula for ED is

$$\text{ED} = \sqrt{(Sum((B1:B63)^2))}$$

B1:B63 is all the rows containing numbers, so to calculate ED for each reading, it changes, e.g., C1:C63, D1:D63, and so on.

	dif1	dif2	dif3	dif4	dif5	dif6	dif7
1	4.30E+00	6.61E+00	1.34E+01	1.42E+01	1.52E+01	1.33E+01	1.38E+01
2	2.87E+00	5.47E+00	1.18E+01	1.46E+01	1.55E+01	1.50E+01	1.51E+01
3	1.34E+01	1.10E+01	1.39E+01	2.11E+01	2.37E+01	2.38E+01	2.41E+01
4	-1.89E+01	-6.82E+00	-4.78E+00	-7.38E-02	3.75E+00	-2.23E+00	-9.61E+00
5	-2.76E+01	-1.03E+01	-6.36E+00	-3.81E-01	9.41E-01	-2.02E+01	-2.75E+01
61	7.98E+00	7.71E+00	5.82E+00	7.56E+00	6.48E+00	6.31E+00	5.07E+00
62	7.16E+00	6.14E+00	5.30E-01	2.18E+00	3.57E-01	1.56E-03	-3.79E-01
63	1.04E+01	4.90E+00	3.70E+00	2.83E+00	1.55E-01	-2.17E-01	8.66E-01
E.D.	107.59878	69.955866	80.584677	92.1056	101.12094	105.47529	115.17902

Figure 2.8 – Data arrangement for ED calculation using RGB difference vector in Microsoft excel of each reading in the same experiment. Each reading gives a single column with 63 rows, and then ED is calculated for all of the readings. ED values are only used in sensitivity and concentration tests.

2.6.2. User Interface

Raw data obtained from python script consisting of difference vector RGB values having 1 column and 63 rows for each reading is imported into the software. Figure 2.9 shows acid combination and corresponding RGB difference data numbers. Each row in Figure 2.9 represents a unique reading during the experiment and is combined here on a single worksheet. Thus, every row has 63 columns from v1-v63, as shown in the figure above.

↓	C1	C2-T	C3	C4	C5	C6	C7	C8
	Reading No	Group	v1	v2	v3	v4	v5	v6
1		water	-4.62E-01	-2.73E-01	1.27E+00	3.71E-01	1.90E+00	1.84E+00
2		70mM Acetic, 2mM propionic	3.56E+00	4.83E+00	5.61E+00	4.85E+00	7.29E+00	1.02E+01
3		70mM Acetic, 6mM propionic	-1.33E+01	-1.20E+01	-8.66E+00	-4.57E+00	-4.53E+00	-4.50E+00
4		70mM Acetic, 10mM propionic	-1.15E+01	-1.20E+01	-9.56E+00	1.24E+00	4.45E-01	-2.11E+00
5		70mM Acetic, 20mM propionic	-1.18E+01	-1.26E+01	-1.04E+01	-3.09E+00	-3.75E+00	-6.53E+00

Acids Mixture   **Raw Data**

Figure 2.9 - Categorized RGB difference data values for acids and their concentration in a Minitab worksheet. Raw data (in green) shows python output of RGB difference and (in red) the acid mixture used.

2.6.3. Software Configuration

PCA

To perform PCA, first import data into software as shown in Figure 2.9. Tap on "Stat" in the menu, select "Multivariate," and click "Principal Components". Doing this will pop a window, as shown in Figure 2.10. Select the data to be computed, i.e., variable "v1" through "v63". Tap on the "Graphs" button to display graphs in the output screen and click "OK" to run the program.
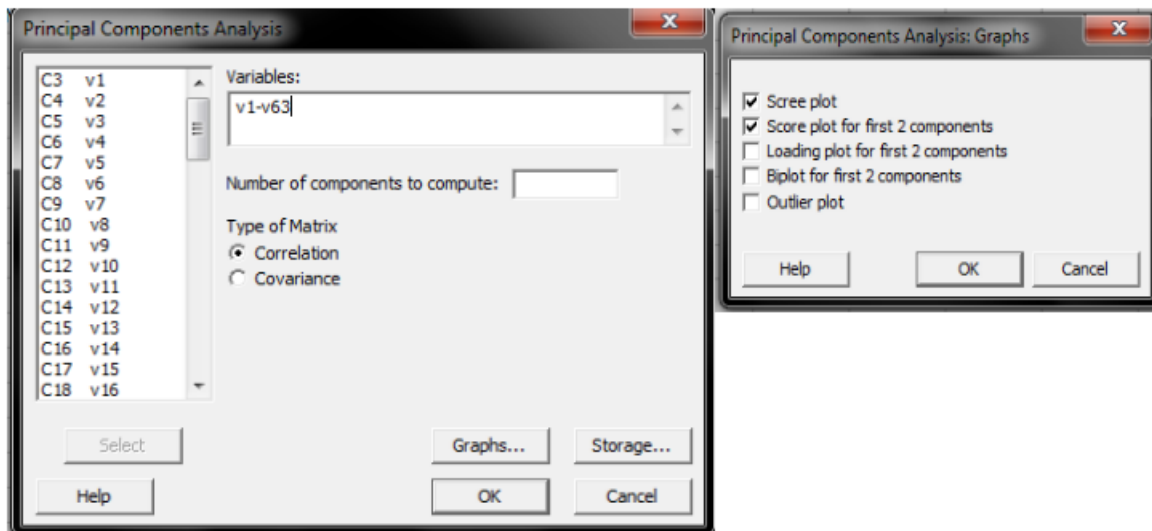
Figure 2.10 - PCA configuration menu.

HCA

Importing data in worksheet and then click "Stat" select "Multivariate" and press "Cluster Observations". A menu window will appear on the screen, as depicted in the figure below. Select the desired options and press "OK". A dendrogram will appear in the output window, which shows a cluster among acid solutions and their similarity percentage.
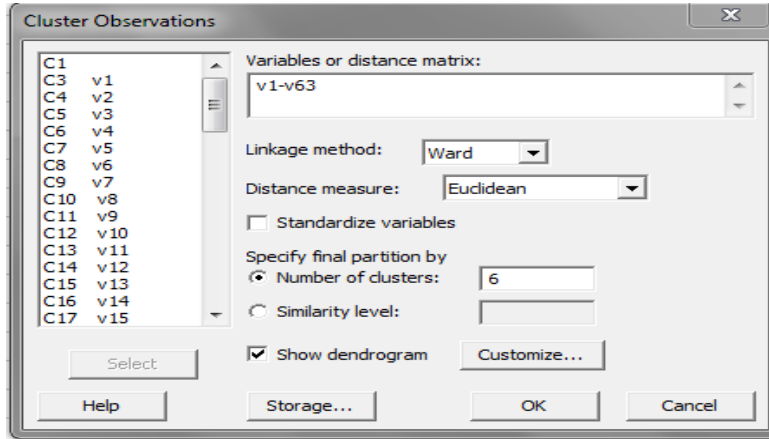
Figure 2.11 - Configuration for HCA.

LDA

Performing LDA was different as compared to PCA, HCA since the predictors from "v1-v63" were highly correlated to each other. So the result data from different experiments were combined to make the sample size larger and calculate new predictors which do not correlate to each other, so LDA could be performed. To perform LDA, click "Stat", then "Multivariate" and select "Discriminant Analysis" in the options menu. The settings used are shown in Figure 2.12. Scikit library for data computation in python is also used for plotting the LDA results as minitab does not give a graphical output of the discriminant analysis.
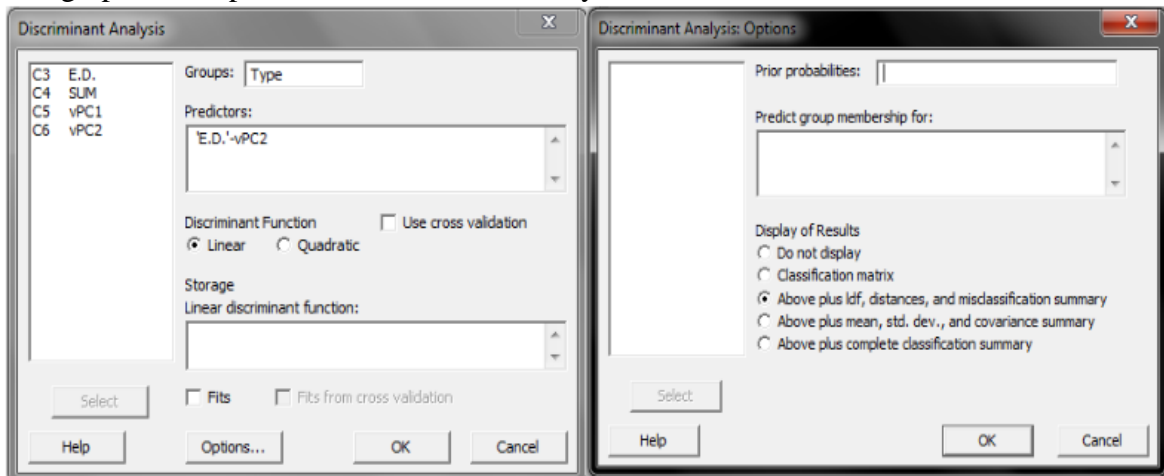
Figure 2.12 - Discriminant Analysis configuration menu used for linear discriminant analysis

Chapter 3

3. Results

This chapter includes details about the output results obtained from experiments performed. Details about the experiments and methodology are discussed in sections 2.4 & 2.5. Results are presented in the form of plots showing different chemicals used and the RGB difference values obtained from the python script.

