2017

Vision based model for identification of adulterants in milk

Jacklyne Atieno Kobek
Faculty of Information Technology (FIT)
Strathmore University

VISION BASED MODEL FOR IDENTIFICATION OF ADULTERANTS IN MILK

Jacklyne Atieno Kobek

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Information Technology at Strathmore University

Faculty of Information Technology

Strathmore University

Nairobi, Kenya

June, 2017

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Jacklyne Atieno Kobek

………………..

………………..

Approval

The thesis of Jacklyne Atieno Kobek was reviewed and approved by the following:

Dr. Joseph Orero,
Supervisor,
Strathmore University

Dr. Joseph Orero,
Dean, Faculty of Information Technology,
Strathmore University

Professor Ruth Kiraka,
Dean, School of Graduate Studies,
Strathmore University

Abstract

Milk adulteration is a social problem that exists in both developed and developing countries. This is due to lack of regulations or enforcement, proper refrigeration techniques, high yields with no market and hence the use of high levels of different adulterants to elongate the shelf life, prevent spoilage, increase thickness and whiteness.

This research proposes the use of a mobile phone application, to determine the intensity and type of adulterant used in milk, specifically water adulterant, by use of back propagation artificial neural network (ANN). A scanned image of milk spiked with acid-base indicator (bromothymol blue) was taken, after it changed color. Using ANN, the image was classified in terms of color descriptors such as mean of red (R), green (G), blue (B), luminosity (L, which is the sum of R, G, and B). After classification, partial least squares regression (PLSR) and principal component regression analysis (PCR) model, was used to predict the adulteration intensity in milk using the intensity of adulteration as a dependent variable.

Key Words: GDP, Milk, Adulteration, Adulterant, Artificial Neural Network, color descriptions

Abbreviations/Acronyms

KDB	Kenya Dairy Board
KEBS	Kenya Bureau of Standards
RGB	Red, Green, Blue
WHO	World Health Organization
FAO	Food and Agriculture Organization of the United Nations
KMDP	Kenya Market-led Dairy Programme
KARI	Kenya Agriculture Research Institute
GDP	Gross Domestic Product
SNF	Solid-no-Fat
ILRI	International Livestock Research Institute
DFID	Department for International Development

Table of Contents

List of Figures

`

List of Tables

Chapter 1: Introduction

1.1 Background of Study

Food safety has emerged as an important global issue with international trade and public health implications (WHO, 2006). In developing countries, the concerns about food safety are increasing due to a variety of factors including increase in the age of human populations, unplanned urbanization and migration and mass production of food due to population growth and changed food habits (Kaferstein & Abdussalam, 1999). Milk is one of the foods that has serious challenges in its handling and distribution (The Organic farmer, 2009). It is food that is consumed worldwide due to its widely known nutritional value and health benefits according to Fox and McSweeney (2009). Apart from being a source of nutrients essential for growth, development and maintenance of health, milk is a major source of protein in the diet of young animals and humans of all ages. According to (Hurley et al., 2006) adults can also benefit from drinking milk, to help keep their bones strong which will reduce the risk of developing osteoporosis.

However the benefits of milk to the populace is becoming a problem, because of the handling and preservation methods being used. In response to the increasing number of food borne illnesses, Governments all over the world are intensifying their efforts to improve food safety. However, food safety issues varies from country to country. Unlike in the west, in Kenya, traditionally adulteration is a major food safety concern. Recently, the occurrence of adulterations and frauds in liquid milk has increased. These adulterations have been practiced in a number of different ways, by the addition of an adulterant agent to increase economic yield (Das et al., 2011).

According to (The Organic farmer, 2009) Kenya's dairy sector has had numerous challenges. Although the sector has witnessed tremendous growth in the last six years, employing over 2 million Kenyans directly or indirectly, quality standards have considerably declined to the point of endangering the lives of consumers. As the samples analyzed from milk bars and processed milk show, milk sold in both the informal and formal market contains a high bacterial load and is therefore unfit for human consumption (Ndungu et al., 2016).Over 800,000 smallholder farmers in Kenya depend on dairy farming for their livelihoods. Small-scale farmers account for 80% of the total milk production and 70% of the total marketed milk in the country. Raw milk safety in Kenya has been disputed over a decade but no measurable data exists despite the fact that it requires monitoring from production to consumption. The regulatory institutions are constrained by lack of resources in terms of personnel and equipment even though, the Kenya Bureau of Standards developed a Hygienic Code of Practice for milk production to assist farmers in producing hygienic milk. The focus of quality milk, should move from big processors to the 80% small scale traders to ensure they produce and sell milk of high quality standards (Muriuki, 2003).

There are several adulterant agents in the market, though some of the adulterants like hydrogen peroxide were advocated for in countries where refrigeration techniques are still not advanced, to help the small scale farmers preserve their milk produce before taking it to the market, its use has been misused and now it’s a concern. Good quality milk should have a pleasantly sweet and clean flavor with no distinct aftertaste. Because of the perishability of milk and the nature of milk production and handling procedures, the development of off flavors/odors is not uncommon. To prevent flavor/odor defects in milk, proper milk handling procedures from the farm to the consumer are essential (Ndungu et al., 2016).

1.2 Problem Statement

Processed milk accounts for only 14 per cent of all milk sold in Kenya. The remaining milk is sold raw through the informal market, which comprises of direct farm sales to household consumers and hotels, through traders, milk bars and cooperatives. This is because raw milk is 25 Kshs. per litre and 50 per cent cheaper than pasteurized milk in rural areas and urban Nairobi, respectively. The raw milk market only exists because the majority of poor Kenyan households simply cannot afford pasteurized milk. When buying milk, most of the consumers check the quality of milk through organoleptic basic tests such as smelling, observing the appearance and tasting which is not able to clearly ascertain if the milk contains adulterants (Muehlhoff, Bennett & McMahon, 2013). This leads to the research problem addressed by this study, which is the fact that, there are no mechanisms or technologies for individual milk consumers to use to determine if water as adulterant has been added, when buying milk, hence the rise of adulterated milk in the market by unscrupulous traders and farmers for profit purpose.

The use of a portable, mobile vision based model, is proposed to aid consumers in identifying adulterants and level of adulterants in milk before buying. The accuracy of the model will lower the error of identifying adulterants in milk and enable consumers to drink milk that is healthy and of good quality. If the problem of adulterated milk is not addressed, it will give rise to risks with potentially large negative implications for livelihoods, human and animal health.

1.3 Research Objectives

i. To establish the characteristics of adulterated milk.
ii. To investigate the current methods used to detect adulteration in milk and challenges associated with the methods.
iii. To review existing models and architectures used in detecting adulterants in milk
iv. To develop a digital based solution for identifying adulterants in milk.
v. To test the accuracy of the proposed solution.

1.4 Research Questions

i. What are the characteristics associated with adulterated milk?
ii. What are the issues associated with the current methods of detecting adulterants in milk?
iii. What are the existing models, mobile applications and architectures used for detecting adulterants in milk?
iv. How can the digital based model be designed?
v. How can the digital based model be validated?

1.5 Justification

The current methods in Kenya for testing the quality and healthiness of milk are cumbersome, time consuming, expensive and can only be conducted by Kenya Bureau of Standards (KEBS). Due to the number of small scale farmers producing high quantities of milk, various cooperatives and small traders selling the milk to consumers and a combination of weak laws, KEBS is unable to test for adulteration, in all the milk sold to consumers directly. This research proposes a vision based model by use of a mobile camera since it is readily available, cheap, simple, rapid and provides accurate inspection rate. The camera will take an image of skewed milk with acid base indicators (bromophenol blue or bromothymol blue for spot test reaction) and analyze the image using artificial neural networks to detect RGB color descriptors (Red, Green and Blue) and other parameters such as hue, saturation and value, to determine the type and level of adulterant in milk. This will give more power to the consumer, since the body that has been mandated to ensure consumer's in Kenya, consume healthy and safe milk, have been unable to regulate the sale of unadulterated milk.

1.6 Scope and Limitation

This research will lays its focus on detection of an adulterant in milk, specifically water adulterant. The study will employ Back Propagation Neural Network algorithm to identify the adulterant based on the pixel features obtained from the images of milk skewed with bromothymol blue.

Chapter 2: Literature Review

2.1 Introduction

This chapter will focus on current challenges consumers have in identifying adulterants in milk with the current methods. It will lay more focus on the common adulterants used and why they are used, challenges in dairy sector contributing to adulteration by small scale farmers, traditional and technological methods and models currently used in detecting adulterants in milk. The chapter will also discuss on vision model and expert systems used in image preprocessing, segmentation, clustering and recognition to identify adulterants in milk.

2.2 Dairy Farming in Kenya

Kenya's dairy industry is private sector driven. It is the largest agricultural sub-sector and contributes 4% to GDP (KDB, MoALF 2012). The sector is dynamic with high growth figures of marketed milk and investments by dairy societies and processors mainly in the cold chain, production of long life milk (ESL) and milk powder (Ettema, 2013). About 80% of Kenya's total milk production (ca. 5 billion litres in 2011, KDB) is produced on small scale farms. This poses huge challenges to the industry in terms of cost of production, collection and cooling, seasonal fluctuations in supply, and the quality of raw milk. There is however a fast growing number of medium-scale farmers/investors who invest in modern and commercial dairy production. Currently, the sector provides food, income and employment for approx. 1.8 million people across the dairy value chain: farmers, transporters, traders and vendors, employees of dairy societies, milk processors, input suppliers and service providers, retailers and distributers. In terms of nutrition and food security, almost all Kenyan's consume milk on a daily basis with an average per capita milk consumption of 115 litres per year (KDB, 2012).

Kenya has about 30 active milk processors, Brookside, NKCC, Githunguri and Sameer being the largest ones processing together ca. 85% of the 1.5 million kilograms of milk processed daily. Market leader is Brookside Dairy Ltd. Although the market for processed milk and milk products strongly grew over the past 10 years, approximately 70-80% of the milk is distributed to the consumer through the raw milk market. The industry's growth and competitiveness are constrained by seasonality in milk production, milk quality issues, a severe lack of knowledge and skills, sub-standard service provision and input supply, as well as high fragmentation of the supply chain and lack of inclusive business models. If these issues can be effectively addressed this will boost further commercialization and growth of the sector, and contribute further to creation of wealth, employment across the value chain, and to food security (Ndungu et al., 2016).

Kenya dairy sector, facts & figures	
Land surface:	583.000 km².
Inhabitants:	44 million
Capital:	Nairobi
Population:	22% Kikuyu, 14% Luhya, 13% Luo, 12% Kalenjin, 11% Kamba, 6% Kisii, 6% Meru, 16% others.
Languages:	English, Swahili
Main trading partners:	Uganda, Tanzania, Great Britain, Germany, South-Africa.
Total milk production:	5 billion kg (2011)
Production by smallholders:	80 %
Milk processed:	30 %
Raw milk market:	70 %
Smallholders:	800.000
Medium / large scale farms:	3500
Milk consumption / capita:	115 litres/year
Active milk processors:	30
Market leading milk processor:	Brookside
Income and employment in the dairy value chain for 1.8 million people	

Table 2.1: Kenya Dairy Sector Facts (Adapted from Ettema, 2013 p.5)

2.2.1 Challenges Facing the Dairy Sector

There are various problems/challenges facing the Dairy sector such as low skills and knowledge level amongst almost all farmers (small, medium and large scale),low level of commercialization by small scale farmers when dairy is not their core business, high cost and seasonality of milk production due to low ability/skills to produce and preserve quality fodder, inefficient and high cost of milk collection and cold chain development (hence: high cost and low quality of milk at factory gate), lack of loyalty between value chain actors and high fragmentation, lack of credible input suppliers and services providers ("pushing products"),large raw milk market and lack of level playing field for the formal sector, oligopolistic nature of the processing industry (Brookside acquiring other brands), lack of clarity on a common vision amongst stakeholders on how to steer the dairy industry into a more sustainable growth path and ineffective sector regulation: policies in the shelves and not enforced on the ground (Mugambi et al., 2004).

Kenyans appear to prefer raw milk. Estimates from various studies indicate that about 85 percent of marketed milk is sold raw. Recently, the Kenya Dairy Board (KDB) and others in the formal milk trade have claimed that the proportion of processed milk has increased to more than 20 percent. An SDP brief provides the following as reasons for unprocessed milk being preferred: It is 20 to 50 percent cheaper than processed milk many people prefer the taste and high butterfat content of unprocessed milk, unprocessed milk is sold in variable quantities, depending on how much money the customer has to spend, it is widely accessible and within the reach of many people, most consumers are accustomed to consuming unprocessed milk. The selling of milk through the unprocessed channel is of concern because of the perceived health risk, particularly owing to its microbial load by the time it reaches the consumer. These challenges compounded with lack of appropriate policies and enforcement has created an increase in adulterated milk in the market (Moturi, Obare, & Kahi, 2014).

2.2.2 Small Scale Traders Market Channels

Before market liberalization in the early 1990s, there was an organized milk collection and bulking system in the formal market, with two types of milk delivery to KCC facilities: by individual dairy farmers; or by dairy cooperative societies. With liberalization and the collapse of KCC, the collection and bulking system also collapsed. At present, collection and bulking is a complex of different systems depending on processors, intermediaries, the road network, milk sheds and many other factors (Muriuki, 2011) . The transportation of milk depends on the amount and the buyer. Major processors have their own collection, bulking and transportation systems. Stainless steel (seamless) cans, and occasionally plastic cans, are used for bulking milk from individual suppliers and delivering it to processors' collection, bulking and cooling centers, from where it is transported in cans or by refrigerated tanks to the main processing plants. In some areas, powerful milk intermediaries (traders) have positioned themselves between the market and the milk producers. Their presence complicates the traceability of milk and brings a risk of cross-contamination and microbial overload.

Figure 2.1 shows a simplification of milk marketing pathways. Most traded milk is sold either directly from farmer to consumer (neighbor) or through unlicensed/informal traders.