3.1.   Raw Data Acquisition

The sensor system's working is discussed in detail in sections 2.3 & 2.4. The camera captures the images, and the python script calculates the RGB values of the dye spot pixels from the taken pictures. The idea is that the before exposure image RGB colors will be different as compared to after exposure RGB colors, and the python program calculates the RGB values. Even though the camera's resolution is high for good quality pictures and a light source is placed for illumination, the images' colors are different from what the naked eye sees. This is a limitation of the camera and due to the color of the light source. The main point here is to see any visual color change in the colorimetric array due to the VFAs, although the color change with the naked eye was not observable during the experiment. To see the color difference map based on the readings from experiments, the RGB values of one reading is taken for depiction. Figure 3.1 shows the colorimetric array from the experiments and the color difference map obtained from RGB values calculated from the python program. The RGB values are also given in 'appendix B' for this example. The sensor system is designed to detect the color change of dyes which is working as expected.

| Before Exposure | After Exposure | Difference Map |
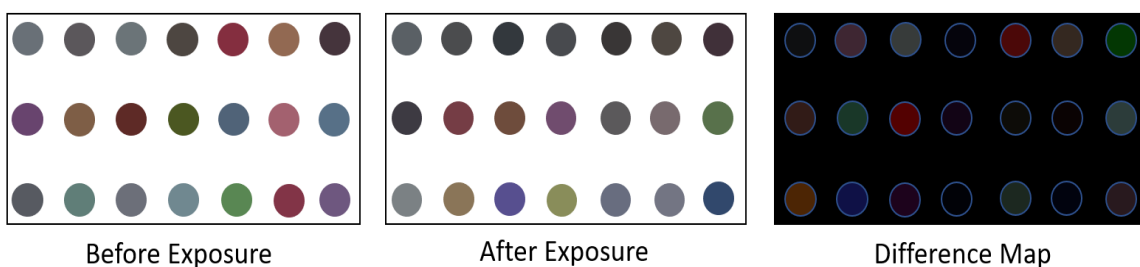
Figure 3.1 – Representation of colorimetric array 3x7 dyes which change color before and after exposure to the VFAs and RGB color difference because of exposure.

3.2. Sensitivity Test Results

The 30mM acetic acid solution was tested to examine the sensor array behavior in short time periods and longer time periods. The resting time of the substance in the dye array has a time gap of 10 minutes to allow equilibrium of the dyes. Figure 3.2 shows the sensor behavior over the time of 20 consecutive readings done and how its response changes when a sample is tested multiple times. Odd numbers on the x-axis are values from water rested for 10 minutes in the sensor container, and even numbers are from 30mM acetic acid rested for 10 minutes in the array. As seen in the graph's first few readings, the sensor array detects changes from a color change in dyes. The Euclidean distance (ED) reduces when tested with an acid solution, whereas ED rises when water is resting inside the array. But a higher number of readings show that the sensitivity reduces when more and more readings are taken, e.g., from 'Reading Number 13', rise and fall are not high, showing a reduction in sensitivity.
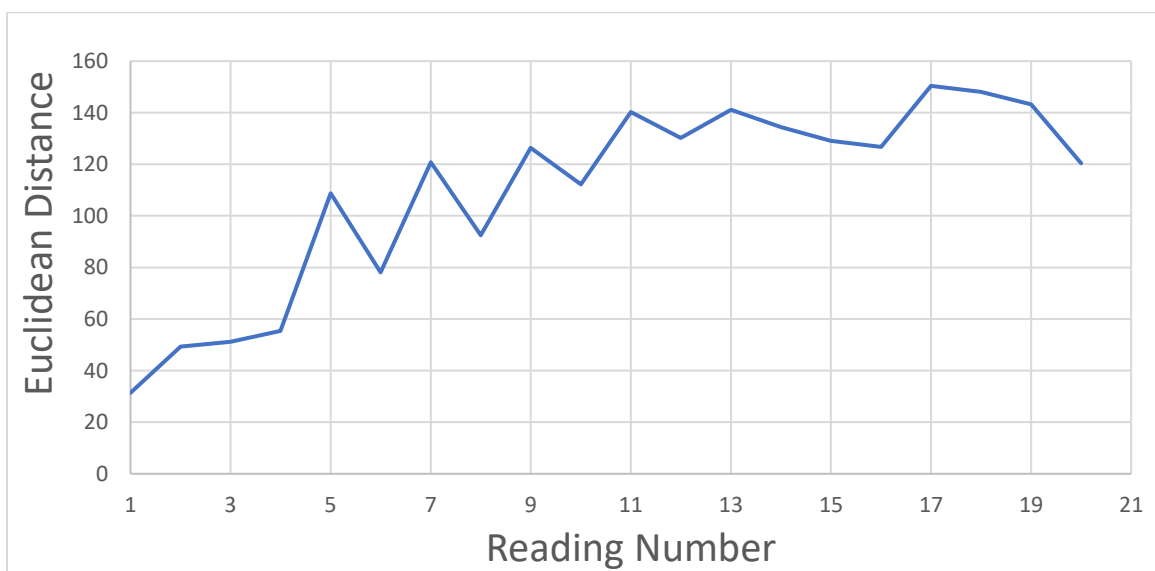
Figure 3.2 - Sensitivity trend of the sensor resulting from alteration between 30mM acetic acid and water. Plot shows variation in ED. Odd reading number (1,3,5,…) represents water resting while even reading numbers (2,4,6,…) represent acetic acid solution resting in sensor array container. Equilibrium time of the dyes for acetic acid and water was 10 minutes.

The second experiment of the sensitivity test is performed over a long period of almost a week, was also done to see how sensitivity is affected and how often new array shows are made for detection. The equilibrium time was 10 minutes same as the first experiment.

A new dye array was made, and first readings 1 to 10 were taken in a newly made array. After a period of 24 hours, readings 11 to 40 were taken and readings 41 to 50 after 5 days. The same

sensor array was used for all 50 readings, which means that by the end of the last readings, dyes were almost *"week old"*. Figure 3.3 shows sensitivity variation over time.

In the plot, it is clear that the array maintains good sensitivity when it is newly made and over 24 hours period when water is left rested in it. But clearly, readings from 41 to 50 sensitivity are significantly reduced, which means it becomes unreliable, and a new array should be made.

Both the sensitivity tests show that the dyes array is able to detect acid and water, shown by changing ED, and sensitivity reduces with a greater number of readings taken and array left for long time periods. However, ED gives no information about the acid group or concentration. It just measures the change in RGB color caused by altercating substances in the sensor array.

Figure 3.3 - Sensitivity trend of the sensor resulting from alteration between 30mM acetic acid and water. Plot shows variation in ED as well as sensitivity change over time, odd reading number (1,3,5,…) represent water resting while even reading numbers (2,4,6,…) represent acetic acid solution resting in sensor array container.

3.3.   Concentration Test Results

The acid concentration ability of the sensor was examined in this experiment. Because in the AD process, the variation is highly likely, so the possible concentration ranges during the AD process are tested in this experiment for all three VFAs.

Firstly, acetic acid with multiple concentrations is investigated. Starting from the lowest concentration, as seen in Figure 3.4 that ED changes at 0.1mM level when water is replaced with acid. This shows that this sensor is able to detect a low level of acetic acid. Further higher levels are depicted in the plot. It is essential to state that the last solution of 1mM is assessed again after

5mM to check whether the sensor is reversible (can sense lower concentrations again) or new dyes array is required to detect lower concentration levels.

The regression analysis between ED values of acetic acid and the various concentrations values like 0.1, 0.5, etc., was done. 'ED' was a "dependent" variable in this case as it represents RGB color change, and the independent variable was the corresponding "acid concentration" value. The ED value of "water" readings was not used for regression analysis just the ED value of a respective acid concentration. The output of the analysis is as follows:

$$R\ Square = 0.6608$$

R square is also known as 'coefficient of determination,' which assesses how accurately independent variables explain and predict future outcomes for a dependent variable. It analyzes how a difference in one variable can explain the difference in another variable. R square ranges between 0 to 1. '0.6608' indicates a good correlation between the dependent and independent variables meaning that ED change is approximately linear (66.08%) with respect to increasing concentration.

The line equation (given in section 1.8.2) of acetic acid regression analysis is:

$$y = 9.18x + 91.14$$
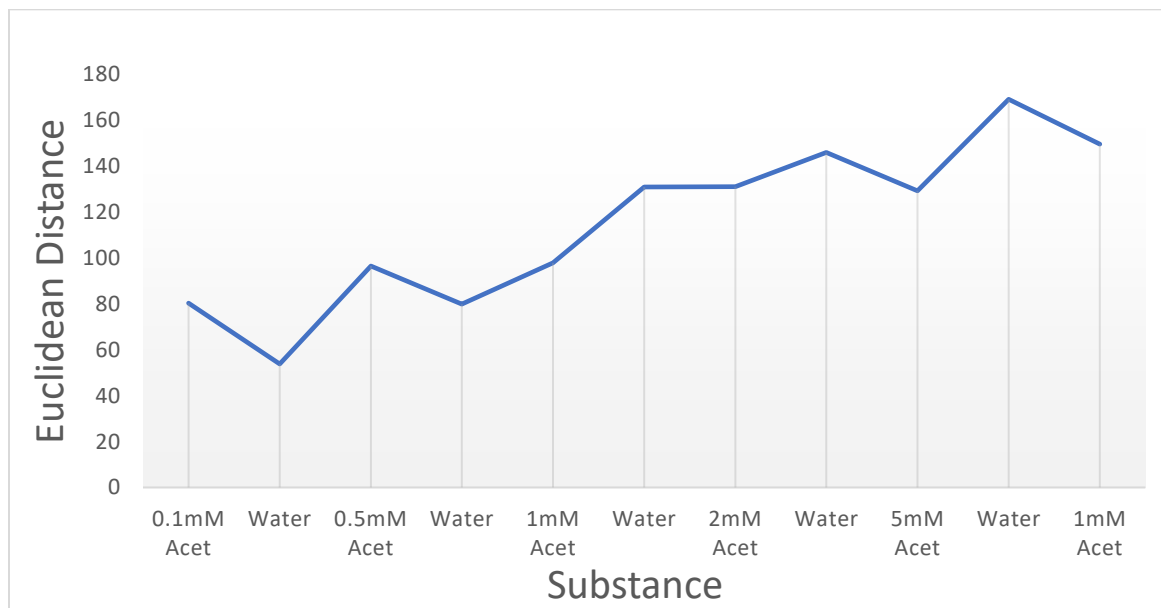
Figure 3.4 - Sensor array response to Acetic acid concentration test.

The second concentration test was done for propionic acid. The same protocol is followed by making a new array and test starting from lower level to higher level of propionic acid. Figure 3.5

41

shows that although the sensor is able to sense the difference between water and propionic acid for 0.5mM and 1mM line is almost straight without any sharp edge, which shows it is not reversible at low concentrations.

Similarly, regression analysis between ED values and corresponding acid concentration values for propionic acid is done. The output results are as follows:

$$R \text{ Square} = 0.5389$$

0.5389 shows medium correlation. The line equation for propionic acid is:

$$y = 6.22x + 95.16$$
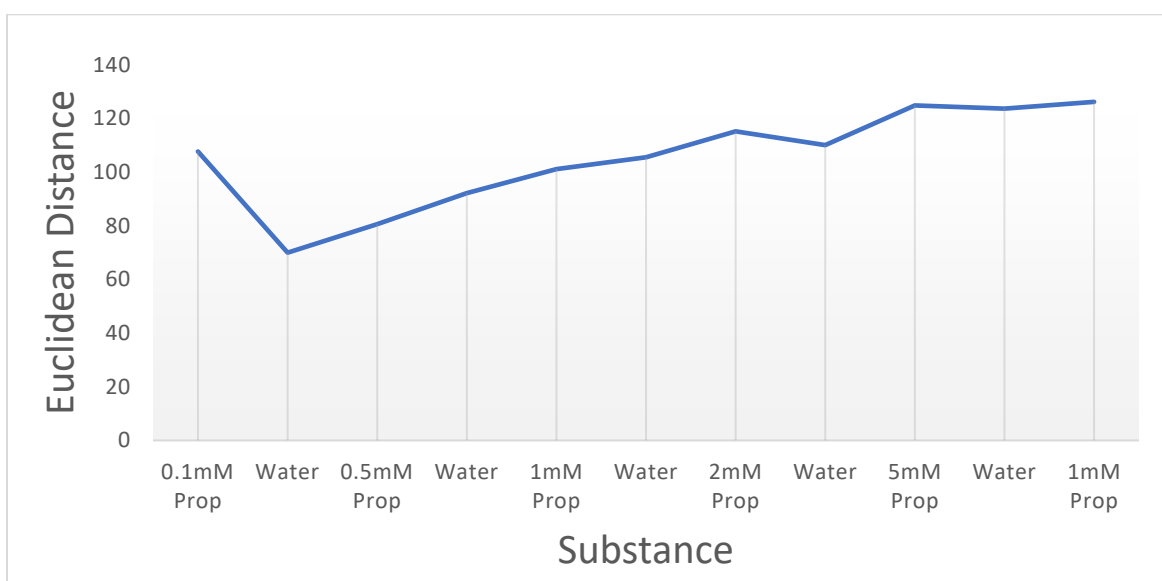
Figure 3.5 - Sensor array response to Propionic acid concentration test.

There is a small rise and fall in the trend line at the last 1mM reading after 5mM showing more reversibility at high concentration.

The last part of the concentration experiment is done with butyric acid. The test was conducted with the same concentrations with a newly made dyes array and following the same steps. Butyric acid, like acetic acid, has a reversible response and shows clear peaks when a substance is changed compared to propionic acid, which had a comparatively flat plotline.

The results of regression analysis carried out for butyric acid are given below:

$$R \text{ Square} = 0.8750$$

'0.8750' indicates a strong correlation between ED and concentration values of butyric acid, meaning change in ED and concentration is at an almost constant rate (linear).

Line equation for butyric acid:

$$y = 23.80x + 63.11$$
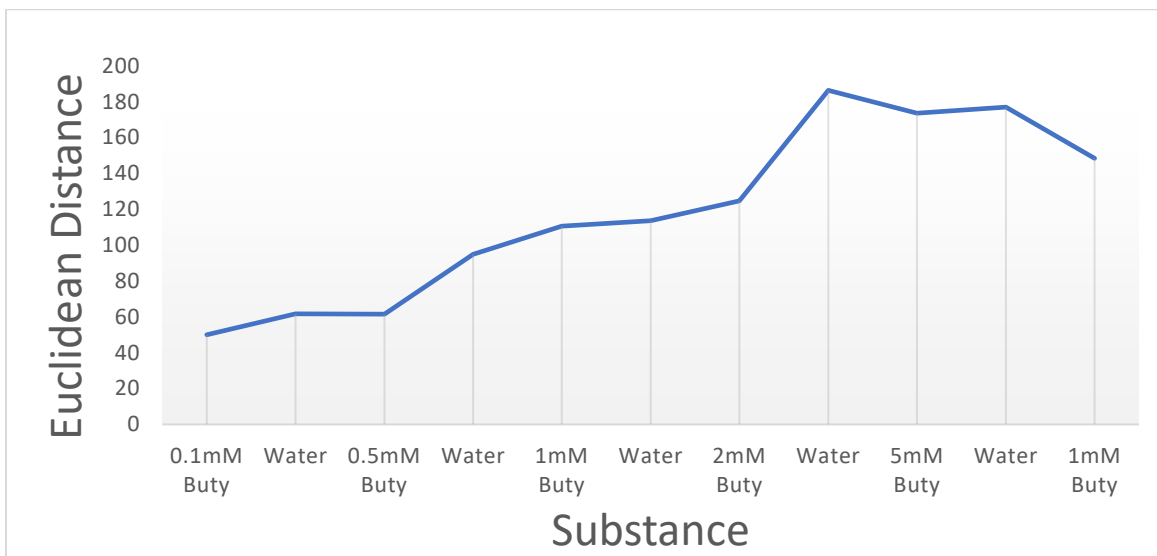
Figure 3.6 - Sensor array response to Butyric acid concentration test.

This whole concentration test gives the idea about the levels which the sensor array can detect. When testing in the field, the concentration range would already be determined, and all the examined values are those that can be expected during the AD process; and this experiment proves that our sensor can sense all possible ranges in AD.

It is important to state here that the ED only tells us that the colors are changing and nothing else. An increase or decrease in ED for acid concentrations is not constant or deterministic. It is variable as seen in Figure 3.4 for water in the first 2 reading ED decreases, and then for the remaining readings when water is added, ED increases. It gives no useful interpretation or information about concentration or classification of the VFAs it just 'measures' change of dye colors in response to change in substance in dye array. Detection and classification of the VFAs is performed with PCA, HCA, and LDA. Therefore, data computation tool is used to classify and group the acid mixtures.

A score plot from PCA is shown in Figure 3.7, displaying all the combinations of acids used in the experiment. The picture of scattered groups gives an idea about the quadrant where each acid group lies after performing PCA. For example, butyric acid in its lower concentration is just in the 4th quadrant close to the y-axis, but with higher concentrations, it moves further to the left into the 3rd quadrant. For the remaining two acids trend with changing levels can be observed in the plot. This result is helpful to observe a pattern in the results of experiments when VFAs are mixed in different proportions to understand a general behavior.
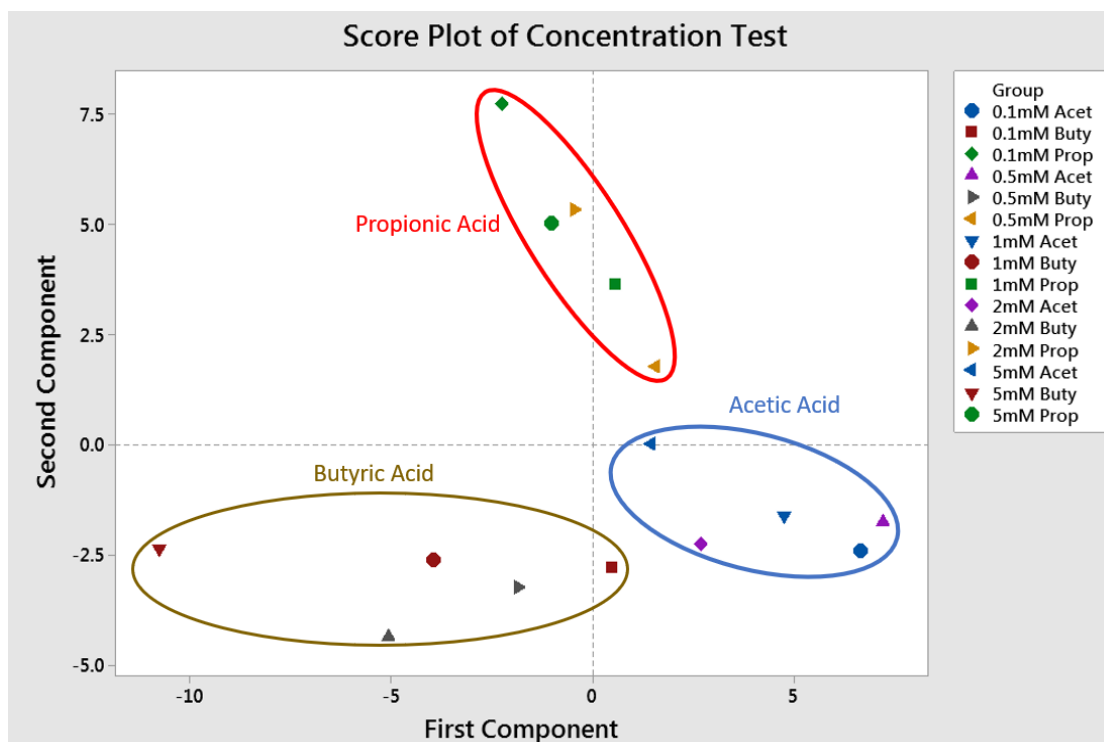
Figure 3.7 - Score plot showing quadrants of individual acids.