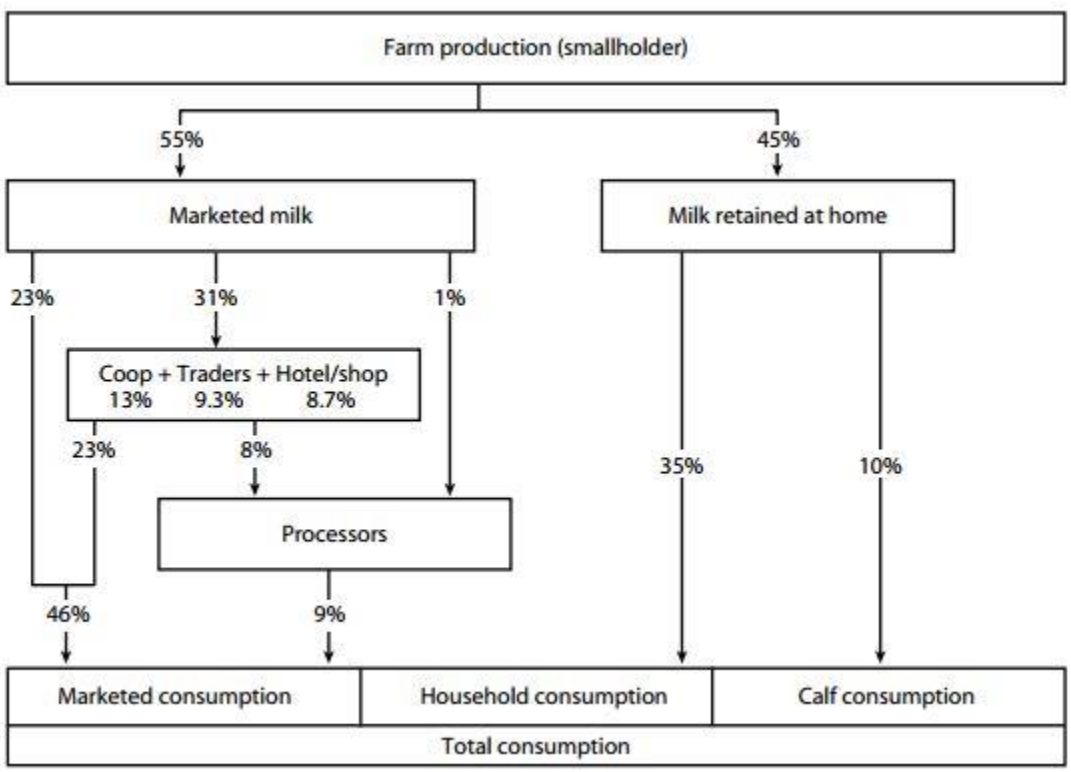

Figure 2.1: Milk Channels (Adapted from Muriuki, 2011 p.13)

Figure 2.2 below shows a simplification of milk chain actors. Most actors in the milk value chain buy directly from farmers and sell it directly to consumers or licensed milk processors.

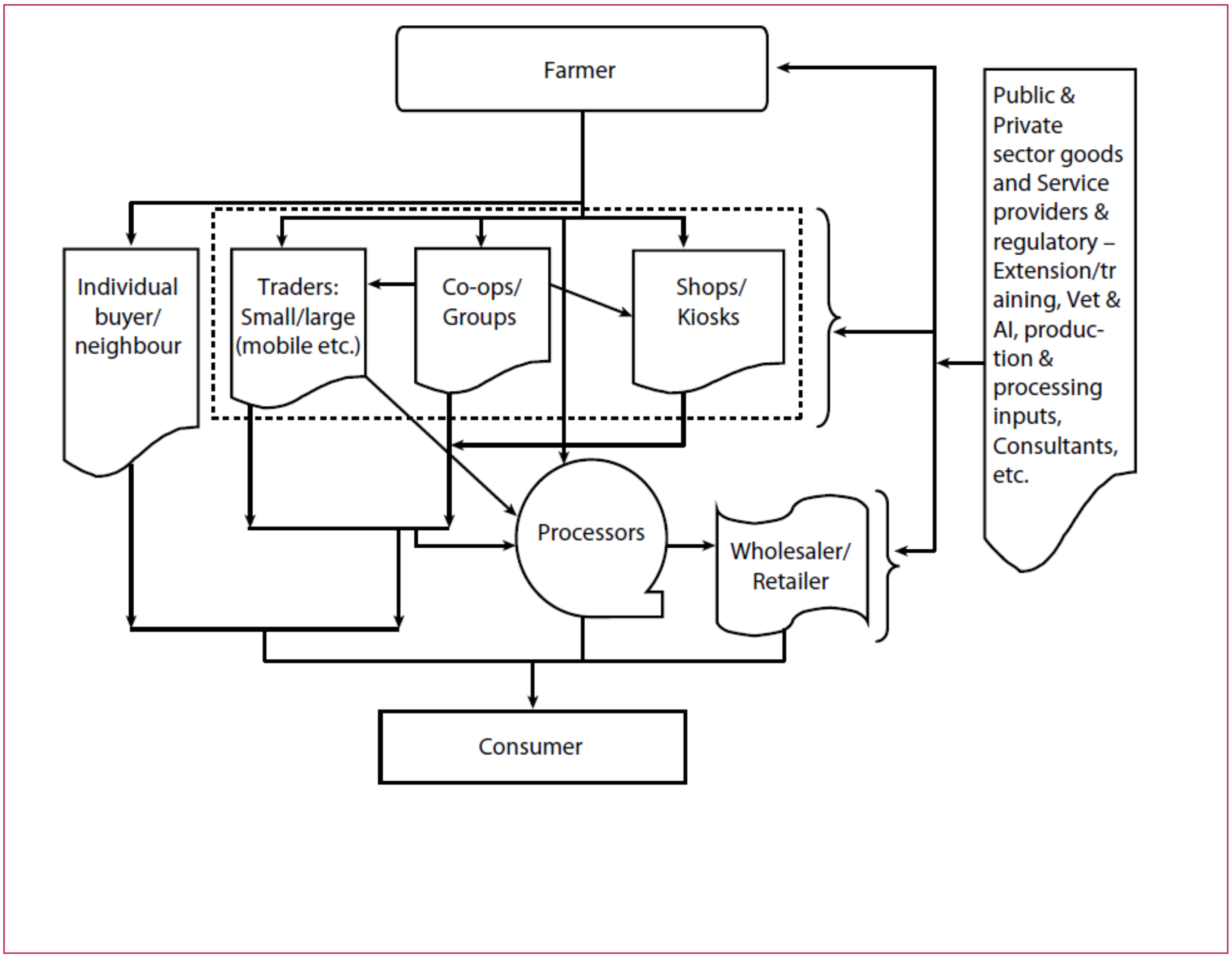

Figure 2.2: Milk Value Chain Actors (Adapted from Muriuki, 2011 p.41)

2.3 Adulteration of Milk

Adulteration of food has been a problem for centuries. Food adulteration has progressed from simple fraud to a lucrative business. Sumar and Ismail (1995), defines adulteration as the act of debasing a commercial product with the object of imitating or counterfeiting a pure or genuine commodity or of substituting an inferior article for a superior one in order to gain an illegitimate profit. He also defines four other identified principal means by which food may be regarded as adulterated as, complete or partial omission or abstraction of valuable constituents, whole or partial substitution of food components replacing the described cheaper product with undeclared alternatives, concealment of damaged or inferior food stuffs and finally addition of undeclared substance or material so as to increase product bulk or weight, reduce food quality or strength, or make product appear more attractive or better value than it is. Adulterated food is therefore foodstuff that has been "interfered" with. "Adulterant" means any material which is or could be employed for making the food unsafe or substandard or misbranded or containing extraneous matter.

2.3.1 Commonly Used Adulterants

One of the foodstuffs that is highly adulterated is milk. Commonly used adulterants in milk are formalin, water, hydrogen peroxide, boric acid, salt, oil, glucose among others. The adulterants are used as follows; (i) Hydrogen Peroxide-used to prolong the life and freshness of milk. Peroxide damages the gastro intestinal cells which may lead to gastritis and inflammation of the intestines, (ii) Urea-added to provide whiteness, increase consistency of milk and for leveling the contents of solid-no-fat (SNF) as are present in natural milk. Urea in milk overburdens the kidneys as they have to filter out more urea content from the body.(iii) Detergents (pulverized soap)-added to milk to emulsify and dissolve the oil in water giving a frothy solution, the characteristic white color of milk. It leads to Gastro intestinal complications. (iv) Starch (cereal flour and arrow root) - improves milk thickness. High amounts of starch can cause diarrhea due to the effects of undigested starch in colon. Its accumulation in the body may prove fatal for diabetic patients. (v) Carbonates and bicarbonates-added to prevent spoilage. Regular intake can cause disruption in hormone signaling that regulate development and reproduction. (vi) Sugar and salt; is added to get the natural taste of milk. Causes irreversible damage in people suffering from diabetes and high blood pressure. It can be fatal for those who have kidney problems.

From the descriptions, adulterated milk is dangerous because it can be toxic and may affect one's health. It could deprive nutrients essential for proper growth and development.

Several things are thought to be correlated with milk adulteration. Some examples include lack of affordable means of preserving milk, increased and available market, easily available preservatives, lack of support for the local dairy farmers, lack of proper and clear regulations, lack of will to enforce regulations and way of making extra profits (McDonald, 2000; Sumar et al., 1995; Barnes, 1996). It is hypothesized that countries with a large number of low scale farmers with no support mechanisms for their produce, tropical weather conditions, high population, widespread lack of cold chains due to lack of reliable electricity and poor enforcement of regulations are likely to have more adulterated milk than countries with effective regulations, support mechanisms for low scale farmers and high end machines. The following four literature reviews, seek to support the above hypothesis and demonstrate how different countries are testing adulteration in milk.

Research article by (McDonald et al., 2000) answers the question of industry regulations and milk quality. He views that for a country to even impose regulatory standards a common understanding of milk quality between the enforcers, processors, farmers, distributors and consumers has to be one. Each group has its own perception and image of milk quality, its own notions about validity of diverse standards of quality and their relative importance, and their own sense of economic cost-benefit involved in assuring "quality" at their particular level in the system of production and processing. The idea that milk quality can vary and that it is often important for milk to meet a particular standard is not new to the different groups. It has been a major issue for the key players in the dairy industry to keep to the stipulated standards and consumer demand for a safe and nutritious product. This has also led to rivalries between the processors, farmers and distributors, with the different group blaming the other for drop in milk quality and high percentage of adulterated milk. The milk industry value chain need to have a common understanding of milk quality and all adhere to stipulated standards to ensure consumers consume safe and high quality milk.

2.3.2 Causes of Adulteration

Muriuki (2003), states that the dairy industry is the most developed of the livestock sub-sectors in Kenya and is comparatively well developed relative to the dairy industries of other countries in sub-Saharan Africa. The dairy sub-sector, like other agricultural sub-sectors, is predominantly smallholder. Dairy production systems in Kenya however can largely be classified as large- or small scale. Small-scale producers (the smallholders) dominate dairy production owning over 80 percent of the 3.3 million dairy cattle, producing 56 percent of the total milk production and contributing 80 percent of the marketed milk (Peeler & Omore, 1997).

Hence the challenges facing the industry is mainly due to the small scale producers who produce and market the highest quantity of milk yet Kenya Dairy Board has been unable to enforce policies and regulations to this group to ensure they produce and sell safe and quality milk.

Some of the contributing factors affecting the Kenya Dairy Sector hence the high quantities of adulterated milk circulating in the Kenyan market are; Low skills and knowledge level amongst almost all farmers (small, medium and large scale),low level of commercialization by small scale farmers when dairy is not the core business, high cost and seasonality of milk production due to low ability/skills to produce and preserve quality fodder, inefficient and high cost of milk collection and cold chain development (hence: high cost and low quality of milk at factory gate),lack of loyalty between value chain actors and high fragmentation, lack of credible input suppliers and services providers ("pushing products"),large raw milk market and lack of level playing field for the formal sector, oligopolistic nature of the processing industry (Brookside acquiring other brands),lack of clarity on a common vision amongst stakeholders on how to steer the dairy industry into a more sustainable growth path, ineffective sector regulation: policies in the shelves and not enforced on the ground, last but not least, poor roads and infrastructure negatively affecting milk price and profitability (Ettema, 2014).

2.3.3 Characteristics of Adulterated Milk

Soomro et al. (2015) states that, the quality and nutritional value of milk changes when different adulterants like urea, water, starch, hydrogen peroxide etc. is used in milk for different reasons. Below are some characteristics of adulterated milk;

i. When boiled on slow heat for 2-3 hours till it solidifies and becomes hard, if the rough residue becomes rock solid, it means the milk is adulterated and if the residue is oily it means it's of good quality.

ii. Bad taste and feels soapy when rubbed and turns yellowish when heated.
iii. If a drop of milk is put on the fist or any slanted surface and you let it flow down, if adulterated it leaves a trail behind, else it's good.
iv. If 2 tablespoons of salt (iodine) is added to 5 ml of adulterated milk, the mixture will turn blue else it remains intact.
v. For formalin existence in milk, put 2-3 drops of sulphuric acid into 10ml of milk, if a blue ring appears at the top, milk is adulterated else not.

2.4 Current Methods of Identifying Adulterants in Milk

Milk detection techniques vary from simple visual methods to complex biological systems. Quality control tests for milk, are a considerable aspect to assure adulterant free milk for consumption. Below are some methods used in Kenya to detect adulterants in milk according to a research by (KDB, ILRI, DFID, Land o Lakes, KARI & FAO, 2014).

1. The Lactometer test

Addition of water to milk can be a big problem where we have unfaithful farm workers, milk transporters and greedy milk hawkers. A few farmers may also fall victim of this illegal practice. Any buyer of milk should therefore assure himself/herself that the milk he/she purchases is wholesome and has not been adulterated. Milk has a specific gravity. When adulterated with water or other materials are added or both misdeeds are committed, the density of milk change from its normal value to abnormal. The lactometer test is designed to detect the change in density of such adulterated milk. Carried out together with the Gerber butterfat test, it enables the milk processor to calculate the milk total solids (% TS) and solids not fat (SNF). In normal milk SNF should not be below 8.5% according to Kenya Standards (KBS No 05-l0:-1976).

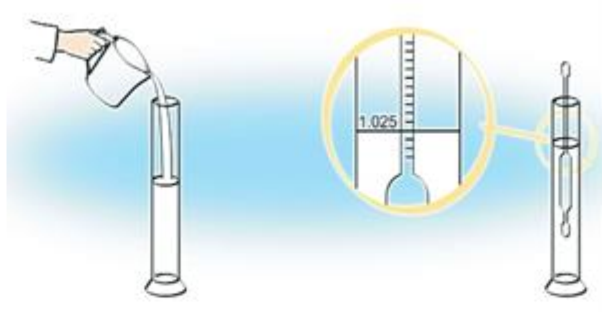

Figure 2.3 Lactometer Test (Adapted from Small Holder Project, 2004)

2. Freezing Point Determination

The freezing point of milk is regarded to be the most constant of all measurable properties of milk. A small adulteration of milk with water will cause a detectable elevation of the freezing point of milk from its normal values of -0.54ºC. Since the test is accurate and sensitive to added water in milk, it is used to detect whether milk is of normal composition and adulterated.

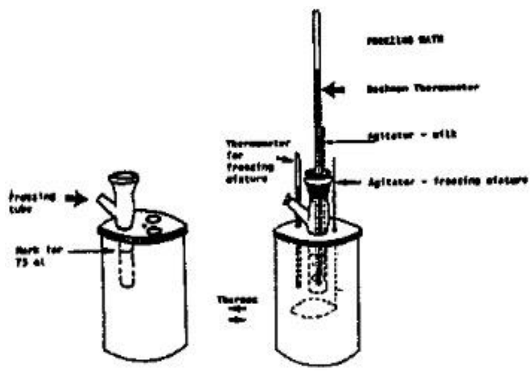

Figure 2.4: Freezing Point Determination (Adapted from Small Holder Project, 2004)

3. Digital Milk Analyzer(Lactoscan SLP 60)

One stop technology that tests various milk parameters such as water, content ,fat, protein and solids to determine the quality of milk delivered by farmers. It is a mobile gadget that rapidly analyses milk and gives a set of result depending on one’s need. Kenya Dairy Board (KDB) in partnership with Agritrace Limited are promoting the new digital milk analyzer machine dubbed Lactoscan SLP 60.