3.4. Mixed VFAs in Water Results

After investigating the sensor's sensitivity, this experiment is done to check the selectivity whether the sensor can distinguish among the multiple solutions of VFAs having multiple concentrations. The experiment details are already discussed in section 2.5.3. Finally, the output data is computed with PCA, HCA to differentiate among the solutions under examination. This experiment is done in water, i.e., 70mM acetic, 1mM propionic, etc., all solutions are made in water. The data obtained from experiments is then used as input to perform data analysis using software "Minitab". The results from the data analysis are given in the Figures below. The score plot shows the scatter of acid solution and water according to their concentrations, and a dendrogram depicts similarity percentages among different groups. The results show that the sensor can distinguish among various VFA mixtures; hence the selectivity of this sensor can be used in the application of the AD process. It is evident from the plots that propionic acid has no significant effect in solutions which is indicated by the grouping (e.g., 70mM Acet, 20mM Buty, 10mM Prop; 70mM Acet, 10mM Buty, 20mM Prop) where Acetic & butyric concentration level is high fall under the same group.

44
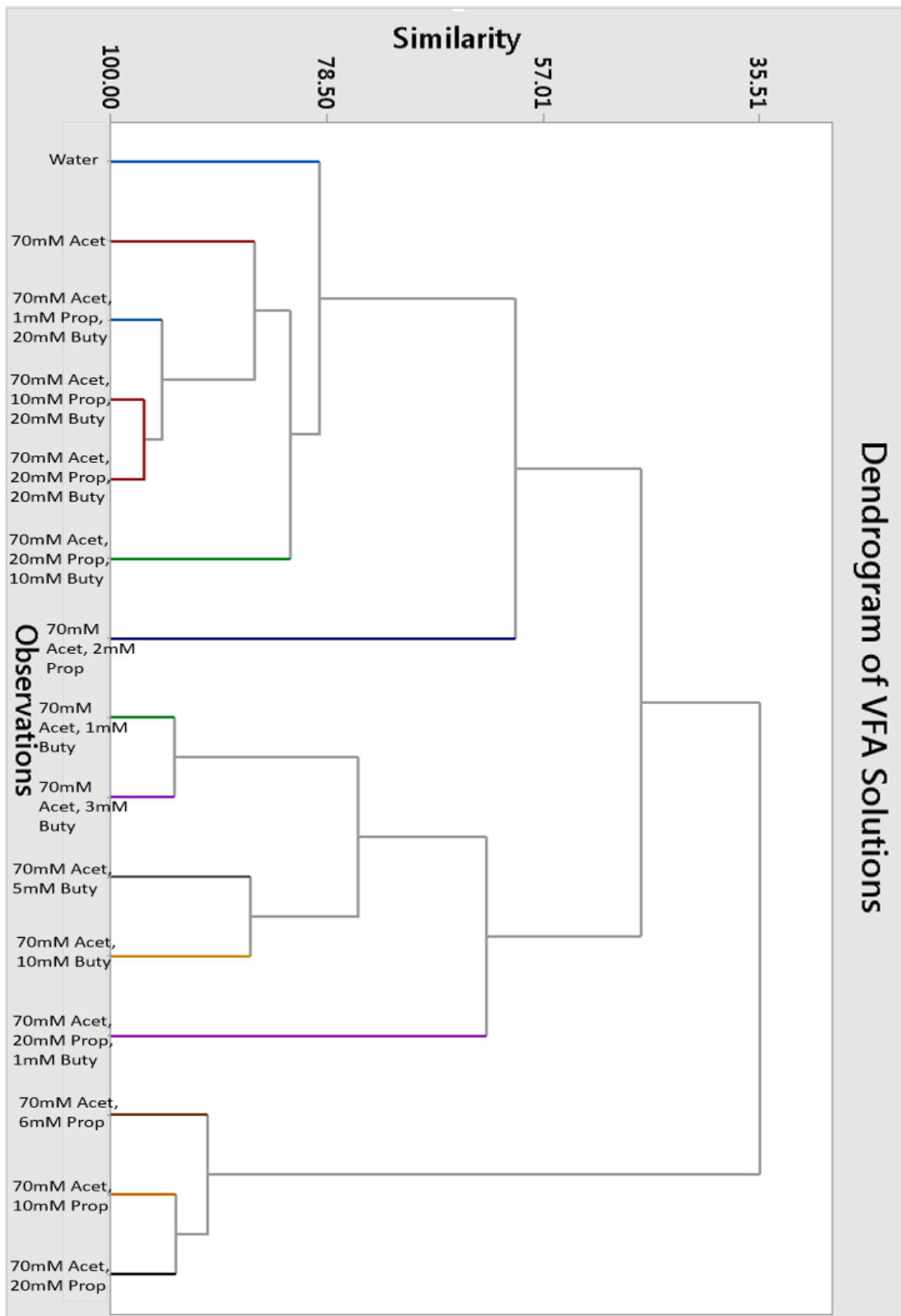
Figure 3.8 - Dendrogram obtained from HCA showing similarity between solutions in water. Diagram shows that no two solutions are 100% similar, which means the sensor can differentiate between different concentrations and mixtures of chemical solutions. Thus, the sensor has desired selectivity for the project.

The (70mM Acet, 1mM Buty, 20mM Prop) mixture is more similar to the Acetic & Butyric mixture, which shows that even though Propionic acid has a higher level of 20mM, its effect in solution is less, making this combination closer to pair of Acetic & Butyric solution shown in the graph. Grouping of the solutions with respect to the VFAs mixture is shown in Figure 3.9 in colors. This could possibly mean that propionic acid has less impact and hence lower detection by the sensor.
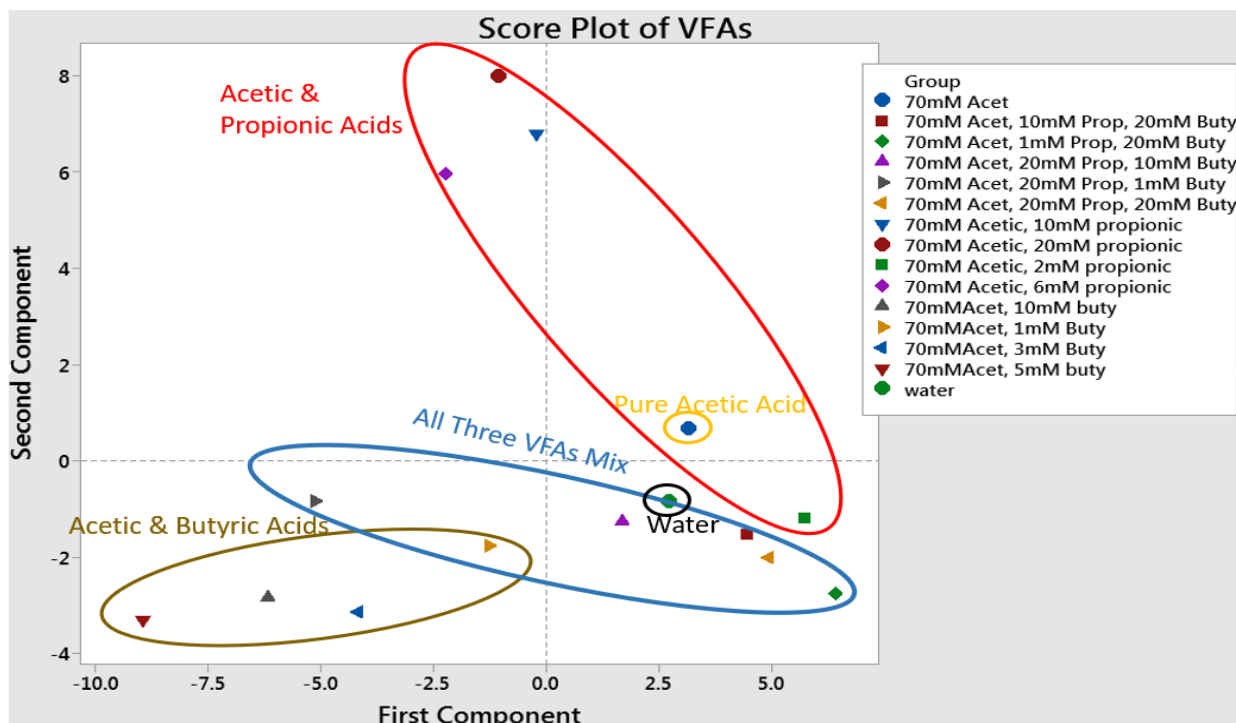
Figure 3.9 - Score Plot of VFAs in water from PCA showing scattered VFA solutions with different concentrations