Figure 2.5: Lactoscan milk analyzer (Adapted from Small Holder Project, 2004)

2.4.1 Challenges of Conventional Methods of Detecting Adulterants in Milk

Conventional methods of testing milk for adulteration have their challenges and may even provide erroneous results (Omore et al., 2005). Some of the challenges described as follows; lactometer test challenge is that, it can tell the purity of milk only if it is adulterated with water. If the milk is skimmed, its density becomes higher, than if water is added to it, the lactometer cannot tell its purity. In other words a lactometer cannot tell the purity of milk if it is skimmed. Zagorska and Ciprovica (2013) states the different challenges of freezing point as, cannot be used for controlling presence of different substances in milk. The thermal treatment of milk has significant influence of freezing point of milk. The presence of different substances in milk has different influence on freezing point and pH of milk. Malali et al. (2016) discusses main challenges of Lactoscan milk analyzer as the cost, requirement of skilled personnel for its operation and data analysis and initial implementation investment is more.

2.4.2. Electronic Methods of Identifying Adulterants in Milk and Challenges

Das, Goswami and Biswas (2016) summarizes different electrical methods used to detect adulteration in milk such as Chlorine, Antibiotics, Water, Non-milk proteins, Color, Preservatives, Urea neutralizers and Whey liquid etc. Electrical methods are accurate and simpler to use than other methods and easy to process because of electrical domain.

Banupriya, Chaitanya, Supriya and Varshitha (2014) illustrates various methods of detecting urea in milk including electrical methods. Their study draws comparison on the merits and demerits of the different electrical methods. From their study it is clear and evident that the conductance measurement method stood out due to its low cost, linear absorption, portability and speed of response.

The different electronic methods with their challenges are:

2.4.1.1 Potentiometric Sensors

Potentiometric allows the determination of a wide spectrum of ions and inexpensive, portable equipment can be developed. Trivedi, Donepudi, Kapse and Panchal (2009) researched on a potentiometric biosensor3 to detect urea adulteration in milk .It uses a NH+4 ion sensitive electrode as the transducer. It is a disposable type urea sensitive enzymatic biosensor system and has been developed by immobilizing the urease enzyme, through entrapping, onto the ion sensitive membrane using a polymer matrix. Challenges of potentiometric sensors are but not limited to; Frequent calibration, less precise and sensitive to changes in ionic strength (Trivedi et al., 2009).

Figure 2.6: Potentiometric Sensor (Adapted from Trivedi et al., 2009)

2.4.1.2 Conductance Measurement

The conductance measurement between two electrodes is a well-known technique to detect adulteration. Most of the times the electrical equivalent model of the electrodes immersed in the sample is evaluated to identify the adulterated milk.

Challenges of conductance measurement are but not limited to: Cannot distinguish between different types of ions, conductivity meters are temperature dependent; conductance increases approximately 2% per °C and does not measure the number of ions in a solution directly (Velasco-Garcia & Mottram, 2003).

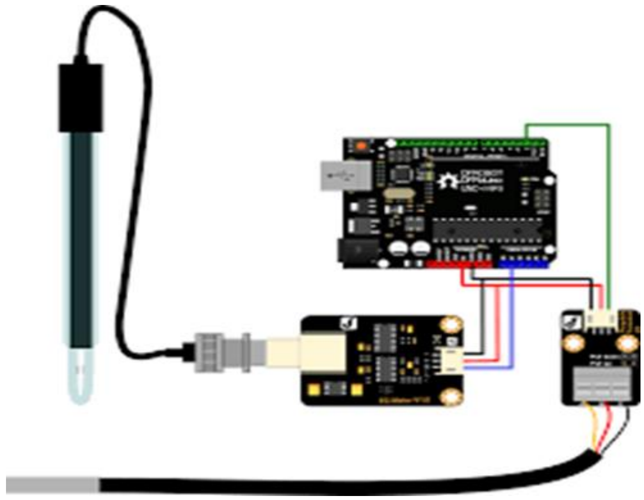

Figure 2.7 Conductance Measurement (Adapted from Velasco-Garcia & Mottram, 2003)

2.4.1.3 E-Nose

There has been extensive research in evaluating electronic noses for monitoring the quality of milk. The two main components of an electronic nose (E-nose) are the sensing system and the automated pattern recognition system. The common pattern recognition systems are either principal component Analysis (PCA), linear discriminant analysis (LDA) or Artificial Neural Network (ANN). Challenges of E-Nose include but not limited to: High cost, complexity of signature detection and matching and broad range of application area e.g. Cyranose 320 used to sniff out explosives, chemicals, food contaminates and even cancers. (Velasco-Garcia & Mottram, 2003)

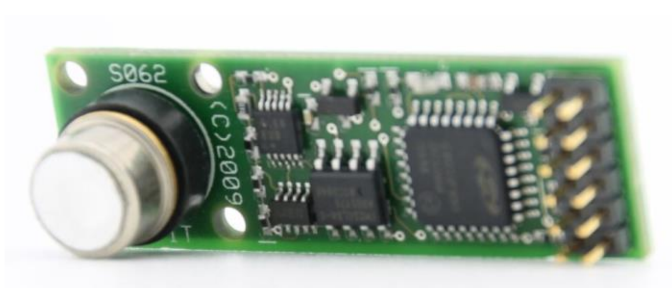

Figure 2.8: E-Nose Sensor (Adapted from Velasco-Garcia & Mottram, 2003)

2.5 Digital Based Solutions Related Applications

Various researches have also been done to detect various compounds in liquids as discussed below:

2.5.1 Furfural determination with disposable Polymer films and smartphone based colorimetry

In 2016 University of Madrid, Ciudad Universitaria, the following students (Rico-Yuste, González-Vallejo, Benito-Peña, Casas Engel, Orellana & Moreno-Bondi, 2016) did a research that developed a method that allows brewers to measure the freshness of beer by using a polymer sensor that changes color upon detecting furfural – a compound that appears when this beverage ages and gives it a stale flavor. The sensor can be controlled from a smartphone app. This development is a simple, low-cost method capable of measuring whether or not beer has gone stale. "Until now, brewers have measured furfural and other freshness indicators using methods based on chromatography techniques. But these methods involve the use of expensive equipment and sample preparation is very time-consuming". The new method consists of sensor discs that detect the presence of furfural in beer. These sensors, made from a polymer similar to the one used to manufacture contact lenses, have been designed to change color (from yellow to pink) when they come into contact with a beer containing furfural.

The sensor material is incorporated with a derivative which reacts with the furfural and produces a pink cyanine derivative that allows one to identify the presence of the marker in the sample. "The intensity of the color increases as the concentration of furfural in the beer rises and as more time passes since the production of the beer", they added in a paper published in the journal Analytical Chemistry. The team has also created a mobile app for Android smartphones that, by taking a picture of the sensor disc, allows for the identification of the amount of furfural present in the beer.

2.5.2 An iPhone-based digital image colorimeter

Research by (Masawat, Harfield, & Namwong, 2015) an iPhone-based digital image colorimeter (DIC) which was fabricated as a portable tool for monitoring tetracycline (TC) in bovine milk. An application named Color-Conc was developed for the iPhone that utilizes an image matching algorithm to determine the tetracycline concentration in a solution. The color values; red (R), green (G), blue (B), hue (H), saturation (S), brightness (V), and gray (Gr) are measured from each pictures of the tetracycline standard solutions. Tetracycline solution extracted from milk samples using solid phase extraction (SPE) is captured and the concentration is predicted by comparing color values with those collected in a database. The amount of tetracycline can be determined in the concentration range of 0.5–10 μg mL^{-1}. The proposed digital image colorimeter (DIC)-iPhone is able to provide a limit of detection (LOD) of 0.5 μg mL^{-1} and limit of quantitation (LOQ) of 1.5 μg mL^{-1}. The enrichment factor is 70 and color of the extracted milk sample is a strong yellow solution after solid phase extraction (SPE).

2.5.3 Digital imaging

Pereira-Filho (2013) proposes the use of digital imaging as an alternative method for the identification and quantification of milk adulteration. Digital images were obtained by a flatbed scanner and the means of ten color descriptors (red, green, blue, hue, saturation, value, luminosity and the three relative colors, i.e. relative red, relative blue and relative green) were used to evaluate the information from the images.

The images is then transferred to matlab (using K-Nearest for quantification models and pattern recognition techniques) to determine the type and level of an adulterant.

2.6 Artificial Neural Networks (ANN) In Image Processing

Alexandrina, Mihai and Gabriela (2015) states that Image processing using artificial neuronal networks (ANN) has been successfully used in various fields of activity such as geotechnics, agriculture, civil engineering, mechanics, industrial surveillance, defense department, automatics and transport. Image preprocessing, date reduction, segmentation and recognition are the processes used in managing images with ANN. An image can be represented as a matrix, each element of the matrix containing color information for a pixel. The matrix is used as input data into the neuronal network.

Vertan (2001) considers images to be abstract (mathematical functions with two variables, continuous or discrete), non-visible (unperceived by naked eye, which imply a sum of bi-dimensional fields of parameters such as temperature, pressure, density, etc.) and visible (perceived by naked eye and generated as distributions of light intensity).

Alexandrina et al. (2015) also states, depending on the type of data that is the matrix, the images are divided into images of intensity scale and indexed (each component being a unique number, a scalar) and vector images (each component being a vector, vector number which in turn splits into several parts). Scalar image intensity is an image where each pixel value (real or natural numbers) is considered a measure of luminous intensity. Scalar indexed image is an image in which the value of a pixel is an index where information can be associated with the color of the pixel in question.

Each color can be represented as a combination of three basic colors: red, green and blue. The array is used as input to the neural networks that are aimed at identifying images or grading.

Each input neuron represents color information in the image, and each output neuron corresponds to an image. All images will be scaled to the same size (width and height) and small to be easy and quick to learn. On the sizes of the images shall be determined on the size of the input vector and the number of neurons. The transfer function for this type of problem is called sigmoid function. The rate of learning has values in the range [0.1] and the error it is recommended to have less than 0.1. Processing of images with ANN involves different processes, such as:

i. Image preprocessing, an operation which shows a picture (contrast enhancement, noise reduction) with the same dimensions as the original image. The objective of images preprocessing with ANN consists in improving, restoring or rebuilding images. The resolved issues are the cartographic types, to optimize a function, an approximation function for the reconstruction of an image.
ii. Data reduction or feature extraction involves extracting a number of features smaller than the number of pixels in the input window. The operation consists in compressing the image followed by extracting geometric characteristics (edges, corners, and joints), facial features, etc.
iii. Segmentation is a division of an image into regions.
iv. Recognition involves the determination of objects in an image and their classification.

2.6.1 Sample Applications of Artificial Neural Networks in Dairy Farming

Snehal, Dahikar, Sandeep and Rode (2014) research used artificial neural network to consider various situations of climatologically phenomena affecting local weather conditions in various parts of the world. These weather conditions have a direct effect on crop yield. Various researches have been done exploring the connections between large-scale climatologically phenomena and crop yield. Artificial neural networks have been demonstrated to be powerful tools for modeling and prediction, to increase their effectiveness. Crop prediction methodology is used to predict the suitable crop by sensing various parameter of soil and also parameter related to atmosphere. Parameters like type of soil, PH, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, sulphur, manganese, copper, iron, depth, temperature, rainfall, humidity.

Expert systems have also been used to predict lifetime milk account in dairy cows. Shailesh, Ram, Dr.Gupta and Dr.Sharma (2013) study, is based on back propagation neural network which has trained and tested based on dataset provided. In modelling the artificial neural networks based predictive model, two-hidden layer network has been constructed. Suitable milk amount predictions can provide farmers and producers with valuable information. For this research Different artificial neural networks had been trained and the best performing number of hidden layers and neurons and training algorithms recorded. The output performance of the artificial neural networks model in simulating cow's performance was compared with actual data as recorded by experimental work. Generally more nearer prediction values can be obtained by a neural network approach. This approach suggests that a non-linear relationship exists among the dependent and independent variables in the data.

Artificial neural network has also been successfully applied to detect mastitis in dairy cattle. - The use of milk sample categorization for diagnosing mastitis using Kohonen's Self-organizing Feature Map (SOFM) is reported. Milk trait data of 14 weeks of milking from commercial dairy cows in New Zealand was used to train and test a SOFM network. The SOFM network was useful in discriminating data patterns into four separate mastitis categories. Several other Artificial Neural Networks were tested to predict the missing data from the recorded milk traits. A multi-layer perceptron (MLP) network proved to the most accurate ($R1 = 0.84$, $r = 0.92$) when compared to other MLP ($R1= 0.83$, $r = 0.92$), Elman

(R2 = 0.80, r = 0.92), Jordan (RI = 0.81, r = 0.92) or linear regression (R' = 0.72, r = 0.85) methods. It is concluded that the SOFM can be used as a decision tool for the dairy farmer to reduce the incidence of mastitis in the dairy herd (Lopez-Benavides, Samarasinghe & Hickford, 2003).

2.6.2 Computer Vision Model

Matsuyama (1989) addresses computer vision (image understanding) as a discipline that studies how to reconstruct, interpret and understand a 3D scene from its 2D images in terms of the properties of the structures present in the scene. The ultimate goal of computer vision is to model, replicate, and more importantly exceed human vision using computer software and hardware at different levels. It needs knowledge in computer science, electrical engineering, mathematics, physiology, biology, and cognitive science.

Limitations of Human Vision;

i. Limited memory-cannot remember a quickly flashed image
ii. Limited to visible spectrum
iii. Illusion

Difficulties of Computer Vision;

i. The problem is ill-posed inverse problem.
ii. Noisy image data or data with uncertainties.

Computer vision overlaps significantly with the following fields: image processing, and pattern recognition.

A step by step description of detecting adulterants in milk using computer vision is demonstrated below in Figure 2.9.