3.5. Mixed VFAs in Digestate Results

This experiment investigates the performance of the dyes sensor in digestate extracted from the digester to study its behavior for real conditions. So, all the chemical solutions were prepared in digestate, a thick black liquid. Section 2.5.4 gives details about how the experiment was performed and information about VFAs. The output is provided by a score plot and dendrogram like the last experiment. Analysis results are given in the Figures below.
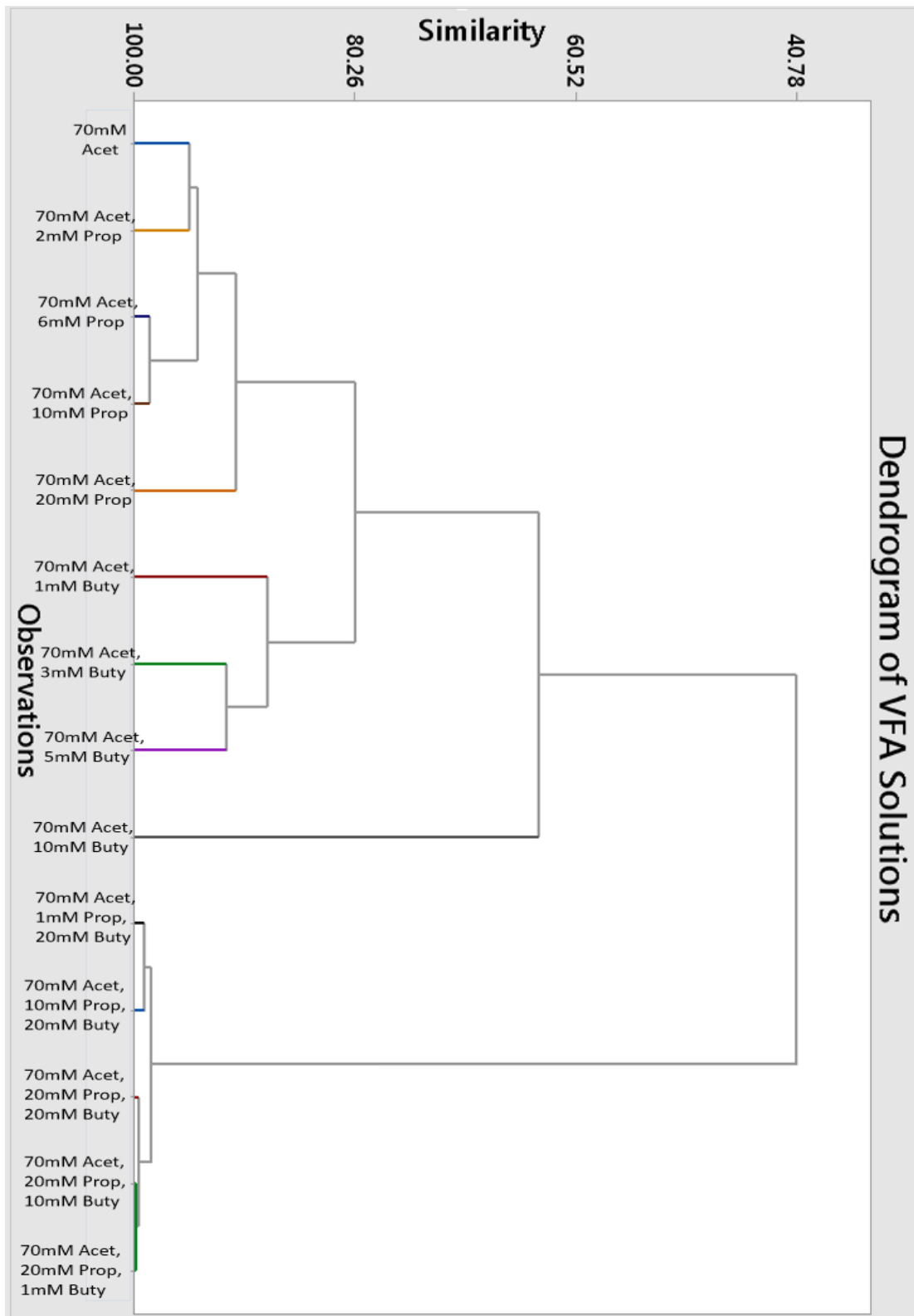
Figure 3.10 - Dendrogram from HCA showing similarity between different VFAs mixtures in digestate. The sensor can differentiate between different concentrations and mixtures of chemical solutions. Thus, the sensor has desired selectivity in digestate.

A similarity percentage is shown in Figure 3.10. It can be seen that the similarity level between some clusters is very high, especially for the solutions where all VFAs are mixed, e.g. (70mM Acet, 20mM Prop, 10mM Buty), (70mM Acet, 20mM Prop, 1mM Buty), etc. which show around 98% similarity. No two clusters are 100% the same, which means that in digestate, the dyes array can distinguish between different chemicals making the sensor's application more practical. The similarity is also reflected in the score plot from PCA shown in Figure 3.11, indicating the groups are closer to each other but not overlapping, meaning the sensor offers both selectivity & sensitivity when used with the digestate.
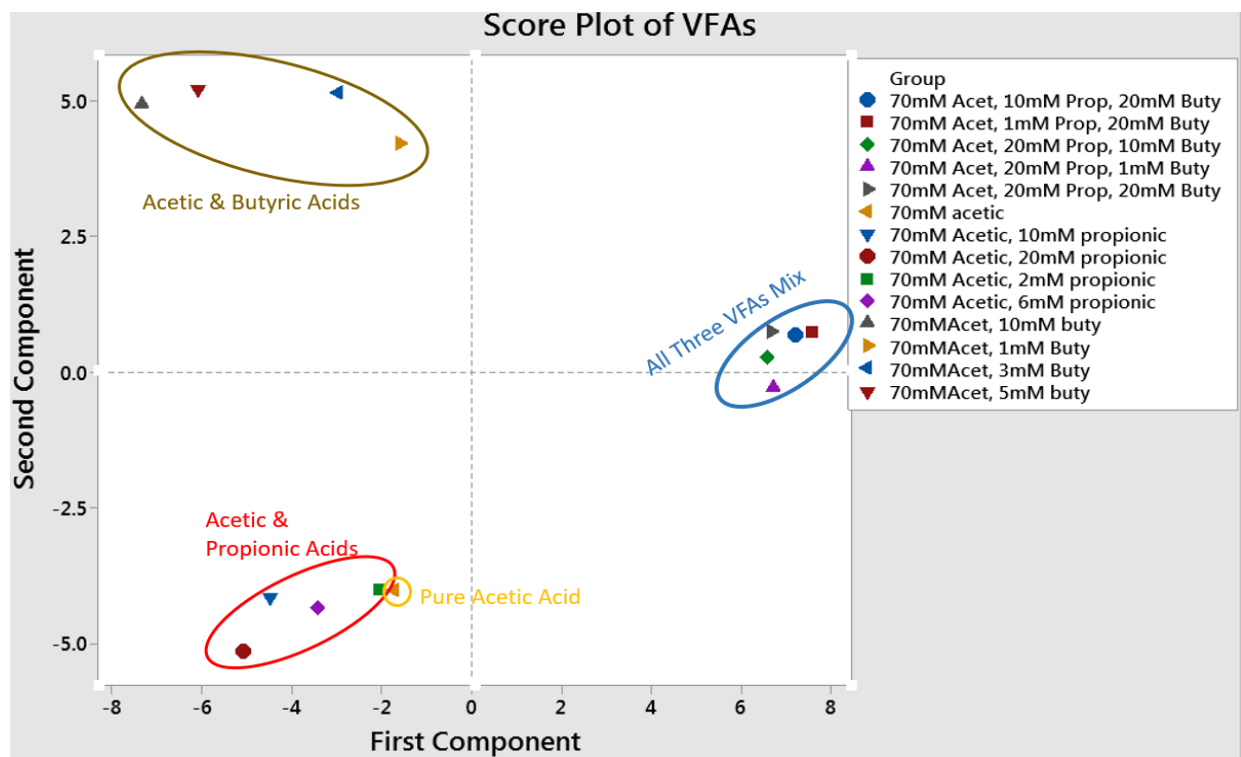
Figure 3.11 - Score plot from PCA of VFAs in digestate showing scattered VFA solutions with different concentrations.

3.6. LDA Results

It is observed in results from PCA that all the components, i.e., VFA solutions with various concentrations, are adequately separated in the score plot. LDA is done to verify the results further, even though there is good separation among principal components. The data values from the same group, e.g., RGB numbers of acetic acid from various concentration levels, were merged to increase the sample size. The different acid groups are separated but not entirely, as shown in Figure 3.12. Although, some of the points from groups like acetic acid at the bottom right corner (in navy blue color) and the mixture of all three VFAs at the top left corner (in purple color) have good separation. Likewise, few points of other groups are also completely distinguishable, like

48

Acet & Prop group at some points. Overlapping and congestion occur with groups of mixed VFAs like Acet & Prop or Acet & Buty when they are tested in various concentrations. The elements are not entirely different, i.e., they have the somewhat same effect on dye array but not wholly the same. Therefore, points are overlapping or closer to components of other groups.

LDA plot obtained from python script shows the discriminant components, which are in some instances close to or overlapping the components of other groups due to low dimension sample size hence less accurate LDA results. Table 7 shows the squared distance (ED) between one group center to another group center. The distance shows how different the groups are compared to each other; acid groups distances are greater than '0' for all groups (except the group itself), meaning that groups are different from each other, and through the discriminant analysis, we can distinguish between groups. It can be stated based on obtained results that although LDA does not give accurate results compared to PCA, HCA in this case but Linear Discriminant Analysis can differentiate between groups of acids, implying that the sensor system can also distinguish among the substance involved in the AD process.
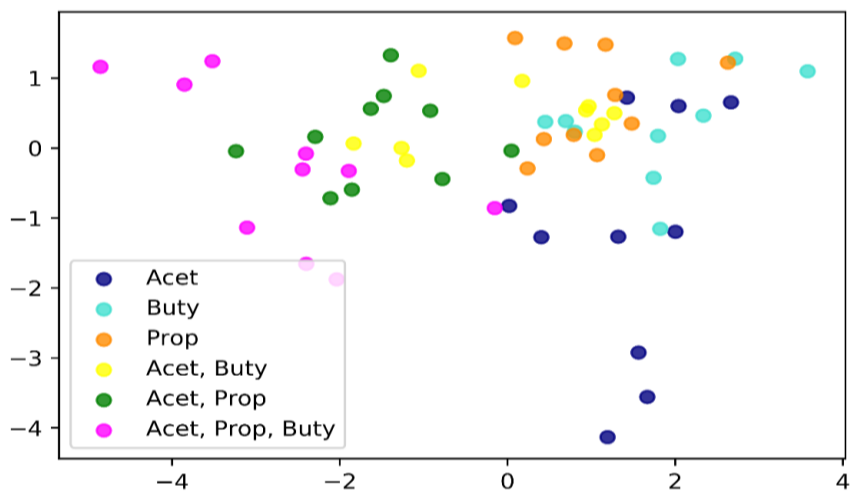
Figure 3.12 - LDA plot of the combined dataset values of multiple experiments.

Table 7 - Euclidean distance between acid groups after performing LDA.

	Acet	**Acet, Buty**	**Acet, Prop**	**Acet, Prop, Buty**	**Buty**	**Prop**
Acet	0.0000	21.6826	7.0690	16.9281	16.1995	12.0543
Acet, Buty	21.6826	0.0000	27.3282	8.1226	5.3749	29.1966
Acet, Prop	7.0690	27.3282	0.0000	14.0734	32.7261	33.7539
Acet, Prop, Buty	16.9281	8.1226	14.0734	0.0000	17.1813	42.0507
Buty	16.1995	5.3749	32.7261	17.1813	0.0000	13.9844
Prop	12.0543	29.1966	33.7539	42.0507	13.9844	0.0000

Chapter 4

4. Discussion

This project aimed to develop a chemical sensor based on the optoelectronic nose concept, mimicking the human olfactory system, to identify the different chemicals in the AD process. The sensor should be robust to perform within harsh conditions of the AD process, and it should have high sensitivity and selectivity (i.e., it can differentiate among different types of substances in their low and high concentrations). In other words, it is similar to a nose which can tell the difference between various smells, whether it is a strong smell or a faint smell. It has been proven with experimental observations that a very low concentration of VFAs is detectable by the sensor array, and it also provides good selectivity.

The sensor is divided into two parts:

1. Designing the experimental setup for stable and high-quality imaging of dye array. Selecting the appropriate dyes that can work with digestate and water and prepare dyes for every new experiment. Making sure the location/pixel of the dyes remain unchanged for RGB difference calculation of before and after images.
2. Making VFA solutions to perform experiments, collect the raw data, and then use it to investigate the sensor's ability, whether it has desired sensitivity and selectivity. Since the sensor has no moving parts and is only a chemical sensor, it can work in harsh conditions of the digester.

Since imaging is an important part of the project, it was crucial that images were clear and stable because all the analysis was based on image difference in RGB colors. The experimental setup consisting of a wooden box with a fixed digital camera and light source incident on the array was constructed to ensure image quality was not compromised.

The difference of the images is computed by using color subtraction of RGB from the new picture and the reference picture. Therefore, to keep the comparison valid, it is imperative that each dye in an array has the same placement while imaging before and after pictures. Misalignment could lead to inaccurate results since the difference image data would be of two different dyes rather than the same one, or the light interacting with the dye may change if the array is slightly moved. The color change of the dyes is not entirely observable through the naked eye. The [x, y] pixel coordinates of the dyes were given input to the script to tell the program where the dyes are located in the glass dye array container. Coordinates shift a little each time it was removed to make new arrays, so it was updated in the script at the start of every experiment.

The RGB difference data is not able to be interpreted alone since it seems a vector of random numbers. Therefore, statistical analysis is performed to analyze the difference data. The score plots of PCA from all experiments show the separation between the acid groups and also the separation between the concentrations of the VFA mixtures. This means that sensor is able to sense and differentiate between VFAs in the biogas AD process having several concentrations. This result is backed by the HCA results, which clearly show that the sensor can differentiate among the VFAs, and that is why no two observations in the HCA plots are 100% similar. Hence it can be concluded that the sensor is applicable for biogas production monitoring. The concentration test in Figures 3.4, 3.5, 3.6 shows that the sensor is sensitive towards lower concentrations and therefore provides an early indication for any impending failure.

Since LDA requires a very high number of samples for each group of VFAs, the LDA analysis was performed by combining the data from the same mixture or pair of VFAs, (e.g., all concentration levels of acetic & propionic) to classify based on the chemical or group of acids. Therefore, the only classification of the main chemical group was performed using LDA. Section 1.8.3 discusses that to get a dependable result, 100 times greater sample size is needed for every individual concentration level of VFAs. Because only 21 dyes are used with a 21x3 output sample size, the sample size is less than required to distinguish every single concentration of VFAs using LDA as done for HCA and PCA.

Chapter 5

5. Future Work

Experiments were carried out to prove that the sensor is able to give desired sensitivity and selectivity. Removing the glass container to add new dyes was time-consuming and required human intervention. Setup could be made such that it is easier to wash out used dyes and add new ones without having to remove the glass container. For example, a pre-made array could reduce the time required in changing dyes, allowing the straightforward replacement of the old array with a new array. The glass container was airtight to prevent leaking of sample or water, but sometimes due to warm water temperature, the glass got condensation on the inside, compromising the image quality of few pictures. The colorimetric array should be such that the picture of the dyes can be captured in the true color to get more accurate results.

A better, improved gel can be used for holding the dyes. It will also evenly spread the dye solution in the dye's well for interaction with VFAs or other substances. More experiments can be carried out to improve the sensor further so that the VFAs can also be quantified based on score plots of many similar experiments if they are consistent. High amounts of data from many experiments will also be useful to train a model for machine learning to make predictions, and computational analysis will give more accurate results. This will help in an understanding even better picture of the AD process.

Moreover, new dyes can be investigated for use in this experiment to obtain ever higher dimensional data, useful for machine learning, which will also help quantify different VFAs and increase and improve sensitivity and selectivity. The high dimension data will also enhance the accuracy of LDA, which requires high sample size data.

Chapter 6

6. Conclusion

This project investigated using the chemical sensor to detect VFAs in order to improve the biogas production process. This is in order to monitor the Anaerobic Digestion process by using the sensor as an early warning indicator to alert plant operators if the process starts to deteriorate so preventive measures can be taken in time. The chemical sensor is based on the optoelectronic nose concept and uses dyes for the detection of VFAs. A setup for capturing high-quality and stable images was built as image acquisition is a fundamental part of the sensor system. Images of the colorimetric array, which is made of many different dyes, are taken before and after the VFAs interact with the array. The interaction changes the RGB color of dyes. A python program controls the camera and other components of the setup. The python script calculates the RGB color values. Based on this difference of the RGB values in images, the data computation analysis is performed. All the sensor system components (i.e., the experimental setup, the colorimetric array, and the script) worked and performed as they were expected. The output from the experiment gives a visual color change, which can be seen through a difference map. The results from the experiments show that the use of the colorimetric sensor for assessing VFAs in anaerobic digestion is a viable method. The sensor was able to detect low concentration VFAs down to around one-tenth of the concentration at which inhibition of the AD process begins. Furthermore, the sensor can successfully classify among various VFAs, providing the desired selectivity. The sensor system was also assessed with the digestate sample from an anaerobic digester. The sensor is robust, has good selectivity, sensitivity and low cost which makes it better suited to be utilized for VFAs detection in the AD process. It performed as expected, implying that the sensor system is ready for industrial trials since all possible concentrations and substances to be expected during the AD process have been tested in the lab and found to be working.

7. References

[1]   Miljødirektoratet, "Klimakur 2030 - Tiltak og virkemidler mot 2030," *M-1625*, 2020.

[2]   R. K. Pachauri *et al.*, *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc, 2014.

[3]   UN, "Sustainable Development Goals," 2020. https://www.un.org/sustainabledevelopment/energy/ (accessed Oct. 19, 2020).

[4]   Government, "Norway Steps Up 2030 Climate Goal," 2020. www.regjeringen.no/en/aktuelt/norge-forsterker-klimamalet-for-2030-til-minst-50-prosent-og-opp-mot-55-prosent/id2689679/ (accessed Oct. 22, 2020).

[5]   J. J. Lamb, O. Bernard, S. Sarker, K. M. Lien, and D. R. Hjelme, "Perspectives of surface plasmon resonance sensors for optimized biogas methanation," *Eng. Life Sci.*, vol. 19, no. 11, pp. 759–769, 2019.

[6]   J. J. Lamb, "Impact Assessment," in *Anaerobic digestion - from biomass to biogas*, vol. 1, SCIO Publishing, 2020, pp. 391–414.

[7]   J. J. Lamb, "Uses of Biogas and Biomethane," in *Anaerobic digestion - from biomass to biogas*, vol. 1, SCIO Publishing, 2020, pp. 333–364.

[8]   T. Bridgwater, "Biomass for energy," *J. Sci. Food Agric.*, vol. 86, no. 12, pp. 1755–1768, 2006.

[9]   IEA, "Outlook for biogas and biomethane: Prospects for organic growth," Paris, 2020. [Online]. Available: https://www.iea.org/reports/outlook-for-biogas-and-biomethane-prospects-for-organic-growth.

[10]   T. Al Seadi, "Biogas handbook," vol. 1, 2008, [Online]. Available: https://portal.findresearcher.sdu.dk/en/publications/biogas-handbook.

[11]   J. J. Lamb, "Industrial Overview and Outlook," in *Anaerobic digestion - from biomass to biogas*, vol. 1, SCIO Publishing, 2020, pp. 415–436.

[12]   J. J. Lamb, "Overview of AD," in *Anaerobic digestion - from biomass to biogas*, vol. 1, SCIO Publishing, 2020, pp. 29–52.

[13]   A. J. Ward, P. J. Hobbs, P. J. Holliman, and D. L. Jones, "Optimisation of the anaerobic digestion of agricultural resources," *Bioresour. Technol.*, vol. 99, no. 17, pp. 7928–7940, 2008.

[14]   J. J. Lamb, "Biogas and the Energy Sector," *Anaerob. Dig. From Biomass to Biogas*, 2020.

[15]   M. Madsen, J. B. Holm-Nielsen, and K. H. Esbensen, "Monitoring of anaerobic digestion processes: A review perspective," *Renew. Sustain. energy Rev.*, vol. 15, no. 6, pp. 3141–3155, 2011.