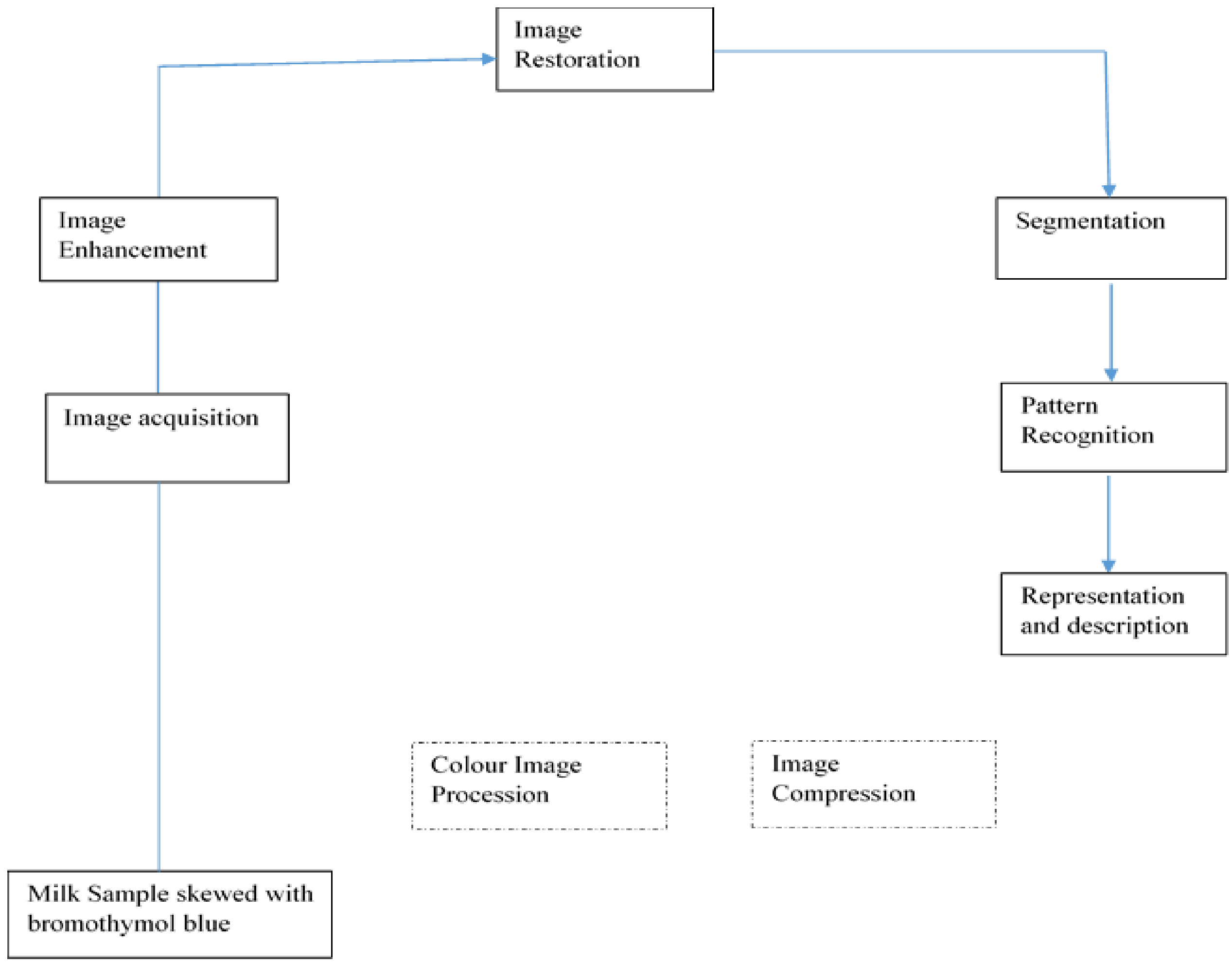

Figure 2.9: Adulterant Detection in Milk

2.7 Conceptual Framework

Figure 2.10 demonstrates how the proposed model will work. Milk traders, buyers and Cooperatives who buy milk from the informal sector make use of their mobile phones to take an image of skewed milk with bromothymol blue .Pixel features of the images will be used as the input for the neural network model. The neural network model will be trained and validated to work. The adulterant identification and levels denotes the output of the model.

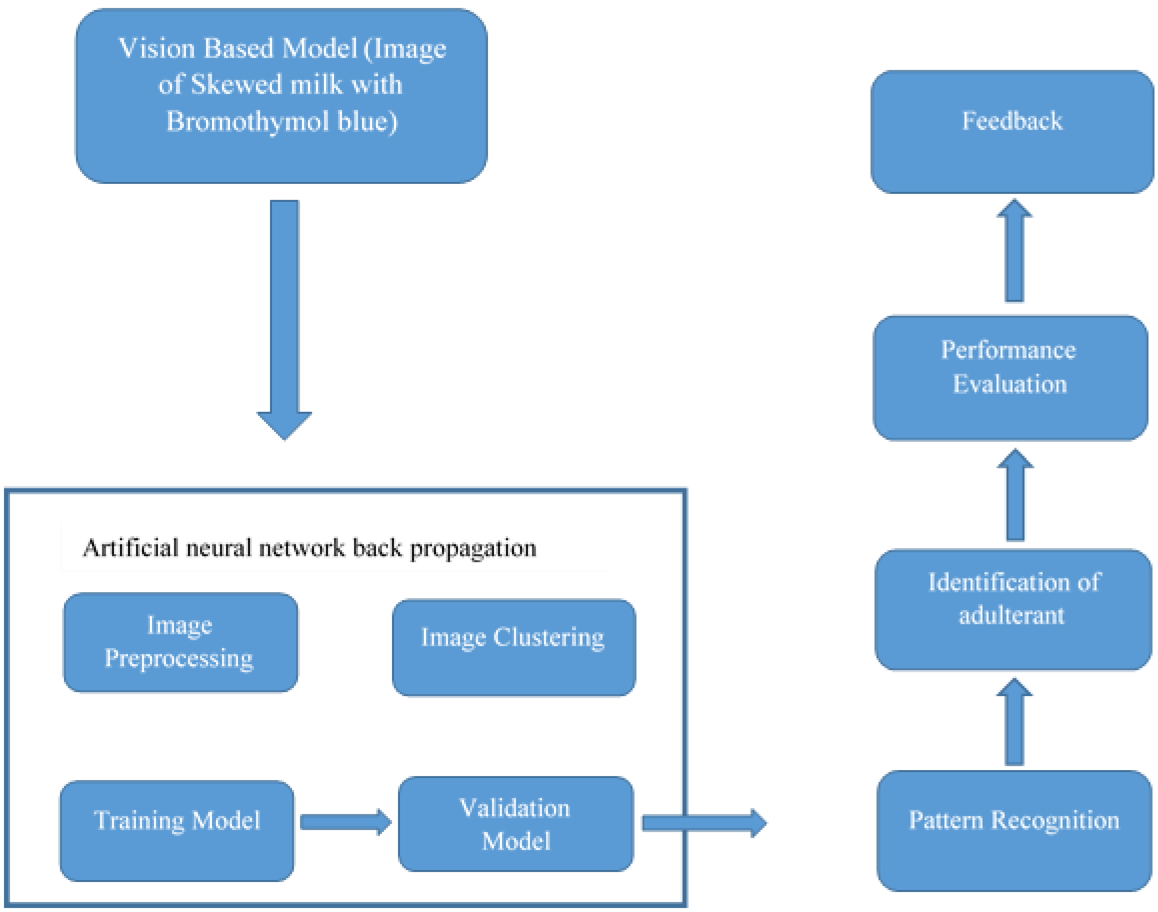

Figure 2.10: Conceptual Framework

Chapter 3: Research Methodology

3.1 Overview

Research methodology is, a systematic way to solve a problem. It is a science of studying how research is to be carried out. It is also defined as the study of methods by which knowledge is gained. Its aim is to give the work plan of research (Rajasekar, Philominathan & Chinnathambi, 2013). This research attempted to give birth to a new efficient and portable system that satisfies the current needs of the user and has scope for future growth within the market constraints. Therefore the research was guided by the objectives that the author was proposing to meet at the end of the research. It was greatly informed by the nature of the problem being studied, research designs that have been used in the related work reviewed in chapter 2.

Of importance was the population that was studied, the sample that was selected based on a certain criteria and the method that was used to obtain the data. A classifier for categorizing the data was also identified and the validity of the classifier determined.

3.2 Rapid Application Software Development

RAD refers to a development life cycle designed to give much faster development and higher quality systems than the traditional life cycle. It is designed to take advantage of powerful development software like CASE tools, prototyping tools and code generators. The key objectives of RAD are: High Speed, High Quality and Low Cost. RAD is a people-centered and incremental development approach. Active user involvement, as well as collaboration and co-operation between all stakeholders are imperative. Testing is integrated throughout the development life cycle so that the system is tested and reviewed by both developers and users incrementally.

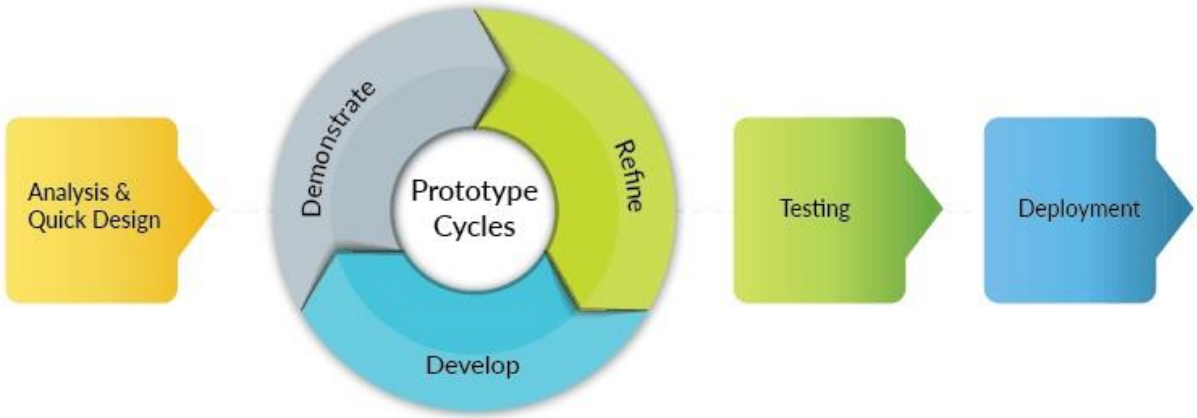

Figure 3.1 Rapid Application Software Development

The prototype was developed using rapid application development (RAD) methodology. This was appropriate due to its shorter development cycle compared to other methodologies and thus suitable for the limited duration of time.

3.3 Research Design

This study utilized both the qualitative and quantitative methods of research. As widely accepted the qualitative research is mainly exploratory. It is used to gain an understanding of underlying reasons, opinions, and motivations. It provides insights into the problem or helps to develop ideas or hypotheses for potential quantitative research while applied research is designed to solve practical problems of the modern world, rather than to acquire knowledge for knowledge's sake.

Two types of data were used during development: the primary and the secondary data. The primary data was derived from the answers respondents gave in the self-administered questionnaires prepared by the researcher. In addition, the information obtained from the interview also provided primary research data that was used to support the study. The secondary data on the other hand, was derived from the findings stated in published documents and literatures related to the research problem. It was based on the recent literatures related to Dairy farming in Kenya and the factors that challenge it, milk handling procedures and practices in Kenya and the concepts cited by the respondents.

The research proposed a neural network model that aided in identifying adulterants used in milk. (Santos, Wentzell, & Filho, 2012) , effectively used scanned digital images with color parameters, feature extraction, classification and calculation of the classification performance to detect water and hydrogen peroxide used in milk as adulterants.

3.3.1 System Analysis

The tools used in gathering the system requirements/data were; functional decomposition diagram that provided a top-down view of the system major functions by breaking it down into major chunks. This helped in validating all the functions that the model will provide. Use case diagrams was used to show the interaction between the user and the system. Sequence diagram was also employed to show interactions of objects over time. It provided a top-to-bottom view with messages being sent back and forth between the different objects.

Data gathering included combining all the data that was used in training the neural network. These were found in databases, spreadsheets, other computer files, or printed matter. The preprocessing performed in this step was only to collect the data into an electronic format that is consistent with the neural network training package being used (data “neutralization,” as it were).

3.3.2 System Design

Android Universal Image Library was used to develop the mobile application. UIL is a library which provides asynchronous out of the box loading and caching of images. With UIL almost everything can be configured, good for fetching and catching of images. Weka was used for backend programming because it is a collection of machine learning algorithms for data mining tasks. The algorithms were applied directly to a dataset and called from Java code. It contained tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It was well-suited for developing new machine learning schemes for the application. Weka was also considered because it is an open source software (Smith & Frank, 2016).

3.3.3 System Implementation

Usability testing was used to test the functionality of the system. Usability testing entailed testing; validation of images taken e.g. quality, pixels, clarity, validation of images, responsiveness and user friendliness (Sivaji, Abdullah, & Downe, 2011).

3.4 Population

This research was carried out in Nairobi County, which has a population of 93 hotels according to the Kenya Bureau of Statistics (2013). The hotels, comprise of both high class and other hotels. Hotels buy milk in bulk from their suppliers for their use. Hotels that have been operational for two or more years are more stable as they have tried, tested different milk suppliers, have more visitors/customers and identified a suitable milk supplier over the years. Therefore, the research assumption was that, it was suitable for hotels that had been operational for more than two years.

3.4.1 Sampling

Stratified sampling was advantageous for the datasets, which were divided into separate clusters (May et al., 2010). In this case, all regions of the input space were adequately covered by the training subset Ttr and the estimate of the model performance measured on Tt was highly precise. One cluster, strata consisted of images that represented healthy milk (unadulterated), the second cluster comprised of adulterated milk images. Different split ratios were then applied to the image feature sets to determine the training data set as well the test data set.

The split ration that produces the highest accuracy was then used in the implementation

$$nh = (Nh / N) * n$$

Eq. (3.2)

nh represents the total sample of the cluster h, Nh is the population size of the stratum, N represents the total population size and n is the total sample size. The proposed sample consisted of 30 images (Reitermanov´a, 2010). Milk consumers (in this case hotels) were interviewed to determine the factors they consider when determining that milk is adulterated and to test the usability of the system.

3.4.2 Sample Split

This process involved identifying the data that was first used for the training set, the second the control (or fine tuning set) and the third and final set, test set. Each of these three data sets contained approximately one third of data with a tendency of the training set to be the smallest and the test set to be the largest one. The control/validation set which is most often ignored was used as a supplier of feed-back information, for correcting the initially selected ANN architecture in cases where the training is performing as expected, but the results on the control set which was, of course, not used in the training, are bad. The sample taken was representative of the identified population.

3.4.3 Data Collection Methods

Various data collection methods were made use of in the research. Under qualitative research method, non-participant observation was used for obtaining the images of the skewed milk. Mobile phone camera was used to capture the image of the skewed milk. In recognition of an adulterant, images were captured and features of the images extracted. Pixel values of the images were obtained for use in training and testing of the model. Classification of the adulterant in the milk was done using a back propagation algorithm as it provided output without bias.

Interviews as a method of collecting data was used to obtain information from hotels on, where they buy their milk, in what quantities, type of milk they buy(raw or processed) , how they test the milk for adulteration, the measures they take and the challenges they face in identifying adulterated milk. The method was appropriate as it enabled the hotels to highlight things that may have been left out when non-participant observation is used. Internet sources and journals were used to gather data on related information to the researcher's area of study. This was useful in identifying the gaps that would be filled in by the research.

3.4.4 Training Data Sets

Training involves providing inputs to the model for processing in order to train the model on the type of input data and the expected output of the training session. The training process was carried out with fast back-propagation until a maximum of 2000 epochs (cycles) or the maximum acceptable sum-squared error was reached. It normally requires at least two layers of neurons: the hidden layer and the output layer. The input layer is non active and is therefore not counted in the scheme .The training data was fed into the neural network model through the identified model neurons by determining the appropriate values for the weight vector *w* and bias vector *b,* using the training data and learning algorithms.

Uniformly scattered training samples, therefore, provided the network with better knowledge to make comprehensive and efficient predictions.

3.4.5 Testing and Validation of Data Sets

The process involved the use of test data to check whether the system was properly trained by observing the actual model output versus the expected output. In using the validation data set, the disparities in the output captured by error performance measures were used to adjust the weights of the neuron for the purpose of fine tuning the model.

The mean squared error (MSE) was adopted to measure the deviation between network outputs and target outputs. For the *j-th* recorded observation, MSE*j* is defined as

$$MSEj = \frac{1}{n}\sum_{j=1}^{n}(\text{q}ij - \text{q}i)2$$

Eq. (3.1)

where η is the number of testing samples (here n=10), q*i j* is the network output for the i-th testing sample in the j-th observation, and qi is the target output of the i-th testing sample (Fei & Xila, 2005).

3.4.6 Presentation of Output

Tables and graphical representations were used to illustrate the model outputs. The tables were used to display the accuracy, precision as well the recall ratio that was obtained during the classification. Graphical representation were used to display the model that was obtained from the implementation.

3.5 Research Quality

Gupta, Rawal, Narashima and Shiwani (2013) in their study, applied Accuracy, Specificity and Sensitivity to carry out their performance evaluation. The same measures were used in this research to carry out the evaluation.

3.5.1 Accuracy

The accuracy (AC) is a measure of the degree of closeness of a measured or calculated value to its actual value. This is the proportion of the total number of predictions that were correct. *tp* represents the true positive, *tn* represents the true negative, f*p* represents the false positive and *fn* represents the false negative in the equations that were used to measure accuracy. Eq. 3.3 below was used to determine accuracy.

Accuracy (AC) = (TP+TN)/ (TP+TN+FP+FN)

Eq, (3. 3)

3.5.2 Specifity

The ability of the system to correctly identify the absence of adulterant in the skewed milk image. To get it, the number of true negatives identified in the images need to be identified and divided by unadulterated milk. Eq 3.4 was used to determine specificity.

Specificity (SP) = (TN)/ (TN+FP)

Eq, (3. 4)

3.5.3 Sensitivity

The ability of the system to correctly identify adulterated milk, as computed using the Eq, 3.5. The number of true positives need to be identified and divided by adulterated milk.

Sensitivity (SE) = (TP)/ (TP+FN)

Eq, (3. 5)

3.6 Ethical Considerations

The researcher ensured that she obtained consent from the participants before she began her interviews and observations. This ensured that both parties were aware of what was happening at any one particular time during the process. The researcher maintained confidentiality of the data obtained and the personal details of the respondents. Data obtained was used for the sole purpose of the research. The researcher cited work obtained from other authors giving them the credit that was due to them.

Chapter 4: System Design and Architecture

4.1 Introduction

This chapter focused on incorporating user requirements, system analysis and design. These three are discussed in details with data analysis focusing on data collected from various users and experts, while system analysis and design takes a look at the structural requirements according to the components interrelationships and providing physical specifics on the system inputs and outputs, databases an computer processes.

The purpose of this study was to come up with an easy and effective way of detecting milk adulterated with water or hydrogen peroxide. Rapid Application Development (RAD) was used to carry out the research.

4.2 Requirement Analysis

This involved the functions and capabilities the application /proposed model was able to exhibit and execute based on the research objectives and user requirements.

4.2.1 Functional Requirements

Describes what the application does and it includes: capturing image (the application will be able to capture an image of the skewed milk using the phone camera), accept image (the application then uploads the captured image. The image should be in .JPEG, .PNG, .TIFF and .TIF. Any other file format should be rejected), analysis (the application extracts the pixel features of the image captured into RGB (Red, Green and Blue) colors/descriptors),it then detects if the milk is adulterated using water or hydrogen peroxide or not adulterated using back propagation neural network, feedback(Application gives correct feedback in terms of the adulterant identified/detected or if not detected based on the pixel image features) and finally the application gives recommendation (Application gives appropriate information to the user on what to do if adulterant is detected and the effects of the adulterant).

4.2.2 Non-Functional Requirements

Describes how the proposed model works and how it does it. These characteristics are important in a system because they make it secure, interactive and easy to use. These include: usability (these include user interface that is easy to navigate, access and has aesthetic value), security, fast response time, performance (the system must be able to process images reasonably fast to enhance the performance time. It must also be able to recover within the shortest time if the system breaks down, reliability (a backup of the entire proposed model should be provided regularly, in the event of failure, the administrator should be able to restore the proposed model and it should have the ability to convert the images provided to it into pixel values.

4.3 Proposed System Architecture

The developed proposed solution has the following main actors: customer (Hotel staff who buys milk) and system administrator (handles overall system management). Figure 4.1 gives an illustration of the proposed model architecture. The process begins by capture of the skewed milk images by the customer. The image is then converted to its pixel values by the application. The pixel values are then converted to Hue Saturation Intensity (HSI) values. The HSI values are then used by the application to train and test the neural network specifically by the back propagation neural network. The back propagation neural network classifies the adulterant and level of to the customer to enable him make informed decision on whether to buy the milk.

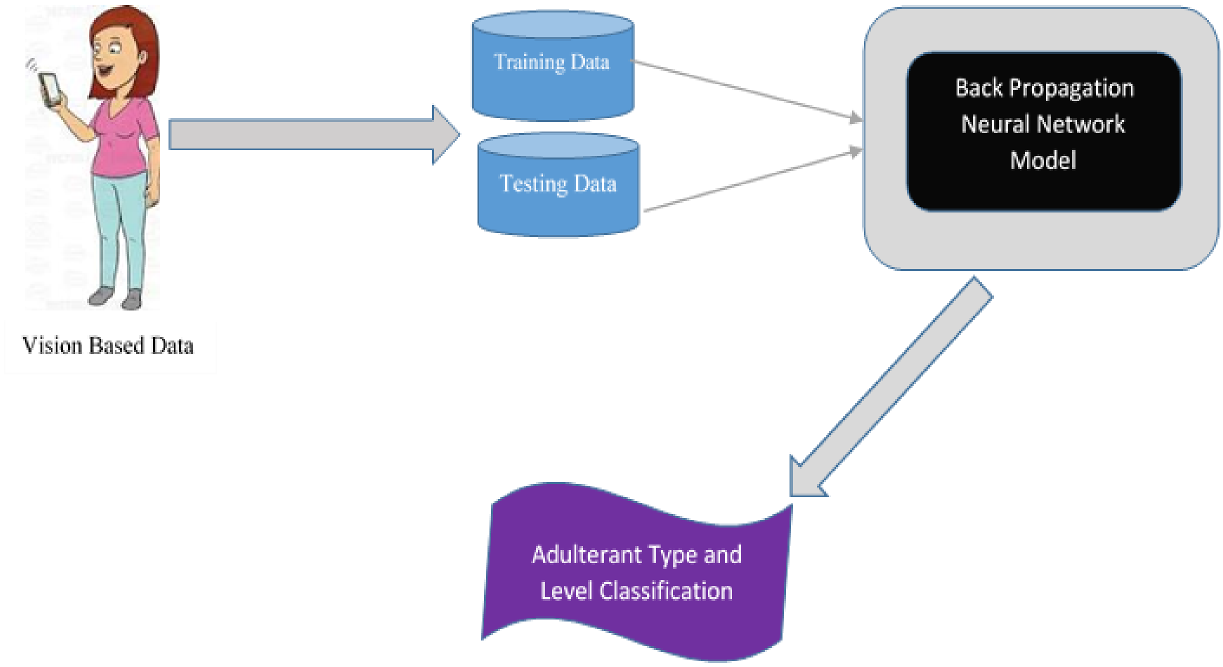

Figure 4.1: Proposed Architecture

4.4 Diagrammatic Representation of the Model

4.4.1 Use Case Diagram

It describes a set of actions that the proposed model performs in collaboration with users of the system who are the customers who will test for water or hydrogen peroxide as an adulterant in milk, boundary which represents the limits to which the system operates and the use case which are a collection of success or failure scenarios. The scope which is represented by the boundary is identification of the adulterant (water) used in milk.

Use Case: Obtain Image of skewed milk

Primary Actor:

Customer

Precondition:

Mobile phone has camera

Milk is skewed with bromothymol blue

Main Success Scenario:

Actor Intention System Responsibility

1. Customer prepares the camera

2. Capture image of milk skewed with bromothymol blue

3. Analyze the image

4. Save image

1. View image obtained
2. View report obtained
3. Exit app

Extension

At any time that the system fails to capture image, restart the camera

At any time the system fails to analyze image captured, restart the app.

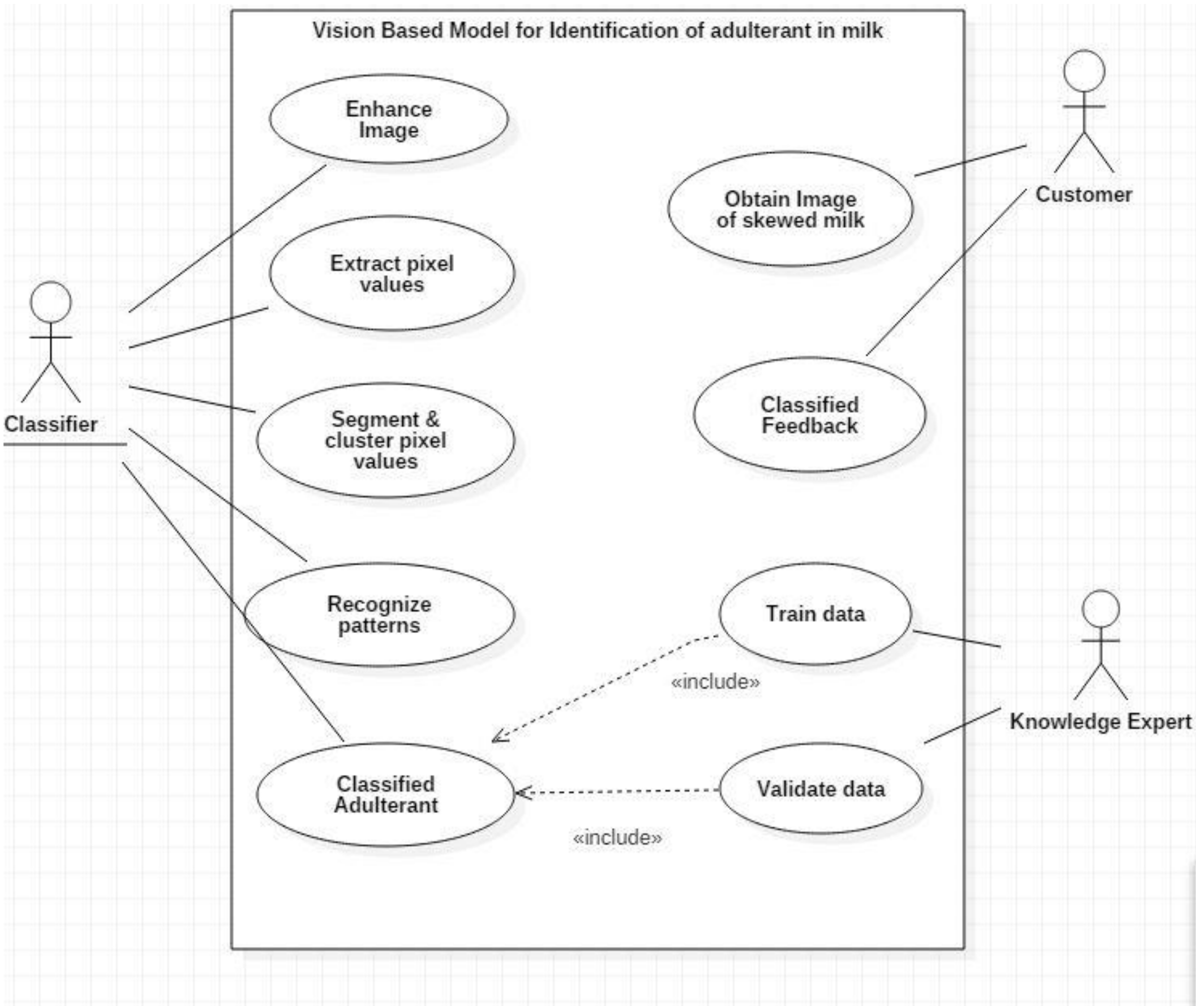

Figure 4.2: Use Case

Use Case: Process the Image

Primary Actor:

System

Precondition:

Images of the skewed milk were successfully captured

Post condition:

Adulterant/s added in the milk was correctly identified.

Main Success Scenario:

i. Image preprocessing-Involves image enhancement and noise reduction. The objective is to improve, restore and rebuild the image.
ii. Extracting pixel values-The operation consists of compressing the image, followed by extracting the geometric characteristics (corners, edges, joints) and finally the pixel RGB colors (red, green, blue).
iii. Segmentation-Involves dividing the image into regions. These regions will be used as input for the neural network as well as the number of hidden values that will be used.
iv. Recognition-Which involves determination of objects in an image and their classification (in this case involves determining type and level of adulterant.
v. Generate Output- System provides the identified adulterant and level based on the images provided to it.

Use Case: Receive Feedback

Primary Actor:

Farmer

Precondition:

Successful identification of the adulterant and level by the system

Post condition:

Identified adulterant name

Main Success Scenario:

Actor	System Responsibility
1. Request for adulterant identification	
	2. Return the output of clustering
2. View analysis	
3. Exit the system	

4.5 Domain Model

The domain model as illustrated in Figure 4.3 is used in illustrating meaningful conceptual classes in problem domain and also represent real world concepts pertinent to the domain that need to be modeled into a software. This is created during object oriented analysis to decompose the domain into concepts or objects of the real world.