[16] J. J. Lamb *et al.*, "Traditional routes for hydrogen production and carbon conversion," *Hydrog. Biomass Bioenergy Integr. Pathways Renew. Energy Appl.*, p. 21, 2020.

[17] J. J. Lamb, "Microbiology of AD," in *Anaerobic digestion - from biomass to biogas*, vol. 1, SCIO Publishing, 2020, pp. 53–88.

[18] J. J. Lamb, O. Bernard, S. Sarker, K. M. Lien, and D. R. Hjelme, "Perspectives of optical colourimetric sensors for anaerobic digestion," *Renew. Sustain. Energy Rev.*, vol. 111, pp. 87–96, 2019.

[19] A. M. Ziganshin, J. Liebetrau, J. Pröter, and S. Kleinsteuber, "Microbial community structure and dynamics during anaerobic digestion of various agricultural waste materials," *Appl. Microbiol. Biotechnol.*, vol. 97, no. 11, pp. 5161–5174, 2013.

[20] G. Tchobanoglous, *Wastewater Engineering: Treatment and Resource Recovery-Vol. 2*. McGraw-Hill, 2014.

[21] H. M. Falk, P. Reichling, C. Andersen, and R. Benz, "Online monitoring of concentration and dynamics of volatile fatty acids in anaerobic digestion processes with mid-infrared spectroscopy," *Bioprocess Biosyst. Eng.*, vol. 38, no. 2, pp. 237–249, 2015.

[22] P. Bajpai, "Process Parameters Affecting Anaerobic Digestion," in *Anaerobic Technology in Pulp and Paper Industry*, Springer, 2017, pp. 13–27.

[23] D. Bolzonella, F. Battista, C. Cavinato, M. Gottardo, F. Micolucci, and P. Pavan, "Biohythane Production From Food Wastes," 2019, pp. 347–368.

[24] B. K. Ahring, M. Sandberg, and I. Angelidaki, "Volatile fatty acids as indicators of process imbalance in anaerobic digestors," *Appl. Microbiol. Biotechnol.*, vol. 43, no. 3, pp. 559–565, 1995.

[25] J. J. Lamb, K. M. Lien, and D. R. Hjelme, "Digitalization of colourimetric sensor arrays for volatile fatty acid detection in anaerobic digestion," *MethodsX*, vol. 6, pp. 2584–2591, 2019.

[26] M. Moo-Young, *Comprehensive biotechnology*. Elsevier, 2011.

[27] O. Bernard *et al.*, "Advanced monitoring and control of anaerobic wastewater treatment plants: software sensors and controllers for an anaerobic digester," *Water Sci. Technol.*, vol. 43, no. 7, pp. 175–182, Apr. 2001, doi: 10.2166/wst.2001.0418.

[28] K. Boe, D. J. Batstone, J.-P. Steyer, and I. Angelidaki, "State indicators for monitoring the anaerobic digestion process," *Water Res.*, vol. 44, no. 20, pp. 5973–5980, 2010.

[29] T. Scheper *et al.*, "Bioanalytics: detailed insight into bioprocesses," *Anal. Chim. Acta*, vol. 400, no. 1–3, pp. 121–134, 1999.

[30] P. Biechele, C. Busse, D. Solle, T. Scheper, and K. Reardon, "Sensor systems for bioprocess monitoring," *Eng. Life Sci.*, vol. 15, no. 5, pp. 469–488, 2015.

[31] R. Belford, "Acid-Base Titrations." May 24, 2021, [Online]. Available: https://chem.libretexts.org/@go/page/60763.

[32] K. Boe, D. J. Batstone, and I. Angelidaki, "Online headspace chromatographic method for measuring VFA in biogas reactor," *Water Sci. Technol.*, vol. 52, no. 1–2, pp. 473–478, Jul. 2005, doi: 10.2166/wst.2005.0555.

[33] J. J. Lamb and B. G. Pollet, "Micro-Optics and Energy."

[34] Z. Li, J. R. Askim, and K. S. Suslick, "The optoelectronic nose: colorimetric and fluorometric sensor arrays," *Chem. Rev.*, vol. 119, no. 1, pp. 231–292, 2018.

[35] S. Salzberg, "Distance metrics for instance-based learning," in *International Symposium on Methodologies for Intelligent Systems*, 1991, pp. 399–408.

[36] J. R. Askim, M. Mahmoudi, and K. S. Suslick, "Optical sensor arrays for chemical sensing: the optoelectronic nose," *Chem. Soc. Rev.*, vol. 42, no. 22, pp. 8649–8682, 2013.

[37] Y. Mori, M. Kuroda, and N. Makino, *Nonlinear principal component analysis and its applications*. Springer, 2016.

[38] R. Asmatulu, "14 - Nanocoatings for corrosion protection of aerospace alloys," in *Woodhead Publishing Series in Metals and Surface Engineering*, V. S. Saji and R. B. T.-C. P. and C. U. N. Cook, Eds. Woodhead Publishing, 2012, pp. 357–374.

[39] Z.-H. Meng, "Chapter 13 - Molecularly Imprinted Sol-Gel Sensors," S. Li, Y. Ge, S. A. Piletsky, and J. B. T.-M. I. S. Lunec, Eds. Amsterdam: Elsevier, 2012, pp. 303–337.

[40] S. H. Lim, C. J. Musto, E. Park, W. Zhong, and K. S. Suslick, "A colorimetric sensor array for detection and identification of sugars," *Org. Lett.*, vol. 10, no. 20, p. 4405, 2008.

8. Appendices

Appendix A — Python Scripts

Appendix B — RGB values table for color difference map

Appendix C — Python script using scikit for LDA

8.1.   Appendix A
Reference.py

```
import os
import matplotlib
import matplotlib.pyplot as plt
import numpy as np
import datetime
import time

from PIL import Image
from pypylon import pylon

print("Taking reference image at:", datetime.datetime.now().time())

timeDistance = []

testtime = []

img = pylon.PylonImage()
tlf = pylon.TlFactory.GetInstance()

cam = pylon.InstantCamera(tlf.CreateFirstDevice())
cam.Open()
cam.StartGrabbing()

#exec(open("/home/pi/pysensor/Water.py").read())
#need to adjust below time for how long between reference and measure liquids
#time.sleep(10)

with cam.RetrieveResult(2000) as result:

    img.AttachGrabResultBuffer(result)

    filename = "/home/pi/Desktop/pysensor/Images/saved_img.tiff"
```

```python
        img.Save(pylon.ImageFileFormat_Tiff, filename)

        time.sleep(5)

        image1 = Image.open('/home/pi/Desktop/pysensor/Images/saved_img.tiff')
        image2 = image1.convert("RGB")
        image2.save('/home/pi/Desktop/pysensor/Images/reference.jpg', "JPEG", quality = 100)

        img.Release()

        time.sleep(1)

cam.StopGrabbing()
cam.Close()
print("Completed taking reference image at:", datetime.datetime.now().time())
```

Water.py

```python
import RPi.GPIO as GPIO
import os
import time
import datetime

GPIO.setmode(GPIO.BCM)
GPIO.setwarnings(False)
chan_list = [22,23,24,27]
GPIO.setup(chan_list, GPIO.OUT, initial=1)
#need to adjust below time for how long the liquid needs to refresh the chamber
pump = 1

print("Water pumping started at:", datetime.datetime.now().time())
#24 is the water solenoid
GPIO.output(24, 0)

time.sleep(1)
#22 is the water pump
GPIO.output(22, 0)

time.sleep(pump)

GPIO.output(22, 1)

time.sleep(1)
```

```
GPIO.output(24, 1)

print("Water pumping completed at:", datetime.datetime.now().time())
```

Sample.py

```
import RPi.GPIO as GPIO
import os
import time
import datetime

GPIO.setmode(GPIO.BCM)
GPIO.setwarnings(False)
chan_list = [22,23,24,27]
GPIO.setup(chan_list, GPIO.OUT, initial=1)
#need to adjust below time for how long the liquid needs to refresh the chamber
pump = 1

print("Sample pumping started at:", datetime.datetime.now().time())
#27 is the sample solenoid
GPIO.output(27, 0)

time.sleep(1)
#23 is the sample pump
GPIO.output(23, 0)

time.sleep(pump)

GPIO.output(23, 1)

time.sleep(1)

GPIO.output(27, 1)

print("Sample pumping completed at:", datetime.datetime.now().time())
```