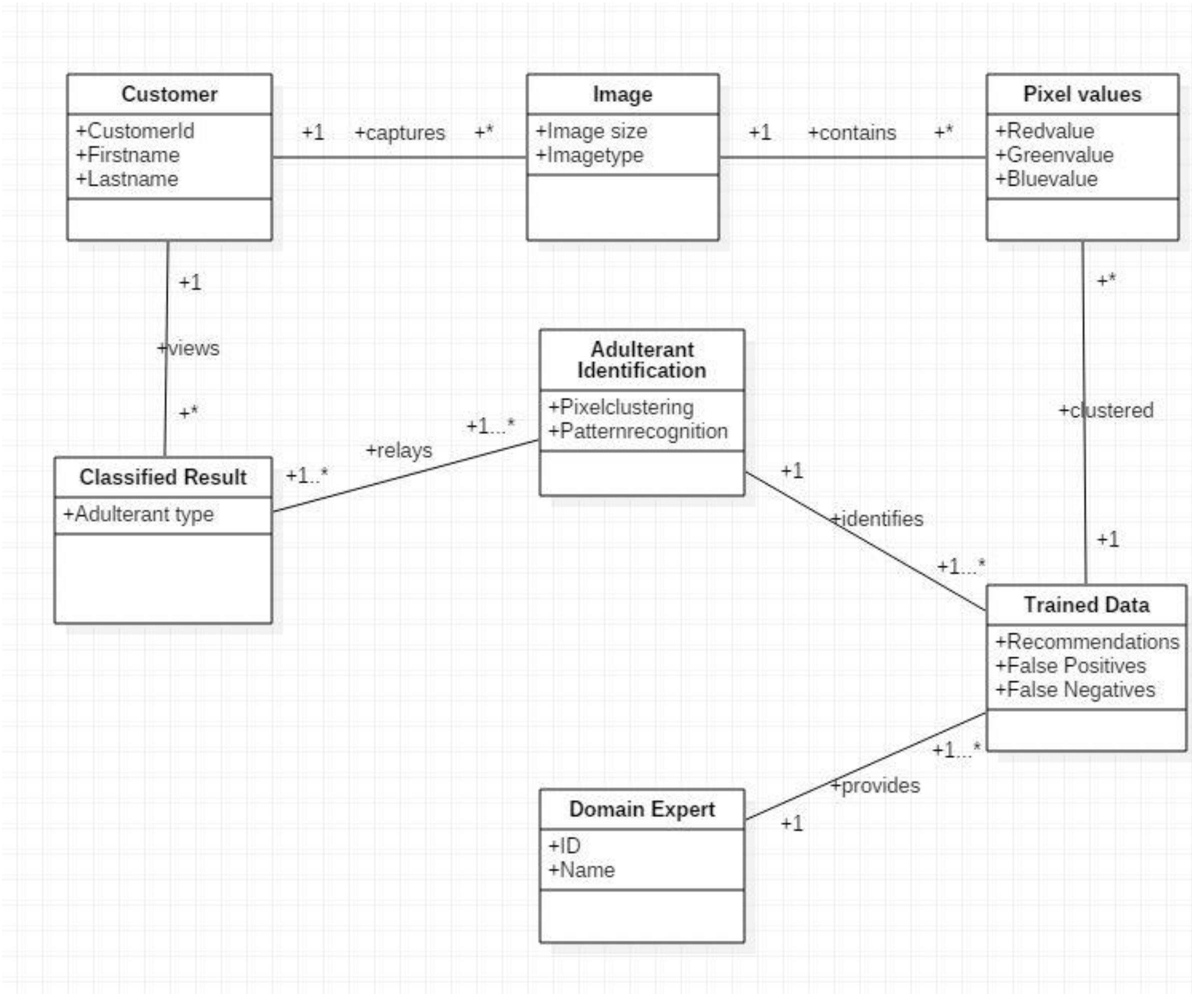

Figure 4.3: Domain Model

4.6 Sequence Diagram

As illsutrated in Figure 4.4 the sequence diagram is used to demonstrate the object interactions arranged in time sequence. It depicts the objects and classes in this scenario(vision based model for adulterant identification) and the sequence of messages exchanged between the objects needed to carry out the functionality of the proposed model.

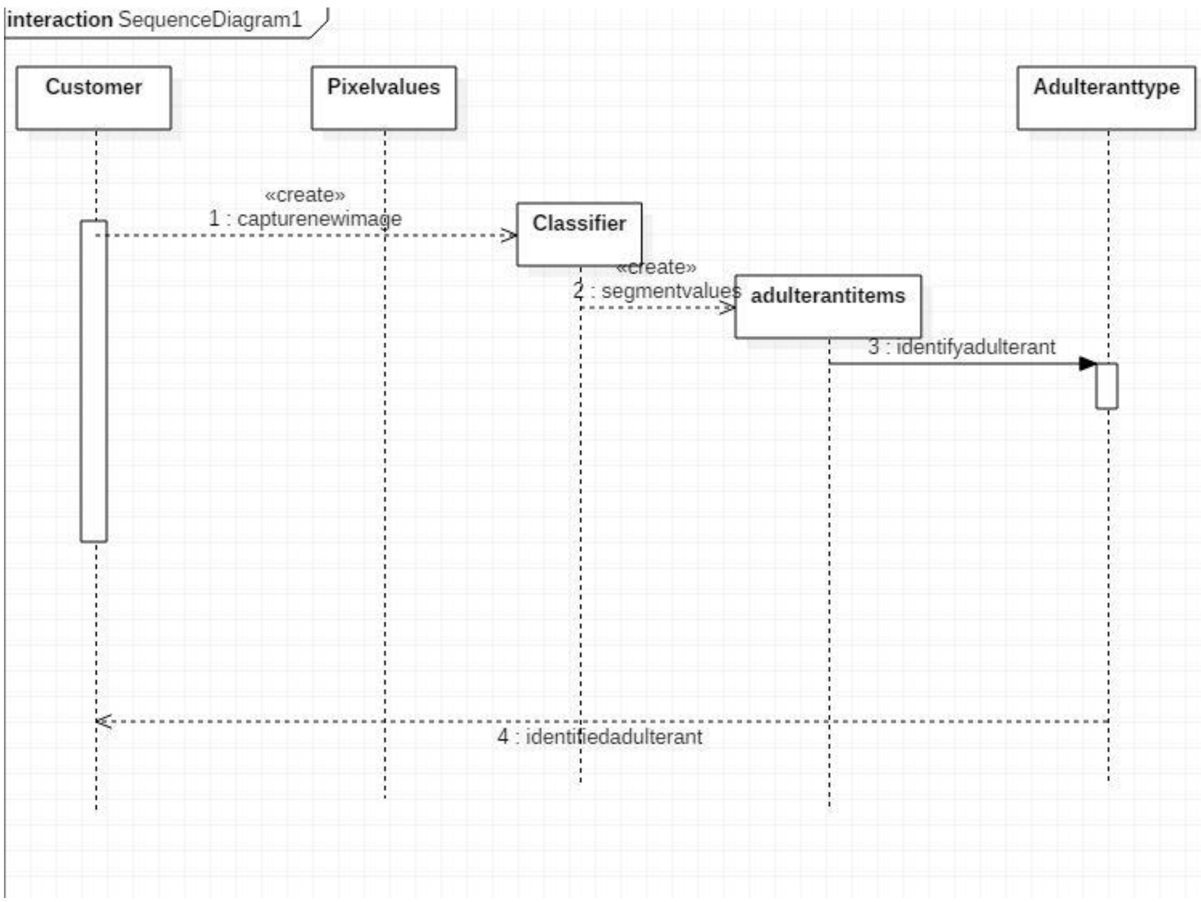

Figure 4.4: Sequence Diagram

4.7 Context Diagram

Figure 4.5 defines the boundary between the proposed model and its environment showing the entities that interact with it. It is a high level view of the proposed model. The main users of the proposed model are the domain expert and the customers. The customer provides an input to the system and receives the output from the system. The domain expert provides the expert knowledge on the adulterants used in milk and the levels.

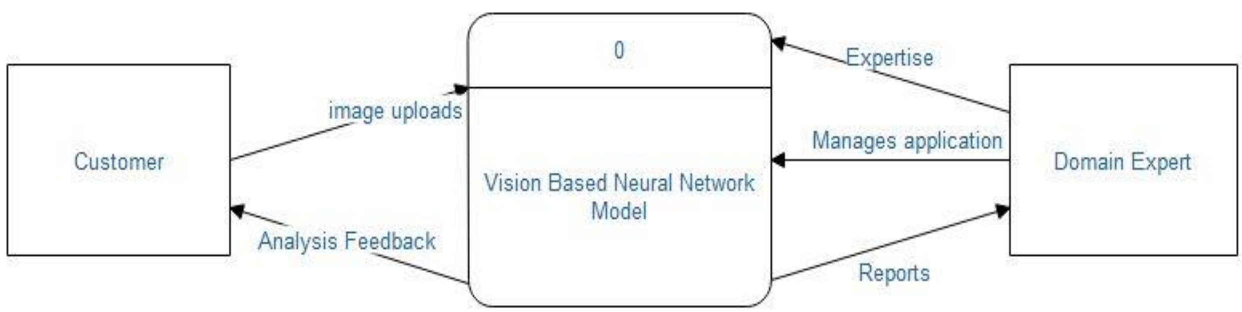

Figure 4.5: Context Diagram

4.8 Entity Relationship Diagram

All the entities used in the database to save and retrieve images that can be accessed by the user through the app and the administrator through the backend are illustrated below in Figure. It graphically represents entities stored in the database and the relationship between the entities. MySQL database was used to store the data for retrieval purposes.

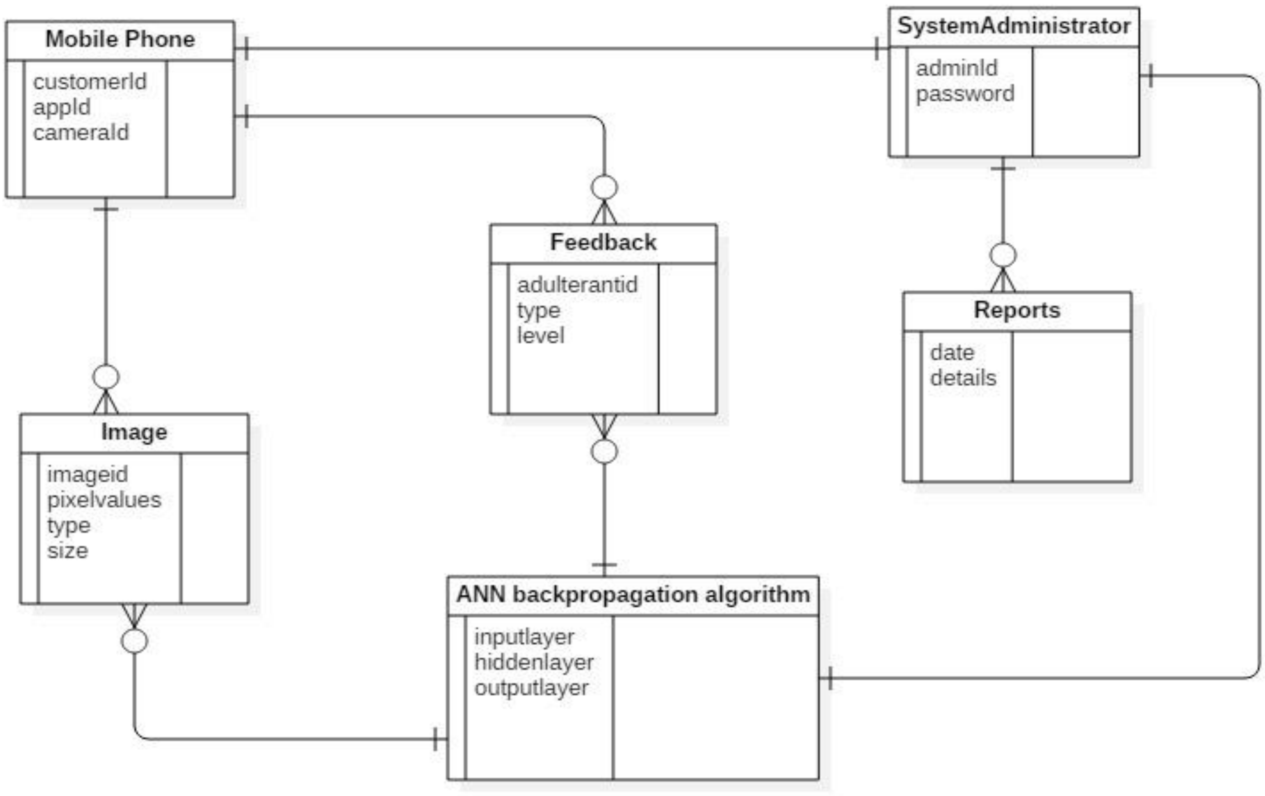

Figure 4.6: ER-Diagram

Chapter 5: Implementation and Testing

5.1 Overview

The chapter focuses on implementation and testing of the proposed model. The implementation part explores how artificial neural network was used in identifying or grading an image. The array of colors (red, green, blue) was used as input in the neural network. Each input neuron represents color information in the image, and each output neuron corresponds to an image. The testing bit focuses on usability and functional testing to verify if the system is attaining its objectives.

5.2 Model Components

The application comprises of a front-end and back-end subsystems: AMI is the front end of application accessed and used by customers and artificial neural network is the back-end subsystem where processing and classification of image is done. Android 8 was used to develop the front end application system, JDK 8 programming language was used to develop the backend and MySQL was used to develop the database. The front and backend models are explained below:

5.2.1 Image Processing

Camera- The model uses a digital camera to obtain images of skewed milk with bromothymol blue. Digital imaging is aimed at recognizing objects of interest in an image by utilizing electronic sensors and advanced computing techniques with the aim of improving image quality parameters. It also involves image processing and pattern recognition techniques. Image processing techniques deal with image enhancement, manipulation and analysis of image.

5.2.2 Neural Network

Artificial neural network back propagation is a prominent algorithm for computing the gradients. Backpropagation procedure can be explained as follows: The network is given the input and the neurons are activated until the highest layer. Using the equations of backpropagation the output error of the highest layer is calculated.

5.2.2.1 Input Layer

This is the first layer of the network. Represents the raw information fed into the network. The input layer nodes are passive doing nothing but relaying the values from their single input to their multiple inputs to one or more hidden layers.

5.2.2.2 Hidden Layer

The hidden layer is where the actual processing is done via a system of weighted connections. It consists of hidden neurons and the number of hidden neurons are critical as this prevents the problem of overfitting or under fitting. Its purpose is to enable the neural network to produce better results of the expected output for the given input. Each neuron in the hidden layers calculates its output using the activation function, which is a function of its weighted sums.

5.2.2.3 Output Layer

This is the last layer on the neural network. The output is provided after all the inputs have been processed and presents a pattern to the external environment. The output of a neuron is the result of the action function, which depends on the weights of the connection and the bias. All images are scaled to the same size (width and height) and small to be easy and quick to learn. The sizes of the images shall be determined on the size of the input vector and the number of neurons. The transfer function for this type of problem is called sigmoid function. The rate of learning has values in the range [0.1] and the error it is recommended to have less than 0.1

5.3 Model Implementation

5.3.1 Image capture of the skewed milk

The image of the skewed milk was captured using a mobile phone camera of 4 megapixels through the app. The image was set as preview to enable user view the image they had taken to confirm it was sufficient before submitting for processing. The images were taken at a close range to avoid a lot of light which would interfere with the intensity of the image.