Control.py

```
import os
import matplotlib
import matplotlib.pyplot as plt
```

```python
import numpy as np
import numpy
import datetime
import time

from skimage.draw import (circle)
from skimage import io
from scipy.spatial import distance
from PIL import Image

waittime = 1 #set the interval after water and sample scripts
waittime1 = 1 #set time between measurements (e.g., a minute or whatever).
#########Counter Code 1 ######
def counter_value(filename="counter.csv"):
    with open(filename, "a+") as f:
        f.seek(0)
        val = int(f.read() or 0) + 1
        f.seek(0)
        f.write(str(val))
        return val
counter = counter_value()
print("This script has been run", counter,"times.")
#####REST OF THE CODE
print("Taking sample image at:", datetime.datetime.now().time())

img = pylon.PylonImage()
tlf = pylon.TlFactory.GetInstance()
cam = pylon.InstantCamera(tlf.CreateFirstDevice())
cam.Open()
cam.StartGrabbingMax(1)
with cam.RetrieveResult(5000) as result:
        img.AttachGrabResultBuffer(result)
```

```python
time.sleep(1)

filename = "/home/pi/Desktop/pysensor/Images/array.Tiff"

img.Save(pylon.ImageFileFormat_Tiff, filename)

time.sleep(1)

image1 = Image.open('/home/pi/Desktop/pysensor/Images/array.Tiff')

image2 = image1.convert("RGB")

image2.save('/home/pi/Desktop/pysensor/Images/array_%d.jpg',counter,"JPEG", quality = 100)

img.Release()

print("Completed taking image at:", datetime.datetime.now().time())

time.sleep(1)

print("Measurement of reference RGB values started at:", datetime.datetime.now().time())

image = io.imread('/home/pi/Desktop/pysensor/Images/reference.jpg')

#the 7 x-coordinates of the dye spots

cox = [780,1200,1620,2050,2480,2910,3340]


#the 6 y-coordinates of the dye spots

coy = [360,795,1220,1660,2090,2520]

#before sample RGB values array

before = []

j = 0

l = 0

#square root of pixel area to be measured for each dye spot

pix = 20

def get_average_color(y, x, n, image):

    r, g, b = 0, 0, 0

    count1 = 0

    for s in range(y, y+n+1):

        for t in range(x, x+n+1):

            pixlr, pixlg, pixlb = image[s, t]

            r += pixlr

            g += pixlg

            b += pixlb
```

```python
        count1 += 1
    return ((r/count1), (g/count1), (b/count1))
for count in range(0, 42):
    r, g, b = get_average_color(coy[j], cox[l], pix, image)
    before.append(r)
    before.append(g)
    before.append(b)
    l = l + 1
    if l > 6:
        l = 0
        j = j + 1
        #if l > 5:
            #l = 0
counter = counter_value()
numpy.savetxt('/home/pi/Desktop/pysensor/Raw_data/Reference_RGB%d',counter,'.csv',before, delimiter=',')
print("Measurement of reference RGB values completed at:", datetime.datetime.now().time())
time.sleep(1)
print("Measurement of sample RGB values started at:", datetime.datetime.now().time())
image = io.imread('/home/pi/Desktop/pysensor/Images/array_%d.jpg' % counter)
#after sample RGB values array
after = []
j = 0
l = 0
#square root of pixel area to be measured for each dye spot
pix = 10
def get_average_color(y, x, n, image):
    r, g, b = 0, 0, 0
    count1 = 0
    for s in range(y, y+n+1):
        for t in range(x, x+n+1):
            pixlr, pixlg, pixlb = image[s, t]
            r += pixlr
```

```
        g += pixlg

        b += pixlb

        count1 += 1

    return ((r/count1), (g/count1), (b/count1))

for count in range(0, 42):

    r, g, b = get_average_color(coy[j], cox[l], pix, image)

    after.append(r)

    after.append(g)

    after.append(b)

    l = l + 1

    if l > 6:

        l = 0

        j = j + 1

        #if l > 5:

            #l = 0

numpy.savetxt('/home/pi/Desktop/pysensor/Raw_data/RAW_RGB_%d.csv' % counter, after, delimiter=',')

print("Measurement of sample RGB values completed at:", datetime.datetime.now().time())

print("Calculating RGB difference values started at:", datetime.datetime.now().time())

n = 0

befR = 0

aftR = 0

difR = 0

befG = 0

aftG = 0

difG = 0

befB = 0

aftB = 0

difB = 0

difference = []

for count in range (0,42):

    befR = before[n]

    aftR = after[n]
```

```
    difR = befR - aftR

    n += 1

    difference.append(difR)

    befG = before[n]

    aftG = after[n]

    difG = befG-aftG

    n += 1

    difference.append(difG)

    befB = before[n]

    aftB = after[n]

    difB = befB-aftB

    n += 1

    difference.append(difB)

numpy.savetxt('/home/pi/Desktop/pysensor/Raw_data/Difference_RGB_%d.csv' % counter, difference,
delimiter=',')

print("Calculating RGB difference values completed at:", datetime.datetime.now().time())

print("Drawing before, after and difference arrays at:", datetime.datetime.now().time())

n = 0

cox = 0

coy = 100

fig, (ax1, ax2, ax3) = plt.subplots(ncols=3, nrows=1, figsize=(10, 6))

imgB = np.full((700, 800, 3), 0.8, dtype=np.double)

imgA = np.full((700, 800, 3), 0.8, dtype=np.double)

imgD = np.full((700, 800, 3), 0.2, dtype=np.double)

for count in range(0, 42):

    difR = difference[n]

    if difR < 0:

        difR = difR * -1

    n += 1

    difG = difference[n]

    if difG < 0:

        difG = difG * -1
```

```
    n += 1

    difB = difference[n]

    if difB < 0:

        difB = difB * -1

    n += 1

    cox = cox + 100

    if cox > 750:

        cox = 100

        coy = coy + 100

    if coy > 650:

        coy = 100

    difR = difR/255

    difG = difG/255

    difB = difB/255

    rr, cc = circle(coy, cox, 30, imgD.shape)

    imgD[rr, cc, :] = (difR, difG, difB)

    if n == 126:

        n = 0

for count in range(0, 42):

    befR = before[n]

    if befR < 0:

        befR = befR * -1

    n += 1

    befG = before[n]

    if befG < 0:

        befG = befG * -1

    n += 1

    befB = before[n]

    if befB < 0:

        befB = befB * -1

    n += 1

    cox = cox + 100
```

```
    if cox > 750:

        cox = 100

        coy = coy + 100

    if coy > 650:

        coy = 100

    befR = befR/255

    befG = befG/255

    befB = befB/255

    rr, cc = circle(coy, cox, 30, imgB.shape)

    imgB[rr, cc, :] = (befR, befG, befB)

    if n == 126:

        n = 0


for count in range(0, 42):

    aftR = after[n]

    if aftR < 0:

        aftR = aftR * -1

    n += 1

    aftG = after[n]

    if aftG < 0:

        aftG = aftG * -1

    n += 1

    aftB = after[n]

    if aftB < 0:

        aftB = aftB * -1

    n += 1

    cox = cox + 100

    if cox > 750:

        cox = 100

        coy = coy + 100

    if coy > 650:

        coy = 100
```

```python
        aftR = aftR/255

        aftG = aftG/255

        aftB = aftB/255

        rr, cc = circle(coy, cox, 30, imgA.shape)

        imgA[rr, cc, :] = (aftR, aftG, aftB)

        if n == 126:

            n = 0

    ax1.imshow(imgB)

    ax1.set_title('Before')

    ax1.axis('off')

    ax2.imshow(imgA)

    ax2.set_title('After')

    ax2.axis('off')

    ax3.imshow(imgD)

    ax3.set_title('Difference')

    ax3.axis('off')

    filename = "/home/pi/Desktop/pysensor/Images/Finger_Print/finger_print_%d.png" % counter

    plt.savefig(fname=filename, format='png')

    print("Completed drawing before, after and difference arrays at:", datetime.datetime.now().time())

    time.sleep(1)

cam.StopGrabbing()

cam.Close()

time.sleep(waittime1)

 cam.Close()
```

8.2.   Appendix B

RGB values (21x3 values) of one reading obtained during experiment shown before & after value changes used for drawing difference map in Figure 3.1

Reference RGB Values	After Exposure RGB Values	Difference RGB Values
105.05	90.75	14.293
112.54	96.68	15.867
119.45	101.74	17.718
91.69	75.54	16.154
87.48	76.59	10.894
91.76	78.75	13.003
107.35	51.47	55.883
116.18	56.57	59.607
120.91	61.94	58.972
77.80	72.34	5.457
70.15	74.25	4.101
65.63	78.28	12.655
132.97	57.07	75.896
46.02	54.75	8.727
63.54	54.61	8.924
130.88	78.68	52.200
112.05	71.22	40.829
98.10	65.47	32.626
66.71	64.28	2.424
52.83	48.05	4.780
59.69	57.49	2.204
93.19	61.07	32.116
84.95	58.67	26.281
92.22	66.33	25.885
117.58	100.63	16.950
61.61	97.87	36.258
69.68	95.85	26.176
110.45	75.03	35.416
76.05	62.14	13.914
60.82	56.05	4.774
82.39	64.13	18.258
64.39	59.39	4.997
77.99	56.64	21.344
91.11	77.79	13.321
89.76	77.61	12.144

91.72	83.28	8.438
107.60	97.44	10.161
94.37	90.45	3.913
98.59	95.29	3.303
83.19	90.24	7.054
85.93	93.47	7.541
82.27	94.32	12.052
49.91	89.27	39.361
72.27	91.32	19.050
108.85	106.54	2.318
123.74	108.69	15.047
129.62	112.59	17.035
132.49	121.74	10.752
128.61	99.16	29.448
108.94	105.22	3.713
81.60	109.91	28.310
109.37	108.54	0.832
113.49	111.94	1.550
128.36	121.12	7.247
137.63	108.96	28.669
141.44	109.40	32.033
147.57	116.07	31.493
104.73	106.21	1.487
109.46	112.95	3.495
127.27	112.69	14.578
115.66	84.79	30.877
117.47	88.16	29.317
131.12	98.23	32.891

8.3.   Appendix C

import matplotlib.pyplot as plt

import pandas as pd

from sklearn import datasets

from sklearn.decomposition import PCA

from sklearn.discriminant_analysis import LinearDiscriminantAnalysis

loc = "C:\\Desktop\\Astart\\raw rgb.csv"

```python
iris = pd.read_csv(loc)

X = iris.iloc[0:, 0:6].values

y = iris.iloc[0:, 6].values

target_names = ['Acet','Buty','Prop','Acet, Buty','Acet, Prop','Acet, Prop, Buty']

lda = LinearDiscriminantAnalysis(n_components=2)

X_r2 = lda.fit_transform(X, y)


# Percentage of variance explained for each components

print('explained variance ratio (first two components): %s'


plt.figure()

colors = ['navy', 'turquoise', 'darkorange','yellow','green','magenta']

lw = 2

n = plt.figure()

for color, i, target_name in zip(colors, [0, 1, 2, 3, 4, 5], target_names):

    plt.scatter(X_r2[y == i, 0], X_r2[y == i, 1], alpha=.8, color=color,

            label=target_name)

plt.legend(loc='lower left', shadow=False, scatterpoints=1)

plt.title('LDA of RAW dataset')

n.savefig('LDA.png',dpi=600,bbox_inches='tight')

plt.show()
```