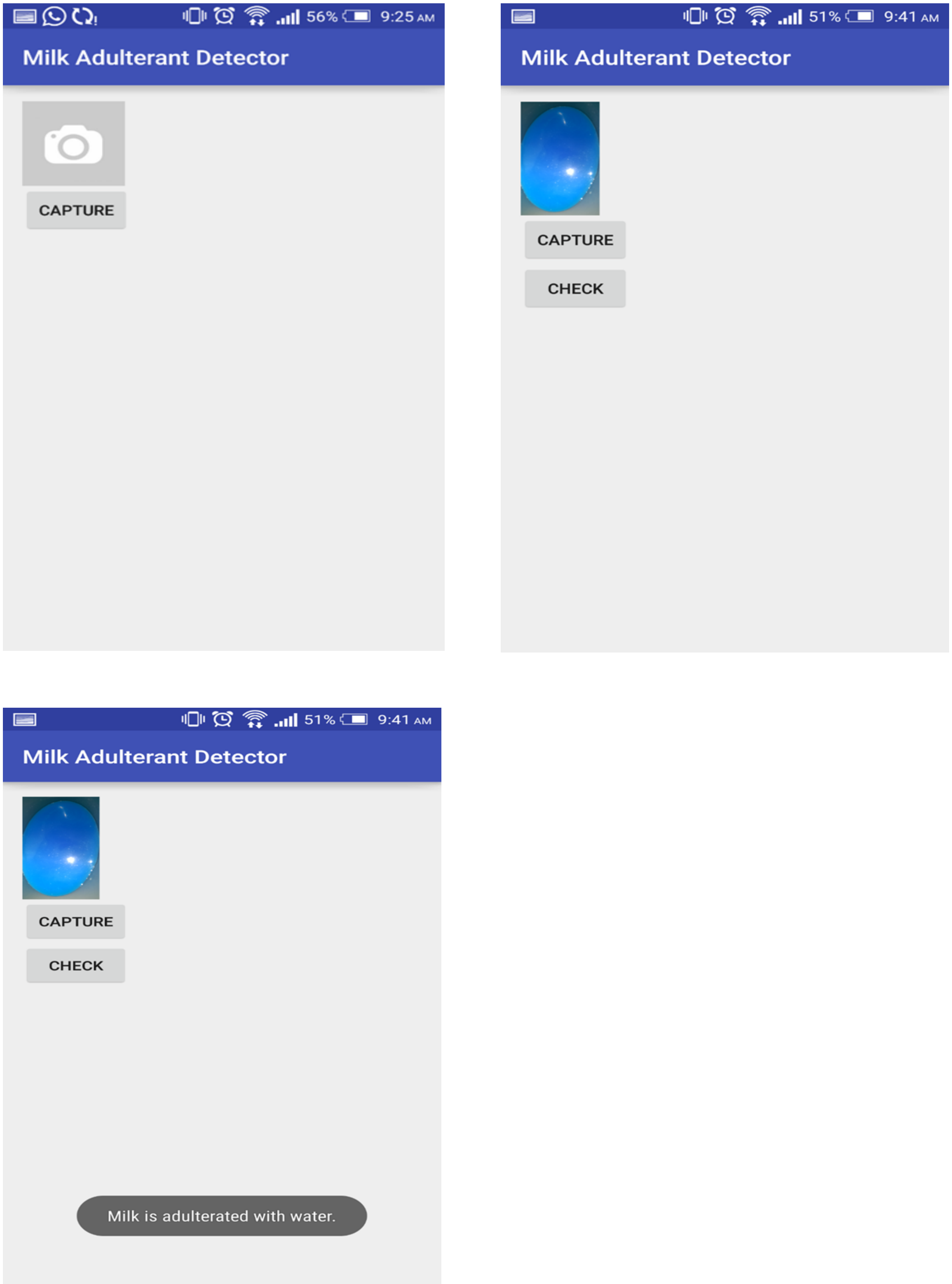

Figure 5.1: Image capture & feedback of Milk Adulterated with Water

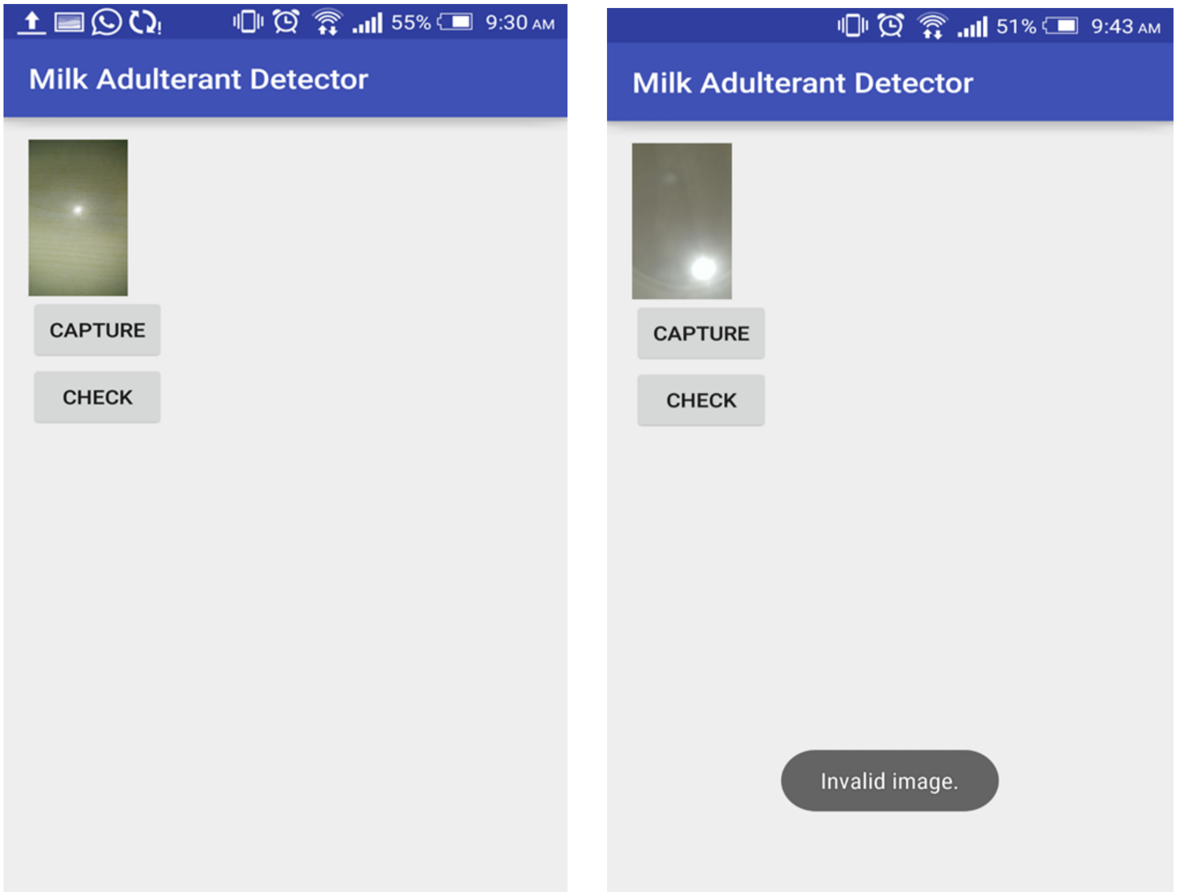

Figure 5.2: Image capture & feedback Water Image

5.3.2 Image Processing

The captured and stored images were then converted in to ten color descriptors (red, green, blue, hue, saturation, value, luminosity and the three relative colors i.e. Relative red, relative green and relative blue) were the used to evaluate the information from the images. RGB (red, green, blue) colors are also converted to HIS (hue, saturation, value) color space as an alternative way of representation and is derived from RGB color space. While RGB is generally related to the spectral response of the human eye, HSV is more representative of the way humans perceive colors, and sometimes it is also more convenient for image processing.

5.3.3 Dataset used

Data was obtained from the images that were captured. The pixel values were extracted from images that were presented after image preprocessing which involves digitization and compression; enhancement, restoration, and reconstruction; matching, description, and recognition. The images were converted into RGB (red, green, blue and relative red, relative blue and relative red) values that were later converted to the Hue, Saturation and Intensity (HSI) values.

5.3.4 Implementation of the artificial neural network

The backpropagation algorithm was implemented in the research to aid in the identification and classification of the resulting adulterant (water) in milk. The algorithm consisted of several elements such as input layer, hidden layer and output layer. The original images are fed into the input layer and the principal components or feedback of the set of images are obtained at the output layer. The neural network approach selected, compressed the images, with little loss of information and finally reconstructed the image for analysis. The perceptron was then trained via the backpropagation algorithm, using a set of images and setting the desired output equal to the input image. Training in this context is when we update the weights and try to find values that give us good results.

5.4 Model Testing

This involves testing the functionality and reliability of the proposed model. There are several parameters to be tested in the model as described below:

5.4.1 Functionality

The functionality of the proposed model is key since the solution was expected to meet the objectives of the study and user requirements. Some of functionalities of the model to be tested are; does the application validate user input to ensure only images are uploaded, Are the different customers who uploaded different images able to receive separate analysis of their respective images, does the application identify the correct adulterant etc.

15 respondents were requested to state the functionality satisfaction level while using the application and their responses are summarized below in Figure: 5.2

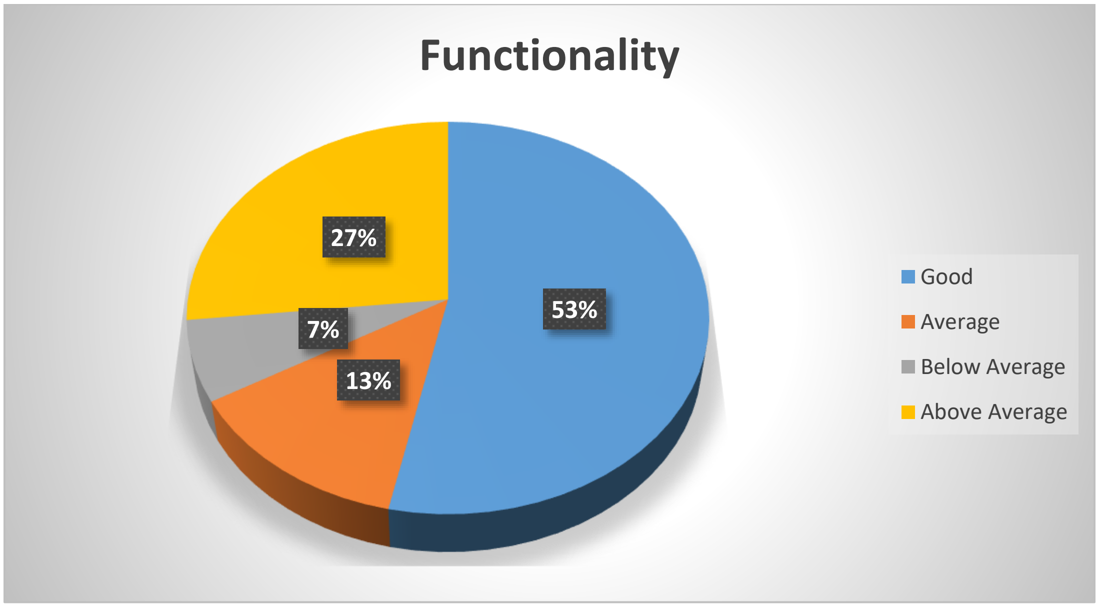

Figure 5.2: Functionality

5.4.2 Reliability

Probability that the system is able to accomplish assigned task successfully within a specified time and is redundant. Is there a difference in the presence of noise between the original image uploaded and the images used for analysis?

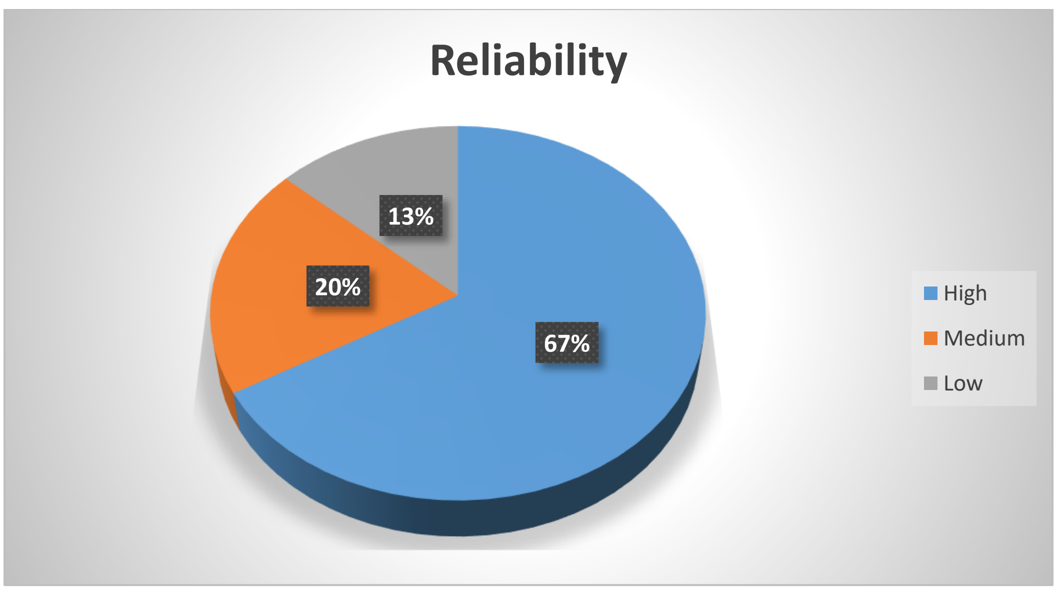

Figure 5.3: Reliability

5.4.3 Usability Testing

All respondents who participated in the usability testing gave their views and evaluations on how easy it was to use the application model. Below are their various responses as outlined in Figure 5.4. It is an important test since the users interacted with the model and were able to give their views on whether it will address the challenges they experience when testing the quality of milk.

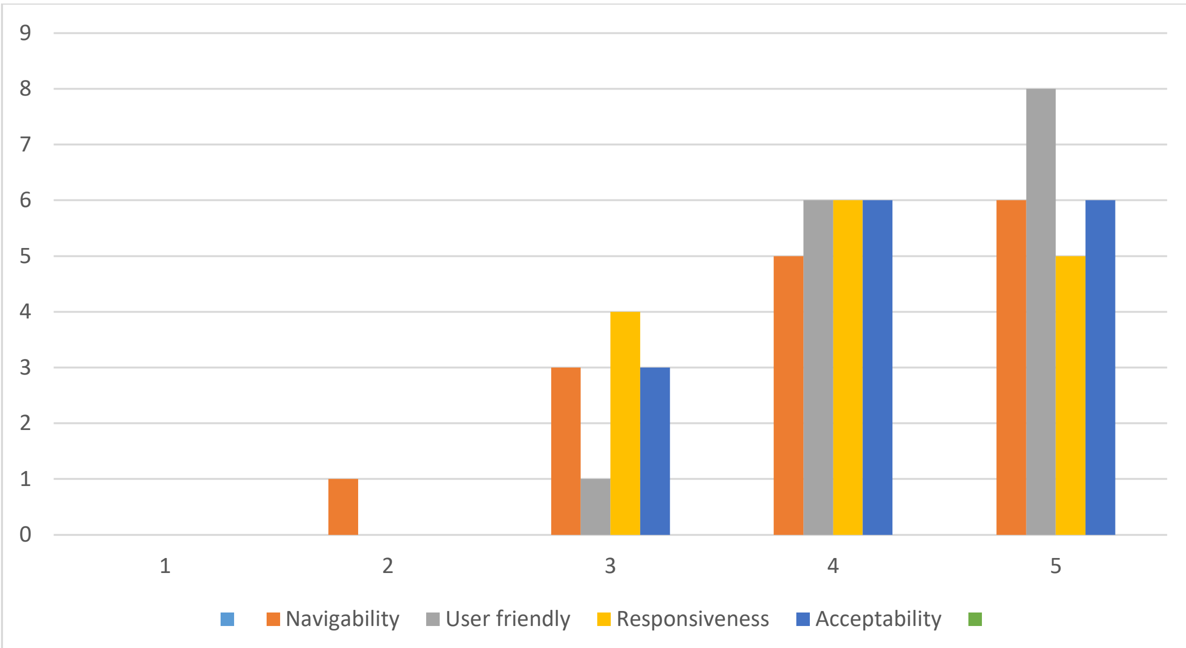

Figure 5.4: Usability Testing

Chapter 6: Discussions

6.1 Overview

This chapter focuses on discussing the proposed model in terms of performance, practicability and relevance by analyzing findings in terms of the research objectives. The proposed model was tested for correct classification on the basis of accuracy, precision, and the error obtained to ascertain its functionality.

6.2 Adulterated milk characteristics.

The first objective in section 1.3 was to determine the characteristics of adulterated milk. From the research findings, it was found that, most respondents had little knowledge on how adulterated milk looked like, tasted and didn't have the basic knowledge of performing simple tests to determine if the milk they have bought is adulterated. This creates a situation where the respondents are unable to make informed decision when buying milk.

6.3 Current methods approaches and associated challenges

The second objective was to investigate the current methods used in testing adulterants in milk and the challenges associated with the methods. From the research findings, it was found that most respondents had limited knowledge on the current methods that are used to test milk for adulteration that range from simple/conventional to complex biological/electrical methods. Respondents also indicated that even if they know how to carry out the tests, they face numerous challenges such as lack of appropriate equipment for testing, unwillingness from the small scale trader/farmer to allow them test the milk, lack of resources to buy the necessary equipment for testing the milk, dependency on visual and taste examination which was prone to errors and inaccuracies.

6.4 Current technologies for adulterant testing

The third objective was to analyze the current models and architectures used in detecting adulterants in milk. Research findings shows that few respondents were conversant with the current methods used, such as freezing point method, lactometer and digital milk analyzer. The literature review discusses vision based model through the use of a mobile and provides more accurate output compared to the human visual examination, easy to use compared to the current technologies and also easy to acquire compared to the current technologies.

6.5 Vision based model application

The fourth objective was to develop a vison based model application for testing of water as an adulterant in milk for Nairobi County. The model developed in this research gave more accurate classification based on the fact that it was implemented based on back propagation neural network algorithm. The use of the machine learning algorithm enabled the model to work much faster thus fast feedback to the customer. Combining machine learning and computer vison strengths, enabled the model provide accurate results. Research findings show that most respondents found it necessary to have an application in their mobile that they would use to test the quality of milk whenever they wanted at no extra cost.

6.6 Vision Based model application testing and validation

The last objective was to carry out testing of the proposed model. The proposed model was validated for accuracy, specificity and sensitivity using the confusion matrix. According to (Kohavi & Provost, 1998) confusion matrix contains information about actual and predicted classifications done by a classification system and it is reliable in assessing the accuracy of an image classification. A total of 20 instances was used to train and test the proposed model and the findings are as illustrated in the below table:

N=20	Predicted: No	Predicted: Yes
Actual: No	8	2
Actual: Yes	2	10

6.7 Proposed Model shortcomings

i. The model did not consider all the features of the image including shape ,texture and edges that would enable the neural network provide more accurate results
ii. The model was limited to water as an adulterant only
iii. The model was limited to mobile phones that had cameras.
iv. The model did not consider the stress and challenges in obtaining the acid base indicator (bromothymol blue) and the technical knowledge on how to use it, especially on rural farmers and customer.

Chapter 7: Conclusions and Recommendations

7.1 Conclusions

This study reveals the importance of consumers (hotels) in Nairobi being able to test quality of milk on their own because of lack of appropriate legislation for regulating small scale milk processing. As highlighted in the interview customers face numerous challenges when buying milk from small scale traders as there are no mechanisms or technologies for individual milk consumers to use to detect adulterants. One of the major problems is that most customers don't have the technical knowledge of how to perform organoleptic tests to determine the quality of milk, lack of appropriate equipment and chemicals to perform the test, last but not least, most small scale farmers wouldn't allow customers to perform the test on their milk before buying hence, the rise of adulterated milk in the market by them for profit purposes.

The research lays its emphasis by taking advantage of computer vision techniques as well as machine learning algorithms for the classification of water and hydrogen peroxide adulterants. Customers are able to feedback on the quality and health of milk and make a decision on whether to buy the milk. If the model is adopted it will help curb the rise of adulterated milk in the market because small scale farmers will be held responsible if found guilty of the practice and consumers will be more knowledgeable on where to get quality, healthy milk.

7.2 Recommendations

The vision based model was appropriate for consumers with smart phones that had digital cameras. The researcher however noted that there was more that could be done and gave the below recommendations: To increase the awareness in use of the system, public awareness should be carried out on the effects of milk adulteration; Due to the realization of the vision model solution, it can be advanced by collaborating with Kenya Bureau of Standards (KEBS); Application can be expanded to include multi-sensors for better classification/segmentation to improve on the output of the model and also remove the aspect of using bromothymol blue as an acid base indicator for the image; The application can also expanded to classify/identify other adulterants apart from water and hydrogen peroxide.

7.3 Future Recommendations

The researcher recommends that the application should include developed to process images from other platforms such as portable cameras, secondly the application can be built to support local languages since most of the milk small scale milk farmers are from the rural areas, the application can be extended to consider other features such as texture and shape of the image to give a more accurate classification.

References

Omore, A. , Staal, S. & Randolph, T. (2004). Overcoming Barriers to Informal Milk Trade in Kenya. *International Livestock Research Institute*, 17-18.

Ahangar, R. G., Yahyazadehfar, M. & Pournaghshband, H. (2010). The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for Prediction Stock Price in Tehran Stock Exchange. *(IJCSIS) International Journal of Computer Science and Information Security*, 38-46.

Anjali, R. M., Bindushree, M., Kumar, K.S. N., Vaishali, S.S. , & Rajashekhar, B.S. (2016). Implementation of Standardized Digital Milk Quality Analyzer. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 1523-1526.

Ayza, A. & Belete, E. (2015). Food Adulteration: Its Challenges and Impacts. *Food Science and Quality Management*, 50-56.

Barnes, P. J. (1996). Adulteration and Fraud Threaten Safety. *Nutrition & Food Science*, 23 - 26.

Chaudhari, S. I., Pathare, M. S., Haral, S., Mule, Y. & Katti, P. (2016). A Survey On Detection Of Unhealthy Region Of Plant Leaves By Using Image Processing. *International Research Journal of Engineering and Technology*, 1318-1320.

Das, S., Goswami, B., & Biswas, K. (2016). Milk Adulteration and Detection:A Review. *Sensor Letters*, 4-18.

Esau, I. & Rieber, G. (2010). On Application of Artificial Neural Network Methods in Large-eddy Simulations with Unresolved Urban Surfaces . *Modern Applied Science* , 3-11.

Ettema, F. (2013). Dairy Development in Kenya. *Kenya Dairy Sector*, 1-5.

Farmer, T. O. (2009). Quality of Milk Worrying. *Bio Vision*, 44.

Fei, T. & Xila, L. (2005). Samples Selection for Artificial Neural Network Training in Preliminary Structural Design . *Tsinghua Science and Technology*, 233-239.

Filho, P. M. & Pereira, E. R. (2013). Digital image analysis – an alternative tool for monitoring milk authenticity. *Analytical Methods,* 3669-3674.

Gupta, N., Rawal, A., Narasimhan, D. V., & Shiwani, S. (2013). Accuracy, Sensitivity and Specificity Measurement of Various Classification Techniques on Healthcare Data. *IOSR Journal of Computer Engineering (IOSR-JCE)*, 70-73.

Hanson, J., Joy, A., Francis, J. & MG, A. (2016). Survey on Image Processing Based Plant Leaf Disease Detection. *International Journal of Engineering Science and Computing*, 2653-2655.

Jooa, D., Hongb, T. & Han, I. (2003). The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors. *Expert Systems with Applications*, 69-75.

Sasithra, S. (2014). Review On Classification Based On Artificial Neural Networks. *International Journal of Ambient Systems and Applications (IJASA)*, 11-18.

Muthukannan, K., Latha, P., Nisha, P. & Selvi, R. (2015). An Assessment on Detection of Plant Leaf Diseases and Its Severity Using Image Segmentation. *International Journal of Computer Science and Information Technology Research I*, 63-69.

Kurwijila, L. R. (2006). Hygienic milk handling, processing and marketing: Reference guide for training and certifi cation of small-scale milk traders in Eastern Africa. *Trainers Guide Volume 1: Hygienic Milk Handling, Processing and Marketing*, 1-128.

Seetha, M., Muralikrishna, I.V. & Deekshatulu, B. (2008). Artificial Neural Networks and Other Methods Of Image Classification. *Journal of Theoretical and Applied Information Technology*, 1039-1053.

Masawat, P., Harfield, A., & Namwong, A. (2015). An iPhone-based digital image colorimeter for detecting tetracycline in milk. *Food Chemistry*, 23-29.

McDonald, H.J. (2000). Milk quality and globalization:metaphors of modernity in Northwestern Michoacan Mexico.

Moturi, W., Obare, G. & Kahi, A. (2014). Establishing the Dominant Milk Marketing Channel in Nyandarua District,Kenya. *IOSR Journal of Agriculture and Veterinary Science (IOSR-JAVS)*, 38-43.

Mugambi, Kimencu, D., Mwangi, M., Wambugu, Kairu, S. & M.M, G. A. (2014). Assessment of performance of smallholder dairy farms in Kenya: an econometric approach. *Journal of Applied Biosciences*, 7891-7899.

Muriuki, H. (2011). Dairy Develeopment in Kenya. *Dairy Reports*, 1-18.

Ndungu, T. W., Muliro, P. S., Omwamba, M., Oosterwijk, G., & Jansen, A. (2016). Quality control of raw milk in the smallholder collection and bulking enterprises in Nakuru and Nyandarua Counties, Kenya. *African Journal of Food Science*, 70-78.

Shetty, P. B. (2014). Comparison of Different Methods Used For Detection of Urea In Milk By Quantification Of Ammonia. *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering*, 7858-7863.

Pandelea, A.-E., Budescu, M. & Covatariu,T.G. (2015). Image Processing Using Artificial Neural Networks. *Buletinul Institutului Politehnic Din IAŞI*, 9-21.

Filho, P., Santos, P. M., (2013). Digital Image Analysis-an alternative tool for monitoring milk authenticity. *Analytical Methods*, 3669-3674.

Reitermanova, Z. (2010). Data Splitting. *WDS'10 Proceedings of Contributed Papers,*, 31-36.

Rodvold, D. M. (1999). A Software Development Process Model for Artificial Neural Networks in Critical Applications. *IEEE* , 3317-3322.

Santos, P. M., Wentzell, P. D. & Filho, E. R. (2012). Scanner Digital Images Combined with Color Parameters: A Case Study to Detect Adulterations in Liquid Cow's Milk. *Food Analytical Methods*, 89-95.

Seetha, M., Sunitha, K. V. & Devi, G. M. (2012). Performance Assessment of Neural Network and K-Nearest Neighbour Classification with Random Subwindows. *International Journal of Machine Learning and Computing*, 844-847.

Singh, R. K., & Prajneshu (2008). Artificial Neural Network Methodology for Modelling and Forecasting Maize Crop Yield. *Agricultural Economics Research Review*, 5-10.

Sivaji, A., Abdullah, A. & Downe, A. G. (2011). Usability Testing Methodology: Effectiveness of Heuristic Evaluation in E-Government Website Development. *ACM Digital Library*, 24-26.

Smith, T. C. & Frank, E. (2016). Statistical Genomics: Methods and Protocols. In *Introducing Machine Learning Concepts with WEKA* (pp. 353-378). New York: Springer.

Soomro, A. A., Khaskheli, M., Memon, M. A. & Soomro, R. N. (2014). Study On Adulteration and Composition of Milk Sold at Badin. *IMPACT Journal*, 57-70.

Soomro, A. H., Nizamani, Z. A., Barham, G. S. & Khaskheli, M. (2015). Surveillance of milk adulteration and its impact on physical characteristics of milk. *Open Access Bioscience Research*, 1-16.

Taylor, B., Darrah, M. & Moats, C. (2003). Verification and validation of neural networks: a sampling of research in progress . *Intelligent Computing: Theory and Applications*, 8-16.

Trivedi, U., Donepudi, L., Kapse, H. & Panchal, C. (2009). Potentiometric biosensor for urea determination in milk. *Sensors and Actuators B:Chemical* , 206-266.

Vishal, V., Srinidhi, S., Srividhya, S., Kumar, K. V. & R.Swathika. (2016). A Survey and Comparison of Artificial Intelligence Techniques for Image Classification and their Applications. *International Journal of Science and Research*, 187-193.

Velasco-Garcia, M. N. & Mottram, T. (2003). Biosensor Technology addressing Agricultural Problems. *Biosystems Engineering*, 1-12.

Zagorska, J. & Ciprovica, I. (2013). Evaluation of Factors Affecting Freezing Point of Milk. *International Journal of Biological, Biomolecular, Agricultural, Food and Biotechnological Engineering*, 106-111.

Appendices

APPENDIX A: QUESTIONNAIRE

Vision Based Model Platform Questionnaire

Dear Respondent,

I am a Masters student in the Faculty of Information Technology, Strathmore University conducting a research entitled VISION BASED MODEL FOR IDENTIFICATION OF ADULTERANT IN MILK. You have been selected to form part of this study. I kindly request you to complete the questionnaire below. The information requested is needed for academic purposes only and will be treated in strict confidence.

Kind Regards,

Jacklyne Atieno Kobek.

* Required

1. **Select your hotel category** *
   Mark only one oval.
   - Town hotel
   - Vacation hotel
   - Criteria restaurant
   - Criteria lodgees/villas

2. **On a scale of 1-5 with 5 being the highest and 1 lowest, rate your experience in using a smartphone** *
   Check all that apply.
   - 5
   - 4
   - 3
   - 2
   - 1

Section A: Purchasing Power

Mark only one

3. **A1: How often do you buy milk?** *
   Mark only one oval.
   - Daily
   - Weekly
   - Monthly

4. **A2: Quantity of milk you buy** *
Mark only one oval.

- ⬭ 10-100 litres
- ⬭ 100-1000 litres
- ⬭ 1000-5000 litres
- ⬭ 5000 and above

Section B: Milk type and testing approaches

5. **B1: Where do you buy your milk?**
Mark only one oval.

- ⬭ Supermarket
- ⬭ Kiosks
- ⬭ Farmer
- ⬭ All of the above

6. **B2:What kind of milk do you buy**
Mark only one oval.

- ⬭ Processed milk
- ⬭ Raw milk

7. **B4: If your answer above is raw milk,how do you know if water has been added to the milk ?** *

____
____
____
____
____

8. **B5: What challenges do you face with the current methods of testing if water is added to the milk?** *

____
____
____
____
____

Section C: Usability

The app is called AMI, kindly install it in an Android smartphone

9. **C1: Were you able to install the app?** *
Mark only one oval.

- Yes
- No
- Maybe

10. **C2: If your answer was No or Maybe, kindly list the challenges experienced?** *

11. **C3: How would you rate the application? Tick all the appropriate attribute** *
Mark only one oval.

- Userfriendly
- Navigability
- Easy to learn
- Reliability
- Responsiveness

12. **C4: Would you recommend the app to be developed for other platforms?**
Mark only one oval.

- Windows Phone
- Blackberry
- IOS

13. **C5: Any comments, suggestions, or recommendation about the application** *

APPENDIX B: ORIGINALITY REPORT

turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Jacklyne Kobek
Assignment title:	2016 Plagiarism Check (GS)
Submission title:	Proposal
File name:	89547-Kobek_Jacklyne_Atieno_-_...
File size:	1.7M
Page count:	62
Word count:	12,187
Character count:	69,513
Submission date:	06-Apr-2017 05:52PM
Submission ID:	795359872

Research

ORIGINALITY REPORT

18%	12%	9%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS