An Intelligent framework for Quality Inspection and Control in Aquaponics, based on Computer-vision, Artificial Intelligence, and Knowledge modeling

by

Rabiya Abbasi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mechanical Engineering
University of Alberta

Abstract

The burgeoning global population and subsequent food security issues have attracted much attention toward sustainable food production systems. As one of the emerging vertical farming methods, aquaponics promises to be a sustainable alternative to food and environmental problems. Aquaponics in integrated production of fish and hydroponic crops with recirculation of the aquaculture effluent used by the plant as fertilizer. This technique offers high water efficiency, a faster growth rate, and high crop yields. Despite all the advantages offered by this technology, its implementation on a commercial scale is hindered by many technical and economic factors, which can be addressed by integrating smart technologies, automation, and control. This thesis, therefore, aims to support research towards developing solutions for crop quality control and viable commercial aquaponic. For this purpose, the status of digitization in the agriculture industry is first investigated, and potential research gaps in aquaponics are identified. Next, an ontology model is formalized to store relevant knowledge pertaining to different domains of the aquaponic 4.0 system, which can be extracted and used to enable data-driven decisions related to crop quality, facility layout, and system operations. An interactive decision support tool is then developed that uses knowledge from the ontology model to automatically determine the design of grow channels in hydroponic units based on crop characteristics for enhanced crop growth and quality. After that, a cloud-based dashboard is developed for the acquisition of sensors' data and crop images from the aquaponic facility, which is also linked with the ontology model and other quality assessment tools developed in this research. A crop disease detection system is then developed to detect and identify diseases in leafy green crops, followed by the development of the model that effectively assesses the quality of lettuce crops based on foliage color. Another model

is then developed to estimate the crop morphological traits in a particular area and plant site spacing for healthy growth of the crop. Finally, a cloud-based application that acts as a decision support system is designed where all the models are deployed. Implementing this decision support system will assist agriculturalists in various decisions related to growing crops in aquaponics and crop quality control and management, thereby paving the way towards developing a smarter and sustainable food production system.

Preface

This thesis is the original work by Rabiya Abbasi. Five journal papers and three conference papers related to this thesis have been submitted, accepted, or published and are listed below. As such, the thesis is organized in paper format by following the paper-based thesis guideline.

1. **Rabiya Abbasi,** Pablo Martinez, Rafiq Ahmad, “Digitization in agriculture – a systematic literature review on agriculture 4.0”, *Journal of Smart Agricultural Technology, 2021:* https://doi.org/10.1016/j.atech.2022.100042.
2. **Rabiya Abbasi,** Pablo Martinez, Rafiq Ahmad, “An ontology model to represent aquaponic 4.0 system’s knowledge”, *Journal of Information Processing in Agriculture, 2021*: https://doi.org/10.1016/j.inpa.2021.12.001
3. **Rabiya Abbasi,** Pablo Martinez, Rafiq Ahmad, “An ontology model to support the automated design of aquaponic growbeds”, *Procedia CIRP, 2021*: https://doi.org/10.1016/j.procir.2021.05.009.
4. **Rabiya Abbasi,** Pablo Martinez, Rafiq Ahmad, “Data acquisition and monitoring dashboard for IoT enabled aquaponic facility”, *The 10th International Conference on Control, Mechatronics and Automation (ICCMA 2022), IEEE Proceedings, 2022.* (Accepted)
5. **Rabiya Abbasi,** Pablo Martinez, Rafiq Ahmad, “Crop Diagnostic System: A robust disease detection system for leafy green crops grown in aquaponic facility”, *Journal of Artificial Intelligence in Agriculture, 2022.* (submitted)
6. **Rabiya Abbasi,** Pablo MartinezRafiq Ahmad, “Non-destructive identification of foliage chlorosis in lettuce crop grown in aquaponic facility using image processing”, *Journal of Engineering, 2022.* (submitted)
7. **Rabiya Abbasi,** Pablo Martinez, Rafiq Ahmad, “Estimation of morphological traits of foliage and effective plant spacing in NFT-based aquaponic system ”, *Journal of Artificial Intelligence in Agriculture, 2022.* (submitted)

Without access to modern farming techniques or machinery, let alone science-based climate and weather data, farmers' livelihoods hinge precariously on a changing environment that they are struggling to understand.

-U.S. AGENCY FOR INTERNATIONAL DEVELOPMENT

Acknowledgements

First and foremost, I would like to praise and thank Almighty God for granting me countless blessings, knowledge, and the opportunity to accomplish this thesis.

I would like to take this opportunity to thank everybody who has been instrumental in completing this thesis. Apart from my efforts, this thesis is a result of the constant guidance and motivation of many others.

I would like to thank my supervisor Dr. Rafiq Ahmad, associate professor of the Department of Mechanical Engineering, University of Alberta, for his constant support and encouragement. This thesis would not have been possible without whom. Thanks to Dr. Pablo Martinez, senior lecturer at Northumbria University, for his persistent guidance and valuable feedback on the research articles published. Furthermore, this project is successful with the assistance of Aquaponics 4.0 Learning Factory (AllFactory), Laboratory of Intelligent Manufacturing, Design and Automation (LIMDA), and the Department of Mechanical Engineering.

Finally, I am forever indebted to my beloved parents, Khuda Dad Abbasi (late) and Ajaz Begum, for their unfailing love, support, and encouragement which have been essential to undertaking this journey and succeeding in the program. To my sister, Zubia Abbasi, for keeping me focused and on track. Lastly, a special thanks to my husband, Nabil Khalid, for always believing in me and giving me extra strength to get things done.

Thank You.

Table of Contents

List of Tables

List of Figures

Acronyms

CNN	Convolutional Neutral Network
R-CNN	Region Convolutional Neutral Network
WSM	Wireless Sensing Module
AI	Artificial Intelligence
DSS	Decision Support System
IoT	Internet of Things
OWL	Web Ontology Language
CM	Confusion Matrix
HSV	Hue Saturation Value
RDF	Resource Description Framework
CAD	Computer-Aided Design
NFT	Nutrient Film Technique
RoI	Regions of Interest

Chapter 1 Introduction

1.1 Background

1.1.1 A global food security problem

Food security is a multidimensional concept that alleviates hunger by ensuring a sustainable, nutritious food supply. It is characterized by a four-pillar model shown in Figure 1-1, with each pillar intrinsic to ensuring food security [1].

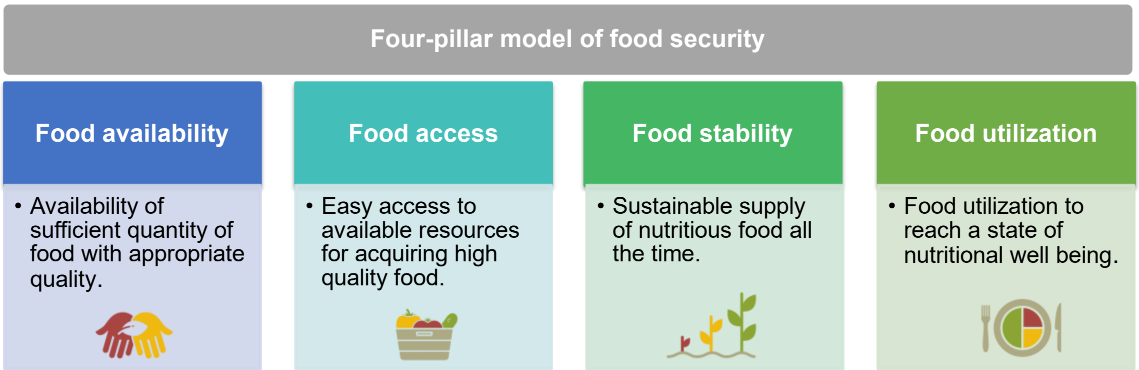


Figure 1- 1. Four-pillar model of food security by Food and Agriculture Organization of the United Nations.

Due to several anthropogenic factors, such as rapid population growth, urbanization, industrialization, farmland loss, freshwater scarcity, and environmental degradation, food security is becoming a serious global issue. This is because these factors are also directly impacting the agricultural industry, which is a primary source of agri-food production around the world. It is anticipated that by 2050 global population will be increased from the current 7.7 billion to 9.2 billion, the urban population will rise by 66%, arable land will be declined by approximately 50 million hectares, and global GHG emissions (source of CO_2 – promote crop disease and pest growth) will be increased by 50%, agri-food production will be declined by 20%, and eventually, food demand will be increased by 59 to 98% – posing an imminent threat to food security and adequate food availability [2]–[4].

1.1.2. Modern farming practices

To meet the current food demands, traditional agricultural methods (open-air fields) are extensively employed, which are labor-intensive and require arable land, a substantial amount of water (for irrigation), and time for agri-food production [5].

Water wastage and water pollution from excess nutrients, pesticides, and other pollutants are some of the other drawbacks of traditional agriculture – making it inefficient to combat the undeniable increase in food demands and the impending threat to food security [3]. Considering these factors, the world can no longer fully rely on traditional agricultural methods, which poses a need to explore modern farming practices such as vertical farming (VF). In principle, VF is a simple concept that involves growing crops in vertically stacked layers. It is essentially a controlled-environment agriculture model, which aims to optimize indoor soilless farming techniques such as hydroponics, aquaponics, and aeroponics [6]. These techniques have the potential to solve food security issues as they offer economical viable and environmentally sustainable food production practices by ensuring faster crop growth regardless of weather and season, high crop yield, less water consumption (around 70% reduction), reduced fertilizer usage, and enhanced crop quality [7]. The soilless farming techniques also offer pesticide-free organic crop production thus paving the way toward a cleaner food production system [8].

1.1.3. Aquaponics – A sustainable food production system

For this research, an aquaponic system is considered, which is the combination of recirculating aquaculture system (RAS) and a hydroponics system (soilless growing of plants), that work together in an integrated environment [9]. This means the aquaponic system consists of 1) a hydroponic unit which consists of grow beds for plant growth; and 2) an aquaculture unit which involves water tanks for fish habitat and biofilters for the breakdown of ammonia. The rationale of this soilless recirculating growing system involves sharing the mutual benefit of the available resources, such as water and nutrients, between aquaculture and plant production. Figure 1-2 illustrates the complete biological cycle of the aquaponics process. Fish eats food and excretes waste which consists of ammonia (NH3+) along with other constituents. This waste is converted by selected microbes to nitrates (NO3-). This enriched effluent is then pumped into the hydroponic component of the system, where the nutrients are readily available for uptake.

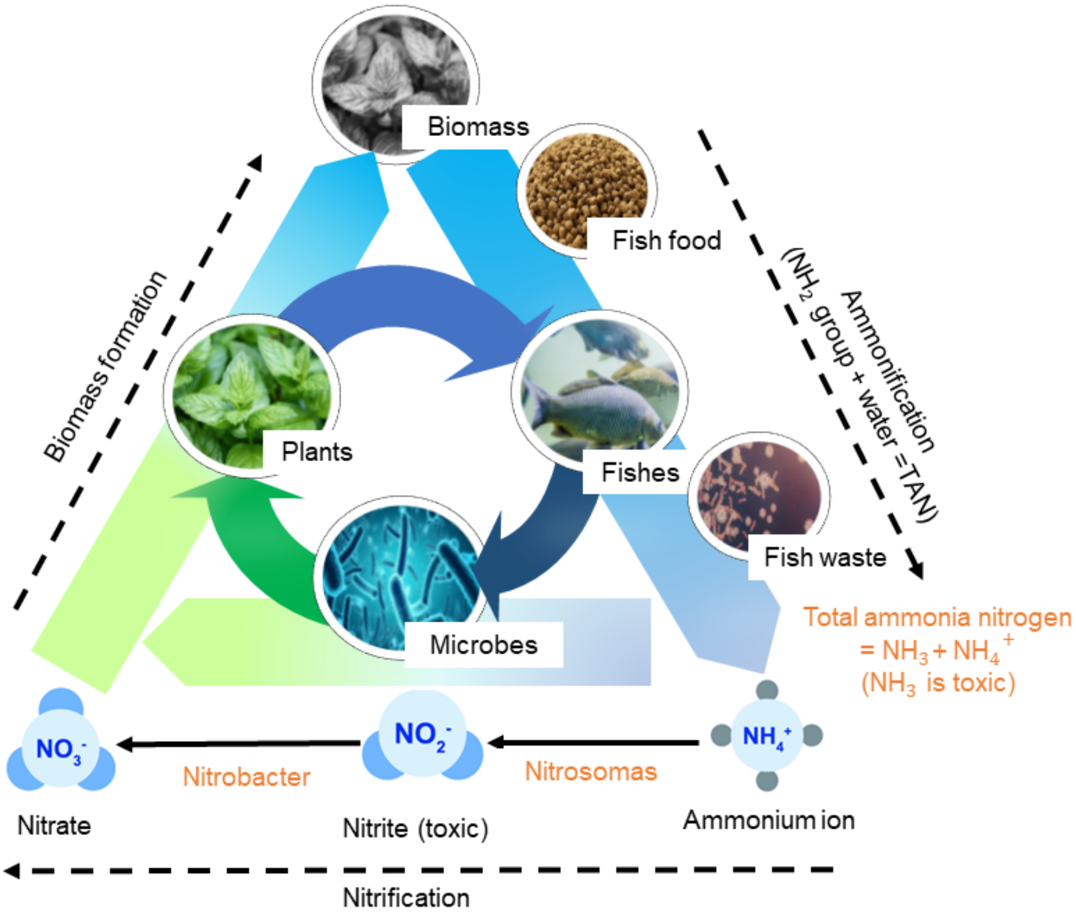


Figure 1- 2. Aquaponics biological cycle.

Hydroponic component eliminates the need for soil and provides the plant roots with direct access to nitrates, oxygen, and water, helping in the rapid growth of plants. An aquaponic system can be categorized as i) coupled system and ii) decoupled system. In a coupled system (research focus), RAS is directly connected with the hydroponic unit, and water is constantly circulated from RAS to hydroponic and back to RAS [10]. In a decoupled system, RAS is connected to the hydroponic unit (with an additional reservoir) via a one-way valve. Water separately recirculates in each system, and it is just supplied on-demand from RAS to the hydroponic unit, but not back [10]. Both systems are shown in Figure 1-3.

Primarily, depending on the design of grow bed and crop type and size, there are three different types of coupled/decoupled aquaponic systems: 1) nutrient film technique (NFT), 2) media-based, and 3) deep water culture (DWC) [11]. Each system has its advantages and disadvantages. In this research, the NFT-based aquaponic system is considered because it has a simple design and is popular for growing leafy green crops. This technique reduces the sheer amount of water required for the bed as compared to DWC and media-based techniques [12].

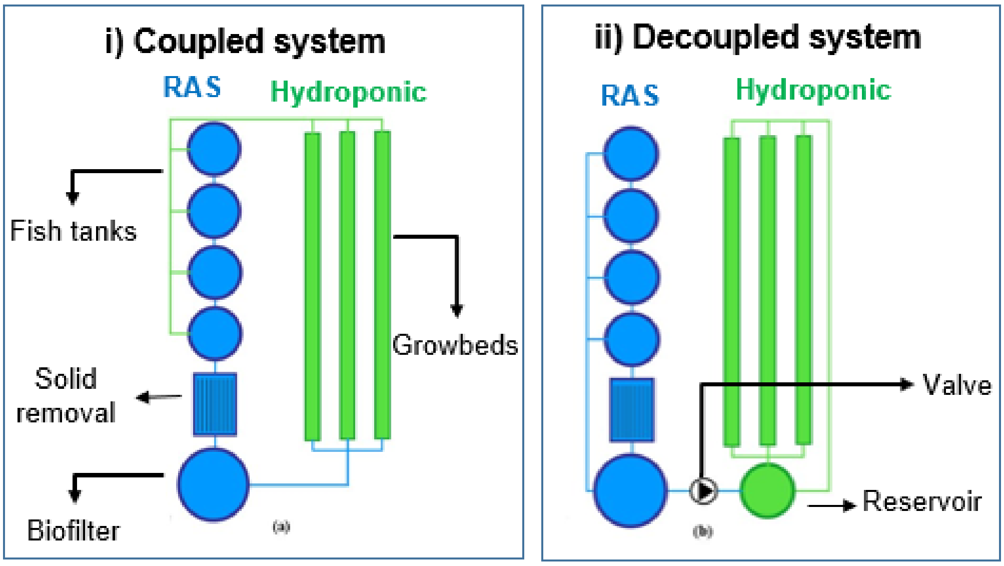


Figure 1- 3. Coupled and decoupled aquaponic systems, adapted from [10].

In NFT systems, a very shallow stream of nutrient-rich water is pumped from the fish tank to enclosed channels (grow bed) sloped slightly at a ratio of 1:30 to 1:40 (that is one inch of drop (slope) for every 30 to 40 inches of the horizontal length of channel) [13]. The top cover of the channel consists of circular or square-shaped pockets known as plant sites where plants sit in small plastic cups. The slope allows the water to flow continuously from one end to another and past the bare roots of plants which absorb nutrients from the water. Dissolved oxygen (DO) is one of the most important indicators for enhancing crop productivity, and flow rate directly affects the amount of DO in the water. In this essence, inadequate oxygenation of the nutrient solution (NS) can occur if the slope is not accurate, which may lead to root hypoxia (lack of oxygen) in crops as a result of low oxygen solubility [14]. Hypoxia results in reduced crop yield as it affects crop nutrients and water absorption. Hence, the slope of the growing channel is a significant parameter for high crop yield. Other parameters that affect crop productivity are planting site spacing, length of the channel, and distance between vertically/horizontally stacked channels. These factors are dependent on the crop type and its characteristics [15]. The general architecture of the NFT-based aquaponic system is shown in Figure 1-4.

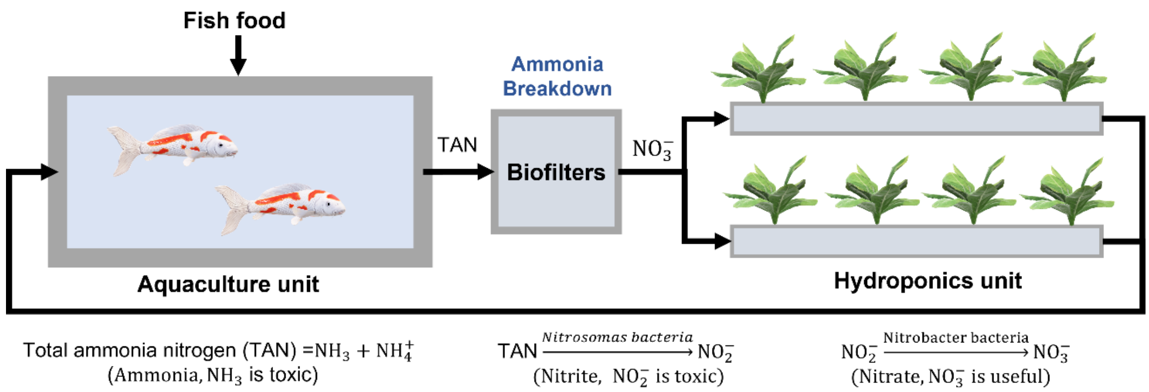


Figure 1- 4. Diagram of a typical aquaponic system.

Aquaponics is a form of sustainable agriculture because it imitates natural systems, where the efficiency of the water is dramatically increased and has fewer environmental impacts [16]. Typically, there is only 0.3% to 5% of the water wastage due to plant transpiration, evaporation losses, scheduled maintenance operations, and fish splashes [17]. On the contrary, in traditional farming, an estimated 10% of water gets absorbed by the plants, and the remaining is lost to evaporation and overflow [18]. Transportation costs and other aspects related to supply chain management can be reduced as it is possible to install aquaponics facilities in densely populated areas which witness high food demands [18].

1.2. Research motivation

Despite all the advantages offered by this imminent and growing technology, a few challenges need special attention, particularly when considering its large-scale implementation. There exists a significant interdependence among various components of an aquaponic system. Hence, designing and managing an aquaponic system is a difficult challenge when trying to achieve high yields and quality. Being a greenhouse and a symbiotic environment, the parameters and factors (light, temperature, pH, moisture, etc.) that need to be controlled are diverse. For the system to be functional and efficient, a delicate equilibrium among these parameters must be established [5]. Optimal conditions must be met for the growth and development of all three varieties of organisms that are present in the system - fish, bacteria, and plants. Another significant challenge is that, just like traditional agriculture, crops grown in aquaponics may also face quality issues resulting from diseases or nutrient deficiency, or inadequate management of the system, impacting crop quality and causing crop wastage [19]. Hence, to witness the efficiency of an aquaponic farm, it is vital to assess the quality of the crop in the early stages as it will allow farm practitioners to take precautionary measures in time if the crop is not healthy, thereby getting both economic and environmental benefits.

Crop quality is characterized by various indicators including:

- Concentrations of essential nutrients such as Calcium, Vitamins, Glucose etc.
- Phytochemical composition (secondary metabolites)
- Health attributes (abiotic and biotic stresses)
- Sensory properties (taste, aroma, texture, and foliage color)
- Morphological traits (crop canopy and its geometric characteristics crop height, width, volume, area, perimeter, fresh weight)
- Safety of a food crop [20].

These quality indicators can be used as the measure of crop quality, health condition, and yield potentiality [21]. With all the multiple components and stated requirements such as disease prevention, water quality, levels, etc., aquaponic systems are complex and require inspections seven days a week, 24 hours per day [5].

With the integration of Industry 4.0 technologies such as the internet of things (IoT), cyber-physical systems (CPS), artificial intelligence (AI), wireless sensor networks (WSN), big data and analytics (BDA), autonomous robot systems (ARS) and ubiquitous cloud computing (UCC) in the aquaponic system, the mentioned challenges can be addressed. This is because these technologies enable intelligent data-driven decisions related to crop quality control, design configuration of the system, and autonomous and robust monitoring and control of the system's operations [22]. But the pace at which Industry 4.0 technologies are being integrated into the farming industry is slow in comparison with other sectors, as shown in the image in Figure 1-5. This graphic indicates that the agriculture industry is far behind in following the pattern of computation and digitization after the construction industry.

In addition, the realization of such a highly digitized aquaponic system requires efficient data integration and information flow among different domains. Recent research has shown that semantic web technology (SWT) plays a key role in ensuring efficient data heterogeneity, interoperability, interpretation, and integration [22]. To enable semantic modelling, ontologies are used, which provide a platform to store information from multiple sources and provide an automatic process known as reasoning to infer new knowledge that has not been explicitly incorporated [23]. This knowledge and different machine learning and deep learning techniques can be utilized to develop a data-driven decision support system.

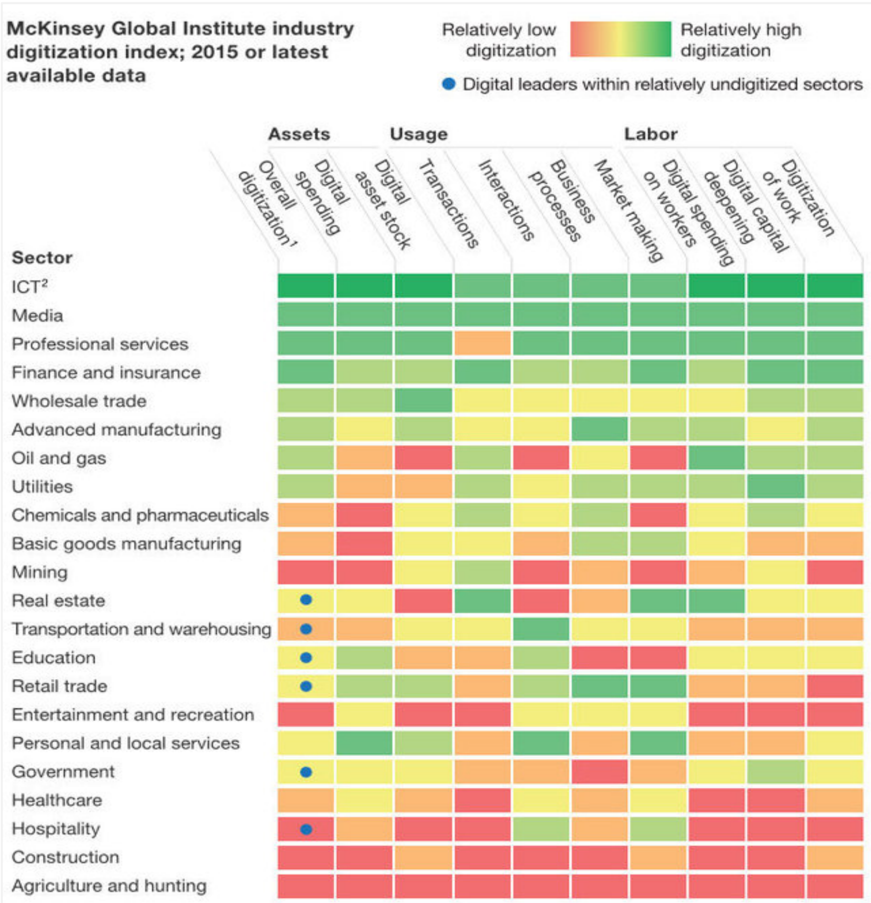


Figure 1- 5. Digitization in different sectors, adapted from [23].

The motivation behind this research stems from the fact that so far, no attempt has been made to develop an ontology model for the aquaponic 4.0 system. Moreover, no unified decision support platform is available that can assist farm practitioners in transparent decision-making regarding crop production, product quality, and facility layout of the aquaponic 4.0 farm. Hence, considering these research gaps and the importance of aquaponics as a future farming method, this research aims to develop an ontology model for the aquaponic system and a decision support system to assess crop quality. In this research, the quality of the crop is evaluated using foliage color, which indicates the chlorophyll content. Generally, if the color of the foliage is green, it represents that the crop is healthy. If it is yellow, it signifies that the crop is suffering from some type of disease or deficiency, causing interference in the production of chlorophyll [24]. Another quality indicator is related to crop biotic and abiotic stresses. Abiotic stress imposed on plants by the environment may be either physical or chemical, whereas biotic stress has resulted from biological units like diseases, insects, etc., [25]. Early and correct detection of these stresses is of utmost importance to reduce yield losses and increase profitability. In this research, a crop disease detection system is developed to detect the diseases and assess the quality of leafy green crops grown in an aquaponic facility. Lastly, crop foliage area is also estimated. All these models are deployed on cloud-based applications and integrated with the ontology model to develop a unified decision support system.

1.3. Research hypothesis, objectives and framework

The primary purpose of this thesis is to widely promote the adoption of aquaponics across the globe, as it offers great potential to overcome food challenges in the upcoming years. Working towards the inclusion of automation and smart techniques will help in reducing the inherent complexity and costly adoption of the aquaponic system. With this aim in mind, along with benefits and impediments to the large-scale implementation of aquaponics technology in commercial facilities, the hypothesis, and six specified objectives are listed below.

Research hypothesis:

"*The integration of knowledge modeling, computer vision, and machine learning technologies in an aquaponic system enhances crop quality and productivity and promotes a reliable, feasible, and sustainable food production*".

Research objectives:

The specific objectives based on the hypothesis are:

- **O1.** Review the status of digitization in the second stage of the agricultural production value chain (in-field) for different types of farms (open-air, soil-based greenhouse, aquaponics, aeroponics, and hydroponics) to present a holistic overview of the dissemination of the tools and techniques employed, the maturity level of the developed systems, along with potential roadblocks or inhibiting factors in the development of agriculture 4.0.
- **O2**. Develop a knowledge model based-on ontology for the aquaponic 4.0 facility to store domain-specific knowledge of an aquaponic system by formalizing the

links between crops/fishes, environmental parameters, production system, contextual data, and product quality.

- **O3**. Develop a decision support framework for parametric design automation of aquaponic's grow bed based on crop characteristics using the ontology model.
- **O4.** Develop a data acquisition and monitoring dashboard to gather data in a real-time setting about six different parameters related to the water quality and environmental condition from a wireless sensing module (WSM) and crop images from a camera module installed at the aquaponic facility.
- **O5.** Develop an automatic inspection system using computer vision and machine learning for evaluating crop quality based on biotic stress such as diseases, foliage color, and morphological traits such as length, width, and area.
- **O6.** Integration of ontology model and automatic inspection system to develop a multi-stage web-based decision support system for decision-making regarding crop quality control.

Research framework:The research framework showing the specific objectives and contributions is presented in pictorial form in Figure 1-6.

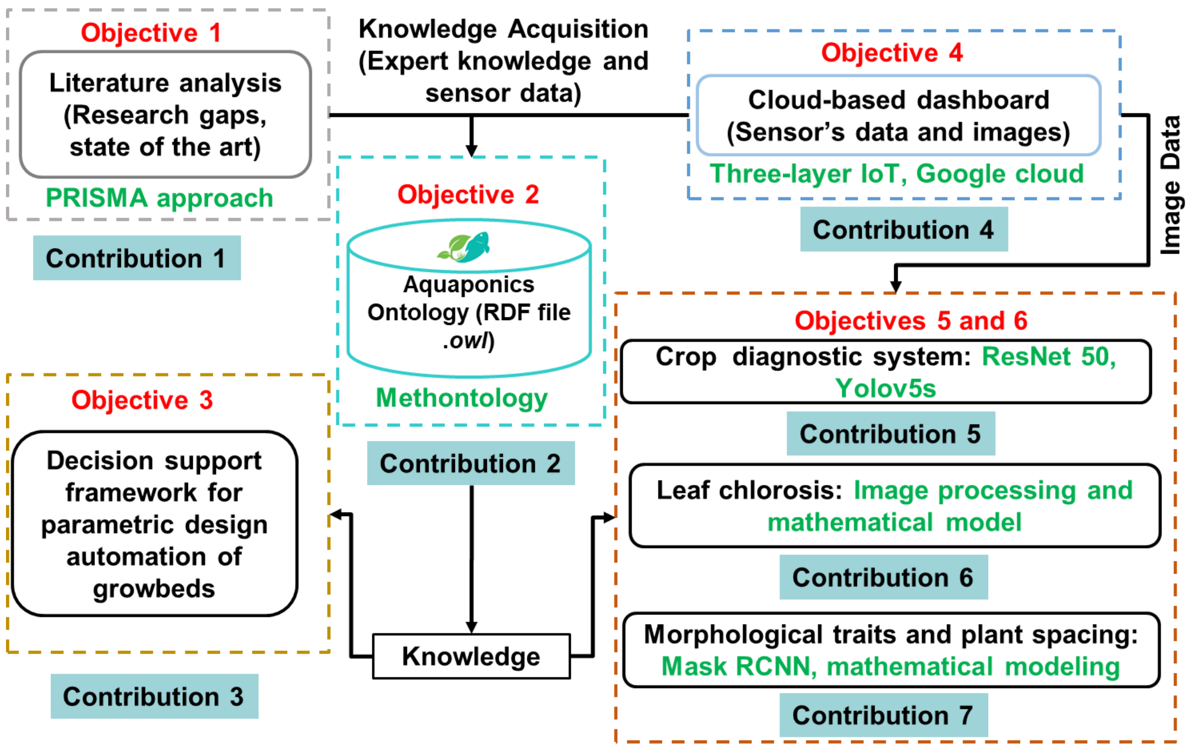


Figure 1- 6. Presentation of objectives in pictorial form.

1.4. Thesis structure

This thesis comprises nine chapters. Chapter 1 presents a brief introduction to research background and motivation along with modern farming practices with a special emphasis on aquaponics. This chapter also frames the research objectives of this thesis. Chapter 2 presents the first research contribution, "The digitization of agricultural industry – a systematic literature review on agriculture 4.0," addressing the first research objective. Chapter 3 discusses the second research contribution, "Chapter 3 An ontology model to represent the aquaponic 4.0 system's knowledge," fulfilling the second research objective. Chapter 4 discusses the third research contribution, "Chapter 4 An ontology model to support the automated design of aquaponics grow beds", addressing the third research objective. Chapter 5 highlights the fourth research contribution, "Data acquisition and monitoring dashboard for IoT-enabled aquaponic facility", fulfilling the fourth research objective. Chapters 6, 7, and 8 address the fifth and sixth research objectives. Chapter 6 "Crop Diagnostic System: A robust disease detection system for leafy green crops grown in aquaponic facility," covers the fifth research contribution. Chapter 7 "Non-destructive identification of foliage chlorosis in lettuce crop grown in aquaponic facility using image processing," presents the sixth research contribution. Chapter 8 "Estimation of foliage area for effective crop site spacing in an NFT-based aquaponic system using Mask-RCNN," covers the seventh and last contribution. Finally, the conclusion of this thesis, summary of research contributions, limitations of this study, and future work directions have been discussed in Chapter 9.

Chapter 2 The digitization of agricultural industry – a systematic literature review on agriculture 4.0

2.1. Introduction

To satisfy the increasing food demands, agricultural practitioners worldwide will need to maximize agricultural productivity involving crop and livestock farming. In this review, the focus is on crop farming, which involves the cultivation of both food and cash crops. A typical agri-food value chain depicting three primary stages, namely pre-field (pre-plantation stage), in-field (plantation and harvesting stage), and post-field (post-harvesting stage), involved in the production of agricultural products is shown in Figure 2-1.

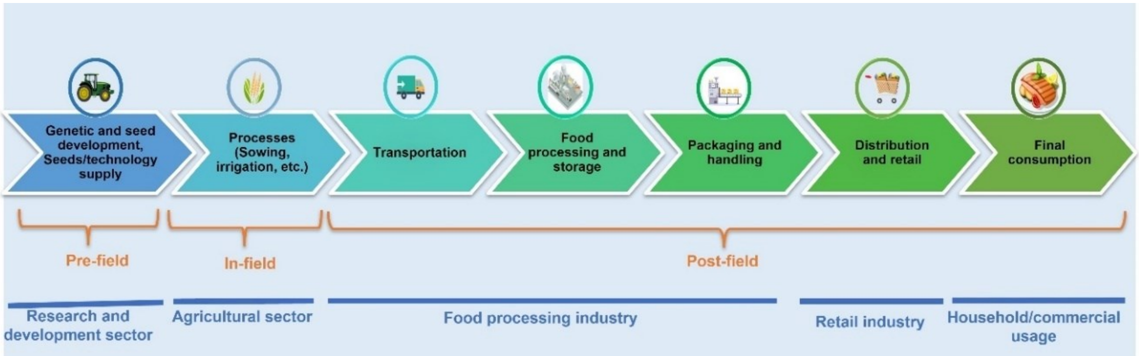


Figure 2- 1. Agriculture value chain: stages and main functions.

All the stages play a vital role in the value chain but, in this review, the second stage ″in-field″ will be considered, which involves several crop-growing processes such as ploughing, sowing, spraying, and harvesting, etc. These processes currently employ traditional agricultural practices that are labor-intensive, and require arable land, time, and a substantial amount of water (for irrigation) – making it a challenge to produce enough agri-food [22]. A part of the problem is also related to the irregular use of pesticides and herbicides and misuse of available technology which cause harm to crops and eventually result in agricultural waste [6]. These issues can be addressed by integrating sophisticated technologies and computer-based applications that ensure high crop yield, less water consumption, optimised pesticide/herbicide utilization, and enhanced crop quality. This is where the smart agriculture concept comes in.

2.1.1 Smart agriculture

Industry 4.0, also known as the fourth industrial revolution, is revolutionizing and reshaping every industry. It is a strategic initiative characterized by a fusion of emerging disruptive digital technologies such as the Internet of Things (IoT), big data and analytics (BDA), system integration (SI), cloud computing (CC), simulation, autonomous robotic systems (ARS), augmented reality (AR), artificial intelligence (AI), wireless sensor networks (WSN), cyber-physical system (CPS), digital twin (DT), and additive manufacturing (AM) to enable the digitization of the industry [26]. The integration of these technologies in agriculture is sparking the next generation of industrial agriculture, namely, agriculture 4.0 – also termed smart agriculture, smart farming, or digital farming [26].

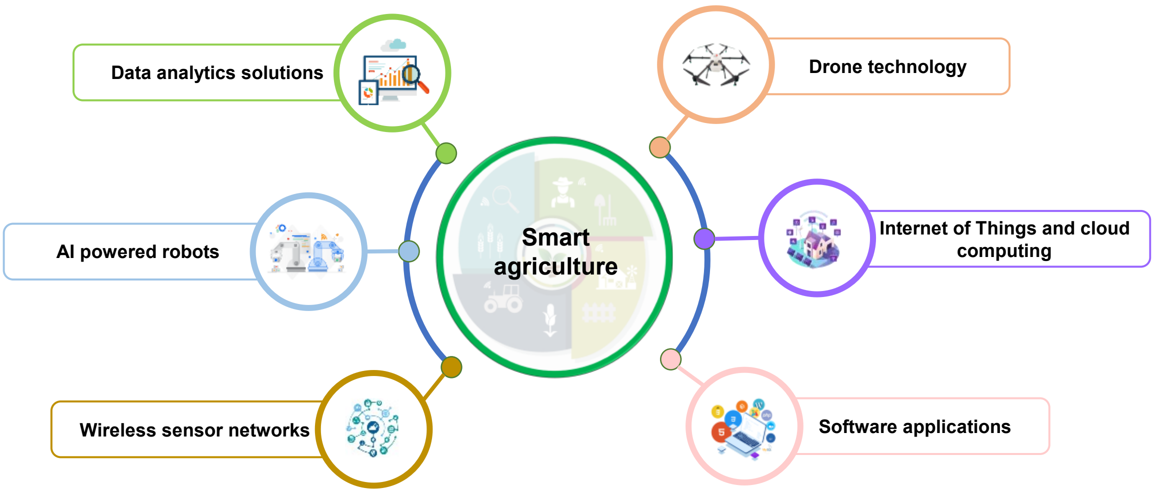


Figure 2- 2. The concept of "Smart Agriculture".

Smart agriculture provides farmers with a diverse set of tools (shown in Figure 2-2) to address several agricultural food production challenges associated with farm productivity, environmental impact, food security, crop losses, and sustainability. For instance, with IoT-enabled systems consisting of WSNs, farmers can connect to farms remotely irrespective of place and time to monitor and control farm operations. Drones equipped with hyperspectral cameras can be used to collect data from heterogeneous sources on farmlands, and autonomous robots can be used to support or accomplish repetitive tasks on farms. Data analytics techniques can be employed to analyze the gathered data with computer applications that can be used to assist farmers in the decision-making process. Likewise, a wide variety of parameters related to environmental factors, weed control, crop production status, water management, soil conditions, irrigation scheduling, herbicides, and pesticides, and controlled environment agriculture can be monitored and analysed in smart agriculture to increase crop yields, minimize costs, enhance product quality, and maintain process inputs through the use of modern systems [27].

2.1.2. Research motivation and contribution

The motivation for preparing this review stems from the fact that digital technologies in agricultural systems offer new strategic solutions for enhancing the efficiency and effectiveness of farms' production. Moreover, digital transformation provides a way forward to implement modern farming practices such as vertical farming (hydroponics, aquaponics, and aeroponics), which has the potential to overcome food security problems. But there is a set of problems and limitations associated with this transformation from the technical, socioeconomic, and management standpoint that must be death to fully exploit the potential of agriculture 4.0 [28]. There are a number of studies that have discussed emerging trends in the development of agriculture 4.0 by providing succinct information on key applications, advantages, and corresponding research challenges of smart farming [28]–[37]. The research focus of these studies is limited to either explaining more generic technical aspects while paying attention to only one or few digital technologies, and/or enhancing agricultural supply chain performance, and/or developing agriculture 4.0 definition, and/or achieving sustainable agronomy through precision agriculture, and/or proposing a smart farming framework. Nevertheless, these studies do not involve explicit discussion on the tools and techniques used to develop different systems and the maturity level of these systems. There is also a lack of studies considering modern soilless farms such as hydroponics, aquaponics and aeroponics (indoor/outdoor) and the implications of digital technologies in these farms. Hence, it is necessary to analyse the evolution of agriculture 4.0 from different perspectives to stimulate discussion in the area. This study aims to present a holistic overview of digital technologies implemented in the second stage of the agricultural production value chain (in-field) for different types of farms as mentioned in section 2.1.1. The main theoretical contribution of the study involves analysis and dissemination of the tools and techniques employed, the farm type, the maturity level of the developed systems, along with potential roadblocks or inhibiting factors in development of agriculture 4.0. The reflections presented in the review will support researchers and agricultural practitioners in future research on agriculture 4.0.

2.1.3. Chapter organization

Following the introduction, the chapter is structured as follows: Section 2 discusses the approach used to gather the relevant literature; then, Section 3 presents the statistical results obtained after a general analysis of the selected research studies; next, Section 4 provides a detailed overview of the core technologies used in the digitization of agriculture; after, Section 5 highlights the technical and socio-economic roadblocks to digital integration in agriculture; next, Section 6 outlines a discussion about the research questions followed by added value, considerations and future prospects related to agricultural digitization, and transition to agriculture 5.0; and lastly, Section 7 concludes the review.

2.2. Research methodology

A systematic literature review (SLR) is a tool used to manage diverse knowledge and identify research related to a predetermined topic [38]. In this study, SLR is conducted to investigate the status of Industry 4.0 technologies in the agricultural industry. Particularly, cases are searched where the term 'agriculture' appeared concurrently in the title, abstract, or keywords of an article with any of the 'Industry 4.0 technologies" mentioned in section 2.1.2. Before conducting the SLR, a review protocol is defined to ensure a transparent and high-quality research process, which are the characteristics that make the literature review systematic [38]. The review protocol also helps to minimize bias by conducting exhaustive literature searches. This includes three steps: the formulation of the research questions, the definition of the search strategy, and the specification of inclusion and exclusion criteria. This work uses a preferred reporting item for systematic reviews and a meta-analysis (PRISMA) approach to conduct SLR. PRISMA is an evidence-based minimum set of items that are used to guide the development process of systematic literature reviews and other meta-analyses [38].

2.2.1. Review protocol

A review protocol (in Table 2-1) is defined before conducting the bibliographic analysis to identify, evaluate, and interpret results relevant to the research scope. First, research questions are formulated to provide insight into the analysis of published studies in the research area of interest from different dimensions. These questions need to be answered in the study. Next, the search strategy is defined, which helps identify appropriate keywords later in the search equation to identify the relevant information sources, such as academic databases and search engines that provide access to a massive amount of digital documentation. Three online research repositories are used to retrieve relevant studies: ScienceDirect[1] , Scopus[2] , and IEEE Xplore[3] . Finally, to refine the search results of each database, boundaries are set by predefining inclusion and exclusion criteria for further investigation and content assessments of selected publications. It involves, for instance, defining the time interval for the research process from 2011 to 2021 to limit the studies to those published in English, disregarding chapters of books and grey literature, such as reports and summaries of events and seminars. These last two steps of the review protocol allow the preliminary filtering of metadata sources and narrow down the scope of research.

Table 2- 1. Review protocol for systematic literature review.

Review questions	RQ1: Which Industry 4.0 technologies have been used in the literature for the digitization of agriculture? RQ2: How and to what extent have these technologies been applied in the context of service type, tools and techniques used, system's maturity level, and farm type? RQ3: What are the primary roadblocks in the implementation of Industry 4.0 technologies for smart farming?
Study selection criteria	Inclusion criteria: • Peer-reviewed journal articles and conference papers. • Studies published during the period between 2011 and 2021.

[1] www.sciencedirect.com

[2] www.scopus.com

[3] ieeexplore.ieee.org

	• Studies should provide answers to the research questions. • The article must include the title, year, source, abstract, and DOI. • Literature focussing on the application of Industry 4.0 technologies in crop plantation and harvesting activities, particularly in-field processes. Exclusion criteria: • Summaries of events and seminars, book reviews, and editorials. • Literature focusing on the application of Industry 4.0 technologies in livestock farming; pre-field processes such as genetic development, seed development and seed supplying; post-field stages such as crop distribution, food processing and consumption; and agri-food supply chain. • Studies published before 2011. • The publication is not available in full text. • The publication is not in English.
Literature search	Sources: Scopus, ScienceDirect, and IEEE Xplore for academic literature, citations in identified literature Search equation: (("agriculture*") AND ("Industry 4.0" OR "Digital Farming" OR "Intelligent Farming" OR "Smart Agriculture" OR " Agriculture 4.0" OR " Smart Farming" OR "Internet of Things" OR "IoT" OR "Cloud Computing" OR "Edge Computing" OR "Wireless Sensor Networks*" OR " Artificial Intelligence*" OR "Big Data*" OR "Data Analytics*" OR " Data Science*" OR "Cyber-Physical System*" OR "Robotics*" OR "Computer Vision*" OR "Machine Learning*" OR "Deep Learning*" OR "Data Integration*"))

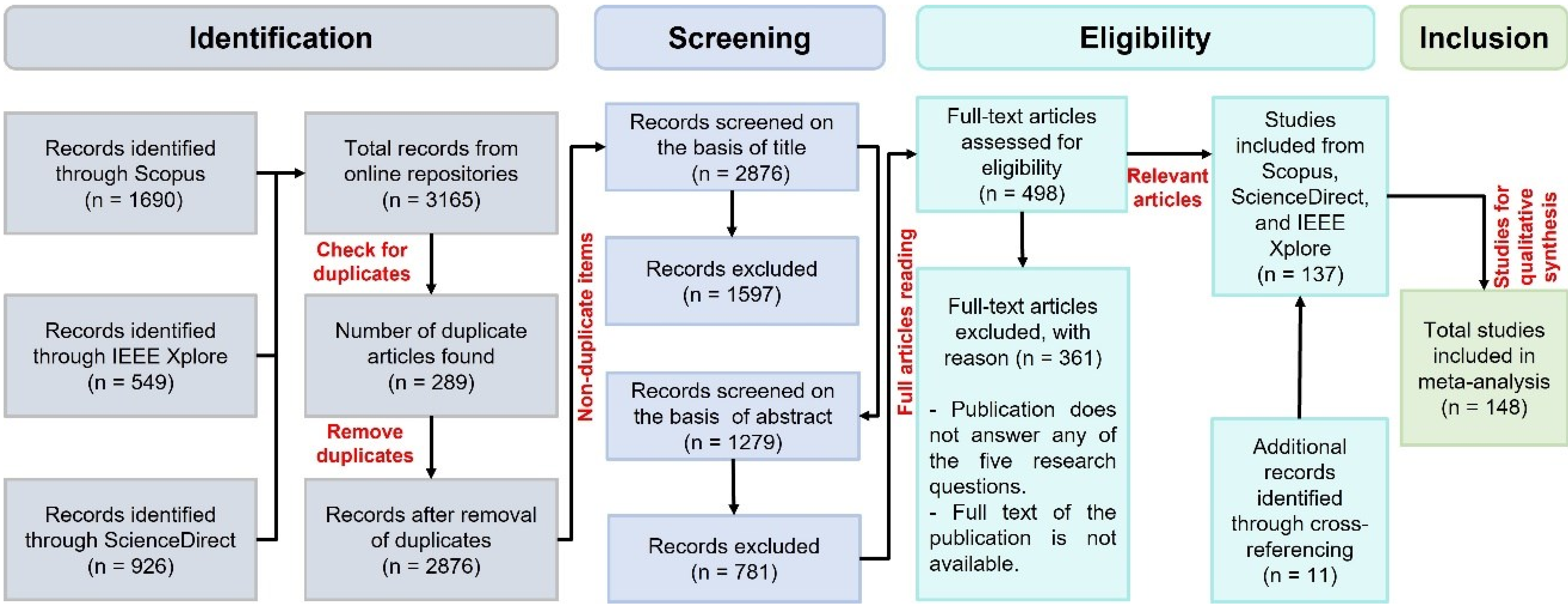


Figure 2- 3. Four-step evaluation of literature search process (PRISMA).

2.2.2. Evaluation process

The evaluation of the literature search process is done in four stages: identification, screening, eligibility, and inclusion, as detailed by the PRISMA flow diagram shown in Figure 2-3. After initial metadata filtering through the application of search expression, a total of 3165 records are found (1690 from Scopus, 926 from ScienceDirect, and 549 from IEEE Xplore), which are then consolidated for the removal of duplicate items in the identification stage. The number of publications after this step is reduced to 2876. In the screening stage, the titles and abstracts of the papers are analysed, and only 498 papers are selected for integral reading. In the third stage, a full-text screening of these articles is performed to verify their eligibility in relation to the objective of this review, which is to answer the research questions mentioned in Table 2-1. Of the 498 papers, 137 are found to be relevant for this review. Another 11 are added through a cross-referencing approach, adding up to 148 papers selected in the final stage for further analysis.

2.2.3. Threats to validity

i. SLR replication: The presented SLR is susceptible to threats to validity because the current search is limited to only three online repositories. More publications could potentially be found if additional sources were explored. The process of SLR is described clearly in sub-sections 2.2.1 and 2.2.2, and hence, validity can be considered well-addressed. However, in the case of replication of this SLR, it is possible that one can find slightly different publications. This difference would result from different personal choices during the screening and eligibility steps of PRISMA, but it is unlikely that the overall findings gathered in relation to different review questions would change.

ii. Search string: the search string used to find the relevant studies cover the whole scope of SLR, but there is a possibility that valuable studies might have been missed. Additional keywords and synonyms with a broader search might return more studies.

2.3. Digitization trends in agriculture

The year-wise distribution of the 148 articles from 2011 to 2021 is represented in Figure 2-4. Around 22% of the scientific publications in the last ten years were published in 2018. This reflects that the agricultural industry is making considerable progress in the context of the implementation of digital technologies, but the pace is still slow as compared to other domains such as healthcare, manufacturing, mining, automotive, energy, etc.,[34].

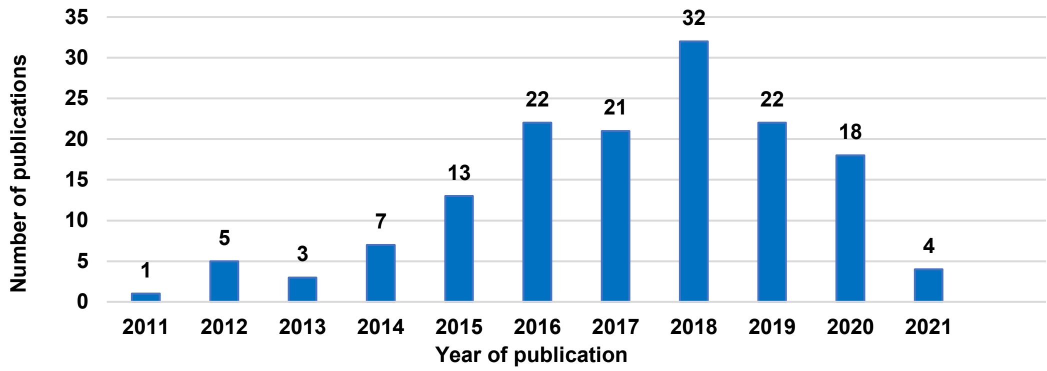


Figure 2- 4. Year-wise distribution of selected research studies from 2011 to 2021.

The breakdown of these publications with respect to digital technologies (mentioned in sub-section 2.1.2) and targeted farm types is represented in Figure 2-5.

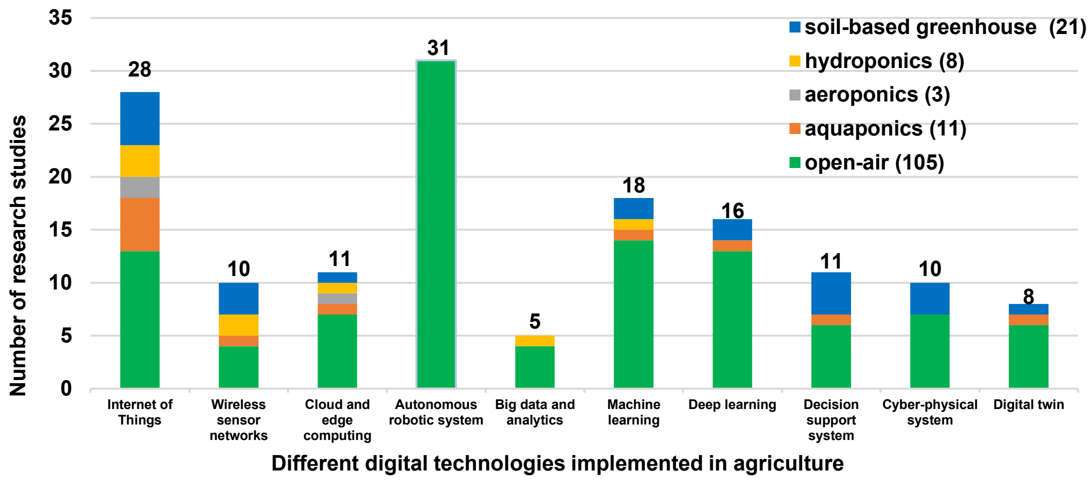


Figure 2- 5. Technology-wise distribution of the 148 selected research studies.

The farm type refers to the crop farming method considered while developing an application or framework. For instance, the farming method can be soil-based or soilless. The soil-based farming category involves open-air fields (traditional outdoor agricultural farms) and greenhouse farms (indoor). On the other hand, the soilless farming category involves modern farming practices such as aquaponics, aeroponics, and hydroponics (mostly indoor). The numbers at the top of the stacked column in Figure 2-6 indicate the total number of studies that have used the particular technology to develop a smart agriculture system, whereas different colors of columns indicate the respective farm types. Use cases from these publications are analyzed, and conclusions are drawn. For instance, it has been found that autonomous robotics systems (including unmanned guided vehicles and unmanned aerial vehicles (drones)), the internet of things, and machine learning appear to be the widely applied technologies in the agricultural domain in the last decade. The same illustration suggests that big data, wireless sensor networks, cyber-physical systems, and digital twins are the emerging areas in agriculture. Moreover, open-air farms are the most frequently considered in research studies (69%), contrary to indoor farms (31%). For soilless farming systems (aquaponics, aeroponics, and hydroponics), only 22 publications are found, which insinuates that these modern farming practices are still in their infancy.

Likewise, services of each use case are identified and are classified under nine different service categories, namely: i) crop management, CM (Estimation/ prediction of crop yield/ growth rate/ harvesting period and seed plantation/ harvesting/ pollination/ spraying (fertilizer/ pesticide)); ii) crop quality management, CQM (fresh weight, green biomass, height, length, width, leaf density, pigment content (chlorophyll) and phytochemical composition); iii) water and environment management, WEM (monitoring and control of flow rate, water level, water quality (nutrients), temperature, humidity, CO2, and weather forecast etc.); iv) irrigation management, IM (water stress detection and scheduling); v) farm management, FM (monitoring of farm operations, tracking and counting products, determining production efficiency, financial analysis, energy consumption analysis, technology integration and decisions implementation); vi) pest and disease management, PDM (pest identification and disease detection); vii) soil management, SM (moisture content, soil nutrients, fertilizer needs and application); viii) weed and unwanted vegetation management, WUVM (weed/unknown vegetation mapping, classification, and herbicides application); and ix) fruit detection and counting, FDC — as shown in Figure 2-6. These categories illustrate the role of different digital technologies in smart farming. Upon analysis, it is found that crop management parameters, such as crop yield prediction, growth rate estimation, or evaluation of harvesting period are the most frequently researched areas for agriculture 4.0 in the last decade (29%). In contrast, tiny heed is paid to soil management (2%), fruit detection and counting (2%), and crop quality management (3%).

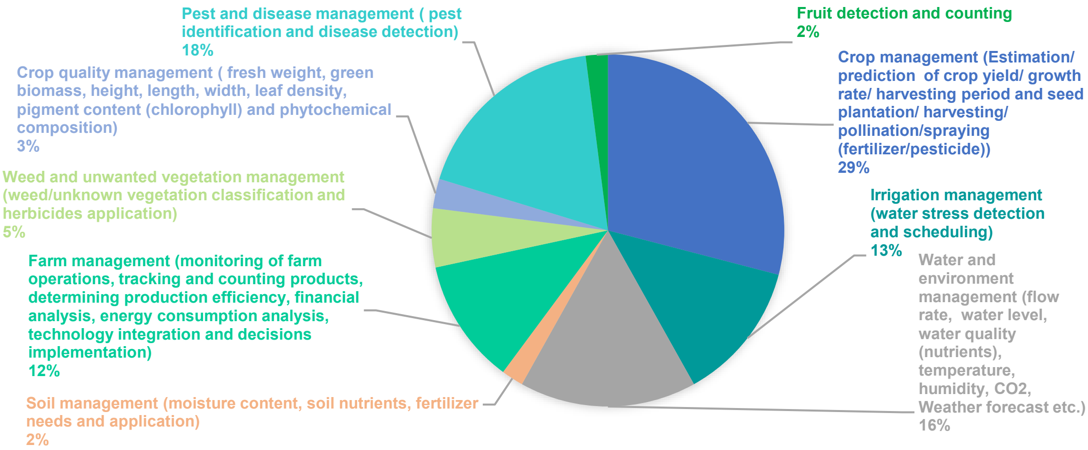


Figure 2- 6. Service-wise distribution of selected research studies.

The technology readiness level (TRL) of all the use cases is examined using European Union's TRL scale that partitions the system's maturity level into three generic levels [39]. The first level is conceptual, which represents European TRL 1–2 (use case is in the conceptual phase), and the second level is the prototype, which means European TRL 3–6 (use case is working even without the complete planned functionality), and the third level is deployed, that includes European TRL 7–9 (use case is mature with all the possible functions). Figure 2-7 depicts the TRL of each use case developed in selected studies. It is observed that little progress has been made in advancing smart agricultural systems beyond the concept and prototype levels to the commercial level. For instance, most use cases (129) are at the prototype level.

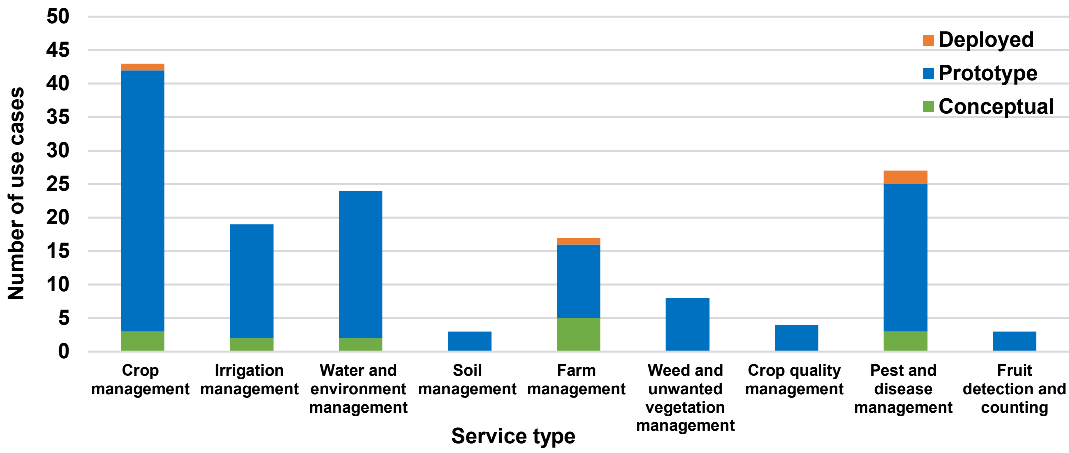


Figure 2- 7. Distribution of studies based on the service category and system's maturity level.

2.4. Agriculture 4.0 enabling technologies

This section provides critical insights towards answering RQ1 and RQ2 from Table 2-1.

2.4.1. Internet of Things driven agricultural systems

Internet of things (IoT) refers to a cosmos of interrelated computing devices, sensors, appliances, and machines connected with the internet, each with unique identities and capabilities for remote sensing and monitoring [39]. The reference architecture of IoT with six layers, namely the perception layer (hardware devices), network layer (communication), middleware layer (device management and interoperability), service layer (cloud computing), application layer (data integration and analytics), and end-user layer (user interface), is shown in Figure 2-8. In the agricultural domain, IoT devices in the physical layer gather data related to environmental and crop parameters such as temperature, humidity, pH value, water level, leaf color, fresh leaf weight, etc. The transmission of this data takes place in the network layer, the design of which depends on the selection of suitable communication technologies relevant to the field size, farm location, and type of farming method. For instance, ZigBee, LoRa, and Sigfox are widely used and employed in outdoor fields because they are cheaper and have low energy consumption and a good transmission range [40], [41] Despite being a secure technology, Bluetooth is only used in indoor farms as it offers a short transmission range [40]. Wi-Fi is not a promising technology for agricultural applications due to its high costs and high energy consumption [40]. RFID (radio frequency identification) and NFC (near field communication) technologies, on the other hand, are increasingly being implemented in agricultural systems for tracking agricultural products [42]. GPRS or mobile communication technology (2G, 3G, and 4G) are used for periodic monitoring of environmental and soil parameters. In addition, communication protocols mostly used in agricultural scenarios are HTTP, WWW, and SMTP. Likewise, to ensure interoperability and system security to their context-aware functionalities, middleware HYDRA and SMEPP are mostly employed in agricultural systems [43]. To store data, cloud computing techniques are employed in the service layer. This data is then used in the application layer to build smart applications used by farmers, agriculture experts, and supply chain professionals to enhance farm monitoring capacity and productivity.

The integration of IoT in agriculture is meant to empower farmers with decision tools and automation technologies that seamlessly integrate knowledge, products, and services to achieve high productivity, quality, and profit. A multitude of studies is performed and put forward concerning the incubation of IoT concepts in the agricultural sector. The main findings of some of the studies are presented in Table 2-2. Multiple technological issues and architectural problems have been addressed through the development of IoT-based agricultural systems. But most of these systems are either in a conceptual stage or in a prototype form (not commercial) at the moment. The focus is mainly laid on-farm management, irrigation control, crop growth, health monitoring, and disease detection. Some of these studies have also explained IoT implementation in modern agricultural systems such as vertical farming (soilless farming - aquaponics, hydroponics, and aeroponics) and greenhouse farming (soil-based). Moreover, most studies have focused on addressing a specific problem.

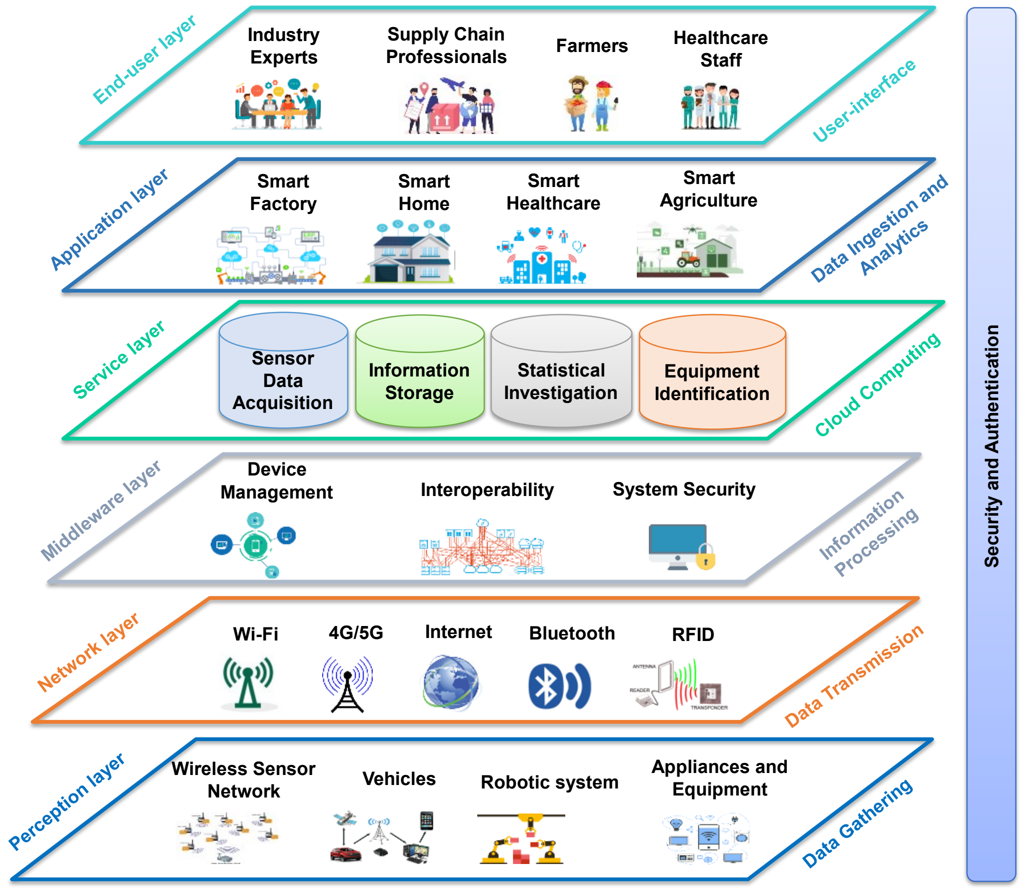


Figure 2- 8. The six-layered architecture of the Internet of Things (IoT), adapted from [44].

Table 2- 2. IoT-driven agricultural systems.

Use case No.	Service category	Tools and techniques	Farm type	Maturity level	Citations
1.		WSN, CC, and reinforcement learning	Greenhouse (soil-based)	Deployed	[45]
2.		Sensors, actuators, and controllers	Open-air	Prototype	[46]
3.		Sensors, controllers, and mobile app	Greenhouse (soil-based)	Prototype	[47]
4.	CM	Sensors, CC, BD analysis, and ML	Greenhouse (soil-based)	Prototype	[48]
5.		Sensors, and CC	Aeroponics	Prototype	[49]
6.		Sensors, actuators, and control system	Aeroponics	Prototype	[50]
7.		Weather boxes, sensors, and camera	Open-air	Prototype	[51]
8.	CQM	IoT devices, LED lights, and software application	Hydroponics	Prototype	[52]
9.	WEM	Sensors, and CC	Aquaponics	Conceptual	[53]
10.		Sensors, Arduino board, and	Open-air	Prototype	[54]

		database			
11.		Sensors, Arduino board, and database	Greenhouse (soil-based)	Prototype	[55]
12.		Sensors, CPS, edge, and cloud computing	Hydroponics	Prototype	[56]
13.		Sensors, electronic components, and network	Aquaponics	Prototype	[57]
14.		Sensors, Arduino, Raspberry Pi3, and deep neural network	Hydroponics	Prototype	[58]
15.		Sensors, and database	Aquaponics	Prototype	[59]
16.		Sensors, actuators, and CC	Aquaponics	Prototype	[60]
17.		Sensors, controllers, and mobile app	Aquaponics	Prototype	[61]
18.	IM	WSN, fuzzy logic and neural network	Open-air	Prototype	[62]
19.		Sensor information unit, MQTT, HTTP, and neural network	Greenhouse (soil-based)	Prototype	[63]
20.	FM	Sensors, controllers, web interface, and CC	Open-air	Conceptual	[64]
21.		Sensors, controllers, cloud, and Android application	Open-air	Prototype	[65]
22.		Sensors, IEEE, and GSM protocols	Open-air	Prototype	[66]
23.	PDM	Sensors, controllers, and image processing	Open-air	Prototype	[67]
24.		Cloud, camera, controllers, and K-mean clustering	Open-air	Prototype	[68]
25.		WSN, controller, and cloud	Open-air	Prototype	[69]
26.		WSN, cloud storage, and agricultural knowledge base	Open-air	Prototype	[70]
27.		WSN, Hidden Markov Model, and SMS module	Open-air	Deployed	[71]
28.		Sensors, Image processing, k-mean clustering, and support vector machine	Open-air	Prototype	[72]

2.4.2. Wireless sensor networks in agriculture

A wireless sensor network (WSN) is regarded as a technology that is used within an IoT system. It can be defined as a group of spatially distributed sensors for monitoring the physical conditions of the environment, temporarily storing the collected data, and transmitting the gathered information at a central location [40]. The general architecture of WSN is shown in Figure 2-9. A WSN for smart farming is made up of numerous sensor nodes connected through a wireless connection module. These nodes have a variety of abilities (e.g., processing, transmission, and sensation) that allow them to self-organize, self-configure, and self-diagnose. There are different types of

WSNs, which are categorized depending on the environment where they are deployed. These include terrestrial wireless sensor networks (TWSNs), wireless underground sensor networks (WUSNs), underwater wireless sensor networks (UWSNs), wireless multimedia sensor networks (WMSNs), and mobile wireless sensor networks (MWSNs) [73]. In agricultural applications, TWSN and UWSN are widely used. In TWSNs, the nodes are deployed above the ground surface, consisting of sensors for gathering the surrounding data. The second variant of WSNs is its underground counterpart – WUSNs, where sensor nodes are planted inside the soil. In this setting, lower frequencies easily penetrate through the soil, whereas higher frequencies suffer severe attenuation [74]. Therefore, the network requires a higher number of nodes to cover a large area because of the limited communication radius. Many research articles are available in the literature that discusses the use of WSN for different outdoor and indoor farms applications such as irrigation management, water quality assessment, and environmental monitoring. A summary of some of these articles is given in Table 2-3. These studies have focused on developing WSNs architectures that are simplified, low-cost, energy-efficient and scalable. Yet, various factors associated with WSNs need further attention, such as minimum maintenance, robust and fault-tolerant architecture, and interoperability.

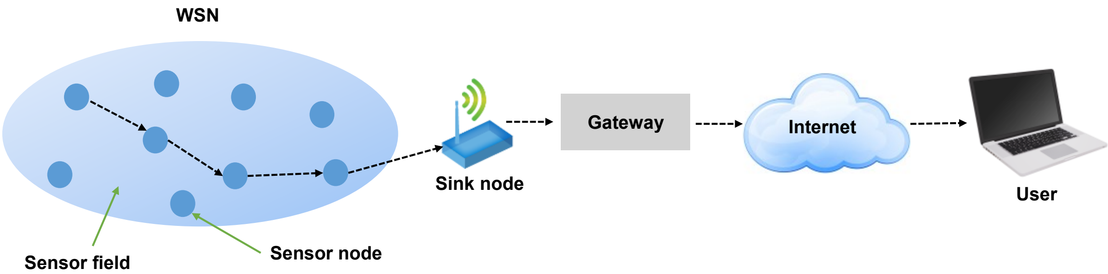


Figure 2- 9. General architecture wireless sensor network (WSN).

Table 2- 3. Use of WSNs in agricultural systems.

Use case No.	Service category	Tools and techniques used	Farm type	Maturity level	Citation
29.	IM	Soil-moisture and temperature sensors, web applications, and photovoltaic panels	Open-air	Prototype	[75]
30.		Electronic board, sensor board and GPRS board.	Open-air	Prototype	[76]
31.		Wireless sensor nodes, and Zigbee	Open-air	Conceptual	[77]

32.		Moisture sensors, actuators, and GUI	Greenhouse (soil-based)	Prototype	[78]
33.	WEM	Wireless communication, temperature, and humidity sensors	Greenhouse (soil-based)	Prototype	[18]
34.		Sensor nodes, gateway unit, database, ordinary kriging spatial interpolation (OKSI) algorithm	Hydroponics	Prototype	[79]
35.		Microcontrollers, wireless radio frequency and sensor nodes	Greenhouse (soil-based)	Prototype	[80]
36.		Wireless sensor nodes, communication networks, and mobile application	Aquaponics	Prototype	[81]
37.		Arduino, a wireless module with temperature, relative humidity, luminosity, and air pressure sensors	Any farm	Prototype	[82]
38.		Zigbee, Wi-fi and sensors	Hydroponics	Prototype	[83]

2.4.3. Cloud computing in agriculture

According to the National Institute of Standard and Technologies (NIST), cloud computing (CC) is defined as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [84]. The main architecture of CC shown in Figure 2-10 is comprised of four layers: data center (hardware), infrastructure, platform, and application [85]. Each of these layers is linked with specific cloud service models, which are classified as software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). Cloud computing has gained great attention over the past decade in the agriculture sector because it provides: 1) inexpensive storage services for data gathered from different domains through WSNs and other preconfigured IoT devices, 2) large-scale computing systems to perform intelligent decision-making by transforming this raw data into useful knowledge, and 3) a secure platform to develop agricultural IoT applications [86]. In combination with IoT and WSN, CC is employed to develop different agricultural applications, most of which are presented in Tables 2 and 3. CC technology is also used to create operational farm management systems (FMSs) to support farmers and farm managers in the efficient monitoring of farm operations. Table 2-4 presents the salient features of some of these FMSs. Another topic of interest that is being explored in global research is related to the traceability of agri-product quality [87]. But only preliminary research has been attempted to explore traceability compliance with standards of food safety and quality.

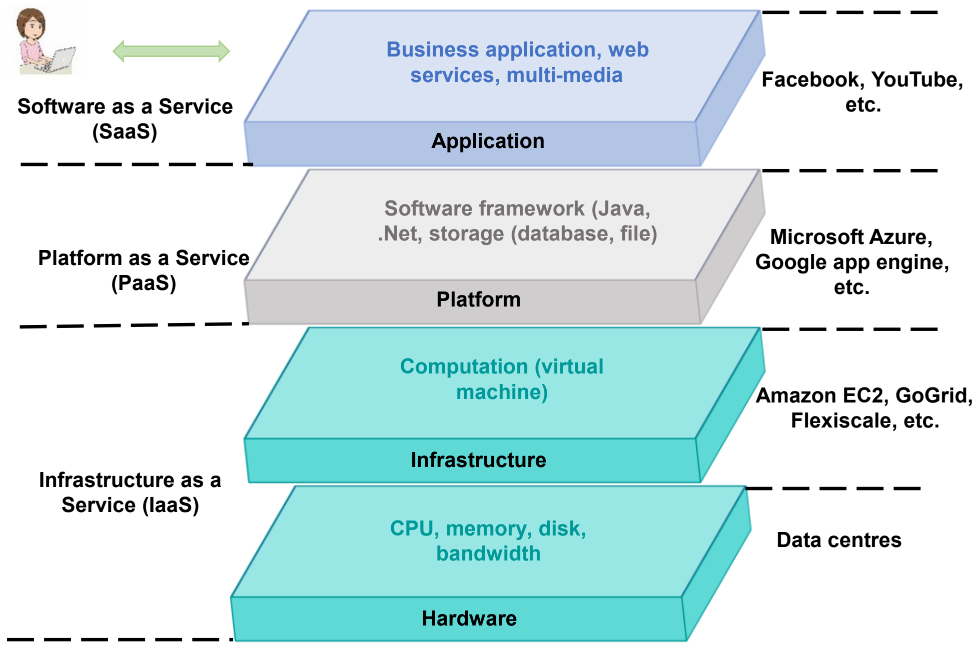


Figure 2- 10. Architecture of cloud computing, adapted from [85].

Table 2- 4. Cloud computing-based farm management systems.

Use case No.	Service category	Tools used	Farm type	Maturity level	Citation
39.	FM	Fuzzy logic, Java, HTML, Apache Karaf, etc.;	Greenhouse (soil-based)	Conceptual	[88]
40.		RFID, and mobile app	Open-air	Deployed	[89]
41.		MySQL, a financial analysis tool and mobile app	Open-air	Conceptual	[90]
42.		Self-leveling scale, control box, LCD, and RFID tags	Open-air	Conceptual	[91]

The cloud-based agricultural systems have the potential to solve problems of increasing food demands, environmental pollution caused by excessive use of pesticides and fertilizers, and the safety of agricultural products. These FMSs, however, do not have the capability to support run-time customization in relation to the distinct requirements of farmers. Moreover, because most farm data is usually fragmented and dispersed, it is difficult to record farm activities properly in current FMSs applications [92].

2.4.4. Edge/fog computing in agriculture

The rapid development of IoT has led to the explosive growth of sensors and smart devices, generating large volumes of data. The processing and analysis of such an enormous amount of data in real-time are challenging because it increases the load on the cloud server and also reduce the response speed. Simply using a cloud server is not able to provide real-time response while handling such a large data set. Additionally, IoT applications are sensitive to network latency because they require a constant exchange of information between devices and the cloud, making CC unfeasible to handle these applications [92]. The emergence of the edge computing concept can resolve the problems associated with CC. This new computing model deploys computing and storage resources (such as cloudlets or fog nodes) at the edge of the network closer to data sources such as mobile devices or sensors. This way, it can facilitate real-time analytics while keeping data secure on the device [92]. Edge computing offers intriguing possibilities for smart agriculture, but the applications of this technology are only in their infancy in agricultural systems. Hence, few research studies are available in this area; see Table 2-5. Most of the edge computing-based agricultural systems discussed in these studies are prototypical and address a limited selection of problems in various agricultural domains. So far, interoperability and scalability issues have not received sufficient consideration.

Table 2- 5. Edge computing-based agricultural systems.

Use case No.	Service category	Edge computing techniques used	Farm type	Maturity level	Citation
43.	FM	Computation offloading	Aeroponics	Prototype	[93]
44.		Computation offloading (automated control)	Hydroponics	Prototype	[94]
45.		Computation offloading (alert generation)	Any farm	Prototype	[95]
46.	PDM	Computation offloading	Open-air	Prototype	[96]
47.	WEM	Latency reduction	Any farm	Prototype	[97]
48.		Computation offloading	Aquaponics	Prototype	[17]
49.	SM	Computation offloading (data analysis)	Open-air	Prototype	[98]

2.4.5. Autonomous robot systems in agriculture

Autonomous robot systems (ARS) are intelligent machines capable of performing tasks, making decisions, and acting in real-time, with a high degree of autonomy (without external influence or without explicit human intervention) [99]. Interest in agricultural ARS (AARS) has grown significantly in recent years because of their ability to automate some practices in outdoor and indoor farms - including seeding, watering, fertilizing, spraying, plant monitoring and phenotyping, environmental monitoring, disease detection, weed and pest controlling, and harvesting [34]. The agricultural robots use a combination of emerging technologies such as computer vision, WSNs, satellite navigation systems (GPS), AI, CC, and IoT, thereby facilitating the farmers to enhance the productivity and quality of agricultural products. AARS in smart farming can be mobile AARS, which can move throughout the working field, or fixed AARS [100]. Mobile AARSs are further classified into unmanned ground vehicles (UGVs) and 2) unnamed aerial vehicles (UAVs), which are explained in the following sections.

2.4.5.1. Unmanned ground vehicles in agriculture

Unmanned ground vehicles (UGVs) are agricultural robots that operate on the ground without a human operator. The main components of UGVs generally include; a platform for locomotive apparatus and manipulator, sensors for navigation, a supervisory control system, an interface for the control system, the communication links for information exchange between devices, and system architecture for integration between hardware and software agents [101]. The control architecture of UGV can be remote-operated (controlled by a human operator via the interface) or fully autonomous (operated without the need for a human controller based on artificial intelligence technologies) [101]. Likewise, locomotive systems can be based on wheels, tracks, or legs [101]. Despite high ground adaptability, intrinsic omnidirectionality and soil protection of legged robots, they are uncommon in agriculture. However, when combined with wheels (wheel-legged robots), these robots offer a disruptive locomotion system for smart farms. In addition to their needed characteristics for infield operations, UGV should fulfill certain requirements such as small size, maneuverability, resilience, efficiency, human-friendly interface, and safety – to enhance crop yields and farm productivity. Table 2-6 summarizes the diverse range of UGVs designed for agricultural operations.

Table 2- 6. Different types of UGVs designed for performing agricultural tasks.

Use case No.	Service category	Primary function	Tools and techniques used	Locomotion system	Farm type	Maturity level	Citation
50.	WUVM	Weed control	Modules (Vision, spray, mechanical weeding), and classification algorithms	Four-wheel-steering system (4WS).	Open -air	Prototype	[102]
51.			Vision system with Kinect v2 sensor, and random sample consensus algorithm	Four-wheel-drive (4WD)	Open -air	Prototype	[103]
52.	PDM	Pesticides spraying	RGB camera, HMI, and LiDAR	Four-wheel-drive (4WD)	Open -air	Prototype	[104]
53.			RGB camera, and laser	Four-wheel-drive (4WD)	Open -air	Prototype	[105]
54.		Crop treatment	Hyperspectral cameras, thermal and infrared detecting systems.	Four-wheel steering system (4WS)	Open -air	Prototype	[106]
55.	CM	Seed sowing	Ultrasonic sensor, and PI controller	Caterpillar treads	Open -air	Prototype	[107]
56.			Ultrasonic sensor, GSM module and actuators.	Four-wheel-drive (4WD)	Open -air	Prototype	[108]
57.		Artificial pollination	Sensing module, pollinator system, RGB camera and	Four-wheel-drive (4WD)	Open -air	Prototype	[109]

		odometry.				
58.	Harvesting	RGB-D camera and RCNN	Four-wheel-steering system (4WS).	Open-air	Prototype	[110]
59.		RGB camera and RCNN	Four-wheel-drive (4WD).	Open-air	Prototype	[111]

Most of the agricultural robotic systems presented above have a 4WD locomotive system because it offers ease of construction and control. The drawback of 4WD is that the wheels are strongly affected by terrains containing stone elements and/or cavities [101]. Hence, it is significant to explore other mechanisms, such as legged or wheel-legged locomotive systems. Some robots have computer vision systems, but due to the difficulty of developing an accurate and reliable system that replaces manual labor, most of these robots are built with a low-cost computer vision system, that is, using conventional RGB cameras. Moreover, most of the systems mentioned above are still in the research phase, with no commercial use on a large scale.

2.4.5.2. Unmanned aerial vehicles in agriculture

Unmanned aerial vehicles (UAVs) or aerial robots are aircraft with no human pilot on board. Depending on the type of technology incorporated to fly (wing structure) and autonomy level, there is a wide variety of UAVs [112]. For instance, according to wing type, UAVs can be fixed-wing (planes), single-rotor (helicopter), hybrid systems (vertical takeoff and landing), and multirotor (drone). Among these, drones (multi-rotor technology) which are lifted and propelled by four (quadrotor) or six (hex-rotor) rotors, have become increasingly popular in the agriculture sector due to their mechanical simplicity in comparison to helicopters, which rely on a much more sophisticated plate control mechanism [113]. Similarly, according to autonomy level, UAVs can be either teleoperated in which the pilot provides references to each actuator of the aircraft to control it, in the same manner, an onboard pilot would, or tele-commanded in which the aircraft relies on an automatic controller on board that is in charge of maintaining a stable flight [112]. Equipped with the appropriate sensors (vision, infrared, multispectral, and hyperspectral cameras, etc.), agricultural UAVs allow farmers to obtain data (vegetation, leaf area, and reflectance indexes) from their fields to study dynamic changes in crops that cannot be detected by scouting the ground [114]. This data permits farmers to infer information related to crop diseases, nutrient deficiencies, water levels, and other crop growth parameters. With this information, farmers can plan possible remedies (irrigation, fertilization, weed control, etc.). Table 2-7 reviews some of the UAV-based systems used for different agricultural operations.

Table 2- 7. Different UAV-based systems developed for performing different agricultural operations.

Use case No.	Service category	Primary function	UAV type	Cameras / sensors	Flight altitude (m)	Farm type	Maturity level	Citation
60.	CQM	Vegetation monitoring	Hexacopter	Hyper-spectral camera	30	Open-air	Prototype	[115]
61.		Biomass monitoring	Octocopter	RGB-sensor	50	Open-air	Prototype	[116]
62.	CM	Real-time growth monitoring	Quadcopter	Digital camera	100	Open-air	Prototype	[117]
63.		Photosynthetic active radiation mapping	Fixed wing	Multi-spectral camera	150	Open-air	Prototype	[118]
64.		Remote sensing	Helicopter	Multi-spectral camera	15-70	Open-air	Prototype	[119]
65.		Remote sensing and mapping	RC plane	Digital camera	100-400	Open-air	Prototype	[120]
66.		Rice pollination	Helicopter	Wind speed sensor	1.15, 1.23, 1.33	Open-air	Prototype	[37]
67.		Droplet distribution estimation	Quadcopter	Digital canopy imager	3.5, 4, 4.5	Open-air	Prototype	[121]
68.		UREA spraying	Quadcopter	Multi and hyper spectral cameras	Few meters	Open-air	Prototype	[122]
69.		Pesticide spraying	Quadcopter	RF module	5, 10, 20	Open-air	Prototype	[123]
70.		Pesticide spray application	Helicopter	Digital camera	3-4	Open-air	Prototype	[124]

71.		Automatic spray control system	Helicopter	Image transmitter	5, 7, 9	Open-air	Prototype	[125]
72.	WUVM	Multi-temporal mapping of weed	Quadcopter	Digital camera	30, 60	Open-air	Prototype	[126]
73.		Weed mapping and control		Digital camera	30	Open-air	Prototype	[127]
74.	IM	Water status assessment	Fixed wing	Multi-spectral camera	200	Open-air	Prototype	[128]
75.		Water stress detection	Fixed wing	Micro-hyper spectral camera	575	Open-air	Prototype	[129]
76.		Water stress investigation	Fixed wing	Digital camera	90	Open-air	Prototype	[130]
77.		Assessing the effects of saline reclaimed waters and deficit irrigation on Citrus physiology	Fixed wing	Digital camera	100	Open-air	Prototype	[131]
78.		Water status and irrigation assessment	Quadcopter	Multi-spectral camera	30	Open-air	Prototype	[132]
79.	PDM	Phylloxera disease detection	Hexacopter	RGB and multi-spectral cameras	60, 100	Open-air	Prototype	[133]
80.		Citrus greening disease detection	Hexacopter	Multi-spectral camera	100	Open-air	Prototype	[134]

Most of the systems mentioned above are still in the research phase, with no commercial use on a large scale. Other problems with these UAVs are associated with battery and flight time [112]. At the moment, lithium-ion batteries are being used because their capacity is larger than that of conventional batteries. But an increase in battery capacity increases the drone's weight, and now research is undergoing to address this issue. In addition, the existing UAVs have complex user interfaces, and only experts can use them to perform agricultural tasks. By improving the user interface and making it human-centered with multimodal feedback will allow people who are older or unfamiliar with UAV technology to control it more easily.

2.4.6. Big data and analytics in agriculture

Rapid developments in IoT and CC technologies have increased the magnitude of data immeasurably. This data, also referred to as Big Data (BD), includes textual content (i.e., structured, semi-structured, and unstructured), and multimedia content (e.g., videos, images, audio) [135]. The process of examining this data to uncover hidden patterns, unknown correlations, market trends, customer preferences, and other useful information is referred to as big data analytics (BDA). Big data is typically characterized according to five dimensions defined by five Vs, which are displayed in Figure 2-11 [136]. The paradigm of BD-driven smart agriculture is comparatively new, but the trend of this application is positive as it has the capacity to bring a revolutionary change in the food supply chain and food security through increased production. Agricultural big data is usually generated from various sectors and stages in agriculture, which can be collected either from agricultural fields through ground sensors, aerial vehicles, and ground vehicles using special cameras and sensors; from governmental bodies in the form of reports and regulations; from private organizations through online web services; from farmers in the form of knowledge through surveys; or from social media [136]. The data can be environmental (weather, climate, moisture level, etc.), biological (plant disease), or geospatial depending on the agricultural domain and differs in volume, velocity, and format [137]. The gathered data is stored in a computer database and processed by computer algorithms for analyzing seed characteristics, weather patterns, soil properties (like pH or nutrient content), marketing and trade management, consumers' behavior, and inventory management. A variety of techniques and tools are employed to analyze big data in agriculture. A summary of some of the studies is given in Table 2-8. Machine learning, cloud-based platforms, and modeling and simulation are the most commonly used techniques. Particularly, machine learning tools are used in prediction, clustering, and classification problems. Whereas cloud platforms are used for large-scale data storing, preprocessing, and visualization. There are still many potential areas that are not adequately covered in existing literature, where BDA can be applied to address various agricultural issues. For instance, these include data-intensive greenhouses and indoor vertical farming systems, quality control and health monitoring of crops in outdoor and indoor farms, genetic engineering, decision support platforms to assist farmers in the design of indoor vertical farms, and scientific models for policymakers to assist them in decision-making regarding the sustainability of the physical ecosystem. Lastly, most systems are still in the prototypical stage.

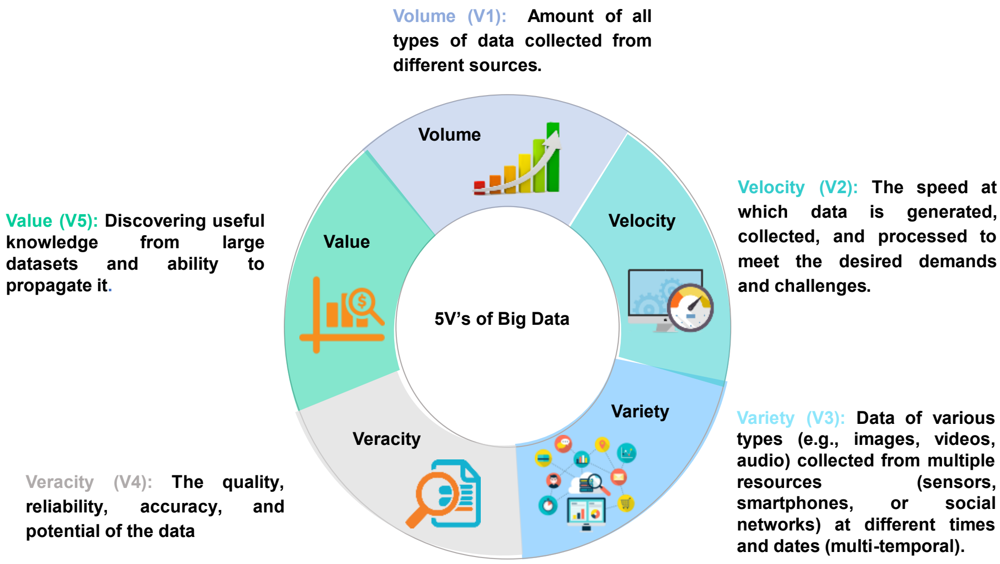


Figure 2- 11. Five dimensions of "Big Data".

Table 2- 8. Big data tools and services in agriculture.

Use case No.	Service category	Tools and techniques used	Big data source	Farm type	Maturity level	Citation
81.	WEM	Crop modelling and simulation, geospatial analysis	Weather station, historical databases	Open-air	Conceptual	[137]
82.	CM	Clustering, prediction, and classification	Sensor, historical, and farmer data	Open-air	Conceptual	[138]
83.		Support vector machine	Sensor data	Open-air	Conceptual	[139]
84.	IM	Cloud-based application.	Sensor data	Hydroponics	Prototype	[140]

85.	Cloud-based platform, and web services	Sensor data, industry standards	Open-air	Conceptual	[141]

2.4.7. Artificial intelligence in agriculture

Artificial intelligence (AI) involves the development of theory and computer systems capable of performing tasks requiring human intelligence, such as sensorial perception and decision-making [142]. Combined with CC, IoT, and big data, AI, particularly in the facet of machine learning (ML) and deep learning (DL), is regarded as one of the key drivers behind the digitization of agriculture. These technologies have the potential to enhance crop production and improve real-time monitoring, harvesting, processing, and marketing [143]. Several intelligent agricultural systems are developed that use ML and DL algorithms to determine various parameters like weed detection, yield prediction, or disease identification. These systems are discussed in the next two sub-sections.

2.4.7.1. Machine learning in agriculture

Machine learning (ML) techniques are broadly classified into three categories: 1) supervised learning (linear regression, regression trees, non-linear regression, Bayesian linear regression, polynomial regression, and support vector regression), 2) unsupervised learning (k-means clustering, hierarchal clustering, anomaly detection, neural networks (NN), principal component analysis, independent component analysis, apriori algorithm and singular value decomposition (SVD)); and 3) reinforcement learning (Markov decision process (MDP) and Q learning) [144]. ML techniques and algorithms are implemented in the agriculture sector for crop yield prediction, disease, and weed detection, weather prediction (rainfall), soil properties estimation (type, moisture content, pH, temperature, etc.), water management, determination of the optimal amount of fertilizer, and livestock production and management [145]. Table 2-9 presents a list of publications where different ML algorithms are utilized for various agricultural applications. From the analysis of these articles, “crop yield prediction” is a widely explored area, and linear regression, neural network (NN), random forest (RF), and support vector machine (SVM) is the most used ML techniques to enable smart farming. The presented use cases are still in the research phase with no reported commercial usage at the moment. Moreover, it is also found that AI and ML techniques are sparsely explored in the greenhouse and indoor vertical farming systems, particularly hydroponics, aquaponics, and aeroponics. There are only a few publications available summarized in the same table where ML techniques are employed. Considering the digital transformation's cyber-security and data privacy challenges, new approaches such as federated learning and privacy-preserving methods are being developed to enable digital farming [146]. These approaches build ML models from local parameters without sharing private data samples, thus mitigating security issues.

Table 2- 9. Machine learning-based agricultural systems.

Use case No.	Service category	Data sources	Algorithms used	Farm type	Maturity level	Citation
86.	CM	Yield maps, climate, and temporal data.	SVM with radial basis functions	Open-air	Prototype	[147]
87.		Vegetation dataset from Landsat 8 OLI.	Boosted regression tree, RF regression, support vector regression, and Gaussian process regression	Open-air	Prototype	[148]
88.		Historical soil and rainfall data	Recurrent neural network	Open-air	Prototype	[149]
89.		Plot-scale wheat data	Multiple linear regression and RF	Open-air	Prototype	[150]
90.		Temperature and rainfall records	Artificial neural network	Open-air	Prototype	[151]
91.		Soil data, and satellite imagery	Counter-propagation artificial neural networks	Open-air	Prototype	[152]
92.		Rainfall records	RF	Open-air	Prototype	[153]
93.		Field survey data of 64 farms	SVM, RF, decision tree	Open-air	Prototype	[154]
94.		Tap water samples	RF	Hydroponics	Prototype	[155]
95.	PDM	Images from a strawberry greenhouse	SVM	Greenhouse (soil-based)	Prototype	[156]
96.		Sensor data	Least squares SVM	Open-air	Prototype	[157]
97.		Sensor data	Decision trees	Aquapon	Prototype	[158]

				ics		
98.	WUVM	Image data	RF	Open-air	Prototype	[159]
99.		Images from a university farm.	SVM	Open-air	Prototype	[160]
100.	SM	140 soil samples from the top layer	Least squares support vector machines	Open-air	Prototype	[161]
101.		Humidity data from Radarsat-2	Extreme learning machine-based regression	Open-air	Prototype	[162]
102.	WEM	Rainfall data	Bayesian linear regression, boosted decision tree and decision forest regression, neural network regression	Open-air	Prototype	[163]
103.		Air temperature, wind speed, and solar radiation data	Artificial neural networks and SVM	Greenhouse (soil-based)	Prototype	[164]

2.4.7.2. Deep learning in agriculture

Deep learning (DL) represents the extension of classical ML that can solve complex problems (predictions and classification) particularly well and fast because more "depth" (complexity) is added to the model. The primary advantage of DL is feature learning which involves the automatic extraction of features (high-level information) from large datasets [165]. Different DL algorithms are convolutional neural networks (CNNs), long short-term memory (LSTM) networks, recurrent neural (RNN) networks, generative adversarial networks (GANs), radial basis function networks (RBFNs), multilayer perceptron (MLPs), feedforward artificial neural network (ANN), self-organizing maps (SOMs), deep belief networks (DBNs), restricted Boltzmann machines (RBMs), and autoencoders. A detailed description of these algorithms, popular architectures, and training platforms is available from various sources [166]. Figure 2-12 illustrates an example of the DL architecture of CNN [167]. In the agriculture sector, DL algorithms are mostly used to solve problems associated with computer vision applications that target the prediction of key parameters, such as crop yields, soil moisture content, weather conditions, and crop growth conditions; the detection of diseases, pests, and weed; and the identification of leaf or plant species [168]. Computer vision is an interdisciplinary field that has been gaining huge amounts of traction in recent years due to the surge in CNNs. It offers methods and techniques that allow the processing of digital images accurately and enables computers to interpret and understand the visual world [169].

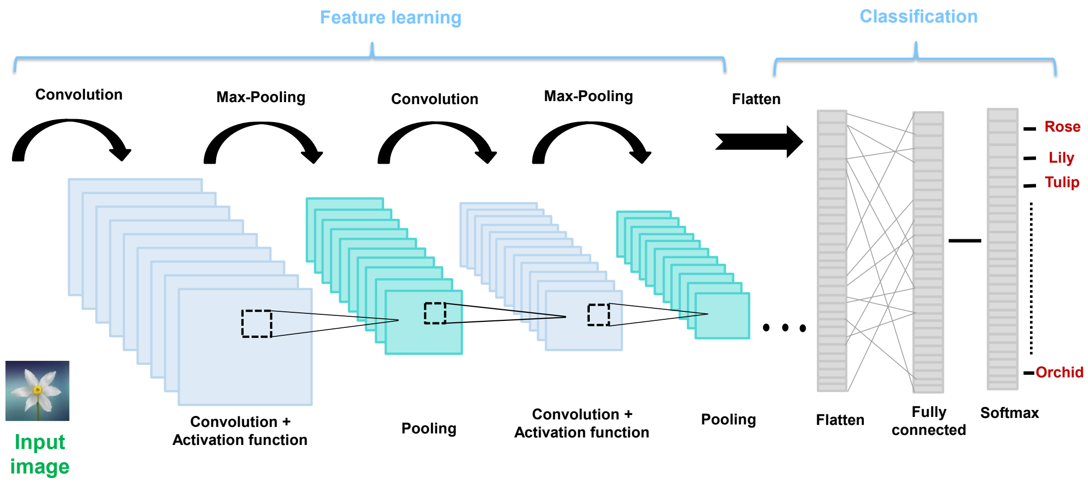


Figure 2- 12. Example of CNN architecture.

A summary of agricultural applications using DL and computer vision techniques is given in Table 2-10. Among all the DL algorithms, CNNs or Convet and its variants are the most used algorithms in agricultural applications. The variants of CNN are region-based CNNs (RCNN), Fast-RCNN, Faster-RCNN, YOLO, and Mask-RCNN, among which the first four are mostly used to solve object detection problems. Mask-RCNN, on the other hand, is used to solve instance segmentation problems. The reader could refer to the existing bibliography for a detailed description of these algorithms and their applications [168]. Few studies have also used other DL techniques. Talking about datasets, most DL models are trained using images, and few models are trained using sensor data gathered from fields. This shows that DL can be applied to a wide variety of datasets. It is also observed that most of the work is done on outdoor farms, whereas next-generation farms (environment-controlled) are not extensively explored. Though DL has the potential to enable digital farming, most systems are still in the prototype phase. Additionally, the new challenges imposed by cyber-security and privacy issues require optimization of current DL and computer vision approaches.

Table 2- 10. Deep learning-based agricultural systems.

Use case	Service category	Data sources	Algorithms used	Farm type	Maturity level	Citation

No.						
104.	CM	Satellite and weather data	LSTM network	Open-air	Prototype	[170]
105.		Rice yield data, meteorology, and area data (81 counties).	Back-Propagation neural networks and RNN	Open-air	Prototype	[171]
106.		Commercial fields' images	CNN	Open-air	Prototype	[172]
107.		Aerial orthoimages	Faster RCNN	Open-air	Prototype	[173]
108.		Historical yields and greenhouse environmental parameters.	Temporal CNN and RNN.	Greenhouse (soil-based)	Prototype	[174]
109.		Lettuce images from the farm.	CNN	Greenhouse (soil-based)	Prototype	[175]
110.	WEM	Soil moisture data, and daily meteorological data	RBMs	Open-air	Prototype	[176]
111.	CQM	Images from the farm and Google search engine	Mask-RCNN	Aquaponics	Prototype	[177]
112.	WUVM	Weed and crop species images from 6 different datasets.	CNN	Open-air	Prototype	[178]
113.	PDM	Images collected from the Internet.	CNN	Open-air	Prototype	[179]
114.		Public dataset	Deep CNN	Open-air	Prototype	[180]
115.		Images from the camera.	Faster R-CNN, and single-shot multibox detector	Open-air	Prototype	[181]
116.		Dataset with images of Walnut leaves	CNN	Open-air	Prototype	[182]
117.	FDC	RGB and multi-modal images	Faster R-CNN	Open-air	Prototype	[183]
118.		Images of oranges and green apples	CNN	Open-air	Prototype	[184]
119.		Images of ripe young and expanding apples.	YOLO-V3	Open-air	Prototype	[185]

2.4.8. Agricultural decision support systems

A decision support system (DSS) can be defined as a smart system that supports decision-making to specific demands and problems by providing operational answers to stakeholders and potential users based on useful information extracted from raw data, documents, personal knowledge, and/or models [186]. DSS can be data-driven, model-driven, communication-driven, document-driven, and knowledge-driven. The salient features of these DSSs are available in the following source [187]. Figure 2-13 presents the general architecture of a DSS, consisting of four fundamental components, each having its specific purpose.

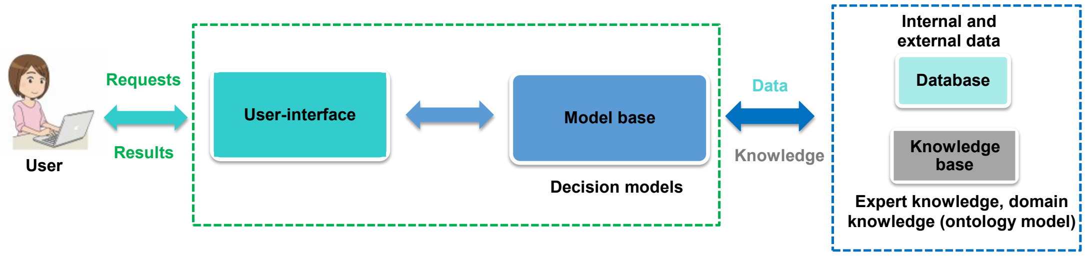


Figure 2- 13. The general architecture of decision support system.

Due to the evolution of agriculture 4.0, the amount of farming data has increased immensely. To transfer this heterogeneous data into practical knowledge, platforms like agricultural decision support systems (ADSS) are required to make evidence-based and precise decisions regarding farm operation and facility layout [188]. Over the past few years, ADSSs has been gaining much attention in the agriculture sector. A number of ADSSs have been developed that focus on a variety of agricultural aspects, such as farm management, water management, and environmental management. Table 2-11 presents a summary of the ADSSs found in the literature. From this analysis, most ADSSs have been found to not consider expert knowledge, which is highly valuable as it allows to development of systems as per user's needs. The other reported issues with some of these ADDSs are complex GUIs, inadequate re-planning components, a lack of prediction and forecast abilities, and a lack of ability to adapt to uncertain and dynamic factors. It is also worth noting that all the ADSSs are for outdoor agricultural systems and are in the research phase. In comparison, the application of ADSS in indoor soilless farming is still very much unexploited.

Table 2- 11. Agricultural decision support systems.

Use case No.	Service category	Data sources	Tools and techniques used	Maturity level	Farm type	Citation
120.	IM	Environmental and crop data	Partial least squares regression and	Prototype	Open-air	[189]

121.		Crop and site data	adaptive neuro fuzzy inference system Fuzzy C-means algorithm	Prototype	Open-air	[190]
122.	WEM	Meteorological and crop data	Geographical information system (GIS)	Prototype	Open-air	[191]
123.		Environmental, economic, and crop data	VEGPER, ONTO, SVAT-CN, EROSION, GLPROD	Prototype	Open-air	[192]
124.	FM	Environmental and crop-related data	B-patterns optimization algorithm	Prototype	Open-air	[193]
125.		Environmental and crop data	Agent-based modeling, SVM and decision trees	Prototype	Aquaponics	[194]
126.		Environmental and crop data	Object-oriented methodology	Prototype	Greenhouse (soil-based)	[195]
127.	PDM	Crop data	Excel based algorithm	Prototype	Greenhouse (soil-based)	[196]
128.		Environmental data	Rule-based approach	Conceptual	Greenhouse (soil-based)	[197]
129.		Environmental data	Rule-based approach	Prototype	Greenhouse (soil-based)	[198]
130.	WUVM	10 years of weather data and a set of vegetation indexes.	Rule-based application	Prototype	Open-air	[199]

2.4.9. Agricultural cyber-physical systems

As one of the main technologies of Industry 4.0, a cyber-physical system (CPS) refers to an automated distributed system that integrates physical processes with communication networks and computing infrastructures [200]. There are three standard CPS reference architecture models: namely, 5C, RAMI 4.0, and IIRA, and their detailed description is available in the following source [201]. Among these, the 5C is a well-known reference model with widespread usage. The architecture of 5C consists of five levels which are represented in Figure 2-14. CPS benefits from a variety of existing technologies such as agent systems, IoT, CC, augmented reality, big data, and ML [202]. Its implementation ensures scalability, adaptability, autonomy, reliability, resilience, safety, and security improvements.

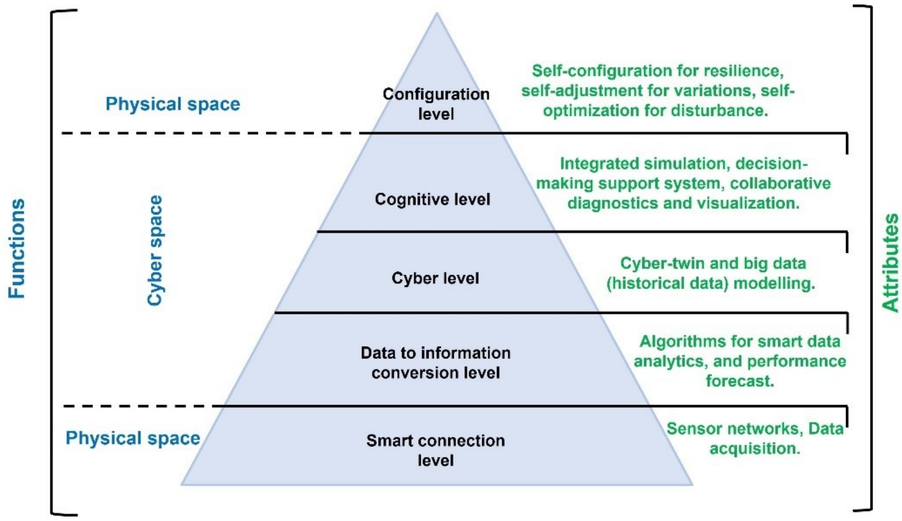


Figure 2- 14. 5C architecture for cyber-physical systems, adapted from [203].

The agricultural field is regarded as one of the complex domains that can benefit from CPS technology. Agricultural cyber-physical systems (ACPSs) use advanced electronic technologies and agricultural facilities to build integrated farm management systems that interact with the physical environment to maintain an optimal growth environment for crops [204]. ACPSs collect essential and appropriate data about climate, soil, and crops, with high accuracy and use it to manage watering, humidity, plant health, etc. A variety of ACPSs has been developed for the management of different services, and their summary is given in Table 2-12. Looking at these ACPSs, most systems are still at the prototype and conceptual level. Moreover, most studies are conducted for outdoor farms, with only a few works published related to soil-based greenhouse systems. No study is found that is relevant to indoor soilless farming systems. ACPSs has attracted significant research interest because of their promising applications across different domains; deploying CPS models in real-life applications is still a challenge as it requires proper hardware and software [205]. Moreover, particular attention should be given to autonomy, robustness, and resilience while engineering ACPSs in order to handle the unpredictability of the environment and the uncertainty of the characteristics of agricultural facilities. There are multiple factors (humans, sensors, robots, crops, and data, among others) that impact ACPSs. To ensure a smooth operation while avoiding conflicts, errors, and disruptions, ACPSs need to be designed carefully and comprehensively.

Table 2- 12. Agricultural cyber-physical systems.

Use case No.	Service category	Tools and techniques used	Maturity level	Farm type	Citation
131.	IM	Integrated open geospatial web service	Prototype	Open-air	[206]
132.		Moisture sensors, and solenoid valves	Prototype	Greenhouse (soil-based)	[207]
133.		Sensor and sink nodes, network, and control centre	Prototype	Greenhouse (soil-based)	[204]
134.		Transceiver modules, multi-sensor array and weather forecasting system	Prototype	Open-air	[202]
135.	PDM	ToxTrac and NS2 simulator	Conceptual	Open-air	[208]
136.		Sensors and cameras	Prototype	Greenhouse (soil-based)	[209]
137.		Unmanned aircraft system	Conceptual	Open-air	[210]
138.	CM	Multispectral terrestrial mobile and autonomous aerial mobile mechatronic systems, and GIS	Conceptual	Open-air	[211]
139.		Edge and cloud computing	Prototype	Open-air	[212]
140.		Sensors, actuators, Arduino, and Raspberry Pi	Prototype	Any farm	[213]

2.4.10. Digital twins in agriculture

A digital twin (DT) is a dynamic virtual replica of a real-life (physical) object of which it mirrors its behaviors and states over multiple stages of the object's lifecycle by using real-world data, simulation, and machine learning models, combined with data analytics to enable understanding, learning, and reasoning [214]. A complete description of the DT concept for any physical system requires consolidation and formalization of various characteristics, including the physical and virtual entities, the physical and virtual environments, the metrology, and realization modules that perform the physical to virtual and the virtual to physical connection or twinning, the twinning and twinning rate, and the physical and virtual processes [215]. The schematic showing the mapping of these characteristics is shown in Figure 2-15. The DT concept has gained prominence due to advances in the technologies such as the Internet of Things, big data, wireless sensor networks, and cloud computing. This is because these technologies allow real-time monitoring of physical twins at high spatial resolutions through both miniature devices and remote sensing that produce ever-increasing data streams [39].

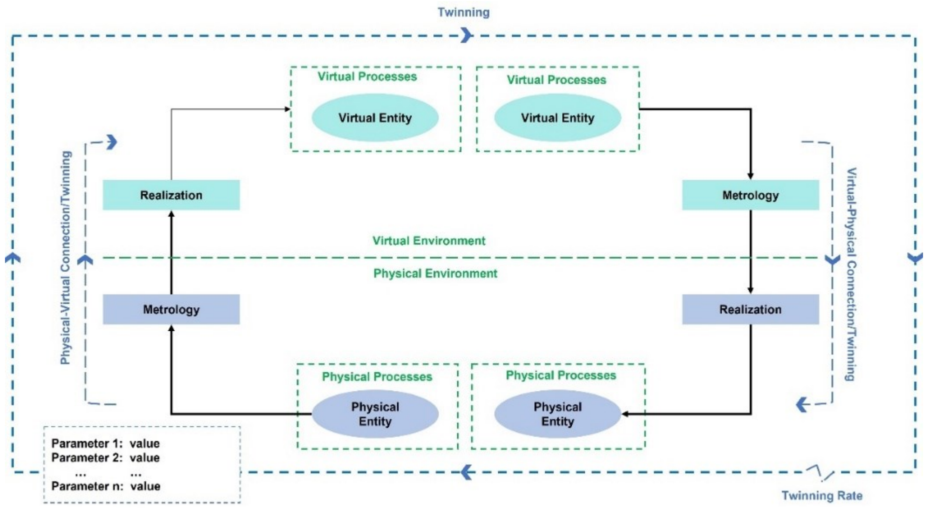


Figure 2- 15. Schematic of a digital twinning process, adapted from [215].

The concept of DT in agricultural applications is rather immature as compared to other disciplines with its first references occurring in 2017; hence its added value has not yet been discussed extensively [39]. This is because framing is a highly complex and dynamic domain because of its dependence on natural conditions (climate, soil, humidity) and the presence of living physical twins (plants and animals) and non-living physical twins (indoor farm buildings, grow beds, outdoor agricultural fields, agricultural machinery). The non-living physical twins interact directly or indirectly with plants and animals (living physical twins), thereby introducing more challenges for DT in agriculture. Whereas in other domains such as manufacturing DTs are mostly concerned with non-living physical twins. Table 2-13 summarizes the agricultural DTs developed in the last 10 years.

Table 2- 13. Digital twins in agriculture.

Use case No.	Service category	Physical twin	Tools and techniques used	Maturity level	Farm type	Citation
141.	WEM	Aquaponic system and building	IoT sensor system, and MQQT broker	Prototype	Aquaponics	[216]
142.	CM	Agricultural product	Sensor, network, and computational units	Prototype	Open-air	[217]
143.	FM	Agricultural machinery	ROS platform, Gazebo 3D and	Prototype	Open-air	[218]

			Open Street Maps			
144.		Farmland	Sensor, network, and computational units	Prototype	Open-air	[219]
145.		Agricultural farm/landscape	Sensors, and PLCs	Conceptual	Open-air	[220]
146.		Agricultural building	Sensors, GUI, and control centre	Prototype	Greenhouse (soil-based)	[221]
147.	PDM	Crops (plants)/ Trees	Mobile application and computational unit	Deployed	Open-air	[222]
148.		Trees planted on orchard	IoT sensors, network, and computational units	Prototype	Open-air	[223]

The analysis shows that most studies have focused on open-air farming systems. Only one study is found that has proposed DT for soil-based vertical farming systems and one study that implemented DT for soilless farming systems (aquaponics). This might be because the design and management of modern farming systems are challenging. Moreover, most DTs are in the research phase with no commercial deployment at the moment. The reported benefits of the DT applications in agriculture are cost reductions, catastrophe prevention, clearer decision-making, and efficient management operations, which can be applied to several agricultural subfields like plant and animal breeding, aquaponics, vertical farming, cropping systems, and livestock farming. While DT technology has great potential, achieving the synchronization between the physical entity and its digital counterpart is challenging. The complexity of this process is further amplified in agricultural systems due to the idiosyncrasies of living physical twins. Hence, implementation of agricultural DT should start with micro-farms, which can then be gradually enhanced to an intelligent and autonomous version by incorporating more components.

2.5. Roadblocks in the digitization of the agriculture industry

This section provides an answer to RQ3 by listing a series of interconnected roadblocks hampering a larger adoption of digital technologies in the agriculture sector. After analyzing 148 articles, 21 roadblocks are identified, which can be categorized at technical and socio-economic levels.

2.5.1. Technical roadblocks

- Interoperability: data is considered a cornerstone for the success of smart systems. Agricultural data usually comes from multiple heterogeneous sources such as thousands of individual farmlands, animal factories, and enterprise applications. This data can have diverse formats, making data integration complex. Hence, data interoperability is essential to enhance the value of this massively dispersed data after systematic data collection, storage, processing, and knowledge mining [224]. Likewise, for establishing effective communication between heterogeneous devices, they need to be interconnected and interoperable. With cross-technology communication, the interoperability of the system can be improved [225].
- Standardization: to fully exploit digital technologies for smart farming applications, standardization of the devices is essential. Output differences can occur because of misinterpretation and alterations from time to time. With standardization, the interoperability issues of the devices, applications, and systems can also be resolved [43].
- Data quality: to produce meaningful results, data quality is also crucial, along with data security, storage, and openness. The lack of decentralized data management systems is another roadblock hindering the adoption of smart farming practices [28]. This issue decreases the willingness of multiple actors to share agriculture data.
- Hardware implementation: the deployment of a smart agricultural setup in large-scale open fields is extremely challenging. This is because all the hardware consisting of IoT devices, wireless sensor networks, sensor nodes, machinery, and equipment is directly exposed to harsh environmental conditions such as heavy rainfall, high/low-temperature levels, extreme humidity, strong wind speeds, and many other possible dangers which can destroy electronic circuits or disrupt their normal functionality [30]. A possible solution is to build an adequate casing for all the costly devices that is robust and durable enough to endure real field conditions [226].
- Adequate power sources: typically, the wireless devices deployed at farms consistently operate for a long time and have limited battery life. A suitable

energy-saving scheme is necessary because, in case of any failure, instant battery replacement is complicated, especially in open-air farms where devices are strategically placed with minimum access [30]. The possible solutions to optimize energy consumption are the usage of low-power sensors and proper management of communication [42], [227]. Wireless power transfer and self-supporting wireless systems are other promising solutions to eliminate the need for battery replacement by recharging the batteries through electromagnetic waves. However, long-distance wireless charging is needed in most agricultural applications [28]. Ambient energy harvesting from rivers, fluid flow, and movement of vehicles and ground surface using sensor nodes offers another viable solution, but the converted electrical energy is limited at present – posing the need to improve power conversion efficiency [228].

- Reliability: The reliability of devices, as well as corresponding software applications, is crucial. This is because IoT devices need to gather and transfer the data based on which decisions are made using several software packages. Unreliable sensing, processing, and transmission can cause false monitoring data reports, long delays, and even data loss – eventually impacting the performance of the agricultural system [43].
- Adaptability: agricultural environments are complex, dynamic, and rapidly changing. Hence, when designing a system, it is pertinent for the devices and applications to proactively adapt to the other entities under uncertain and dynamic factors - offering the needed performance [5].
- Robust wireless architectures: wireless networks and communication technologies offer several benefits in terms of low cost, wide-area coverage, adequate networking flexibility, and high scalability. But dynamic agriculture environments such as temperature variations, living objects' movements, and the presence of obstacles pose severe challenges to reliable wireless communication. For instance, fluctuations in the signal intensity occur due to the multipath propagation effects – causing unstable connectivity and inadequate data transmission [229]. These factors impact the performance of the agricultural system. Hence, there is a need for robust and fault-tolerant wireless architectures with the appropriate location of sensor nodes, antenna height, network topology, and communication protocols that also require minimum maintenance [230].

- Interference: another challenge is wireless interference and degradation of the quality of service because of the dense deployment of IoT devices and wireless sensor networks. These issues can be mitigated with efficient channel scheduling between heterogeneous sensing devices, cognitive radio-assisted WSNs, and emerging networking primitives such as concurrent transmission [231]. Since agriculture devices are distributed in indoor greenhouses, outdoor farmlands, underground areas, or even water areas, cross-media communication between underground, underwater, and air is also required for the complete incorporation of smart technologies [232].
- Security and privacy: the distributed nature of smart agricultural systems brings potential vulnerabilities to cyber-attacks such as eavesdropping, data integrity, denial-of-service attacks, or other types of disruptions that may risk the privacy, integrity, and availability of the system [233]. Cyber-security is a major challenge that needs to be addressed within the context of smart farming, with diverse privacy-preserving mechanisms and federated learning approaches [146].
- Compatibility: to achieve the standards of fragmentation and scalability, the models or software applications developed should be flexible and run on any machine installed in the agricultural system [32].
- Resource optimization: farmers require a resource optimization process to estimate the optimal number of IoT devices and gateways, cloud storage size, and amount of transmitted data to improve farm profitability. Since farms have different sizes and need distinct types of sensors to measure different variables, resource optimization is challenging [234]. Secondly, most of farm management systems do not offer run-time customization in relation to the distinct requirements of farmers. Hence, complex mathematical models and algorithms are required to estimate adequate resource allocation [92].
- Scalability: the number of devices, machinery, and sensors installed at farms is increasing gradually due to advancements in technologies. To support these entities, gateways, network applications, and back-end databases should be reliable and scalable [235].
- Human-centered user interfaces: complex user interfaces of existing agricultural applications and devices are impeding smart farming practices. Most GUI is designed in a way that only experts can use to perform agricultural tasks.

Improving the user interface by making it human-centered with multimodal feedback will allow a larger group of people to use it to perform different agricultural operations [112].

2.5.2. Socio-economic roadblocks

- Gap between farmers and researchers: the participation of farmers is a key factor in the success of the digitization of the agricultural industry. Farmers face a lot of problems during the agri-food production process, which smart technologies could fix, but agricultural experts are not usually aware of these issues [35]. Moreover, to devise an adequate smart solution, first, it is important to fully understand the nature of the problems. Hence, it is essential to bridge the gap between farmers, agricultural professionals, and AI researchers.
- Costs associated with smart systems: the costs associated with the adoption of smart technologies and systems are the major deterrent to the digitization of the agricultural sector. These costs usually involve deployment, operating, and maintenance costs. The deployment costs of smart systems are usually very high as they involve; i) hardware installation such as autonomous robots and drones, WSNs, gateways, and base station infrastructure, etc., to perform certain farm operations, and ii) hiring skilled labor [236]. Likewise, to facilitate data processing, management of IoT devices and equipment, and knowledge exchange, subscription to centralized networks and software packages is required, which ultimately increases the operating costs [237]. Though sometimes service providers offer free subscription packages with restricted features, the amount of storage capacity is limited. To ensure the adequate operations of the smart system, occasional maintenance is required, which then also adds up to total costs. Other types of costs associated with smart systems deployment could be environmental, ethical, and social costs. To overcome cost-related roadblocks, initiatives focusing on cooperative farming are needed that provide; i) support services for better handling of costs and needed investments and ii) hardware solutions to transform conventional equipment into smart farm-ready machinery to reduce high initial costs [237].
- Digital division: another factor that is slowing the digitization of the agricultural sector is the lack of knowledge of digital technologies and their applications. The

majority of farmers have no idea about the significance of digital technologies, how to implement and use them, and which technology is suitable for their farm and meets their requirements [33]. Hence, it is essential to educate farmers about modern farming technologies and systems. Moreover, different strategies are needed to build tools using natural language that farmers with low education levels can easily understand [238].

- Return on investment: in agriculture, the profit margin is very important, like in other sectors. When it comes to the implementation of advanced technologies, farmers have concerns related to the time to recover the investment and to the difficulties in evaluating the advantages [31].
- Trust building: unlike in other disciplines, building trust regarding the effectiveness of smart technologies in agriculture is difficult because many decisions affect systems that involve living and non-living entities, and consequences can be hard to reverse [35]. Additionally, insufficient proof of the impact of digital tools on-farm productivity further intensifies the current challenges.
- Laws and regulations: different regions and countries have different legal frameworks which impact the implementation of digital technologies in the agriculture sector, particularly in monitoring and agri-food supply [87]. Likewise, regulations related to resource allocation (spectrum for wireless devices), data privacy, and security also vary from one country to another [87].
- Connectivity infrastructure: most less-developed countries usually have insufficient connectivity infrastructure that limits access to advanced digital tools that would help to turn data from heterogeneous sources into valuable and actionable insights [29].

2.6. Discussion

This section discusses the main conclusions of RQ1, RQ2, and RQ3. In addition, added value, considerations, and future directions are also presented to ensure higher accuracy and great advancements in the agricultural industry.

2.6.1. RQ1, RQ2 and RQ3

The present study tried to articulate the emerging digital technologies being implemented in the agricultural industry to anticipate the future trajectories of agriculture 4.0. By looking at Tables 2-2 to 2-13 in section 2.4, it can be seen some technologies such as big data and analytics, wireless sensor networks, cyber-physical systems, and digital twins are not significantly explored in agriculture. A reason for this gap could be that implementing advanced technologies with more complex operations can be expensive, at least in the early experimental phase of their adoption. Hence, the development of these technologies in the agricultural industry should increase in the coming years. The results of SLR also show that IoT is significantly implemented in farms. This is due to the broad functionality of IoT such as in the monitoring, tracking, and tracing of agriculture machinery, and precision agriculture [39]. It can be said that IoT is one of the main research objectives within the agriculture 4.0 approaches. Nevertheless, only a few studies have considered data security and reliability, scalability, and interoperability while developing an intelligent agricultural system.

The research findings also demonstrated that most use cases are still in the prototype phase. The possible reason could be that most agricultural operations have to do with living subjects, like animals and plants or perishable products, and developing systems is harder than non-living human-made systems. Another reason might be that agriculture is a slow adopter of technology because of the transdisciplinary nature of this field, and hence to develop intelligent systems, the agricultural community must become familiar will all the digital technologies. Lastly, variations in plant/crops species and growth conditions also make the digitization of agricultural systems complex [204]. The SLR findings also show that most of the systems are developed for open-air soil-based farms contrary to indoor farms (soilless and soil-based). This is due to the complex design and management of indoor farms, particularly soilless farms where parameters and factors (pH, air temperature, humidity, etc.) to be controlled are diverse [22]. But with the introduction of digital technologies and data-driven computer applications in indoor farms, more robust control of the process can be achieved. Furthermore, it is also revealed from SLR that limited research is conducted in three (soil management, fruit detection, counting, and crop quality management) out of nine different service categories mentioned in section 2.3. This corroborates that substantial research and development are needed in some areas to ensure the successful digitization of the agriculture industry in developed countries as well as in developing countries.

The complexity of the agriculture ecosystem presents a series of interconnected roadblocks that hinder the full integration of digital technologies for agriculture 4.0 realization. Hence, it is essential to identify potential roadblocks in order to come up with strategic solutions to overcome them. This study is an attempt to explore what these roadblocks are. Based on the analysis, 21 roadblocks were identified and classified at technical and socio-economic levels. These roadblocks are listed in section 2.5, which suggests what needs to be done for the digitization of the agricultural industry on a larger scale. But it is still not known to what extent the elimination or mitigation of these roadblocks assists in the successful integration of digital technologies.

2.6.2. Added value of agricultural digitization

Based on the analysis, several benefits that can motivate framers and other actors to support the digitization of the agricultural industry are identified and summarised below. The presented benefits have the potential to maximize the farm's productivity and enhance product quality, but they should not be considered a panacea for challenges associated with smart agriculture [237].

- Improved agility: digital technologies improve the agility of farm operations. Through real-time surveillance and forecast systems, farmers or agricultural experts can rapidly react to any potential fluctuations in environmental and water conditions to save crops [236].
- Green process: digital technologies make the farming process more environmentally friendly and climate-resilient by significantly reducing the usage of in-field fuel, nitrogen fertilizers, pesticides, and herbicides [239].
- Resource use efficiency: digital platforms can improve resource use efficiency by enhancing the quantity and quality of agricultural output and limiting the usage of water, energy, fertilizers, and pesticides [3].
- Time and cost savings: digital technologies enable significant time and cost savings by automating different operations, such as harvesting, sowing, or

irrigation, controlling the application of fertilizers or pesticides, and scheduling the irrigation [240].

- Asset management: digital technologies allow real-time surveillance of farm properties and equipment to prevent theft, expedite component replacement and perform routine maintenance [29].
- Product safety: digital technologies ensure adequate farm productivity and guarantee a safe and nutritious supply of agri-food products by preventing fraud related to adulteration, counterfeit, and artificial enhancement [233].

2.6.3. Considerations and future prospects

The upcoming initiatives would result in significant improvements in the agricultural sector. But in order to make things sustainable for small and medium-scale growers, roadblocks mentioned in section 2.5 need to be addressed first. Awareness campaigns highlighting the significance of smart agriculture at every level of the agricultural value chain and promoting innovative ways (such as gamification) to encourage stakeholders to take on an active role in the digital revolution can mitigate some of the mentioned roadblocks [28]. Government-level initiatives, grants and endowments, public-private partnerships, the openness of data, and regional basis research work can also assist in coping with potential roadblocks. Lastly, a roadmap can be adopted while developing a smart agriculture system, starting from basic architecture with few components and simpler functionality, gradually adding components and functionality to develop a complex system with the full potential of digitization [39]. These considerations can pave the way for the successful implementation of agriculture 4.0.
The future prospects of digital technologies in smart agriculture involve using explainable artificial intelligence to monitor crop growth, estimate crop biomass, evaluate crop health, and control pests and diseases. Explainable AI fades away the traditional black-box concept of machine learning and enables understanding the reasons behind any specific decision [34]. Description of big data through common semantics and ontologies and the adoption of open standards have great potential to boost research and development toward smart farming. Similarly, to ensure enhanced connectivity and live streaming of crop data, 5G technology need to be extensively explored [6]. 5G technology will minimize internet costs and augment the overall user experience of farm management and food safety by performing accurate crop

inspections remotely [241]. Furthermore, it will significantly bridge the gap between stakeholders by keeping them well informed on produce availability. Lastly, blockchain in combination with IoT and other technologies can be implemented to address the challenges related to data privacy and security [242].

2.6.4. Transition to Agriculture 5.0

Industrial revolutions have always brought a breakthrough in the agricultural sector. As formally discussed in previous sections, agriculture 4.0 has great potential to counterbalance the growing food demands and prepare for the future by reinforcing agricultural systems with WSN, IoT, AI, etc. While the realization of agriculture 4.0 is still underway, there is already talk about agriculture 5.0. Agriculture 5.0 extends agriculture 4.0 with the inclusion of industry 5.0 principles to produce healthy and affordable food while ensuring to prevent of degradation of the ecosystems on which life depends [243]. The European Commission formally called for the Fifth Industrial Revolution (industry 5.0) in 2021 after observing that industry 4.0 focuses less on the original principles of social fairness and sustainability and more on digitalization and AI-driven technologies for increasing efficiency and flexibility [244]. Industry 5.0 complements and extends the industry 4.0 concept to recognize human-centricity, sustainability, and resilience [245]. It involves refining the collaborative interactions between humans and machines, reducing environmental impact through a circular economy, and developing a high degree of robustness in systems to achieve an optimal balance between efficiency and productivity. The enabling technologies of industry 5.0 are Cobots (collaborative robots), smart materials with embedded bio-inspired sensors, digital twins, AI, energy efficient and secure data management, renewable energy sources, etc., [244]. In agriculture 5.0 settings, farm production efficiency and crop quality can be enhanced by assigning repetitive and monotonous tasks to the machines and the tasks that need humans' critical thinking. For this purpose, similar to the manufacturing sector, cyber-physical cognitive systems (CPCS) that observe/study the environment and take action accordingly should be developed for the agricultural sector. This may include collaborative farm robots which will work in the fields and assist crop producers in tedious tasks such as seed sowing and harvesting etc. Likewise, digital twins in agriculture 5.0 can also offer significant value by identifying technical issues in agricultural systems and overcoming them at a faster speed, detecting crop diseases, and making crop yield predictions at a higher accuracy rate. This shows that agriculture 5.0 has potential to pave a way for climate smart, sustainable, and resilient agriculture but as of now, it is in the developing phase.

2.7. Conclusions

Increased concerns about global food security have accelerated the need for next-generation industrial farms and intensive production methods in agriculture. At the forefront of this modern agricultural era, digital technologies offered by Industry 4.0 initiative are suggesting a myriad of creative solutions. The scientific community and researchers integrate disruptive technologies in conventional agriculture systems to increase crop yields, minimize costs, reduce waste, and maintain process inputs. An SLR discussing the prevailing state of these technologies in the agriculture sector is presented in this study. After applying the SLR protocol, 148 articles were considered from the time frame of the year 2011 to 2021. Various research questions pertaining to i) current and continuing research trends, ii) functionality, maturity level, farm type and tools and techniques used, iii) primary roadblocks, and iv) added value of digital technologies; were put forward and answered. Several conclusions are drawn, such as the integration of big data and analytics, wireless sensor networks, cyber-physical systems, and digital twins in agriculture is only in its infancy, and most use cases are in the prototype phase. Likewise, 21 roadblocks are identified and classified at technical and socioeconomic levels. To ensure the digitization of the agricultural industry, these roadblocks must be analyzed and overcome. The added value of digital technologies in the agriculture industry is also identified and presented in the study. Overall, this study contributes to the research being carried out around agriculture 4.0. The primary limitation of this review is twofold: firstly, only three online repositories are considered for literature search (Scopus, IEEE, and Science Direct), and secondly, additional keywords and synonyms might return more studies. In both scenarios, it is highly unlikely that the overall findings would change. For future work, additional research databases and aspects can be considered to provide a holistic overview of the agricultural industry in terms of digitization. Moreover, studies targeting agriculture 5.0, in general, will also be included.

Chapter 3 An ontology model to represent aquaponic 4.0 system’s knowledge

3.1. Introduction

The gradual decrease in farmlands due to ongoing trends of increasing population, rapid urbanization, anomalous environmental changes, diminishing water supply, and resulting food security issues have attracted much attention towards vertical farming (VF) practices [8]. In principle, VF is a simple concept that involves growing crops in vertically stacked layers. It is essentially a controlled-environment agriculture model, which aims to optimize indoor soilless farming techniques such as hydroponics, aquaponics, and aeroponics. As one of the modern VF methods, aquaponics has the potential to be the future of agriculture as a sustainable farming method with high yield and low water consumption. As discussed, being a symbiotic process, the design and management of an aquaponic system are challenging, when scaling it up to a commercial level. However, through the introduction of automation, smart strategies, and connectivity, the aquaponic system's feasibility can be strengthened.

With the advent of agriculture 4.0—the agricultural counterpart of Industry 4.0—modern vertical farms can leverage disruptive digital technologies such as the internet of things (IoT), cyber-physical systems (CPS), artificial intelligence (AI), wireless sensor networks (WSN), big data and analytics (BDA), autonomous robot systems (ARS) and ubiquitous cloud computing (UCC) to achieve sustainable intensification. Aquaponic 4.0 system is a digital farm based on a smart farming concept that uses all these technologies to bring improvements in systems' design and operation by ensuring autonomous monitoring and control and intelligent data-driven decisions in the fast-processing pervasive environment [246]. The realization of aquaponics 4.0 brings flexibility and adaptability to the system; however, it requires efficient data integration and information flow among different domains. Data can be defined as a value (measurement or descriptor) that by itself has no meaning [247]. Data can be of two types: i) data created by people, which are mainly distributed through the Web (social networks, emails, online television, online broadcasting, etc.) or available in form of books, documents, and computer files; and ii) data generated by multiple heterogeneous sources such as sensors, IoT devices and suggested services [247]. This mixed traffic of data needs to be stored, categorized, mined, and processed to extract useful knowledge and utilize it to solve complex real-world problems such as managing complex processes. When this data is placed in context, it acquires meaning which then provides information about a certain object [247]. Information that is structured and organized as the result of cognitive processing and validation becomes knowledge [247]. The continuous evolution of digital technologies, however, has led to complex systems' architectures – generating enormous volumes of data with diverse formats. The exponential increase in data size is causing interoperability issues making data integration and knowledge extraction complex and difficult [248]. Recent research has shown that semantic web technology (SWT) plays a key role in solving the problems of data heterogeneity, interoperability, interpretation, and integration [224], [249]. To ensure reliable semantic modeling, knowledge management, and data integration, ontologies are considered promising tools. Ontologies are used to store information from multiple sources and provide an automatic process known as reasoning to infer new knowledge that has not been explicitly incorporated [250]. They are commonly employed in the development of a knowledge base - one of the building blocks of a decision support system. Besides, ontologies have appeared as an alternative to relational databases (RDB) and are considered more powerful than RDB because; i) accessibility of the data is simple in ontologies – it is easy to define changes, and 2) Inference of new knowledge from existing models is carried out using reasons in ontology, whereas in RDB one needs to create every single link manually to infer new knowledge – making RDB difficult to manage in case of knowledge modeling of big data [251].

3.1.1. Research motivation and contribution

So far, no attempt has been made to develop an ontology model for the aquaponic 4.0 system. Therefore, this study aims to create an ontology "AquaONT" for the aquaponic 4.0 system by utilizing web ontology language (OWL). AquaONT stores aquaponics 4.0 knowledge gathered from domain experts, literature, databases, and IoT devices installed at the farm. It also builds the semantic relevance among fundamental elements of an indoor soilless vertical farm, such as optimal environmental parameters, system configuration, and product qualitative aspects. It can provide the optimal solution for the operation of IoT devices based on contextual data received from the farm, updates on crop quality along with respective causes and treatments, and design configuration of grow beds concerning crop characteristics, when integrated with the suitable interface. The practical implementation of AquaONT in the context of engineering design (parametric design automation of aquaponics grow beds based on crop characteristics) can be found in a study which is the extension of this work and is covered in the next chapter (Chapter 3). Overall, the current work can be used as a semantic framework to build agricultural applications that will allow vertical farming practitioners to access each dimension of aquaponics knowledge for more precise decision-making regarding crop production and facility layout in aquaponics farms. In summary, the main contributions of this work are listed as follows.

- Review of the knowledge modeling approaches and current state-of-the-art ontology models in the agriculture sector.
- Description of the domain-specific concepts and sub-concepts of an aquaponic 4.0 farm and the relationships between them.
- Populating the respective knowledge domains with data and information from multiple sources to enable automatic decision-making related to various aspects such as process, design, environment, and quality control.

3.1.2. Chapter organization

The rest of the chapter is organized as follows. Section 2 presents a theoretical analysis focusing on the general concepts of an aquaponic 4.0 system, knowledge modeling, and research contributions related to ontology-driven smart systems and agricultural ontologies. The research methodology used to develop AquaONT is discussed in Section 3 followed by its detailed formulation in Section 4. The implementation and validation of AquaONT are presented in Section 5. In Section 6, a detailed discussion of the work is provided. Finally, the concluding remarks and future directions are presented in Section 7.

3.2. Theoretical analysis and state of the art

3.2.1. Aquaponics 4.0 historical landscape

The historical interaction between the industrial revolution and aquaponics evolution is depicted in Figure 3-1. Aquaponics technology was first implemented towards the end of the 19th century, when indigenous tools were used to develop the system and can be referred to as aquaponics 1.0 [5]. Intensive research was conducted afterwards to determine the efficiency of aquatic plants at filtering and consuming the nutrients in wastewater from aquaculture farms, and several electrical devices such as water pumps, aerators, and fish feeders were installed, which can be marked as aquaponics 2.0 [252]. Around the 1970s, technologies like robotics, information technology (IT), embedded systems, and software engineering were integrated with the aquaponic system to enable precision farming, which can be marked as aquaponics 3.0 [12]. The research regarding the implementation of industry 4.0 concepts in an aquaponic system started towards the end of 2016 [253]. This can be marked as the beginning of aquaponics 4.0—a digital aquaponics farm that involves remote monitoring and control of ecosystem parameters, a high degree of automation, and intelligent decision-making to ensure high crop yield and quality.

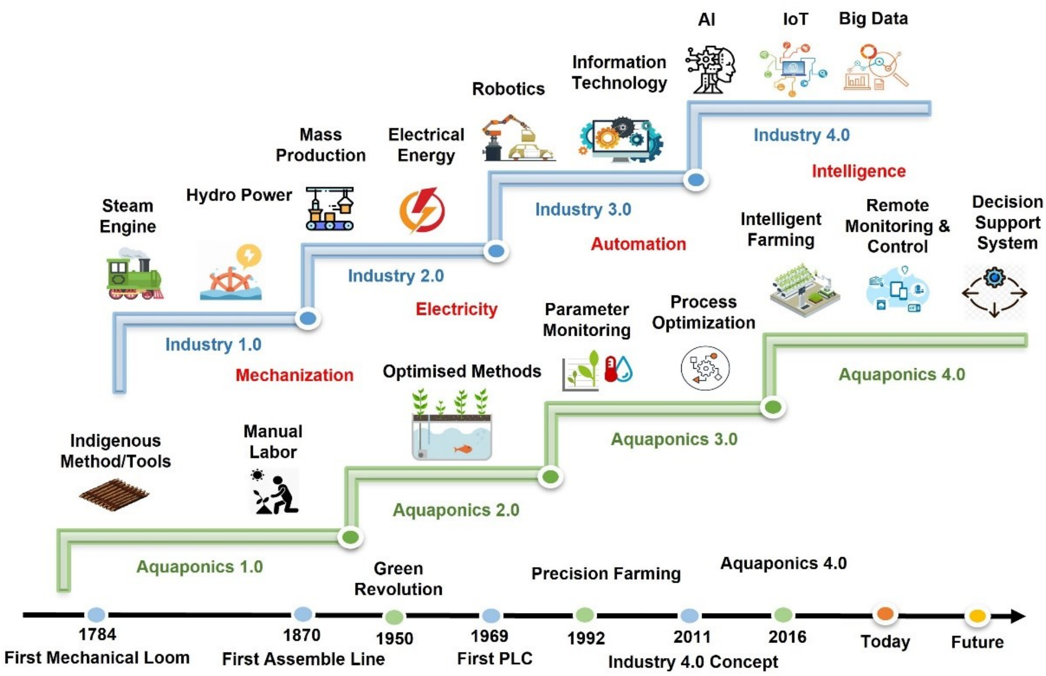


Figure 3- 1. Roadmaps describe the interaction between the industrial and aquaponics evolution.

Applications of various Industry 4.0 technologies in an aquaponic system are reflected in Figure 3-2. The realization of such a high level of digitization requires efficient data integration and information flow along with knowledge management so that the system can vary and adapt its behaviors to different scenarios based on past experiences and learning capabilities [254].

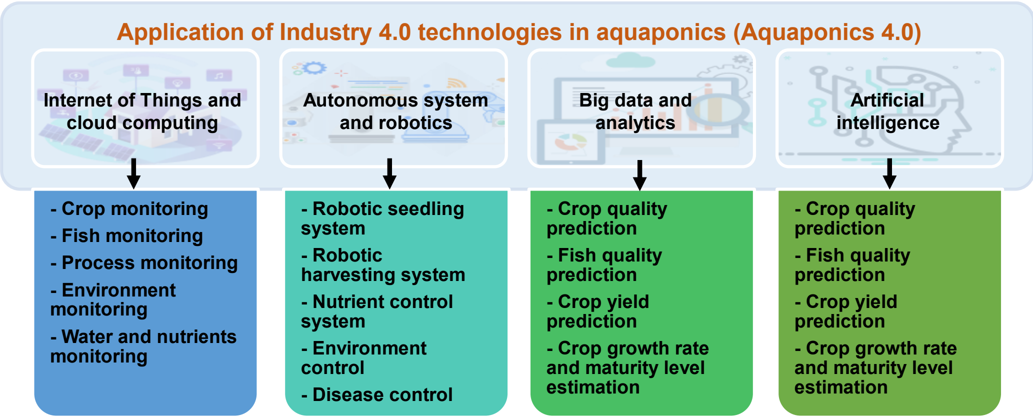


Figure 3- 2. Aquaponics 4.0—Applications of Industry 4.0 technologies in an aquaponic system.

3.2.2. Knowledge modeling conceptual overview

According to Davenport et al., knowledge is a high-value form of information combined with experience, context, and interpretation that is applied to decisions and actions [255]. To store and represent this knowledge, modeling techniques are employed [256]. Based on fundamental theories of knowledge base technology, knowledge modeling and manipulation techniques can be classified into four groups such as 1) linguistic knowledge base; 2) expert knowledge base; 3) ontology; and 4) cognitive knowledge base, and their detailed description can be found at reference [257]. For situations in which large data acquisition systems are used, i.e., aquaponic 4.0 system, ontologies are employed as they support the creation of customized rich web-based data platforms and ease data accessibility to interested parties. Therefore, for this study, an ontology-based knowledge modeling and manipulation technique is employed to model knowledge of the aquaponic 4.0 system.

3.2.3. Overview of ontology modeling

Ontology is a branch of metaphysics that is related to the philosophy of the "being". Ontology can be defined as a formal, explicit specification of a shared conceptualization, where “conceptualization" refers to an abstract model of some phenomena in the world that can be identified by its relevant and explicitly defined concepts and constraints [258]. Ontologies provide mechanisms to represent shareable heterogeneous data among domains, in the form of knowledge models that vary due to the complex and dynamic nature of a system. An ontology is typically formulated as a tuple, O={C,I,OP,DP}, where C (concept) is a set of instances, I (instance) is the object in the domain, OP (object property) is the relationship between two concepts or instances, and DP (datatype property) links instances with literals (integer or string) [259]. Most researchers classify ontologies into four categories: 1) application ontologies; 2) domain ontologies; 3) representation ontologies, and 4) generic ontologies based on generality levels and conceptualization [260]–[263]. In this study, a domain ontology is proposed, representing the concepts that are valid only in a specific domain—the aquaponics domain. Detailed insights into ontology, its architecture, and related computational tools such as SWRL and SPARQL can be found at [264]–[268]. To develop an ontology model, Protégé[4] is employed, which is an open-source ontology editor and framework developed at Stanford University [269].

3.2.4. Ontology-driven smart systems

The recent advancements in big data, IoT, and cloud computing have spurred the rise of artificial intelligence (AI) in various sectors such as agriculture, aquaculture, manufacturing, healthcare, etc. AI is all about data analysis, which is roughly classified into two categories: data-driven and rule-based [270]. Traditionally, many smart devices and services use a set of rules for situation recognition and inference. These rules are extracted from the long-term experiences and knowledge of human experts (human learning). Semantic modeling (ontology) is a representative technique for this approach - performed on already-built semantic models (a set of rules in semantic language) and new data [271]. Today, the focus is more on data-driven approaches such as machine learning and deep learning - extensively used in image recognition and natural language processing because of their high accuracy. Both approaches, however, have their drawbacks and benefits. In semantic modeling, rules should fit the real phenomena consistently. The rules may become obsolete if circumstance changes dynamically and drastically. In machine learning, assumptions are made on the data reflecting the whole phenomena we are interested in. The analysis becomes meaningless if the assumptions made are not satisfied [272].

[4] https://protege.stanford.edu/products.php#desktop-protege

Recently, ontologies are increasingly being used in combination with machine learning and deep learning techniques to model smart ecosystems such as smart homes, smart farms, smart factories, and power grids [36], [260], [273]–[288]. The main applications include solving interoperability issues, detecting the cyber-attacks, remote monitoring and controlling of system's parameters and entities, analysis of big data, creating and analysis of digital twin models, predicting patient disease, checking the accountability of AI systems, improving the monitoring of industrial operations, enhancing the flexibility of control solutions in human-robot collaborative cells, and optimizing the design parameters for maximum reliability and minimum cost. Examples of how ontology and machine learning or deep learning technologies are used in these applications are context modeling, semantic filtering, automatic ontology population, utilization of background knowledge stored in ontology models, developing intelligent decision support systems, and ontology-based learning and applications. Table A.1 given in Appendix A provides a summary of relevant publications, where ontologies are used in combination with digital technologies.

The idea of combining machine learning, AI, and ontology modeling techniques is relatively new in the agricultural field, and hence, there are only a few studies available, which are mentioned towards the end of section 3.2.5. Most of these studies have either focused on enabling smart services (monitoring and control) in IoT-based farming systems or detection of cyber-attacks using ontology models. None of these studies has directly used machine learning in combination with ontology modeling. There could be various reasons for that: i) Agricultural data is vast and scattered, and no unified model is available to keep that data in one place, ii) Agricultural sector is a slow adopter of technology, owing to the increasing complexity of IT (information technology), and iii) Presence of complex biological processes, environmental parameters, and living organisms. Therefore, the notion of presenting examples of different domains (Table A.1 in Appendix A), where ontology is used in combination with smart technologies is to highlight the significance of the ontology modeling approach and how it can be used in different capacities to bring improvements in a variety of systems.

3.2.5. Agricultural ontologies

Ontologies have been a dominant research area for the representation, storing and management of agricultural knowledge. For instance, AGROVOC, initially published in the early 1980s is the renowned agricultural thesaurus [289]. AGROVOC is the amalgamation of a controlled RDF vocabulary with around 32000 concepts covering several areas of interest, such as food, nutrition, farming, fisheries, forestry, and the environment. Another prominent ontology model is crop ontology (CO), which was designed to provide a structured and controlled vocabulary for significant crops' phenotypes for food and agriculture research [290]. Several ontologies are developed to represent specific aspects of a crop or a system. For example, Aree et al. proposed an ontology model for Thai rice, aiming to present a plant ontology prototype and specify crop growth data [291]. In the study by Maleerat et al., domain ontology based on the agricultural expertise retrieval framework "ARGIX" was constructed in Protégé with the idea of improving the performance of information retrieval using simple query terms and association rules mining method for inference [292], [293]. Suresh et al. developed a farming ontology with extensible vocabulary to support the dataset with agricultural aspects related to production, geography, and meteorology [294]. Hifza et al. extended this ontology to include several other classes, such as water, pesticide, nutrients, and seed, to assist farmers in decision-making related to rice crops [295]. To represent the technical knowledge of the agriculture operations field, Elcio et al. presented a formal task ontology model [296]. The field operations associated with task agents, agent roles, input resources, task and sub-task decomposition, control flow, task concepts, attributes, and relations were defined. Most recently, Aydin et al. proposed a generic ontology-based data acquisition model to create data acquisition forms based on a model-view-controller (MVC) design pattern, with the notion of publishing and using agricultural open data platforms [297]. A tool OWL2MVC was developed that integrates the hazelnut ontology to illustrate the proposed model's effectiveness for generating data acquisition forms.

Ontologies are also developed to enable smart farming services. For instance, Chukkapalli et al. developed a member farm and co-op ontologies for a connected cooperative, a smart farming ecosystem to provide a more accurate and data-driven dimension to precision agriculture [283]. AI applications are employed to use the information gathered from the cloud for diagnosing the critical conditions of the farm, such as crop diseases, soil conditions, water levels, etc. Sai et al. developed a smart farm ontology (incorporates users, sensors, and systems in a farm) for implementing an attribute-based access control (ABAC) to evaluate access control requests in farms dynamically [285]. Sivamani et al. proposed a vertical farm ontology (VFO) model to enable a smart service based on ubiquitous sensor networks [298]. Their research concentrates more on the monitoring and controlling of the internal and external environment parameters. This work was extended to link VFO with web services aiming to assist different entities related to hardware, user, service, and environmental factors [299]. In the study by Kim et al., an ontology using a context modeling technique is designed for an intelligent service in a vertical farm by integrating several environmental and control factors, which were validated for different scenarios [300].

The contributions mentioned above have established the philosophical foundations in representing agricultural knowledge, but most of the models are designed for soil-based methods. For soilless VF, such as hydroponics, aquaponics, and aeroponics systems, limited or no research is conducted in relation to knowledge representation through ontology models. Moreover, the current models are limited to the representation of knowledge concerning a single product type and its growth data, monitoring, and controlling of environmental parameters, usage of pesticides, and seed plantation. Nevertheless, no heed has been paid to integrating and specifying the heterogeneous metadata related to crop quality and soilless VF design. With these observations in mind, this work aims to present an ontology model for representing and storing multidimensional knowledge of an aquaponic 4.0 system with a notion to use it for developing further applications using machine learning models, which will assist aquaponics practitioners in decision-making related to aquaponics farms.

3.3. Methodology for ontology development

To represent and model the essential knowledge of the aquaponic 4.0 system, an ontology is developed. There are multiple methodologies to create an ontology from scratch. In this study, the "methontology" approach is employed to formulate and evaluate AquaONT. This approach presents a well-structured framework to develop ontologies from scratch by enlisting and tracking the activities necessary in an ontology development process [301]. These activities are classified into six stages:

ontology specification, knowledge acquisition, knowledge conceptualization, knowledge formalization, ontology evaluation, and ontology verification and validation, which are presented in Figure 3-3.

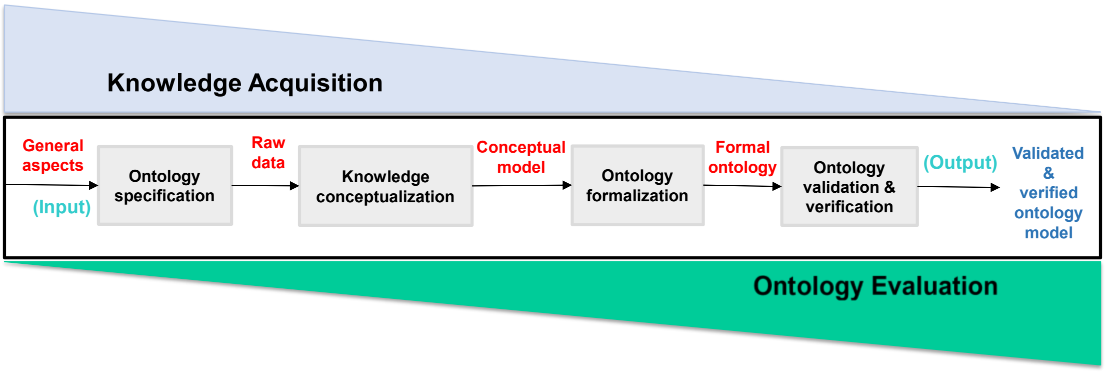


Figure 3- 3. Methontology approach for ontology development, adapted from [301], [302]

First, ontology's scope is specified in the ontology specification stage by describing and assembling the general aspects such as the ontology domain, the purpose of ontology, its intended users, the level of formality, and several key terms. A conceptual model is then developed in the second phase – involving organization and modeling of the raw data gathered in accordance with the scope during the knowledge acquisition stage in a more formal and structured form. A glossary of terms (GT) describing physical and conceptual objects related to each knowledge domain is created in tabular or graphical form. A conceptual model (taxonomy) is converted to a formal model (computable) in the ontology formalization phase, using an ontology editor such as Protégé. Domain-specific concepts and sub-concepts are defined in this stage. Finally, to test the correctness and coherence and detect incompleteness, inconsistencies, and redundancies, formal ontology is evaluated. The evaluation process is carried out during each phase, between the phases, and at the end, and it involves validation and verification [302]. Verification refers to "building the system right" and involves a technical process that ensures that the ontology is built correctly as per the requirements established during the specification phase [301]. Whereas, validation refers to "building the right system." It ensures that the ontology corresponds to the system that it represents and guarantees that the designed ontology performs correctly with an acceptable level of accuracy by checking the quality of the solutions when the system is queried [301]. The validated and verified ontology model is then used in different applications for automatic decision-making.

3.4. Formulation and evaluation of AquaONT

Following the steps mentioned in section 3.3, formulation and evaluation of AquaONT are carried out. Each stage is comprehensively described in the following sub-sections.

3.4.1. Ontology specification

An informal ontology specification document written in natural language and describing the clear and concise purpose and scope of an ontology was generated for AquaONT before its formalization, see Table A.2 in Appendix A. The purpose of AquaONT was specified in this document, which is to structure, model, and store the aquaponic 4.0 system's knowledge and use it to enable data-driven decisions for farmers by developing a functional decision support system. These decisions will be related to determining the optimal growth environment, assessing the system configuration based on product characteristics, and evaluating the quality of products based on the environment-based contextual data.

3.4.2. Knowledge conceptualization

To organize and store the gathered knowledge, a conceptual model representing the upper-level hierarchy of AquaONT is developed, see Figure 3-4. The resources used to gather the knowledge for AquaONT include i) literature, which provided information about optimal environmental parameters, optimal growth parameters, qualitative aspects of the product, and standard operation of the aquaponics farm; and ii) the aquaponic 4.0 system established as a learning factory (AllFactory) at the University of Alberta, which provides real-time contextual data [303]. For the AquaONT, six knowledge domains are identified for creating a conceptual model such as consumer product, production system, production facility, ambient environment, product quality, and contextual data. The common dependencies among and within knowledge domains are identified as the relationships and are illustrated in the model. Several GT e.g., Crop, Fish, Digital_System, Qualitative_Value_Assessment, and relationships between these terms, e.g., Product_Quality "is_Determined_by" Quality_Aspects, are identified for AquaONT.

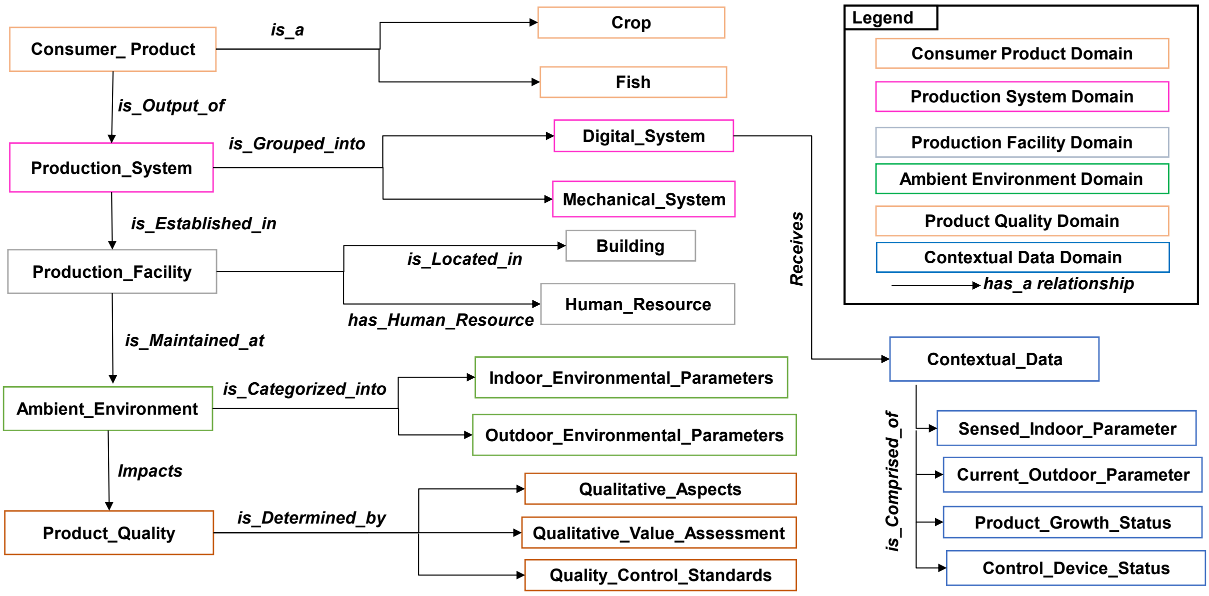


Figure 3- 4. Conceptualization tree for the upper-level ontological model of AquaONT.

3.4.3. Ontology formalization

The upper-level ontological knowledge model developed for AquaONT during the conceptualization stage is formalized and implemented using Protégé 5.5. Six "classes" or "concepts" were created for the six knowledge domains mentioned in the previous subsection, and accordingly, "subclasses" were formed, see Figure A.1 of Appendix A. The relationships between these classes and subclasses were specified using "object properties." Instances of classes are modeled using "individuals," and attributes are stipulated using "data properties." In the next sub-sections, all the contents are distinctly presented and explained.

3.4.3.1. Domain-specific concepts

Consumer product concept

A product is the outcome of any production system. In an aquaponic 4.0 system, the notion of the Consumer_Product class is to provide an abstract view of the type, growth status, and growth parameters of ready-to-harvest crops and fish, which are primary products in this case [304], [12]. The hierarchical tree-like structure of this class is shown in Figure 3-5. Crop as an indispensable part of any aquaponic system is further defined to include subclasses such as Crop_Growth_Parameters (optimal humidity, light intensity, water temperature, air temperature, pH, carbon dioxide

(CO_2), etc.), Crop_Type (leafy green vegetables such as lettuce, basil, mint, cabbage, and cilantro), and Crop_Growth_Status (the growth rate of a specific crop from seed to ready-to-harvest crop). The second product in the aquaponic system is fish, which plays a vital role in supplying nutrients for the healthy growth of plants in the RAS. The subclasses Fish_Species (fish used in an aquaponic system), Fish_Growth_Parameters (water pH level, ammonia content, amount of dissolved oxygen, etc.), and Fish_Growth_Factors (growth rate, feeding rate, stocking size, etc.) are specified under the class Fish.

Ambient environment concept

For the healthy growth of crops and fish in an indoor aquaponic 4.0 system, it is necessary to ensure that environmental parameters (water and atmospheric) are within the optimum range [305]. The Ambient_Environment concept is modeled in this regard to specify the threshold of indoor influenceable environmental parameters according to the tolerance range of crops and fish considered under the subclasses Optimal_Water_Parameters and Optimal_Atmospheric_Parameters. The hierarchical diagram of this concept is given in Figure 3-6. In an aquaponic system, living organisms can exist only if water and atmospheric parameters such as temperature, pH, electroconductivity, ammonia, dissolved oxygen, nitrate and nitrite level, water hardness, light intensity, humidity, and CO_2, water level, water flowrate, alkalinity, and salinity are within the optimal range or else they may perish [305].

Contextual data concept

The IoT-based system has many heterogeneous environments that consist of several devices generating context information/data. Therefore, it is necessary to integrate, store and share this information between system entities for which ontologies are used [306]. The contextual information is usually gathered from sensors through proper connectivity channels and is utilized to enable data-driven decisions [254]. Hence, contextual information plays a vital role in a ubiquitous environment. The Contextual_Data concept in AquaONT specifies the set of environmental parameters and crop growth status at a particular location at a scheduled time interval. The hierarchical structure of this concept is shown in Figure 3-7. Three subclasses are identified for this concept include Sensed_Indoor_Parameter, Current_Outdoor_Parameter and Current_Product_Status. The instances of these subclasses are populated by importing cloud data of farm in ontology model using transformation rules.

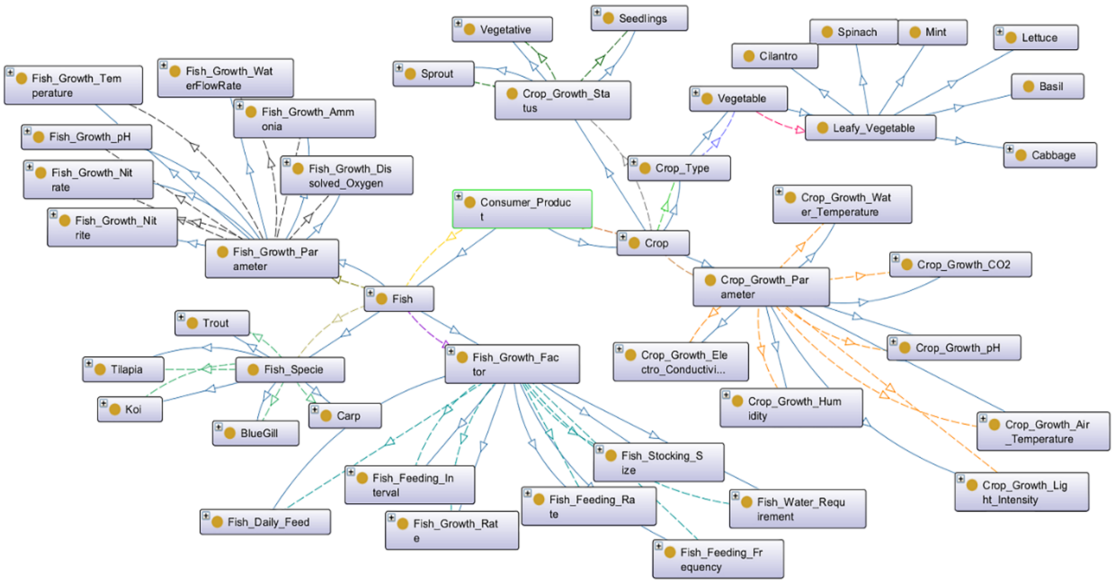


Figure 3- 5. Consumer product concept showing knowledge elements of crop and fish.

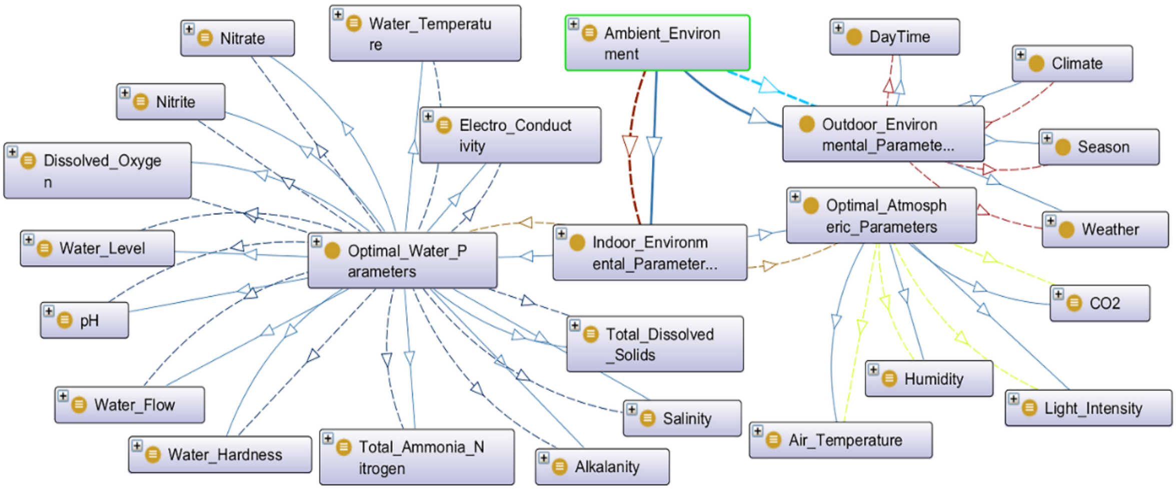


Figure 3- 6. Ambient environment concept showing a variety of atmospheric and water parameters.

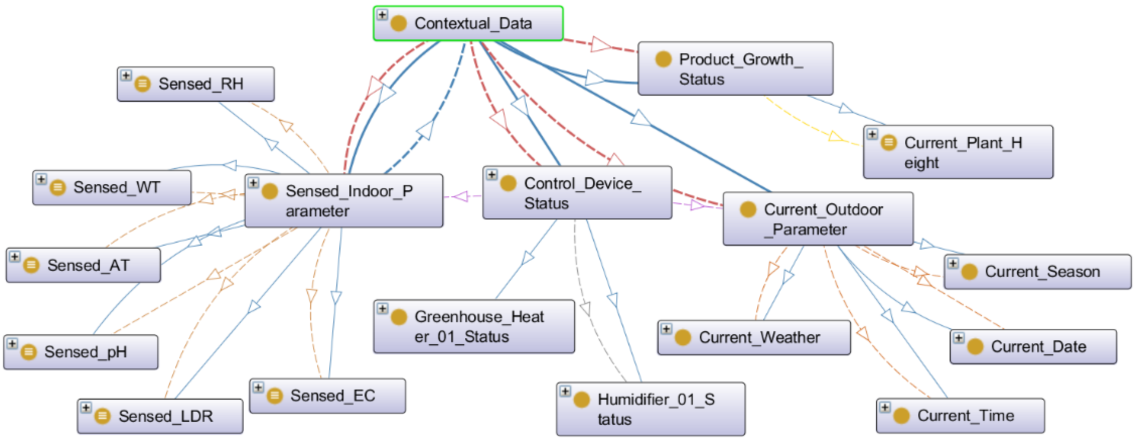


Figure 3- 7. Contextual data concept showing real-time data that is imported in AquaONT.

Production system concept

An indoor aquaponic 4.0 system consists of various mechanical and electrical components. The Production_System concept is devised to model the knowledge of these components under the subclasses Digital_System and Mechanical_System, shown in Figure 3-8. The mechanical system of the aquaponic 4.0 farm is comprised of NFT (nutrient film technique) grow channels for plant growth represented under subclass Hydroponic_Unit and the fish tanks for fish habitat and biofilters for the breakdown of ammonia mentioned under subclass Aquaculture_Unit. The design of these systems is dependent on the physical characteristics of plants and fishes, such as height and width [6]. The subclass Digital_System lists sensors, controllers, and other electronic and network devices that are employed in an aquaponic 4.0 farm to achieve autonomous services, such as remote monitoring and control [241].

Product quality concept

The Product_Quality concept models and stores the qualitative product aspects, quality control standards, and quality assessment criteria and links these attributes with the knowledge represented under Consumer_Product, Production_System, and Ambient_Environment classes. The diagram showing the hierarchical structure of this concept is depicted in Figure 3-9. The Subclass Qualitative_Aspects represents the quality attributes of both fish and crops, such as products' physical aspects (size, shape, color, structure, transparency, or turbidity, etc.) and nutritional value (well-balanced ratio of carbohydrates, fats, proteins, minerals, etc.) as standardized by local and international food associations [307]. The Quality_Control subclass defines the set of procedures adopted to ensure that product meets the desired quality standards set in subclass Qualitative_Aspects. It involves two main areas; product inspection and product handling. The last subclass defined under this concept is Qualitative_Value, which classifies the product as a good or bad quality product by comparing the information from product inspection with qualitative aspects.

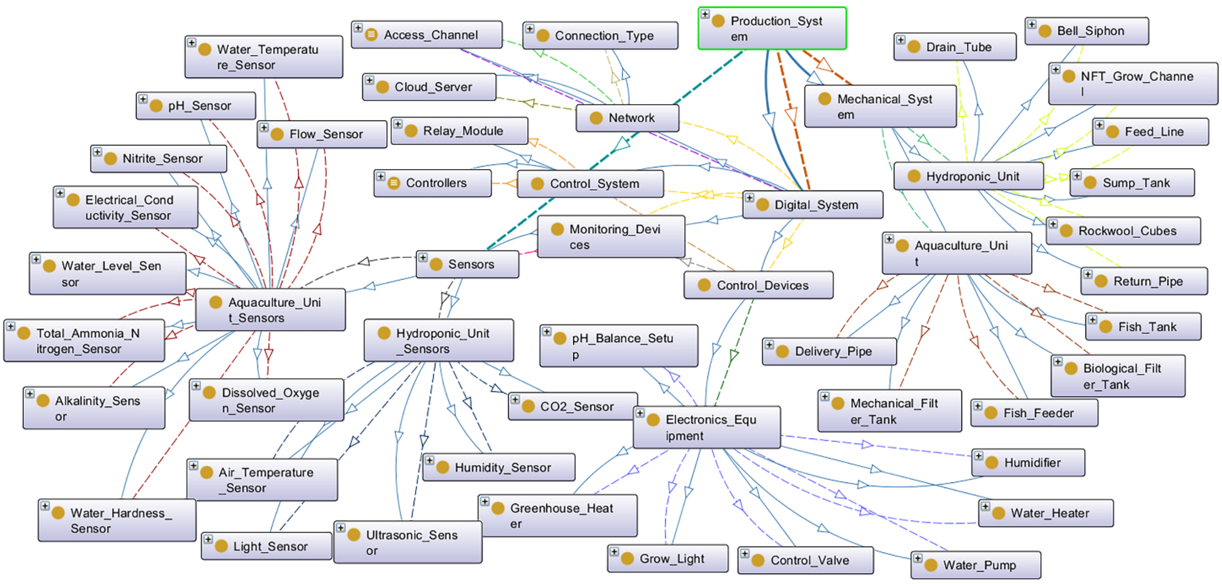


Figure 3- 8. Production system concept with digital and mechanical components in aquaponic 4.0 farm.

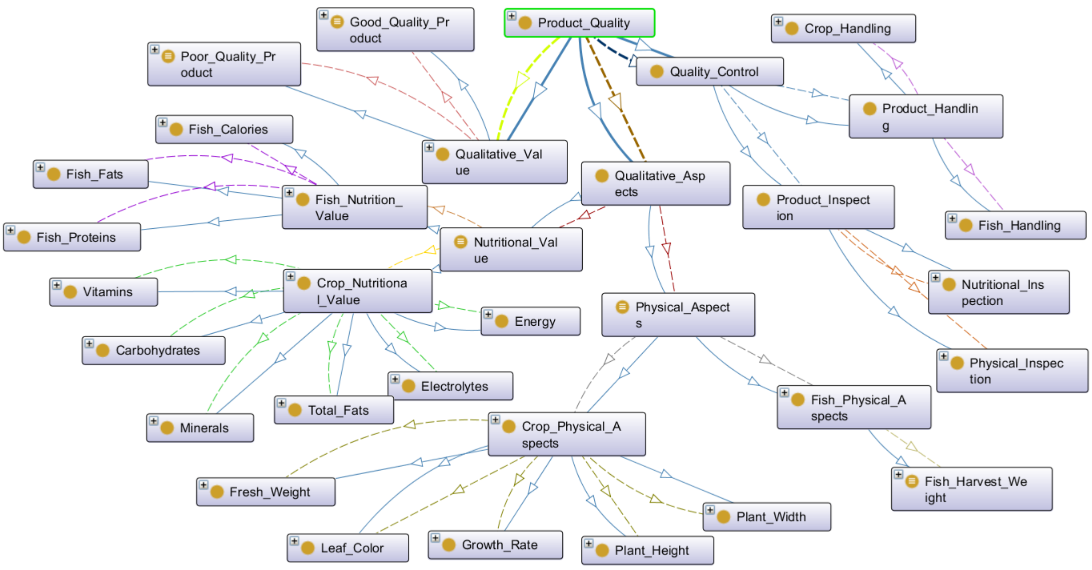


Figure 3- 9. Product quality concept showing potential elements related to qualitative aspects.

Production facility concept

The main purpose of an indoor aquaponic 4.0 system is to increase crop yields by maintaining optimal conditions, even in the middle of urban areas [2]. Production_Facility concept shown in Figure 3-10 is designed to specify the location of the aquaponic 4.0 farm.

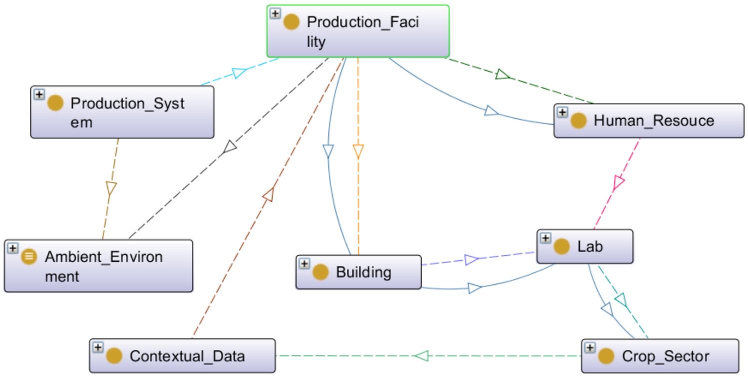


Figure 3- 10. Production facility concept showing the interactions of farm facility and other domains.

In this study, the production facility is the AllFactory. The crop grow area in the facility is divided into multiple sections referred to as crop sectors to grow a variety of crops. For instance, Sector_01 is allocated to grow lettuce, and Sector_02 is assigned to grow cabbage. Each crop sector has its digital system composed of sensors, control devices, controllers, and network access points.

3.4.3.2. AquaONT instances

For the AquaONT, a total of 310 instances/individuals are defined for different classes and subclasses. For example, the instances defined for Ambient_Environment class are Optimal_Light_Intensity, Optimal_Humidity, Optimal_Temperature, and so forth. Likewise, to classify the product quality, the instances defined are Good_Quality_Crop (crop meeting the required qualitative aspects) and Poor_Quality_Crop (crop lacking the desired qualitative aspects). To further express the outstanding quality issues, potential causes, and recommended solutions in the context of poor-quality lettuce, instances are specified, as shown in Figure 3-11. Instances are also defined for subclasses of Production_Facility, Consumer_Product, Production_System, and Contextual_Data concepts. For the Contextual_Data concept, the instances are real-time data coming from sensors and control devices. For the

Production_System concept, the defined instances give information about operating voltages, equipment identifiers, suppliers along with a sector-wise description of all the sensors, devices, and control architecture installed in the AllFactory. The list of instances of some other classes is given in Figure A.2 of Appendix A.

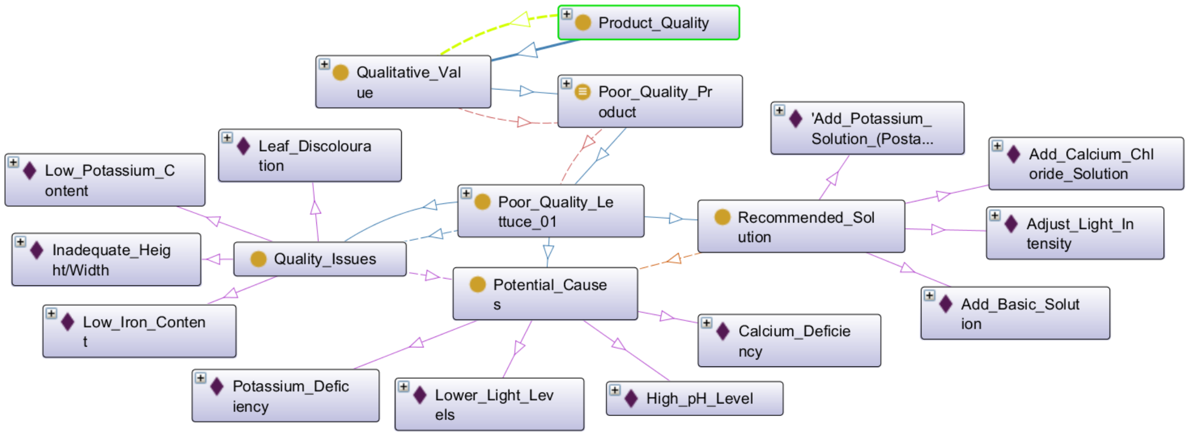


Figure 3- 11. Instances for subclasses of Product_Quality concept specifying qualitative requirements.

3.4.3.3. Object and datatype properties

The domain-specific concepts defined for AquaONT are related to each other, and their relationships are represented through the property module, which consists of two types of properties viz object property and the datatype property. Object properties represent the ontological relationships that link different classes (concepts) together [308]. In AquaONT, 139 object properties are defined to represent all the interactions between classes and subclasses. Figure A.3 and A.4 in Appendix A show object properties of AquaONT and relationships between classes and subclasses, respectively. Object property assertions are also applied to develop relationships between different instances. For example, in AquaONT, to determine the solution, causes, and quality issues of a poor-quality product, the instances of subclass Recommended_Solution are linked with the instances of Quality_Issues through object property is_Recommended_for, and instances of Quality_Issues are linked with instances of Potential_Causes through is_Consequence_of. The second type of property is the datatype property that links an instance to an RDF literal. RDF literals can be in the form of Boolean or integer values, as well as string variables [308]. For the AquaONT, 53 datatype properties are created, and they are used with 575 different literals. The list of datatype properties of AquaONT is given in Figure A.5 of

Appendix A with Figure A.6 shows the relationship between some of the instances and attributes.

3.4.3.4. Standards, restrictions, and rules

After creating the instances for subclasses of domain-specific concepts, these are assigned numerical and qualitative values through datatype properties. These values conform to the standards defined in literature or local bodies. For example, in AquaONT, the optimal values for instances of Ambient_Environment class were taken from the literature [5]. The value of Optimal_Light_Intensity is PPFD (photosynthetic photon flux density) 600-900 nm wavelength, and in terms of PAR (photosynthetic active radiation), it is 5 to 17 mols/m2/day. For other indoor environmental parameters, the values of instances are defined in the same way. Likewise, for assessing the qualitative aspects of the products, the data was taken from the Health Canada database [309]. To enable the knowledge retrieval process, the real-time data from AllFactory is then compared with these standard values using a data query engine. Restrictions refer to the constraints in the ontology model. In ontologies, there are three main types of restrictions that can be placed on classes: quantifier restrictions, cardinality restrictions, and "hasValue" restrictions [269]. In AquaONT, only cardinality constraints are imposed using the object restriction creator to limit the number of sensors employed to get variable data. In Protégé, rules are usually written in the SWRL editor, which is a built-in development environment to work with SWRL. As multiple scenarios occur at the aquaponics farm, for which rules are created within the AquaONT model.

3.4.4. Ontology verification and validation

Different types of built-in reasoners are available to evaluate an ontology model in Protégé for its consistency and coherence. The most used reasoners for ontology evaluation are "Pellet" and "HermiT" [310]. For the AquaONT, both reasoners are tested with selection based on the empirical results obtained. After testing, HermiT v.1.4.3.456, a Protégé built-in reasoner, is employed for verification and validation of AquaONT because it processed the ontology in 2098 ms - lower than for "Pallet", which is 3450 ms. The computation is done with no errors, showing the consistency and coherence of AquaONT. To further verify it, the DL queries were created and executed after the reasoner classification [311]. Examples of DL queries and corresponding results for AquaONT are shown in Figure 3-12.

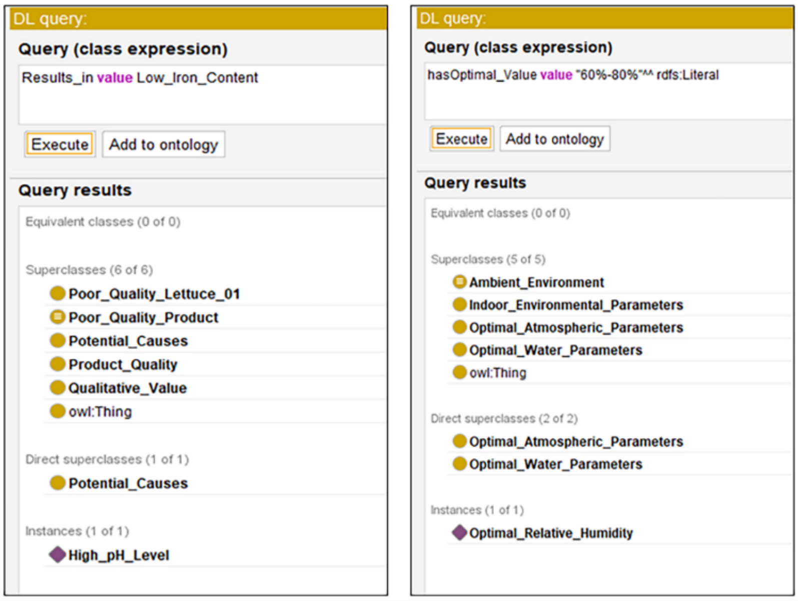


Figure 3- 12. DL queries for verification of AquaONT.

The first DL query is about finding the potential cause of one of the quality issues mentioned in the query (class expression) tab. The result shows that the potential cause of the Low_Iron_Content of lettuce is High_pH_Level, which is an individual in AquaONT. The second DL query is related to checking whether the given value of humidity is optimal or not. The result of this search shows that 60%–80% is the Optimal_Relative_Humidity, which is an individual of humidity. Nevertheless, it is also possible to search for AquaONT according to other scenarios. The results of these queries depict the correctness of the captured knowledge and requirements set during the specification stage. Likewise, the validation of AquaONT was performed, which is discussed in section 3.5.

3.5. Validation and Implementation of AquaONT

As stated, AquaONT is developed to capture and store the essential knowledge of the aquaponic 4.0 system, which then can be retrieved and used in other applications to make informed decisions for a variety of applications related to an aquaponic system, from systems' design and preparation to systems' control and monitoring. A framework in which AquaONT is currently utilized is shown in Figure 3-13. This framework has two primary purposes: 1) The first one is to structure how the data is acquired and stored in the ontology model as it populates its instances and to validate that the appropriate knowledge is being retrieved; and 2) The second one is to give an insight that how AquaONT can be integrated with the aquaponic 4.0 farm through an interface (decision support system) to control operations and ensure intelligent decision-making regarding design, quality, among others as required by the user. The framework consists of four building blocks, including 1) data generation and communication, 2) knowledge representation and modeling, 3) knowledge extraction and validation, and 4) knowledge application and decision-making. An aquaponic 4.0 system reflects a context-rich environment that has entities that generate data related to indoor environmental parameters and crop growth characteristics. Hence, the real-time data from these entities is gathered and transmitted to an accessible cloud database, which along with other databases such as for product quality, populate the instances of AquaONT concepts. The existing and inferred knowledge can be extracted from AquaONT and applied to enable autonomous decision-making and control of farm operations through an interface, which is part of future work (the fourth building block of the framework). A variety of test cases in relation to the aquaponic 4.0 farm and the capacity of AquaONT mentioned in section 3.1 were then considered to validate the accuracy of ontology. To achieve this, AquaONT was uploaded in Apache Jena Fuseki 3.16.0, which is a SPARQL server to retrieve the desired information using the SPARQL query language. These test cases are explained below.

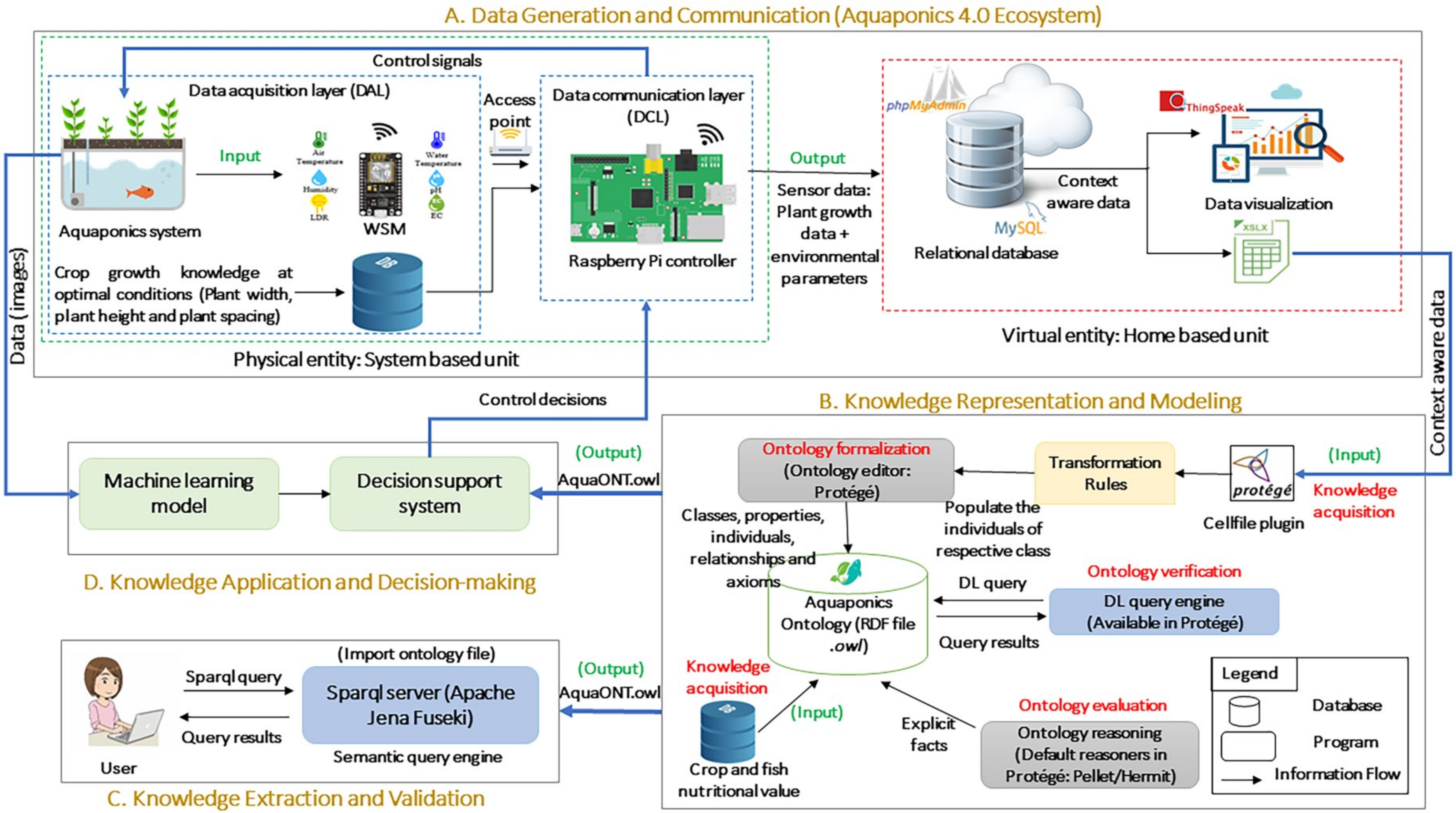


Figure 3- 13. Overview of the implementation of AquaONT in the aquaponic 4.0 system.

Test case 1

One primary function of AquaONT is that it will allow users to keep track of indoor environmental conditions at the farm. In case, if conditions are not adequate, AquaONT will tell users what to do in that scenario. To validate this, we have performed a simple test case using the historical data of the aquaponic 4.0 farm (AllFactory) available on the IoT-enabled dashboard, the layout of which is given in Chapter 5. The farm has different crop sectors, with each sector has a distinct crop with its particular atmospheric parameters and water conditions. This means each crop sector has its set of wireless sensors and control equipment. The historical data used is related to the growth of Little Gem Romaine Lettuce, which is located in sector 1. A sector can be defined as a location where vertically stacked NFT grow channels are being kept, and sector 1 is the place where Lettuce is being grown. The cloud data is imported in AquaONT, which is then compared to the knowledge stored in it, and the required information is retrieved by running SPARQL queries. One simple example of this test case is shown in Figure 3-14. Let's say a user wants to know what to do under certain indoor temperature conditions. For instance, the temperature value in the farm has exceeded 18°C. The query and corresponding results for this case are shown in Figure 3-14.

From the results, it can be seen that AquaONT has given several suggestions to a user on what to do in a particular scenario. For example, it has been suggested a user turn off the heater in sector 1, which is represented by GHH01:OFF. This is because the optimal temperature to sustain the aquaponic 4.0 ecosystem is 18°C–30°C, and for the healthy growth of the lettuce crop, the temperature should be kept between 16°C–19°C. Whereas, the temperature on the farm at the moment is 24°C, higher than needed for the growth of Lettuce. A similar pattern is followed for other temperature values (lower and optimal temperature) to see if the AquaONT is giving the correct information that corresponds to real-world phenomena. Every time, the query produces the same results for the given conditions, proposing the validation of AquaONT. These results, later on, can be integrated with some applications such as expert systems or decision support systems to control the environmental conditions at the farm. In addition to that, these results can also allow new farmers or aquaponics startups to get information on which type of sensors and electronic equipment is most suitable for their aquaponic 4.0 farms before building the physical system.

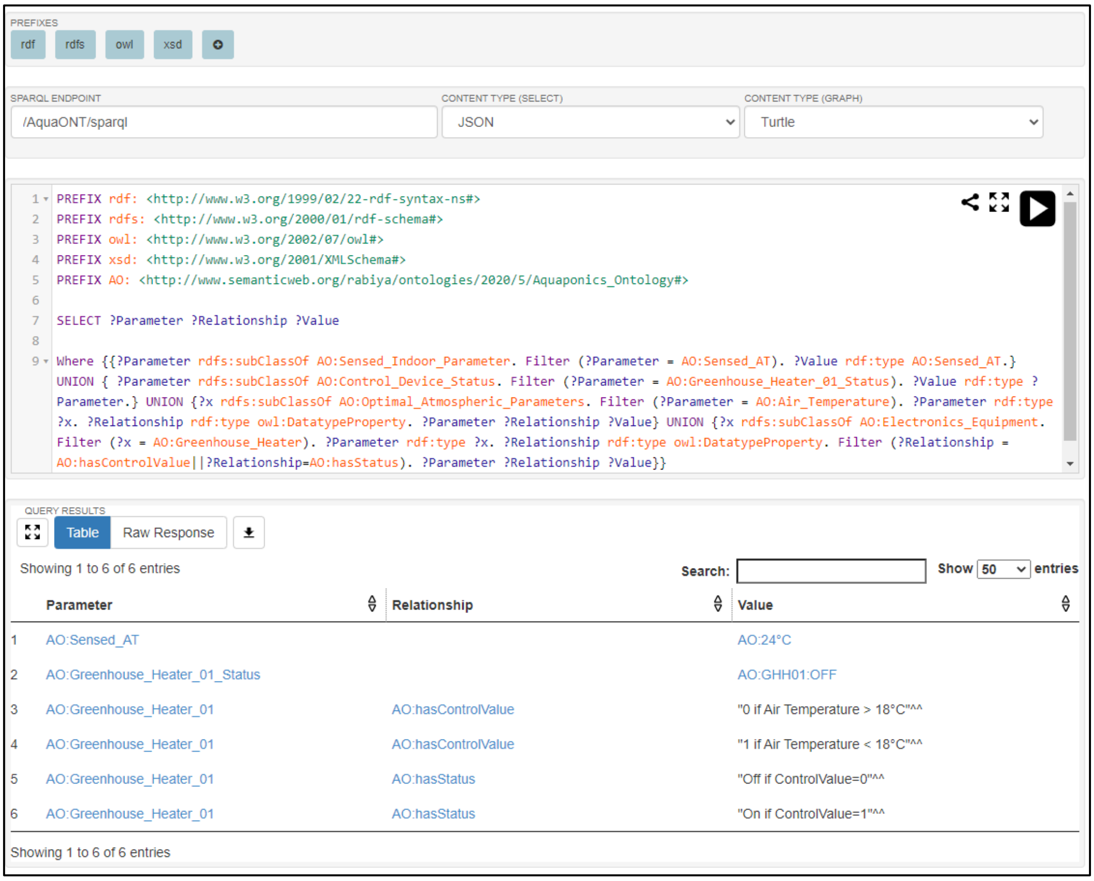


Figure 3- 14. Test case 1—SPARQL query 1 and results (Temperature variations in the aquaponic 4.0 farm).

Test case 2

To further validate AquaONT, another simple test case is conducted considering a different class and scenario. This test case reflects another primary function of AquaONT, which is to allow users to get information on different qualitative aspects of crops or fish involved in the process. A simple example of this test case is shown in Figure 3-15. When considering growing a certain crop in an aquaponic 4.0 farm, a few elements are important to be known beforehand: i) what quality issues in lettuce crop any practitioner can encounter; ii) what can be the causes of the particular quality issue, and iii) what treatments should be employed to solve this particular quality issue. Each crop has standardized qualitative aspects in terms of nutritional value and physical characteristics recommended by local food authorities that need to be met to ensure a marketable product. Hence, the answers to all these questions can be found through the knowledge stored in AquaONT, which is imported from literature, databases, and food agencies. Running simple queries, results show the potential quality issues for selected crops, i.e., quality issues for the lettuce crops in Figure 3-15. For each potential quality issue, AquaONT provides the user with information on the cause and treatment of that issue. For instance, one potential quality issue of the lettuce crop is inadequate length or width at a certain period of its growth cycle or at the time of harvesting. The primary cause of this issue could be lower calcium content in the nutrient solution, for which the suggested treatment is to add calcium chloride. This kind of information will guide a user on what necessary steps should be taken before start growing the lettuce crop.

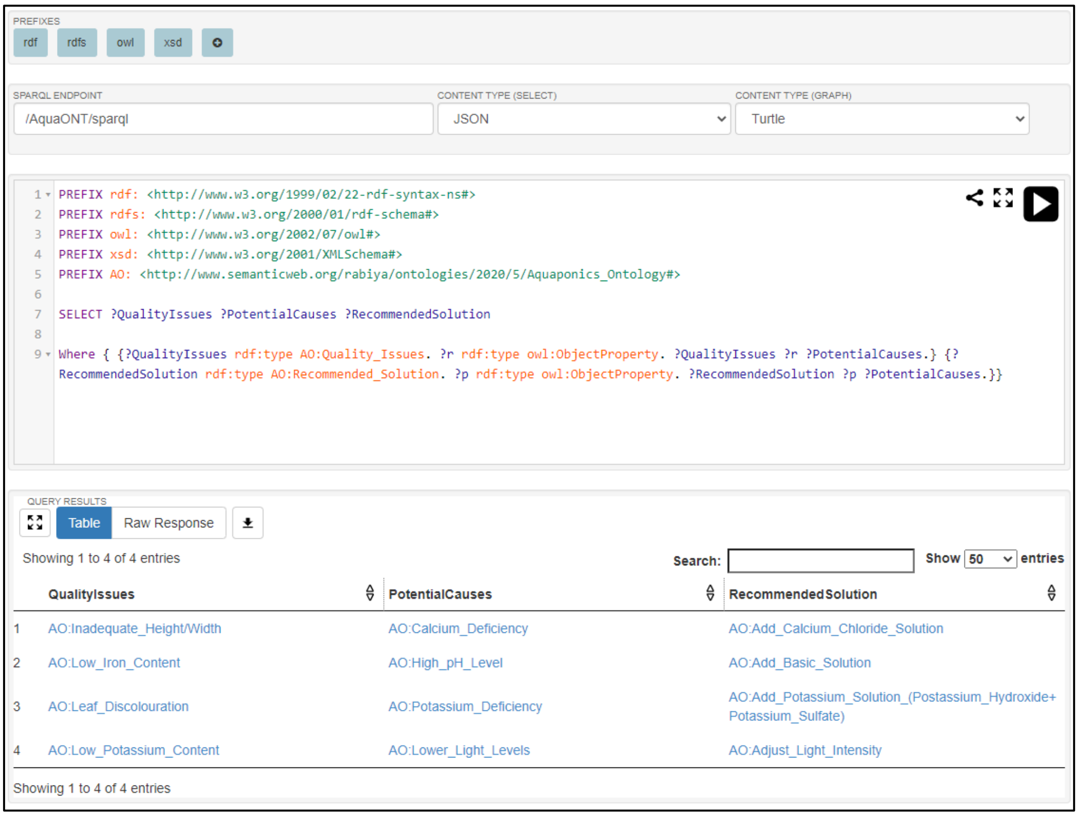


Figure 3- 15. SPARQL query 2 and results for test case 2.

Another scenario for this case study could be if a user is already growing a certain crop in an aquaponic 4.0 farm and wants to compare the quality of the crop with the standards available. Let's say that this user has a crop inspection system installed at the farm, from where it is gathering real-time data on crop height and width. This data can be populated in the instances of AquaONT. Upon running the query, the user will get information on whether the crop is of adequate dimensions or not along with causes and treatments. For this scenario, a proper interface is needed with which AquaONT needs to be integrated.

The aforementioned test cases and queries are created considering the capacity of AquaONT mentioned in section 3.1 in order to search and extract useful knowledge of real-time events happening at the farm. Upon analysis of the final results after query execution, it was observed that AquaONT provides accurate information each time the particular query is executed, which proposes the validation of AquaONT. Further, these results can be utilized to reconfigure and supervise the aquaponic 4.0 system accordingly by integrating AquaONT with an external interface or application.

3.6. Discussion

Driven by rapid advancements in AI due to big data, IoT, cloud computing, machine learning and deep learning, the agriculture sector is shifting towards a smart farming ecosystem to balance the increase in food demands. Not only that, modern farming techniques such as hydroponics, aquaponics, and aeroponics are being employed to increase crop yield, optimize plant growth, and improve crop quality. Within the scope of this study, we focused on the aquaponic system, that couples RAS with hydroponics. Research has shown that the design and management of an aquaponic system are complex due to the presence of living organisms such as plants and fishes, complex biological processes, and diverse environmental parameters [5]. To deal with these issues, the concept of aquaponics 4.0 is introduced in this chapter. Aquaponic 4.0 system is a digital farm based on Industry 4.0 technologies. It consists of smart sensors and IoT devices that bring automation to the system and provide data-driven applications to improve farming practices with minimal human and natural resources and enable farmers to make optimal decisions for the farms.

However, developing such a highly digitized system requires efficient data integration and knowledge management. With its growing number of devices and their diversity, along with the big data from divergent sources, the reality of IoT is challenging current approaches and technologies for smarter integration of data, applications, and services in the agricultural domain. Providing interoperability among IoT devices and other heterogeneous big data sources is one of the most fundamental requirements to support object addressing, tracking and discovery, and information representation, storage, and exchange. While the Web is seen as a convenient platform for integrating things, Semantic Web Technologies (SWT) can further improve its capacity to understand things' data and facilitate their interoperability along with enabling knowledge management and data exchange in a machine-interpretable way. This synergy between SWT and IoT domains gives rise to the birth of a new appellation; known as the semantic web of things (SWoT) [248]. To develop this semantic web stack for IoT, metamodels such as ontologies play a key role in facilitating semantic integration and aggregation of data generated by different sources. Moreover, ontologies can connect and qualify unstructured, semi-structured or structured data formats without any need for standardization. They streamline the process of identifying core concepts and improving classification results to collate critical information.

Considering the complexity of the aquaponic 4.0 system with different data sources and the unavailability of a unified knowledge model, AquaONT — aquaponic 4.0 ontology — is proposed in this work. AquaONT is designed to store and model the knowledge of the aquaponic 4.0 system. It gives a full description of concepts concerning remote monitoring and controlling of environmental parameters, production facility layout based on crop selection, and product quality to analyze the quality issues and suggest the desired treatments. AquaONT also enables semantic interoperability among multivariate data sources. The knowledge from AquaONT can be retrieved and used to make intelligent decisions regarding farm operations and system design by integrating it with an external interface (decision support system). The final service and quality of application, however, depends on the quality of the knowledge base, which is usually constructed from ontology models [312], [15].

With a unified knowledge model for the aquaponic 4.0 system, such as AquaONT, it is possible to get insights on what are the optimal environmental parameters for growing different types of crops, what are the optimal growth parameters for fish, and what are the optimal water and atmospheric parameters for sustaining an aquaponic 4.0 farm. This information will assist farmers in deciding which sensors and IoT devices are most suitable for their farm and their choice of the crop before building it. This will, in turn, allow farmers to avoid wasting money on the wrong sensors and IoT devices. Even if the farm is already running and growing a certain crop, AquaONT will assist in finding out if the parameters are within the range by collecting the data from sensors and comparing it with standard values. If integrated with some external interface, the knowledge from AquaONT can be exploited to control parameters and processes in the aquaponic 4.0 farm.

In an aquaponic system, the design of grow bed is directly dependent on the type of crop to be planted [6]. Each crop has certain width and height at optimal water and environmental conditions that impact the design of grow beds in terms of the spacing between plants and between channels. Hence, the same configuration of grow bed cannot be used for different crops. The correct design configuration of grow bed enables plants to absorb the right amount of nutrients, ensuring healthy crops, high crop yields, and enhanced farm productivity [313]. With the AquaONT-driven decision support platform, it is possible to visualize the impact of crop characteristics on grow bed design that assists farmers in deciding on what design configuration of an aquaponics grow bed is suitable for their crop choice and space availability before building a physical system. Such an ontology-driven platform can also enable parametric design automation by retrieving the data from the ontology model and integrating it with some CAD software. This kind of platform can save a lot of money by preventing farmers from choosing the wrong materials and saves a lot of time as a new design of the system can be built with just one click due to parametric modeling. Another significance of AquaONT is that it can provide information on quality issues, causes, and recommended treatments for different crops and fish species, which will assist farmers in taking necessary steps before building a system to avoid quality issues.

At the moment, AquaONT can only be applied to an NFT-based aquaponic 4.0 farm that grow leafy green vegetables such as lettuce, spinach, parsley, basil, and cabbage. But it can easily be expanded to include aspects and knowledge of other aquaponic 4.0 systems such as deep-water culture (DWC) and media-based aquaponics. For this purpose, the "Production_System" class will be modified to include subclasses, instances, and attributes related to the design configuration of aquaculture and hydroponic units. For instance, DWC uses more water for the hydroponics portion of the equipment with a ratio of about seven times as much water for plants as for fish [314]. Hence, the design of grow bed, the capacity of fish tanks and the specifications of biological and mechanical filters will be different from that of an NFT-based aquaponic 4.0 system. This information is required to be modeled in AquaONT before using it for different aquaponic systems. The rest of the AquaONT classes for different aquaponics setups will remain the same if leafy green vegetables are to be grown. In case other crops such as fruits, roots, and flower vegetables are to be considered, then all other classes of AquaONT will have to be expanded. Overall, the concepts of AquaONT can be extended and reused in different aquaponic 4.0 systems. Other approaches, such as empirical and theoretical, can be utilized to model the aquaponic 4.0 system [315]. These modeling techniques serve different objectives. For instance, an empirical approach that uses statistical models can be employed to perform an analysis of historical data from previous experiments. This approach is useful to estimate potential factors affecting fish and crop production in the aquaponic system, which can further be used in future experiments - making the utilization of costly research assets more effective. A theoretical approach such as mathematical modeling or system dynamic (SD) modeling can be employed to understand and optimize the farm for effective management and control of complex processes. The theoretical approach offers various benefits over the statistical approach. For instance, with theoretical models, it is possible to perform process simulation considering the hypothesis, comparative analysis of simulation results with observed data, evaluation of proposed model and hypothesis, and optimization of the system, whereas statistical models just confirm the hypothesis with no clear evidence of underlying processes [315]. Developing a theoretical model, however, requires different parameters that usually come from the empirical model [315]. A theoretical approach such as SD modeling is widely employed in literature to model the aquaponic system with many aims. These include understanding the dynamic behavior of different aquaponics processes such as crop and fish growth, waste production and filtration, greenhouse climate and hydroponics; evaluating the performance of multis-stage RAS; and estimating the energy-efficient greenhouse parameters [316], [317]. SD modeling has proven an effective tool as it provides a simulation platform to represent real-world entities in the form of equations, which allows us to understand and anticipate changes in complex systems. There are also a few drawbacks to this modeling approach. For instance, to develop a complete dynamic model for a system, sub-models of all the contributing entities have to be built and then integrated to ensure process synchronization, which is a time-consuming process. In case a new entity must be added later on, a new dynamic model has to build for this entity and linked with other sub-models. Moreover, different sub-systems and their dynamic models introduce complexities driven by time and spatial scales and multiple interactions among the factors [318]. However, with ontology modeling, it is easy to expand ontologies by incorporating new entities as classes and linking them with existing ones. But to exploit the full potential of the ontology model, it needs to be integrated with an external interface.

While ontologies offer various advantages concerning storing knowledge and solving interoperability issues, they also face various challenges. For instance, knowledge is described as a priori in ontology models, making them less adaptable to systems where the notion is to predict and analyze behaviors of different environments and users [294]. These challenges can be resolved by integrating machine learning techniques with ontology models in different applications. This is because machine learning supports the prediction and analysis of systems using classification, clustering, and association identification techniques – bringing adaptability in ontology-driven systems. However, one problem with machine learning models is that they struggle to explain the rationale for decision-making, where multi-domain semantic modeling and rule-based reasoning can excel [272]. The idea of using cloud data and underlying essential semantic knowledge with learning algorithms preserves the interoperability and re-usability of classification processes and brings intelligence to systems.

However, combining machine learning and ontology modeling is not sufficiently addressed in the agricultural domain, but there are several applications available in other domains, which are explained in section 3.2.4. These applications focus on improving cybersecurity, patient monitoring, design improvement, digital twin analysis, etc. Regarding ontology-driven IoT and AI systems for the agriculture sector, only a few studies are available, which are explained towards the end of section 3.2.5. These applications mainly focus on cyber-attack detection and monitoring and controlling of IoT-based farms. The reason behind limited research work in the agricultural sector in the context of ontology-driven AI systems or a combination of ontology modeling and machine learning approaches could be the complex dynamics of agricultural operations and the presence of complex biological processes. In this essence, AquaONT is a first step toward introducing an ontology model that can be integrated with machine learning models to bring intelligence in aquaponic 4.0 farms by autonomously managing farm operations and providing solutions for farm design.

3.7. Conclusions and future work

An ontology model, “AquaONT” is presented to model and store the knowledge of the aquaponic 4.0 system – a digital farm that uses smart technologies to improve the system’s design and operations. AquaONT provides information for the optimal operation of IoT devices by comparing contextual data coming from a farm with standard/ideal data from experts, taking corrective actions on qualitative issues of crop and fish, and designing the configuration of grow beds based on crop characteristics when integrated with the suitable interface. This information can assist farmers and users in clear decision-making regarding IoT devices, sensors, and other components necessary for farm development. To achieve remote control of the aquaponic 4.0 farm’s operation, AquaONT can be integrated with the external interface, which will exploit the AquaONT knowledge to control the electronic devices installed at an aquaponic 4.0 system such as heaters, lights and fish feeders or humidifiers, etc. The knowledge of AquaONT can be extracted and used to get an insight into crop characteristics and their impact on the design of aquaponics’ grow beds, which can be used to make a decision support platform for parametric design automation.

In future work, an autonomous decision support system will be developed by integrating AquaONT and intelligent techniques such as machine learning, deep learning, and computer vision for controlling farm operations and crop quality. Moreover, it is also possible to investigate the monetary benefits of parametric design automation of grow beds.

Chapter 4 An ontology model to support the automated design of aquaponic grow beds – Application of AquaONT

4.1.Introduction

Traditional agriculture methods employed for crop production require vast amounts of land, time, and manpower and hence are not very efficient in meeting the growing food demands. The current paradigm, therefore, poses a need to explore new farming practices such as aquaponics mentioned in previous chapters to develop and achieve economically viable and environmentally sustainable food production [319]. An aquaponic system is comprised of two integrated units: 1) a hydroponic unit that consists of grow beds for plant growth, and 2) an aquaculture unit that involves water tanks for fish habitat and biofilters for the breakdown of ammonia [320]. These units work together in a symbiotic environment to enable plant and fish growth. Primarily, depending on the structure of the plants' grow bed and crop type and size, there are three different types of aquaponic system designs: nutrient film technique (NFT), media bed, and deep water culture (DWC) [11]. In this work, the NFT-based aquaponic system is considered because it is the most popular type of aquaponic setup used. Moreover, it uses less water and is suitable for growing leafy green crops. In NFT systems, a very thin film of nutrient-rich water is pumped to enclosed channels. The top cover of the channel consists of circular or square-shaped pockets known as plant sites where plants sit in small plastic cups, allowing their roots to access the water and absorb the nutrients [321].

The design and management of an NFT-based indoor aquaponic system present several challenges when scaling it to a commercial level [5]. These challenges are mainly attributable to the design of growing channels based on crop selection. Each crop has a certain width and height at optimal environmental conditions that impact the design infrastructure of the aquaponic system in terms of plant site spacing and distance between grow channels [15]. This in turn affects the system productivity which involves crop yields and product quality. Hence, to ensure high system productivity, the proper design and placement of grow channels are significant. To achieve this, the complex and heterogeneous existing links between grow bed design and crop characteristics need to be formally described by appropriately capturing the data and managing the knowledge related to these entities. In this essence, ontology is regarded as one of the normative knowledge modeling tools that provide semantic interoperability and a general understanding of specialized multidimensional knowledge in various domains that is cognitively transparent and accessible to human experts and software agents [288], [322], [323]. The ontology models, in combination with rule systems, act as strong candidates to construct a decision support platform for the representation of different knowledge sources and the facilitation of knowledge-driven decisions in a reusable and modular manner [312].

4.1.1. Research motivation and contribution

So far, no attempt has been made toward knowledge modeling of the aquaponicsystem particularly for the representation of the grow bed design knowledge based on crop selection. Therefore, the purpose of this work is to provide a knowledge model in the form of an ontology model to support the parametric design automation in an indoor aquaponic system with the notion of automatically determining the design parameters of grow bed based on crop selection. This ontology model stores knowledge gathered from the farm, domain experts, and databases. The inferred knowledge is then extracted and used to calculate grow bed design parameters for a specific crop. To streamline the decision-making process, a graphical user interface (GUI) is developed. This research study allows aquaponic's practitioners to visualize the impact of crop selection on aquaponic system design, which eventually will facilitate better decision-making regarding crop production in aquaponic farms.

4.1.2. Chapter organization

This chapter is structured into 8 sections. Section 2 introduces a knowledge-based decision support framework for the parametric design automation of aquaponic grow beds based on crop selection. Section 3 provides an overview of the main classes of aquaponic ontology, namely, AquaONT, and the relationships between them. Governing equations devised to determine the design features of grow beds are described in section 4. The user interface developed to visualize the behavior of AquaONT is presented in section 5. Section 6 presents a use case considering the basil crop. The analysis of the results obtained in section 6 is covered in section 7. Finally, section 8 concludes the chapter by addressing the efficacy of this study.

4.2. Decision support framework for automated design of aquaponic grow beds

The effective decision-making related to the design of grow beds based on crop selection in aquaponic farms is contingent upon the representation, extraction, and usage of available knowledge about contributing entities. For this purpose, a decision support framework is proposed, the layout of which is shown in Figure 4-1.

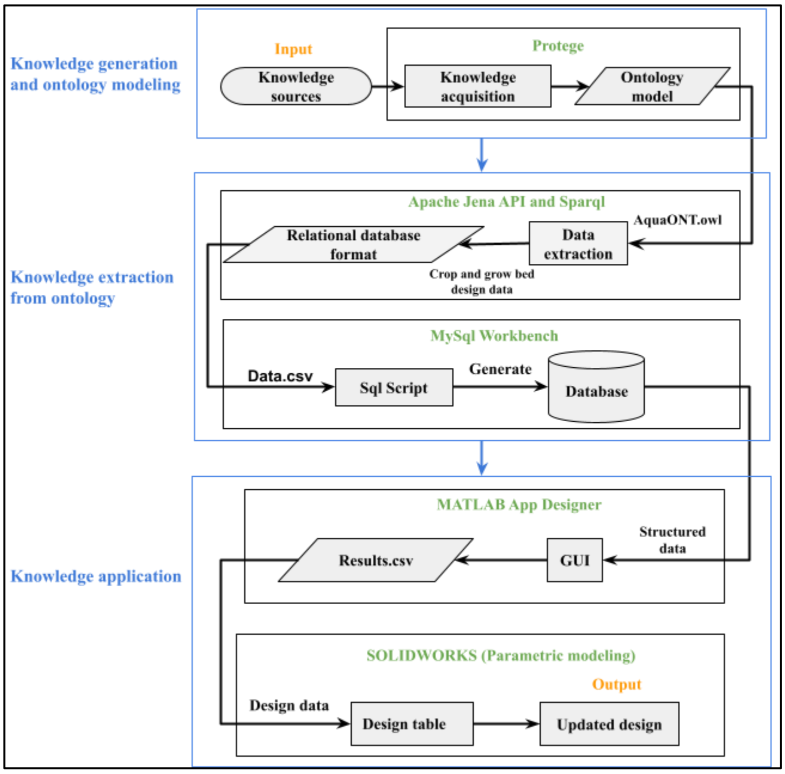


Figure 4- 1. Decision support framework for automated design of grow beds.

The proposed framework consisting of three primary stages depicts the complete lifecycle of the decision-making process based on the knowledge extracted from the ontology model. To represent the aquaponic's knowledge, first, an ontology model is developed by acquiring knowledge from farm and domain experts and unifying it as domain concepts. Then, the existing and inferred knowledge about crop characteristics and grow bed design features are extracted from the ontology model using Apache Jena API[5] and SPARQL query language. Then, MySQL Workbench[6], is used to import and organize extracted knowledge into a database. MATLAB

[5] https://jena.apache.org/tutorials/rdf_api.html

[6] https://www.mysql.com/products/workbench

database explorer toolbox is employed to link this database with the MATLAB app designer module, which along with various mathematical equations, is utilized to develop a graphical user interface (GUI). Finally, the results (design features) obtained from GUI are exported to SOLIDWORKS for parametric modeling of the final grow bed design.

4.3. AquaONT: an ontology model for the aquaponic system

In this section, AquaONT developed in Chapter 2 is used, which is an OWL ontology developed to represent and model the essential knowledge of the aquaponic system. This ontology model is created in Protégé 5.5, which is an open-source ontology editor developed by Stanford University. First, the upper-level ontological knowledge model known as base ontology is presented which provides the domain-specific concepts related to the aquaponic system. Then, product and production system concepts are presented that define the crop characteristics and grow bed features, respectively.

4.3.1. Upper-level ontological knowledge model

An ontology model, O, represents the dimensions of domain-specific knowledge in terms of four fundamental elements referred to as a tuple: O = {C, I, OP, DP}, where concept (C) is a set of instances, the instances (I) are the objects in the domain, the object property (OP) is the relationship between two concepts or instances, and the datatype property (DP) links instances with literal variables (integer or string) [259]. Figure 4-2 shows upper-level ontological model of AquaONT, also known as the base ontology model. Six "classes" or "concepts" are created to represent the six knowledge domains. These concepts are related to each other through object properties, which are given in Table 4-1. The class Ambient_Environment specifies the optimal ranges of environmental parameters that ensure the healthy growth of crops and fish in an indoor aquaponic system. These parameters are classified into two categories: 1) indoor environmental parameters, which include water temperature, pH, electroconductivity, ammonia, dissolved oxygen, nitrate, and nitrite level, water hardness, water level, water flowrate, alkalinity, salinity, air temperature, light intensity, humidity, and CO2; and 2) outdoor environmental parameters which involve the daily weather conditions, routine climatic changes, day-night times, and seasons. The notion of product in any production system refers to the outcome of the process [324].

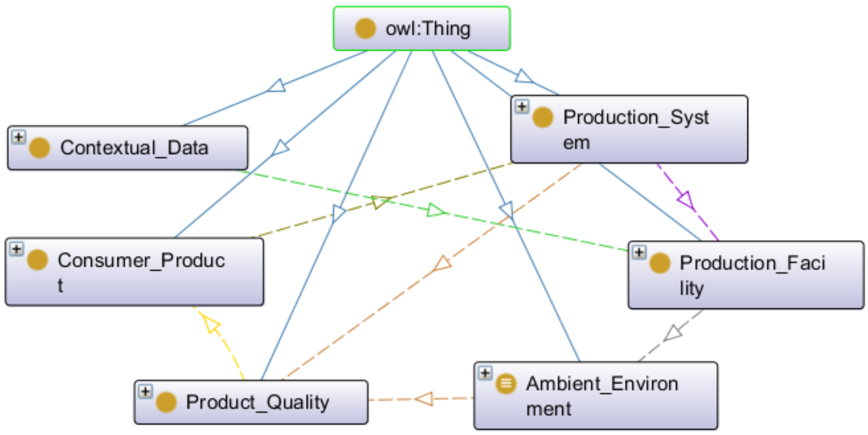


Figure 4- 2. Upper level ontological model of AquaONT.

Table 4- 1. Relationships between classes/concepts.

Domain	Object property	Range
Ambient_Environment,	have_Impact_on	Product_Quality
Production_Facility	is_Maintained_at	Ambient_Environment
Contextual_Data	is_Received_from	Production_Facility
Consumer_Product	is_Output_of	Production_System
Product_Quality	is_Characteristic_of	Consumer_Product
Production_System	is_Established_in	Production_Facility

In an aquaponic system, there are two primary products: ready-to-harvest crops and fish. Consumer_Product class represents the product knowledge in terms of crop and fish type, crop and fish growth status, and crop and fish optimal growth parameters. A wide variety of crops can be grown in an aquaponic system, but each crop needs a distinct environment to thrive and has its own standard height and width at the maturity stage or at the time of harvesting. These aspects are significant in determining the design of grow beds and, therefore, are also represented under this class. Besides biological components, an indoor aquaponic system consists of various mechanical and electrical components. Production_System class models the knowledge about these components under the subclasses digital system and mechanical system. The digital system is further categorized to include sensors, controllers, and other electronic or network devices. Whereas the mechanical system subclass represents design features of grow beds, fish tanks, and biofiltration tanks with respect to crop and fish type. In an indoor aquaponic system, the idea is to control and maintain the optimal environmental conditions to enhance crop yields, for which the location of the system plays a significant role. Production_Facility class, therefore, specifies the location where the aquaponic system is located and managed. This class also represents the workers that are responsible for managing each part of the aquaponic system through a centralized system. For remote monitoring and control of the aquaponic system, context information is obtained from sensors through proper connectivity channels and is utilized to enable data-driven decisions in the knowledge model. This context information is related to real-time data of surrounding conditions in aquaponics farms and is therefore represented under the class Contextual_Data. The Product_Quality concept models the qualitative product aspects, quality control standards, and quality assessment criteria and links these attributes with the knowledge represented for a consumer product, production system, and ambient environment covered in previous concepts.

To verify AquaONT, Protégé built-in reasoner, HermiT was used. The computation was done successfully without errors, showing the accuracy of ontology. Similarly, to validate ontology, SPARQL queries were developed and executed. Every time, these queries produce the same results for the given conditions, representing the consistency and coherence of ontology.

4.3.2. Consumer product and production system concepts

Ontologies enable the interoperability of autonomous agents and support the design of production systems [325]. In this study, AquaONT is used to enable parametric design automation – involving the determination of design features of aquaponic grow beds pertaining to each crop. To achieve this, two concepts, namely, Consumer_Product and Production_System are employed and extended to include several sub-concepts, which are then populated with the knowledge of grow bed design features and crop characteristics gathered from domain experts and farms. The detailed hierarchical architecture of these two concepts, along with significant sub-concepts and instances, is shown in Figure 4-3.

The different types of crops are defined as instances (Icrop) under the sub-concept Crop_Type. The crops considered in this study are leafy green vegetables: basil, chard, lettuce, parsley, and spinach. The characteristics of these crops involve standard plant spacing (PS), width (Wi), and height (H) as recommended by aquaponics professionals. PS is defined as the distance between two consecutive plants on the same channel. These characteristics are the attributes of respective crops represented as literals and linked with instances through corresponding datatype properties: "hasPlantSpacing", "hasPlantWidth", and "hasPlantHeight" respectively.

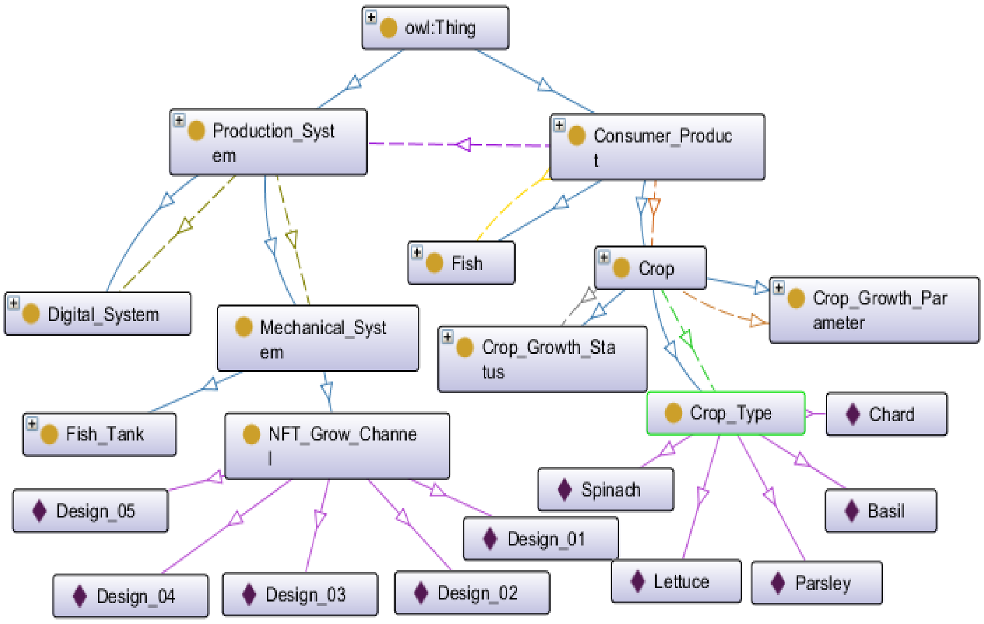


Figure 4- 3. Architecture of Consumer_Product and Production_System class.

Likewise, the design parameters of the grow channels are modeled under the instances (Idesign) of the sub-concept NFT_Grow_Channel. These instances represent different design categories, and each category specifies a certain width (W), length (L), depth (D), plant site spacing (S), plant site size (SS), vertical channel spacing (VCS), and horizontal channel spacing (HCS) of an NFT grow system. These parameters are the attributes represented as literals and linked with the design categories through datatype properties "hasWidth", "hasLength", "hasDepth", "hasPlantSiteSpacing", "hasPlantSiteSize", "hasVerticalChannelSpacing", and "hasHorizontalChannelSpacing" respectively. Figure 4-4 shows the crops' basic dimensional characteristics and generalized design features of an NFT grow channel.

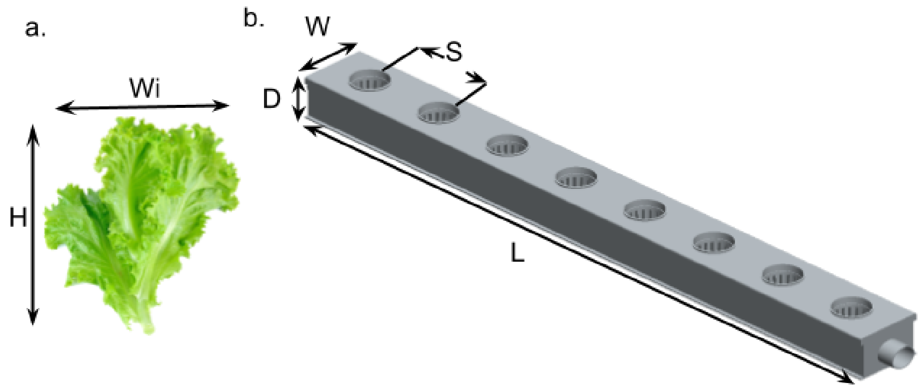


Figure 4- 4. a) Crop characteristics; b) Grow channel design features.

4.4. Calculation of grow bed design parameters

Using the attributes specified for instances of sub-concepts- crop type and NFT grow channel given in section 4.5 - equations are developed to calculate the design parameters of grow bed. For instance, PS and L are used to determine the number of plant sites per channel (NPSC). NPSC is defined as the capacity of each channel to grow a number of plants. In Figure 4-4b, NPSC is 8, which implies that in this particular channel, only 8 plants can be grown. The S on the grow channel is directly related to PS and is essentially important to ensure high crop yields. Other yield parameters that are impacted by PS in the aquaponic system are plant height, leaf area, and leaf number. The general rule of thumb in this essence is to build plant sites on each channel and keep the spacing of channels according to the expected width of the plant at its maturity stage [177]. NPSC, along with the total number of channels (NC) needed to build the complete hydroponic unit, determines the production capacity (PC) of the aquaponic system, which is equivalent to the maximum possible crop yield. Equations (1) and (2) are developed for computing NPSC and PC, respectively.

$$NPSC = L/S \quad (1)$$

$$PC = NPSC \times NC \quad (2)$$

The grow channels can be stacked horizontally or vertically or both by maintaining the recommended HCS and VCS. Moreover, the farm space must also be taken into consideration while choosing NC and respective stacking setups. With horizontally stacked NFT channels, the length of the fully developed hydroponic unit is the same as the length of the grow channel, L, whereas the width of (WHU) is equivalent to the sum of widths of all channels and horizontal spacings between channels. Equation (3) is formulated to determine WHU.

$$\mathrm{WHU} = (\mathrm{NC} \times \mathrm{W}) + ((\mathrm{NC} - 1) \times \mathrm{HCS}) \qquad (3)$$

Another significant agronomic factor that enhances crop yield is plant density or plant population (PD). PD measures the number of plants per unit area, and its optimum value varies with the genotype and geographic location [326]. In aquaponic systems, the number of plants to be grown refers to the production capacity of the system, whereas the unit area is related to the area of the hydroponic component. To compute PD, Equation 4 is devised.

$$\mathrm{PD} = \mathrm{PC} \,/\, (\mathrm{L} \times \mathrm{WHU}) \qquad (4)$$

These equations use the existing and inferred knowledge from AquaONT to determine mentioned design features and to visualize this, GUI is developed, which is explained in the next section.

4.5. AquaONT application: Graphical User Interface

To visualize the behavior of AquaONT, a GUI is developed using the MATLAB app designer tool which is shown in Figure 4-5. This GUI uses inferred knowledge from AquaONT, and equations developed in section 4.4. It allows users to make a crop and a channel length selection and observe the impact on design parameters in terms of numerical value. For better visualization of design variations in the grow channel as a 3D CAD model, these numeric values are sent to SOLIDWORKS, where they are applied to the already-built design, referred to as default parametric design.

Five fields are created on the GUI to represent the knowledge of the ontology model: 1) Crop Field, 2) Grow Bed Design Field, 3) Environmental Parameters, 4) NFT Channel Selection, and 5) NFT-based Crop Production System. The first four fields are populated with existing and inferred knowledge from AquaONT - acquired directly through the SQL database, whereas the last field is linked with the set of equations created in section 4.4. The Crop Field describes the five leafy green crops along with their characteristics, such as H, Wi, and PS [12]. The Grow Bed Design Field gives information about the grow bed type, PS, HCS, and VCS of each crop. The Environmental Parameters field specifies the optimal growth conditions for these crops. The entries of fields 2 and 3 are auto-populated once the crop is selected. For the selection of the right NFT channel, the NFT Channel Selection field is incorporated, where the length of the channel is the deciding factor. The channel lengths considered are 6 feet, 8 feet, 10 feet, and 12 feet.

The other parameters under this field, such as the width and depth of the channel, are kept constant for the sake of simplifying the model. Moreover, the shape of the plant site is chosen to be circular with a diameter of 2 inches. The plant site can also be squared in shape. The last field on the GUI is the NFT-based Crop Production System. This field uses entries of previous fields and governing equations given in section 4.4 in order to calculate parameters. This field is important as it gives information about the production capacity of the system along with the length and width of the complete hydroponic unit once the user selects the number of channels. In addition, three auxiliary fields are created on the lower side of the GUI window, which displays the total area of the hydroponic unit, total growing area, and plant density (plant population).

4.6. Use Case - Grow bed design for basil crop

The use case presented here aims to illustrate the feasibility of AquaONT and GUI. For this purpose, the basil crop is considered, which is one of the most common economically viable products in aquaponic systems. The optimal environmental conditions to grow basil in indoor farms, standard height and width under these conditions, and HCS, VCS, and PS are shown in Figure 4-5. These values are extracted from AquaONT. Assuming the user selects 6 feet long NFT channel for its aquaponic system and his/her farm can accommodate a maximum of four channels. After entering these values in the relevant fields in GUI, the design parameters under the fifth field are automatically calculated.

For the given inputs, such as L = 6ft and NC = 4, the results show that only 7 basil plants per channel can be grown, and these plants must be placed 10 inches apart on each channel.

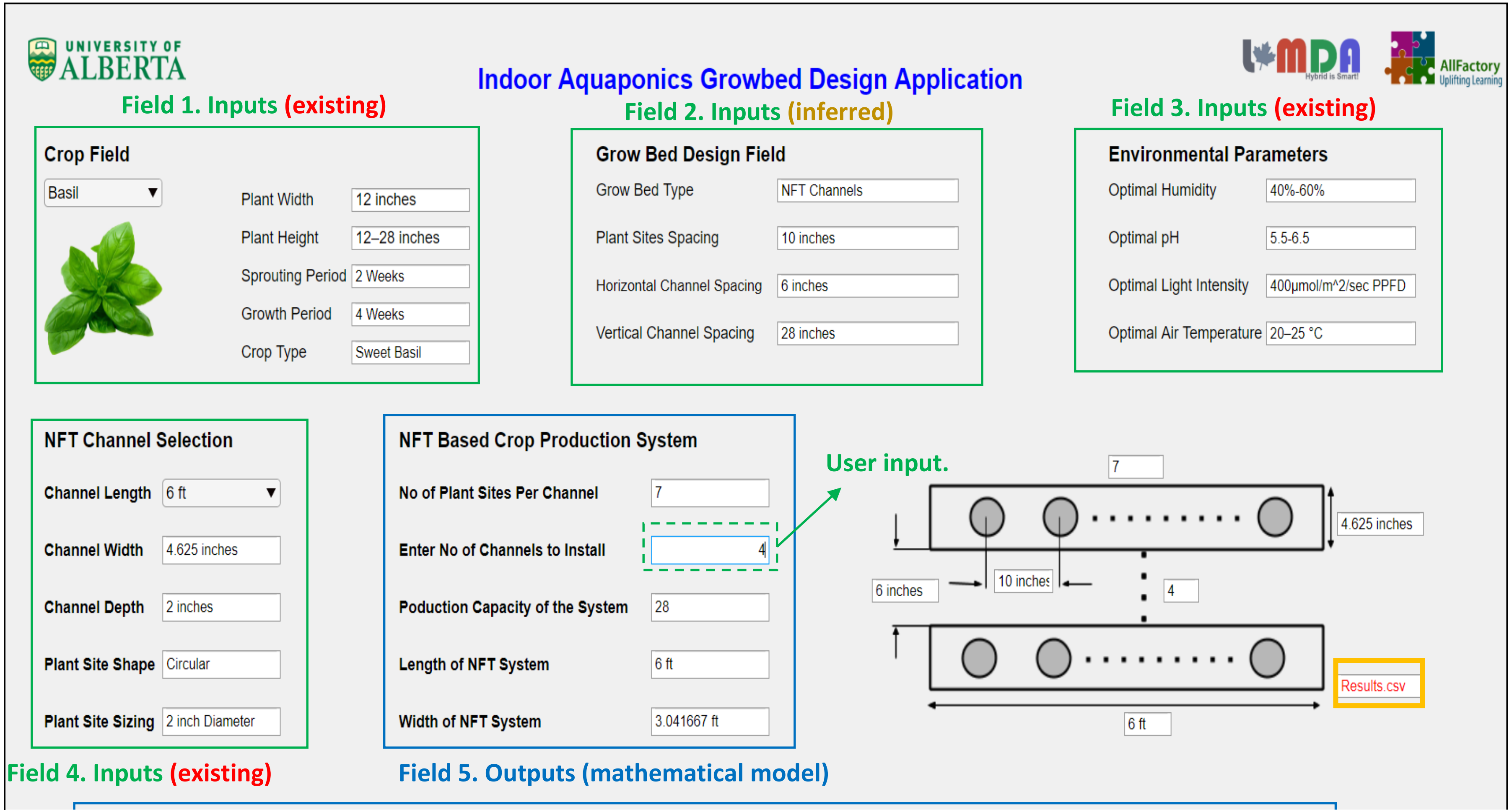


Figure 4- 5. Graphical User Interface for AquaONT design application developed by LIMDA, University of Alberta.

In addition, each channel must be placed at a distance of 6 inches from the other. The application also calculates the total area, the effective growing area, and the PD of the hydroponic unit, which in the case of basil are: 18.25 ft2, 0.61ft2, and 2/ft2, respectively, see Figure 4-5.

Finally, to visualize the CAD model of the NFT grow system for basil, the calculated design parameters from MATLAB are imported into SOLIDWORKS. These parameters are saved in a design table which enables parametric modeling. The idea is to develop a default design of a grow system in CAD software and automatically update it with a single click without designing the entire part or assembly again by using the new design details stored in the design table. This process is showcased by presenting the basic case of the basil crop. The default and updated grow channel design for basil is shown in Figure 4-6. Before implementing the parameters saved in the design table, L = 96in with NPSC = 8 for default design but after application, L becomes 72in with NPSC reduced to 7 – showing the updated design configuration for basil. The process is repeated for basil, lettuce, and parsley for different input values. The results obtained are explained in the next section.

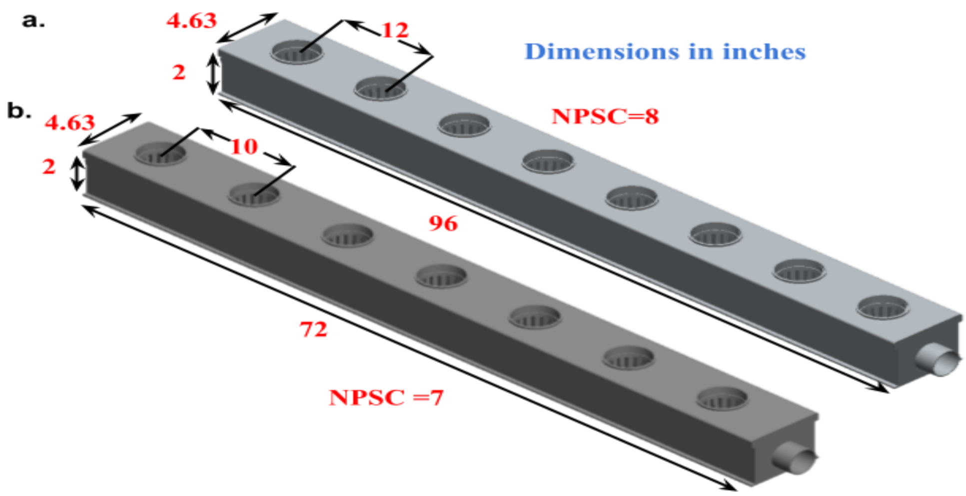


Figure 4- 6. a) Default grow bed design. b) Updated grow bed design for basil.

4.7. Results and discussion

The proposed system is simulated for all the crops mentioned in section 4.3.2. Figure 4-7 shows the design configurations of the hydroponic unit for three crops with two different input sets – including {L, NC} = {72in, 4} and {96in, 6}. The results show that for the same channel length, NPSC is different for each crop. This is due to the distinct requirement of plant site spacing (S) for each crop such as {Sbasil, Slettuce, Sparsley} = {10,8,12}. Similarly, the production capacity of the hydroponic unit is also different for each crop. For the same NC, it is observed that the PC of the system for lettuce is 22.22% and 33.33% higher than for basil and parsley, respectively. If L is increased from 72in to 96in and NC is increased from 4 to 6, the resulting NPSC and PC will also be increased. For instance, in Figure 4-7(e, f) NPSC and PC for parsley are increased from 6 and 24 to 8 and 48, respectively. With these visualization results in place, crop characteristics such as PS, Wi, and H significantly impact the design parameters of grow channel in an aquaponic system. Having a correct grow bed design in an aquaponic system for crop growth is crucial because it ensures high yields. Moreover, it also ensures the right amount of water and nutrient absorption, which eventually leads to high crop quality with the right nutritional value. In this essence, a quick knowledge-based virtual tool assists in decision-making related to the proper design of grow bed based on crop characteristics.

For future work, intelligent techniques such as machine learning, deep learning, and computer vision will be incorporated to make the system smart and autonomous. Moreover, a cost model will also be integrated to optimize the aquaponic grow beds based on market demand.

4.8. Conclusions

Aiming at providing a knowledge-based system for automated decision-making regarding crop production and respective grow bed design in aquaponics farms, this work has proposed a decision support framework. An ontology model, AquaONT, is developed to assist in decision-making process, which can be extended to include other elements and tested against robust case studies. GUI is developed that uses inferred and existing knowledge from AquaONT and mathematical equations to calculate design parameters. To visualize the impact of crop selection on the design of grow beds, parametric modeling is performed. The analysis of results shows that the correct design of grow bed ensures high crop yield and quality.

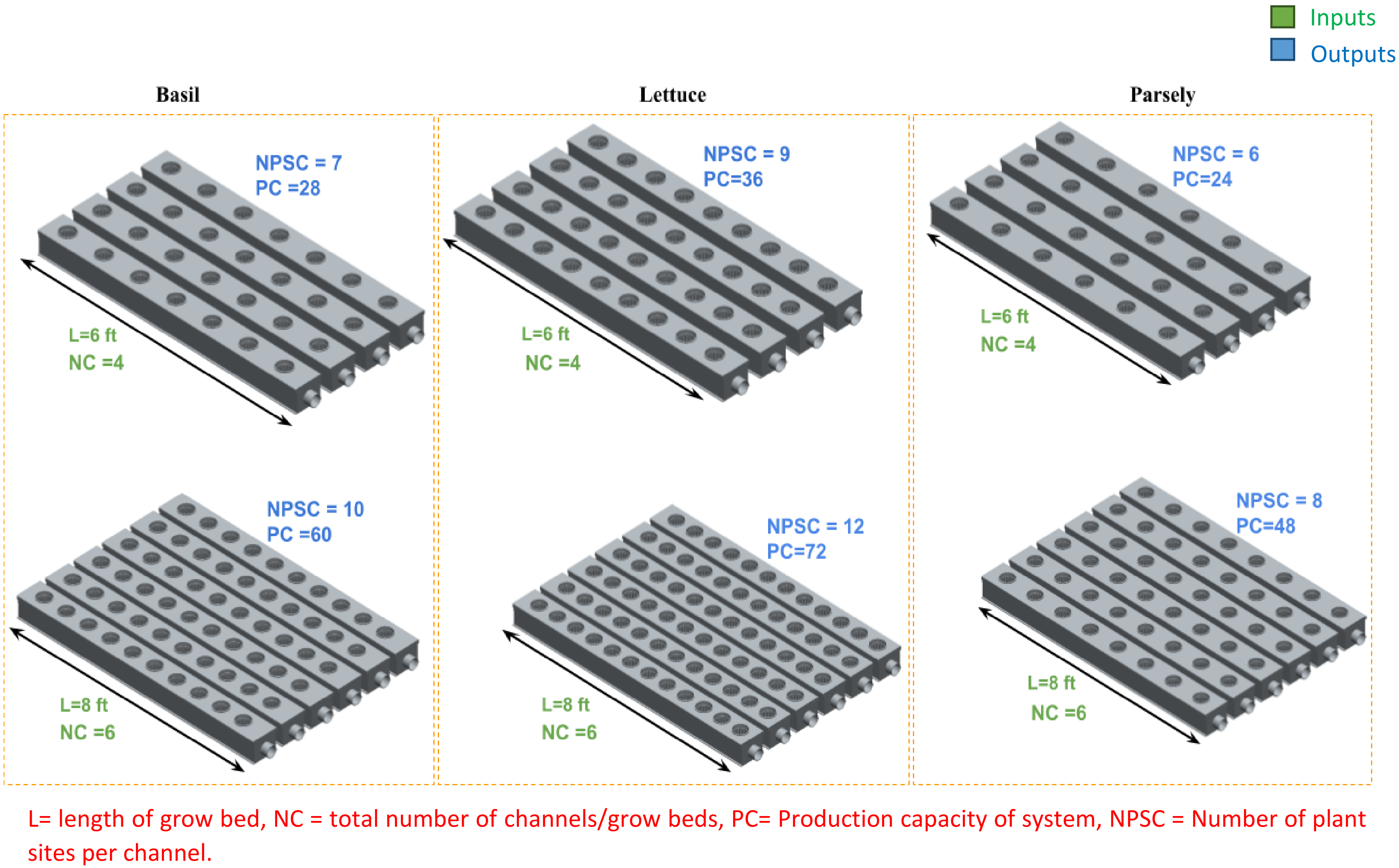


Figure 4- 7. NFT grow bed design configurations for different crops: (a, d) Basil; (b, e) Lettuce; (c, f) Parsley.

Chapter 5 Data acquisition and monitoring dashboard for IoT-enabled aquaponic facility

5.1. Introduction

Aquaponics offers several benefits as mentioned in previous chapters and hence, can be scaled to commercial levels to meet ongoing food demands. The presence of diverse parameters in an aquaponic system such as dissolved oxygen (DO), electroconductivity (EC), pH, water temperature, light intensity, humidity, and air temperature impact the growth rate, yield, and quality of the crops and fish. For instance, a low pH level decreases the nitrification process causing stressful conditions for fish that often results in fish diseases or death, whereas, high pH levels stop the nitrification process which can affect plant growth [5]. Similarly, if the water temperature goes out of optimal range, the productivity of the bacteria will tend to decrease and the nitrification process will not be successful. Another important parameter is DO, which is the amount of oxygen in the water and determines the ability to support aquatic life [177]. If the Oxygen level is low, the bacteria will stop breaking down the ammonia and nitrite, increasing potential health risks for fish and plants. The fish population is also affected by the changes in the EC which is related to how fresh the water is. Low levels of EC indicate an unbalanced system and high levels indicate that water is polluted, and it may cause the death of the fish population [177]. Air temperature and light intensity also affect plant growth. At higher temperatures and inadequate lighting, leafy greens can bolt, flower, seed and become bitter [5]. The mentioned challenges require keeping these parameters within their optimal ranges, which poses a need to monitor the system 24 hours per day and seven days a week, which is challenging if done manually.

The emergence of industry 4.0 technologies such as the internet of things (IoT), cloud computing, wireless sensor networks and artificial intelligence (AI), etc., has allowed intensive automation, monitoring, and control in the aquaponic system - giving rise to the concept of the smart aquaponic system [177]. With smart aquaponics design, monitoring and controlling essential parameters can be regulated remotely through IoT and cloud computing technologies. In contrast, processing and analysis of parameters can be achieved using machine learning and deep learning models.

Extensive research is carried out to develop monitoring systems and provide clear visualization of relevant aquaponics' parameters. Still, only a few studies have addressed the modeling of each parameter and future autonomous and smart applications. Naser et al. presented a smart aquaponic system capable of controlling and monitoring the essential aquaponics' parameters, such as degree of acidity, water level, water temperature, and fish feed [59]. An internet-based mobile application was also developed to visualize the collected data. Haryanto et al. designed an IoT-based aquaponics monitoring system in a real-time setting that uses NodeMcu to collect data (dissolved oxygen, pH, temperature, etc.) from sensors and send it to a web server [327]. Pasha et al. developed an IoT-based monitoring and control system for aquaponics that measures and displays three parameters: pH, temperature, and water level [328]. A web interface using web socket protocol is also created to transfer information, provide secure server connections, and keep the system running in a real-time. Odema et al. developed an IoT-based system that allows remote monitoring and control of the aquaponics' parameters such as DO, temperature, pH, etc. A Modbus TCP standard protocol is used to pull data from the sensing nodes of a supervisory computer [60]. Nagayo et al. created a GSM and Arduino-based monitoring and control system that sends alerts to the users when parameters' values reach dangerous levels, such as temperature, relative humidity, light, pH, water level, DO, EC, total dissolved solids (TDS), and salinity [329]. A Graphical User Interface (GUI) is also designed to display the information which can be extracted using NI LabView. Wang et al. designed a data acquisition sensor module consisting of different sensors to provide real-time data on temperature, humidity, light, water level, and dissolved oxygen in an aquaponic system [330]. Arduino and WRTnod transmit data wirelessly to the control and management platform, which stores the data, processes it, and sends it to the server for further analysis and data-driven decision-making to control the aquaponics' system. Vernandhes et al. constructed a real-time monitoring and control system that uses an Arduino connected to a web server through an ethernet shield [61]. A GUI is also developed that enables users to remotely switch on or off different devices.

5.1.1. Research motivation and contribution

The studies mentioned above have contributed to the enhancement of an aquaponic system. But none have focused on constructing smart decision-support models capable of predicting and correlating parameters. The key to creating such models is the availability and robustness of well-structured and well-defined data platforms that showcase the accurate representation of the system in a real-time. The current work aims to develop a cloud-based data acquisition and a monitoring dashboard. The data about six different parameters related to water quality and environmental conditions is gathered from a wireless sensing module (WSM) that was designed and implemented for an aquaponic system in previous work [331]. The data is then uploaded wirelessly to a dashboard, providing real-time insights into the monitored aquaponic system. The successful deployment of this work will stimulate the building of data-driven models that will autoregulate the parameters within optimal ranges and promote the development of decision support systems for accurate system design, focusing on maximizing yield and quality control.

5.1.2. Chapter organization

The remainder of the chapter is structured as follows: Section 2 will present the architectural design employed to develop a data acquisition and monitoring dashboard, Section 3 will discuss the results and discussion, and finally, Section 4 will discuss conclusions and prospects.

5.2. Data and monitoring dashboard development

To develop a data acquisition and monitoring system for an aquaponic facility, this study uses the three-tier IoT architecture [332]. The three layers are perception, network, and application, as shown in Figure 5-1. In general, the IoT architecture refers to a framework that defines the physical components, the functional organization and configuration of the network, operational procedures, and the data formats to be used [332].

In an aquaponic system, the perception layer, also known as a physical layer, is used to sense and collect relevant information related to changes in the environmental conditions, water quality, and crop and fish growth status and automatically control the parameters to maintain an equilibrium of the system. In this study, the perception layer is comprised of a WSM and a camera module, which is installed on the hydroponic unit, a nutrient film technique (NFT) based system located at the Aquaponics 4.0 Learning Factory (AllFactory, University of Alberta, see Figure 5-2). The primary components of WSM and the camera module are listed below:

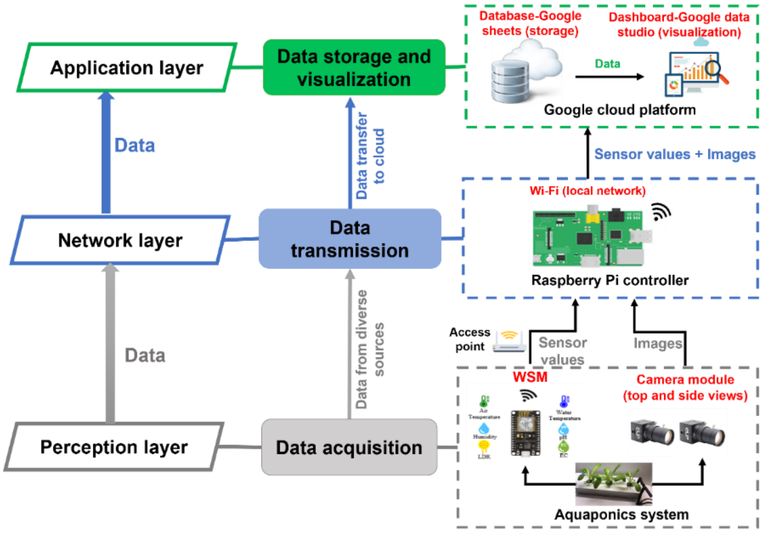


Figure 5- 1. Architectural design of the data acquisition and monitoring system.

Wireless sensing module:

- 1× PH-4502C liquid pH value detection sensor
- 1× Gravity analog electrical conductivity sensor
- 1× DS18B20 water temperature sensor
- 1× DHT22 air temperature and humidity sensor
- 1× TEMT6000 LDR sensor
- 1× ESP8266 wireless sensor
- 1× 2-channel relay module
- 1× 5V power supply
- 1× Arduino UNO USB microcontroller

Camera Module:

- 2× ELP 1080P webcam (2.8–12 mm HD Varifocal Lens)

The complete development and working of WSM are detailed in a previous work by the authors [331]. For this system, the leafy vegetable Little Gem Romaine Lettuce is chosen. A Raspberry Pi 4 (Model B Rev 1) controller is programmed to receive the data from the WSM (wirelessly) and camera module (through USB) simultaneously after 30 min from 6:00 h to 18:00 h. In total, around 1000 data points and 1000 images (500 top views and 500 side views) with three plants each are collected over the period of one month (December 1st – December 31st, 2021). One set of growth lights (T5 high output bulb, full daylight spectrum, 24 W) is also installed, which are programmed to work at on/off intervals of 12 h. The temperature and humidity levels of the lab are maintained to meet the optimal ranges of essential aquaponics parameters.

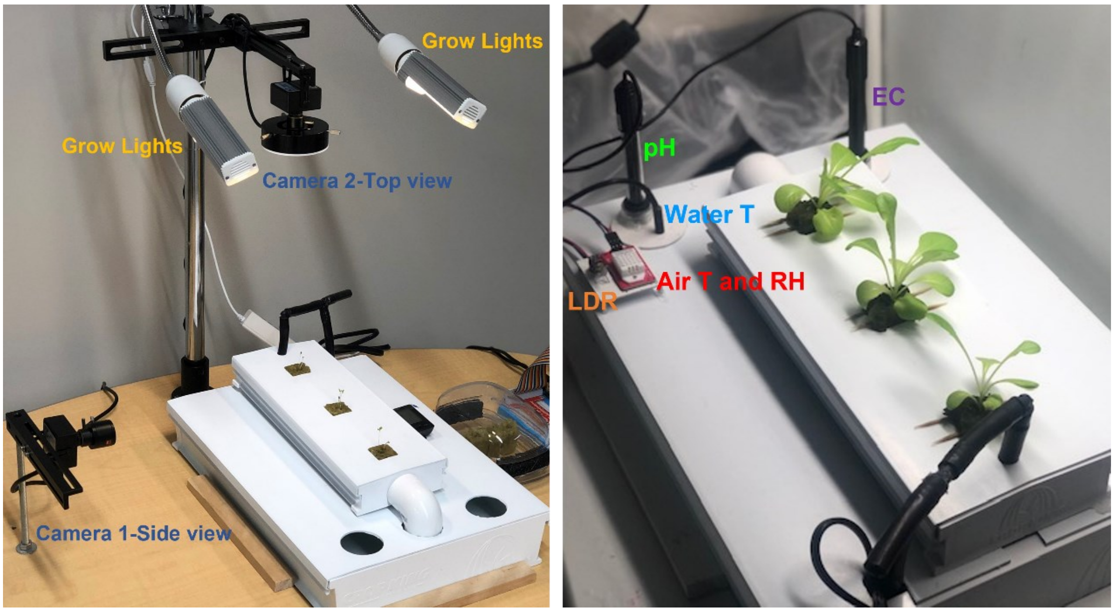


Figure 5- 2. Physical system with cameras, sensors, and grow lights.

The network layer realizes the data transmission, routing, and control [1]. The communication technology (equipment and programs used for data transmission) and communication protocol (communication rules and unified formats) are the primary elements of this layer. In this study, Wi-Fi technology is used for local network connection via the device's local Internet Protocol (IP) address to transmit data from the perception layer to the last layer.

The application layer provides a platform built on the cloud for data storage, visualization, and analysis [241]. Task-specific applications and dashboards can be developed in this layer to monitor and control the system, make predictions on data, and generate data-driven decisions. This study uses the Google cloud platform to develop a monitoring dashboard. In particular, Google sheets are used to store data, and Google data studio is used to build a dashboard for data visualization and analysis in a real-time setting. The complete layout of the dashboard is shown in Figure 5-3 as well as in Figure B.1 to Figure B.5 in Appendix B. It consists of four tabs: data acquired, data visualization, parameter monitoring, and crop monitoring. The `data

acquired` tab imports data related to six parameters (air temperature (air T), relative humidity (RH), light intensity, pH, water temperature (water T), and electroconductivity (EC)) from Google sheet and displays concerning the date and time it is taken. The `data visualization` tab displays the time series plots of all six parameters. The `parameter monitoring` shows the parameter's measurements in the form of a gauge meter, along with top and side images of the lettuce crop.
Lastly, the `crop monitoring` tab displays the links (respective Google drive folder) for top and side images of the lettuce crop. In the future, the dashboard will be updated to include parameters related to the aquaculture unit, and a new tab, `fish monitoring`, will also be added. Currently, the data from the aquaponic system is sent to the cloud every 30 minutes; hence, the refresh rate of the dashboard is set at 30 minutes, which means it is automatically updated with new data every 30 minutes.

5.3. Results and discussion

The process starts with raspberry pi sending the commands to WSM and the camera module. After receiving commands, both modules retrieve the required data and send it back to a raspberry pi. The data is comprised of sensor values and top and side view images of the lettuce crop. Upon receiving the data, raspberry pi sends it directly to the Google cloud platform, where the first data is stored in a Google sheet, see Figure 5-4. The data from the Google sheet is then imported into the dashboard developed in Google data studio. Figure 5-5 shows an example of time series plots of six parameters displayed on the data visualization tab of the dashboard. To balance the aquaponic system for healthy growth of crops, the optimal ranges of the six parameters are: air T = 18℃ – 30℃, RH = 60% – 80%, light intensity = 400 lux – 500 lux, water pH = 6.5 – 7.0, water T = 17℃ – 30℃, and water EC = 100 – 3000 µSiemens/cm. Figure 5-5 shows the variations of these parameters over time. Measurement of ambient light sensor shows values greater than 400lux for 12-hour intervals between 6:00 h to 18:00 h and lower than 50 lux at all other times. Tab parameter monitoring, shown in Figure 5-3 and Figure B.4 of Appendix B, also provides the path (URL link) where images of crops are stored, which are captured at the same instance as sensor values. These images can be accessed by clicking the buttons, namely, Crop-Top-View and Crop-Side-View, available at the bottom of the tab. The dashboard can be accessed by inserting the provided credentials (username and password). The monitoring dashboard ran smoothly throughout the experimental phase without experiencing data loss and stability issues. The proposed IoT-based monitoring dashboard provides a way forward to integrating smart technologies and prediction tools in an aquaponic system.

Apart from basic implementations such as remote monitoring and data visualization, it can be used as a robust database to work with deep learning and machine algorithms for future smart applications. Moreover, the well-structured and well-defined data acquisition and monitoring system is also vital to constructing accurate knowledge-based decision support systems. Based on this monitoring dashboard, the authors have developed a knowledge-based application for the automated design of aquaponics' grow beds which is presented in Chapter 4. The proposed dashboard can further be extended to integrate heterogeneous data from multiple modules installed at different locations in a commercial-level setting.

5.4. Conclusions

The dashboard is developed using the Google cloud platform to import and display data from six sensors and two cameras installed at the hydroponics unit of the aquaponic facility in a real-time. The monitoring dashboard is available on the cloud, which extends the remote monitoring capability of the facility while eliminating the need to perform on-site parameter monitoring manually.

Currently, the dashboard only considers parameters' values and images from the hydroponics unit. No sensors are installed at the aquaculture unit because of the limitations of internal regulations about animal experimentation. In future work, WSM will be upgraded to include aquaculture sensors such as dissolved oxygen, ammonia, nitrites and nitrates, salinity, and dissolved solids, among others. Likewise, the dashboard will be upgraded accordingly.

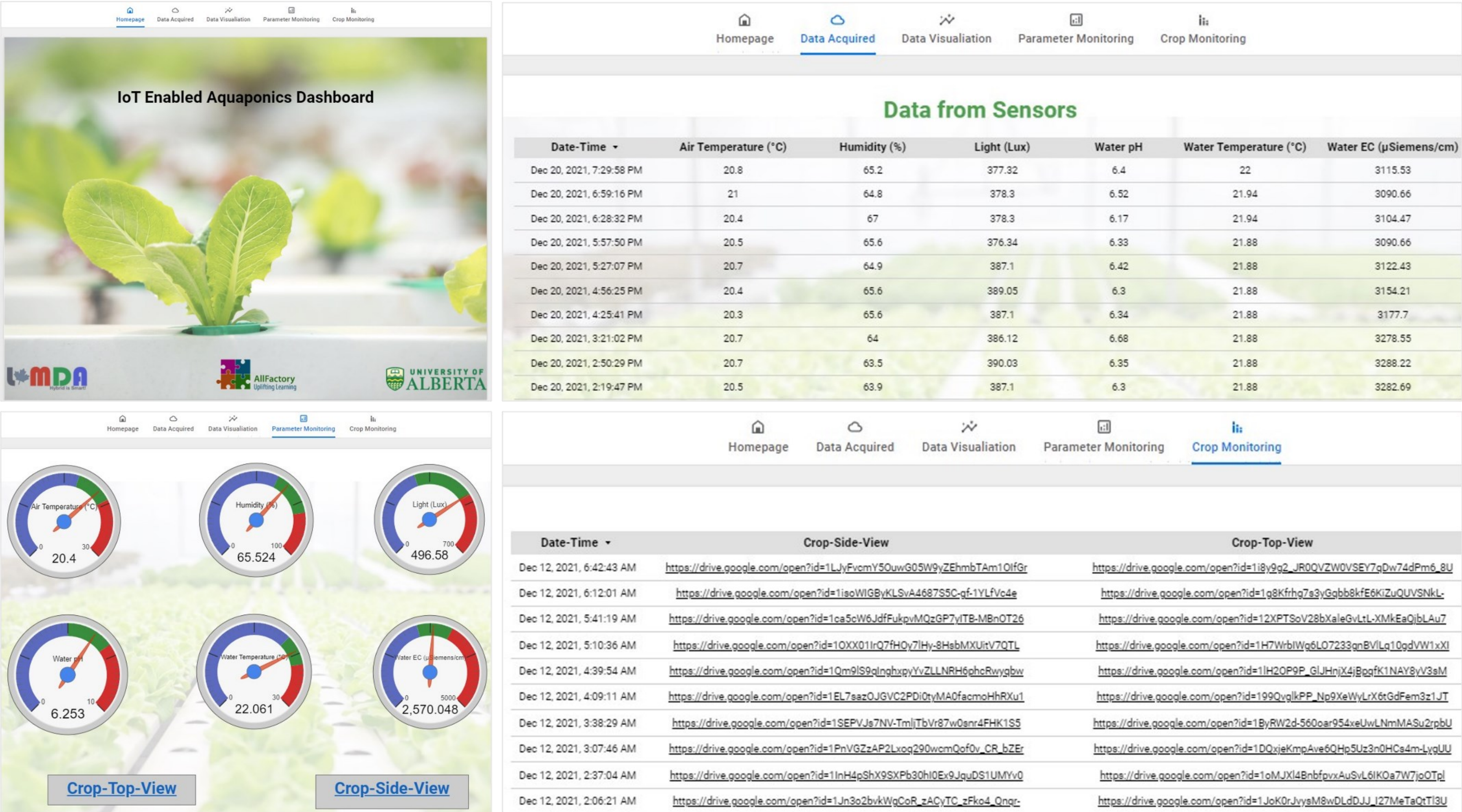
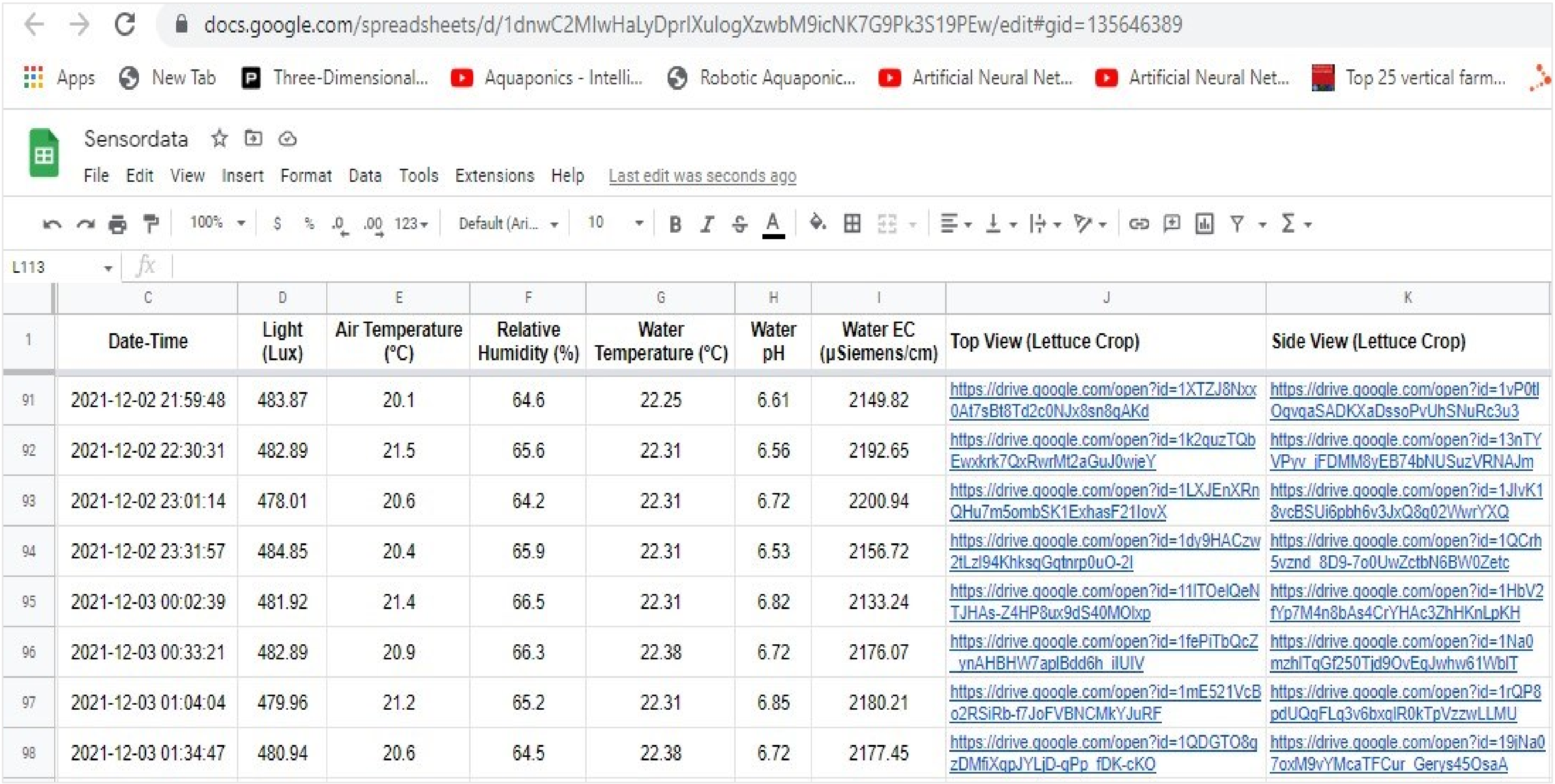


Figure 5- 3. Layout of dashboard for IoT-enabled aquaponic system.



	Date-Time	Light (Lux)	Air Temperature (°C)	Relative Humidity (%)	Water Temperature (°C)	Water pH	Water EC (μSiemens/cm)	Top View (Lettuce Crop)	Side View (Lettuce Crop)
91	2021-12-02 21:59:48	483.87	20.1	64.6	22.25	6.61	2149.82	https://drive.google.com/open?id=1XTZJ8Nxx0At7sBt8Td2c0NJx8sn8qAKd	https://drive.google.com/open?id=1vP0tlOqvqaSADKXaDssoPvUhSNuRc3u3
92	2021-12-02 22:30:31	482.89	21.5	65.6	22.31	6.56	2192.65	https://drive.google.com/open?id=1k2quzTQbEwxkrk7QxRwrMt2aGuJ0wjeY	https://drive.google.com/open?id=13nTYVPyv_jFDMM8yEB74bNUSuzVRNAJm
93	2021-12-02 23:01:14	478.01	20.6	64.2	22.31	6.72	2200.94	https://drive.google.com/open?id=1LXJEnXRnQHu7m5ombSK1ExhasF21IovX	https://drive.google.com/open?id=1JlvK18vcBSUi6pbh6v3JxQ8q02WwrYXQ
94	2021-12-02 23:31:57	484.85	20.4	65.9	22.31	6.53	2156.72	https://drive.google.com/open?id=1dy9HACzw2tLzI94KhksqGqtnrp0uO-2I	https://drive.google.com/open?id=1QCrh5vznd_8D9-7o0UwZctbN6BW0Zetc
95	2021-12-03 00:02:39	481.92	21.4	66.5	22.31	6.82	2133.24	https://drive.google.com/open?id=11ITOeIQeNTJHAs-Z4HP8ux9dS40MOlxp	https://drive.google.com/open?id=1HbV2fYp7M4n8bAs4CrYHAc3ZhHKnLpKH
96	2021-12-03 00:33:21	482.89	20.9	66.3	22.38	6.72	2176.07	https://drive.google.com/open?id=1fePiTbQcZ_ynAHBHW7aplBdd6h_iIUlV	https://drive.google.com/open?id=1Na0mzhlTqGf250Tjd9OvEqJwhw61WbIT
97	2021-12-03 01:04:04	479.96	21.2	65.2	22.31	6.85	2180.21	https://drive.google.com/open?id=1mE521VcBo2RSiRb-f7JoFVBNCMkYJuRF	https://drive.google.com/open?id=1rQP8pdUQqFLq3v6bxqIR0kTpVzzwLLMU
98	2021-12-03 01:34:47	480.94	20.6	64.5	22.38	6.72	2177.45	https://drive.google.com/open?id=1QDGTO8qzDMfiXqpJYLjD-qPp_fDK-cKO	https://drive.google.com/open?id=19jNa07oxM9vYMcaTFCur_Gerys45OsaA

Figure 5- 4. Example of data stored in Google sheet.

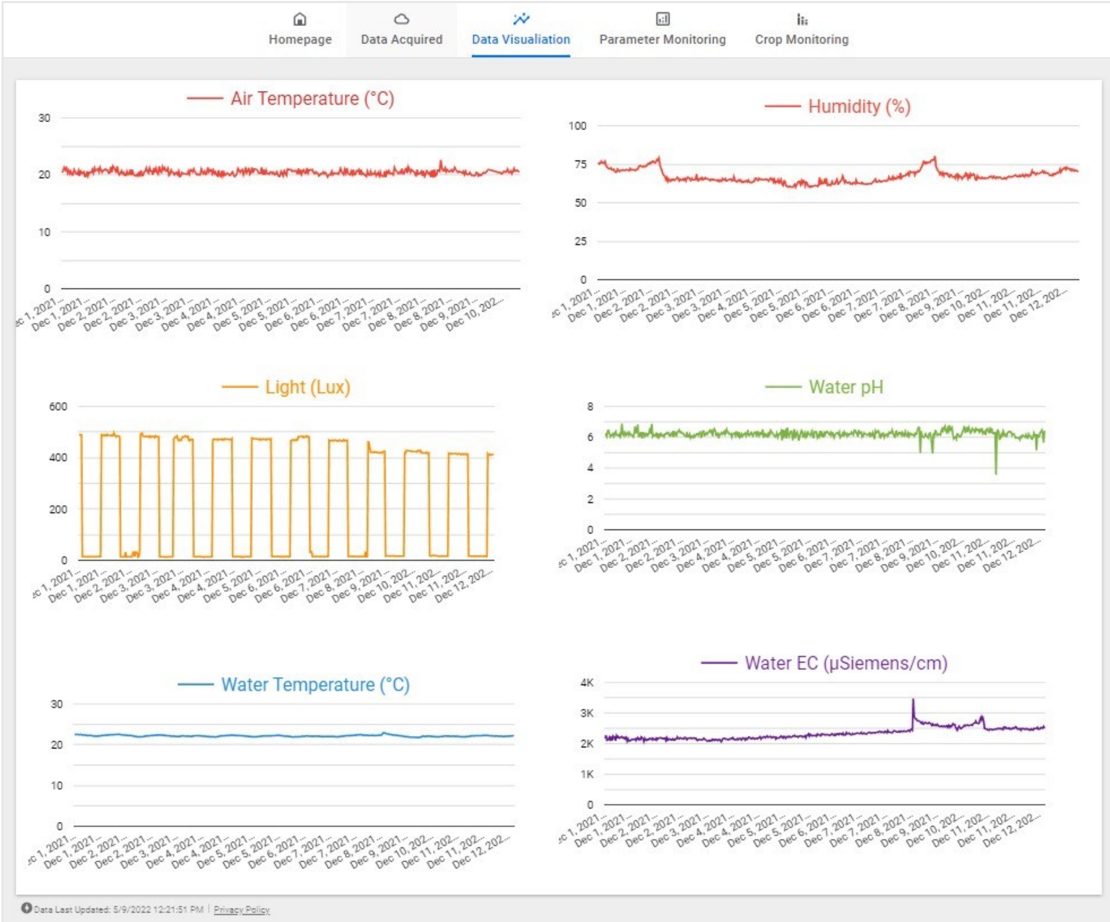


Figure 5- 5. Time series plots of six parameters in dashboard.

Chapter 6 Crop Diagnostic System: A robust disease detection and management system for leafy green crops grown in an aquaponic facility

6.1. Introduction

6.1.1. Research motivation and contribution

Despite all the advantages offered by aquaponics technology, a few challenges mentioned in Chapter 1 need special attention, particularly considering its large-scale implementation [333]. One of these challenges is related to crop diseases resulting from an either nutrient deficiency or inadequate management of the system, impacting crop quality and causing crop wastage [19], [334]. As Khirade and Patil pointed out, identifying crop diseases and applying disease management practices are key to preventing losses in the yield and quantity of agricultural products [335]. For this reason, early detection of disease outbreaks is crucial for the progress of aquaponics farms. Traditionally, crop diagnostic is performed by agricultural specialists who visually examine the plant leaves. This practice, however, is subjective, destructive, time-consuming, and labor-intensive [336]. Moreover, it also requires the experts to be proficient with extensive knowledge of various diseases, their symptoms, and treatments [337]. Other methods include chemical analyses, leaf color chart (LCC) matching, soil plant analysis development (SPAD), hyperspectral imaging, and spectral remote sensing, which again are either time-consuming or costly or destructive techniques [338]. To address these problems, different automatic crop disease detection systems based on artificial intelligence (AI) techniques such as machine learning and deep learning are developed as they offer contactless, rapid, environmental-friendly, and accurate methods for performing a non-invasive evaluation of crops' health and quality [339], [340]. Deep learning techniques offer two significant advantages over machine learning techniques. First, the feature extraction process is automatic, and second, the time to process large datasets of high dimensions is significantly reduced [339].

In addition to disease detection, it is also paramount that farm practitioners and researchers have access to relevant information about crop management strategies that allow them to pick up methods and treatments appropriately to prevent diseases, thereby gaining both economic and environmental benefits [158]. In most cases, such information is dispersed throughout multiple heterogeneous data sources — posing a need for a unified model that contains knowledge about the causes and treatments of different crop diseases. Semantic technologies such as ontologies have proven effective for data integration in multiple domains [341]. An ontology is a formal and explicit specification of a shared conceptualization [342]. The logical formalisms behind ontological models allow autonomous agents to interpret the information that is being processed [268]. Ontology can be used to construct a knowledge base containing relevant information about causes and suggested treatments of crop diseases, which can be extracted upon disease detection [341]. With this information, farm practitioners are able to get clear guidelines to effectively perform crop monitoring and disease management.

In this study, an automatic system based on deep learning techniques is presented for the detection and classification of diseases in four leafy green crops, lettuce, basil, parsley, and spinach, grown in an aquaponic facility. Taking advantage of semantic technologies, an ontology model, 'AquaONT' developed by authors in chapter 3 is used that contains knowledge about causes and treatments of different diseases. This ontology model is integrated with a disease detection system through an interface established on a cloud-based application.

6.1.2. Chapter organization

The remainder of the chapter is structured as follows: Section 2 summarizes the most recent literature related to crop disease detection systems, Section 3 presents the methodology used to design the proposed system, Section 4 discusses the experimental results and findings, and finally, Section 6 concludes the chapter and presents the future prospects.

6.2. Related work

The rapid developments in AI have made a major breakthrough in deep learning (DL) and computer vision (CV) technologies by solving complex problems like image classification, object detection, speech recognition, voice recognition, natural language processing, and medical imaging, among others [21]. In particular, convolutional neural networks (CNNs) in combination with computer vision have proved their efficiency in these fields and are widely being integrated into agriculture for automatic crop disease detection — presenting a reasonable alternative to traditional practices. In recent years, several models and applications have been developed for crop disease identification and diagnosis. This section investigates some latest works present in the literature.

Fuentes et al. combined ResNet with Faster R-CNN, R-FCN, and SSD. They proposed a method to detect the diseases and insect pests of tomato plants, achieving the effective identification of nine different types of diseases and insect pests [181]. Chen et al. proposed a method to detect rice plant diseases using the DenseNet model of deep transfer learning [343]. To identify the cucumber disease spots in greenhouses, Ma et al. developed a CNN-based system, combining a compound color feature with a region-growing algorithm [344]. A disease recognition algorithm based on VGGNet and InceptionV3 with reduced model size and improved recognition accuracy is proposed by Rahman et al. for rice plants [345]. Oppenheim et al. proposed a disease classification algorithm based on an improved VGG network for accurate and quick identification and classification of spots on potato crops [346]. A method based on an improved CNN is proposed by Fan et al. to identify nine kinds of common corn diseases from images with a complex background [347]. Khan et al. proposed an apple disease detection system that works in two stages [337]. Based on the Xception model, the first stage classifies whether the leaf is healthy or diseased, and the second stage, based on Faster-RCNN, performs disease detection. Qi et al. developed a disease recognition system based on an improved YOLOv5 (squeeze-and-excitation (SE) module is added) model to identify the tomato virus diseases in the greenhouse [348]. Nandhini et al. proposed a deep learning model that combines RNN and CNN for disease classification and early prediction in the Plantain tree [349]. Abbas et al., 2021 proposed a deep learning-based method for tomato disease detection that utilizes the Conditional Generative Adversarial Network (C-GAN) to generate synthetic images of tomato plant leaves [350]. A DenseNet121 model was then trained on synthetic and real images using transfer learning to classify the tomato leaves images into ten categories of diseases. An efficient detection model (EFDet) consisting of an efficient backbone network, a feature fusion module, and a predictor is proposed by Liu et al. for the detection of cucumber leaf diseases in complex backgrounds [351]. (Mathew and Mahesh, 2022) proposed a YOLOv5-based disease detection model to detect bacterial spot disease in the bell pepper plant from the symptoms seen on the leaves. A framework is proposed by Barosa et al. for an aquaponic system based on image processing and decision tree methodology that performs disease detection of four leaf species, eggplant, chili, citrus, and mandarin. It automatically generates a report which is sent to the owner through the mobile application if the disease is detected [158]. Musa et al. presented a CNN-based approach for detecting plant disease in smart hydroponics, providing a tool to the farmers capable of doing the task of an agricultural extension worker with even better accuracy [352]. Lisha et al. developed an application based on image processing and SVM to classify apple diseases [353]. Yudha Pratama et al. proposed a model based on Faster R-CNN with Inception V2 algorithm to recognize the diseases in hydroponic lettuce [354].

The aforementioned studies have significantly contributed to the scientific and research community. However, the analysis shows that most disease detection systems are developed for open-air farms. There are only a few systems that are developed for modern farming systems, such as aquaponics or hydroponics. Most models are developed considering multiple diseases of only one crop. Moreover, to the best of the authors' knowledge, no unified disease detection system is proposed for identifying diseases of multiple leafy green crops grown in aquaponics facilities. Disease detection in leafy green presents various challenges. For instance, sometimes, a strong resemblance exists among the foliage of different leafy green crops that might impact the performance of the detection system. Secondly, due to differences in light illumination during imaging, the visual symptoms of different diseases may appear similar. Another challenge is the availability of a dataset of leafy green crops that can be used for disease detection. Deep learning models require a huge amount of data for training, and to the best of the authors' knowledge, there is no sufficient-sized large-scale open-source dataset available that can be utilized for this research. There are few datasets such as PlantVillage, PlantDoc and CropDeep [355], [340], [356]. PlantDoc and PlantVillage are open-source datasets with no categories of leafy green crops. CropDeep dataset contains images of some of the leafy green, but it is not open-source. Lastly, none of the aforementioned models provides information related to the causes and treatments of detected diseases.

Apart from AI techniques, ontology-based systems are also developed over the years for plant disease diagnosis and treatment recommendations. Jearanaiwongkul et al.

developed an ontology-based expert system called 'RiceMan' for disease identification and control recommendation in rice crops [357]. Likewise, Rodríguez-García et al. proposed a decision support system based on an ontology model for crop pests and disease recognition [341]. It also provides information on agriculture practices and permitted pest control measures. In these systems, users are required to select crop and observed symptoms from the list for further processing, which is a time-consuming process. Whereas, in deep learning models, this information can be obtained by using crop images. Deep learning techniques can be combined with ontology models to develop efficient decision support systems for disease management in crops. The idea of combining the two techniques is relatively new in the agriculture sector, and hence, limited work is done in this regard that primarily focuses on enabling smart services (monitoring and control) in IoT-based farming systems or detection of cyber-attacks [22].

Considering the research gaps and potential opportunities, this study aims to create a dataset consisting of high-quality RGB images (healthy and diseased) of four leafy green crops: little gem romaine lettuce, spinach, parsley, and basil. This study also aims to develop a crop diagnostic system based on deep learning models and ontology models for detecting diseases and identifying causes and potential treatments in stated crops, respectively.

6.3. Research methodology

The block diagram illustrating the three sequential modules of the research methodology is shown in Figure 6-1. The first module involves the preparation of the dataset and training of classification and object detection models. The disease detection model works in three phases. The first and second phase uses lightweight classification models to classify the type of crop and identify whether the classified crop has a disease or not, respectively. Phase 3 is the detection stage which uses an object detection model to detect and localize the diseased and non-diseased spots in the crops. The third phase also tells the class of the diseased spots. The purpose behind adding two classification phases before the detection phase is three-fold. First, to improve the detection performance by reducing the number of wrong detections which could arise as the model has to identify and localize different disease spots of varying sizes. Second, to determine the characteristics of the crop identified in the first phase in relation to aquaponics' system design by linking it with the knowledge model. Lastly, to reduce the overall processing time by filtering out invalid inputs in the second phase. The second module aims to extract the instances of relevant classes such as potential causes and treatments of detected diseases from the ontology model 'AquaONT' developed by authors in chapter 3. In the third module, a cloud-based application is developed using Streamlit, where a pre-trained disease detection model and ontology model are deployed to obtain a complete crop diagnostic system. Upon identification of the crop in phase 1, its characteristics in relation to optimal environmental (pH, temperature, illumination, etc.), growth (width, height, area, etc.), and grow bed design (plant site spacing) parameters for an aquaponic facility are extracted from ontology model using OWLready2 (ontology-oriented programming package in Python). The authors have conducted a study in chapter 4 that identified design parameters as vital knowledge in ensuring high crop yields and product quality in an aquaponic facility. Likewise, once the disease and its type are detected in phase 3, the potential causes and recommended treatments are extracted from the ontology model. Each element of each module is presented in detail in the following subsections.

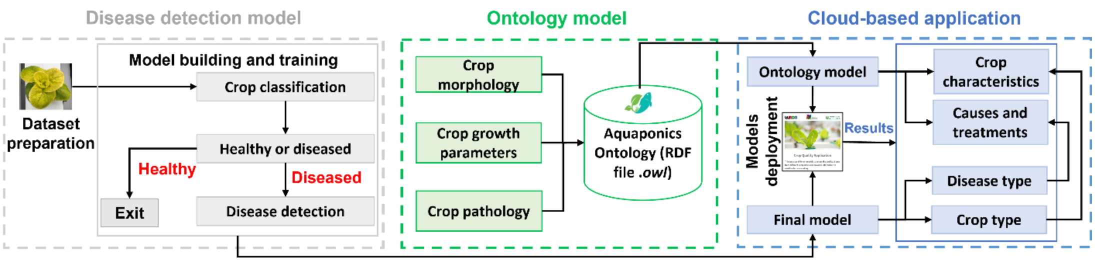


Figure 6- 1. Proposed methodology for disease detection and control recommendation system.

6.3.1. Dataset preparation

The dataset preparation involves three steps, i) data acquisition, ii) data annotation, and iii) data augmentation, which are detailed below.

6.3.1.1. Data acquisition

This study considers four leafy green crops, lettuce, basil, parsley, and spinach. The dataset consists of healthy and diseased images of these crops, which are acquired from different sources such as NFT-based aquaponic facility built in the Allfactory

4.0 Lab (University of Alberta, Canada), the Google search engine, and Ecosia. The diseases considered for the four crops while developing the dataset are given below.

- Lettuce: Bacterial leaf spot and Downy mildew
- Basil: Downy mildew
- Parsley: Septoria leaf spot
- Spinach: Downy mildew and Stemphylium leaf spot

These six diseases are considered as they are common in mentioned leafy green crops when grown in a greenhouse environment [358]. To enhance the flexibility of the model to correctly classify and detect disease, it is ensured that images have non-homogeneous backgrounds, different illumination conditions, and disease maturity stages. A total of 2000 images are gathered from all the resources. Among these images, 800 images belong to a healthy category of each crop (4×200), and 1200 images belong to diseases mentioned above (5 ×240). Figure 6-2 shows some of the sample images from the dataset.

6.3.1.2. Data annotation

Data annotation is one of the vital steps for the successful development of object detection models. The process is manual and involves labeling the desired objects in an image with a label or tag that refers to a particular class. The labeled data is used during the training of the model. There is a number of open-source annotation tools, but in this study, LabelImg[7] is used. LabelImg is a python based graphical annotation tool that supports a variety of deep learning algorithms. For instance, it generates annotations in COCO JSON format, XML files in PASCAL VOC format, and YOLO Darknet TXT format with the addition of a YAML file containing model configuration and class values [348]. In this study, the annotations are generated in COCO JSON and YOLO Darknet TXT formats because, in the disease detection phase, two object detection models are tested to design the final system.

[7] https://github.com/tzutalin/labelImg

Figure 6- 2. Samples from leafy green image dataset.

6.3.1.3.Data augmentation

Next, a data augmentation process is performed to supplement and enrich the dataset. This helps increase the model's generalizability and overcome the problem of overfitting. Moreover, it also allows the model to learn as many relevant features as possible. This study uses Albumentations, a Python library, for fast and flexible image augmentations [359]. The different augmentation techniques applied are flip, rotation, noise, blur, and brightness. Figure 6-3 shows examples of different augmentation operations. After applying the data augmentation, the final dataset comprises 2640 images with their annotations. The final distribution of the dataset is presented in Table 6-1. For all three models in three phases, 75% of the dataset is used for training, 20% for validation and 5% for testing.

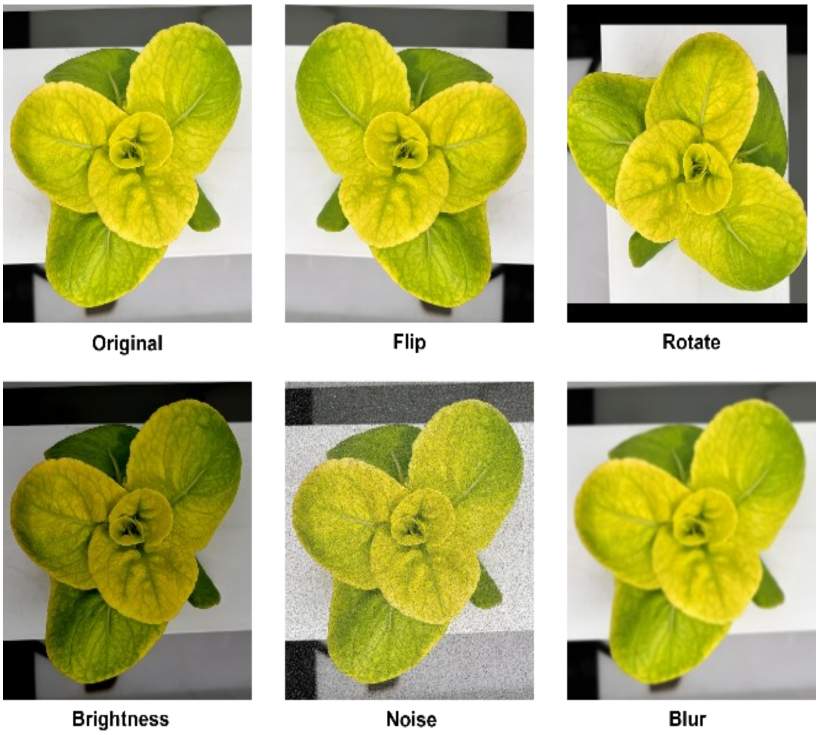


Figure 6- 3. Example of different augmentation operations applied on original image.

Table 6- 1. Distribution of dataset among four crops.

Crop	Healthy	Diseased		Total
		Disease 1	Disease 2	
Lettuce	240	280	280	800
Basil	240	280	-	520
Spinach	240	280	280	800
Parsley	240	280	-	520

6.3.2. Disease detection model development

Object detection is a complex task, and disease detection of leafy green crops comes with its own set of challenges. To overcome these challenges, the detection process in this study is divided into three primary phases, which are discussed in the next subsections. Figure 6-4 shows the full protocol followed to develop the proposed disease detection system.

6.3.2.1. Phase 1: Crop classification model

The first phase of the proposed system uses a lightweight CNN architecture to classify input images into one of the four types of crops: lettuce, basil, parsley, and spinach.

Recent years have witnessed the birth of numerous CNN architectures such as AlexNet, VGG, Inception, Xception, ResNet, Inception-ResNets, etc.; each offers several advantages and disadvantages listed in [360]. The crop classification model developed in this study is based on ResNet-50 (Residual Network). ResNet-50 is one of the variants of ResNet having 50 deep layers. It has a simple design and high accuracy and is suitable for smaller datasets [361]. ResNet-50 is similar to the typical deep CNN structure with an additional identity mapping capability shown in Figure 6-5 [361]. ResNet predicts the delta that is required to reach the final prediction from one layer to the next and hence reduces the vanishing gradient problem by allowing this alternate shortcut path for the gradient to flow through. The identity mapping used in ResNet allows the model to bypass a CNN weight layer if the current layer is not necessary, which helps in avoiding the overfitting problem of the training [362].

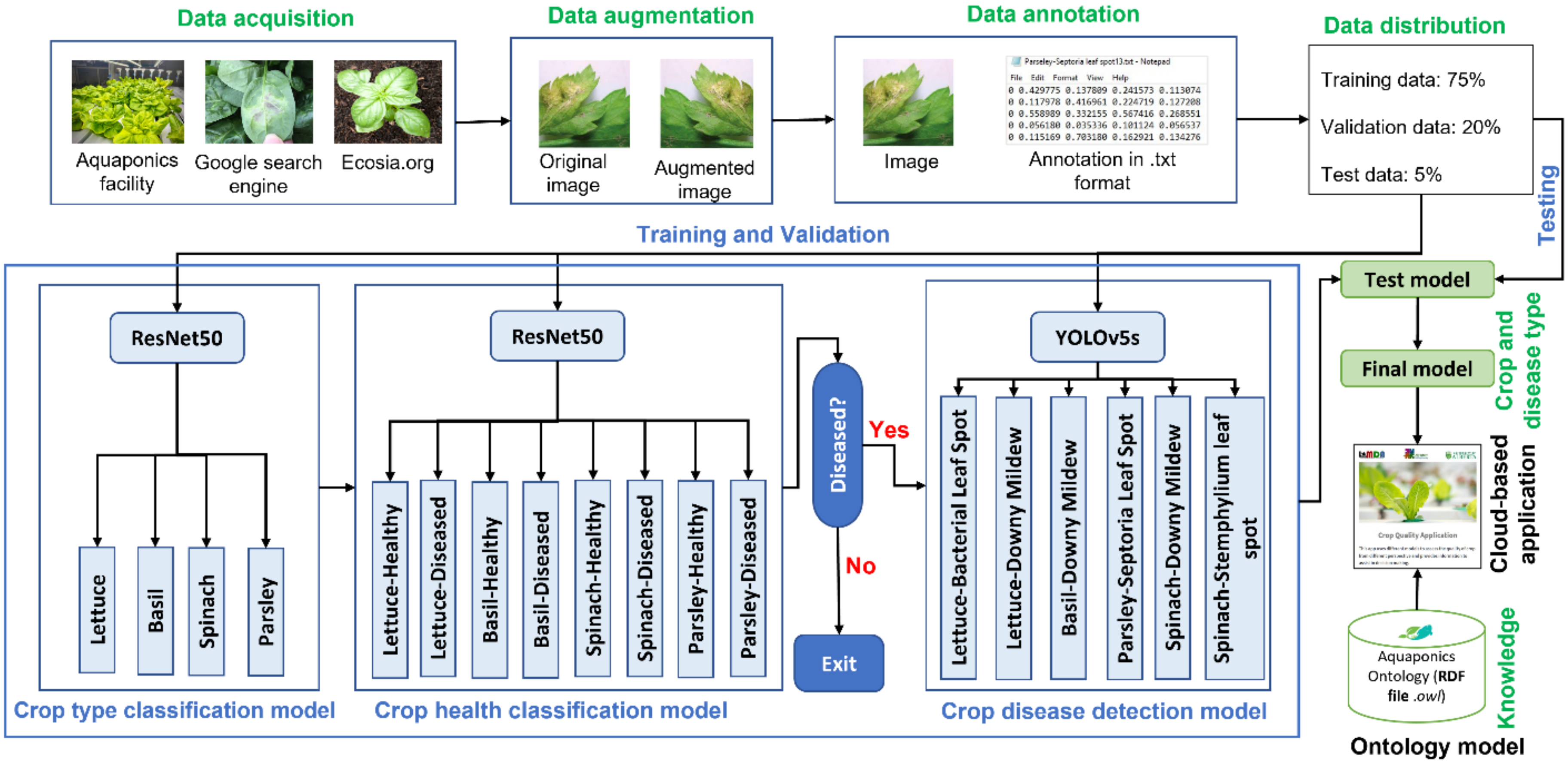


Figure 6- 4. Protocol for disease detection process.

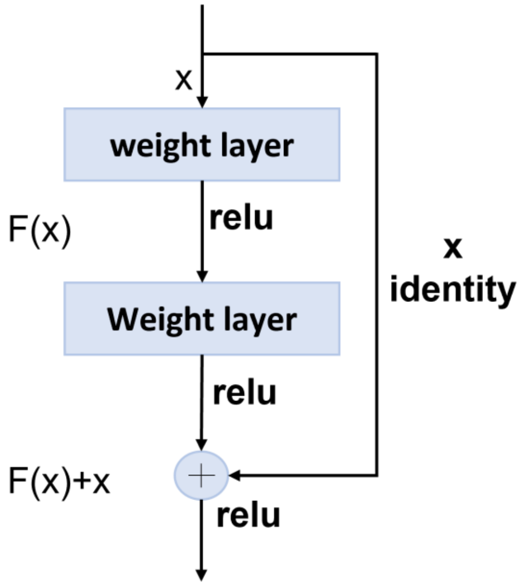


Figure 6- 5. A residual building block for ResNet-50 [361].

In this study, ResNet-50 is used as the base model, and its last layer is replaced with one Global Average Pooling layer followed by one Dense layer (fully connected layer) of size 1024 and activation function ReLu followed by Output layer for making final predictions, and it uses Softmax for the classification task. The crop type identified in this stage saves to a folder and also acts as an input to the next phase.

6.3.2.2. Phase 2: Crop health classification model

Phase 2 of the system also uses ResNet-50 and classifies the input from phase 1 into one of the following eight classes.

i) Lettuce-Healthy
ii) Lettuce-Diseased
iii) Basil-Heathy
iv) Basil-Diseased
v) Spinach-Healthy
vi) Spinach-Diseased
vii) Parsley -Healthy
viii) Parsley-Diseased

A similar architectural design of ResNet-50 is followed in phase 2 as discussed in section 6.3.2.1 except for the Output layer, which now has eight classes. If the input image is classified into one of the 'Diseased' crop categories, it goes to phase 3. On the other hand, if any of the 'Healthy' crop categories are identified, the process ends, and the classified image does not go to the next phase for further processing.

6.3.2.3. Phase 3: Crop disease detection model

The third phase of the proposed system is disease detection, which involves classifying and localizing the diseased spots in an image and classifying them into one of the disease classes mentioned below.

i. Lettuce-Bacterial leaf spot

ii. Lettuce-Downy mildew

iii. Basil-Downy mildew

iv. Parsley-Septoria leaf spot

v. Spinach-Downy mildew

vi. Spinach-Stemphylium leaf spot

This phase activates only when the input from the previous phase is one of the 'Diseased' categories. To develop a disease detection model, object detection algorithms are used. In the past recent years, advances in deep learning and computer vision have greatly accelerated the momentum of object detection [337]. Numerous object detection algorithms (object detectors) are developed and used in the disease detection of crops. These detectors are broadly classified into two categories: i) two-stage detectors based on region proposal and ii) one-stage detectors based on regression or classification [363]. The popular two-stage detectors are Fast-RCNN, Faster-RCNN, and Mask-RCNN, and one-stage detectors involve YOLO (You Only Look Once) family [351]. Khan et al. conducted a research where they ran three different models Faster-RCNN, YOLOv4, and EfficientDet, to solve a similar problem for apple crops [337]. It has been observed that Faster RCNN with mAP (mean average precision) of 42.1% outperformed YOLOv4 (mAP of 41.4%) and EfficientDet (mAP of 38%). As per these results, Faster-RCNN seems the right choice for this study. But the YOLOv5 model developed by Ultralytics has substantially improved the detection speed while maintaining the detection accuracy [364]. Therefore, both approaches are tested in this study.

Faster-RCNN

Faster R-CNN is an improved version of Fast-RCNN and is a region-based object detector [365]. The first image runs through a backbone network (CNN) which creates feature maps. On the last feature map of Convolution layers, a fully convoluted network called region proposal network (RPN) is trained, which outputs a set of bounding boxes along with their scores which determine the likelihood of an object [365].

YOLOv5

The YOLOv5 network algorithm is an improved algorithm based on YOLOv3 as it proposes a method of multi-scale prediction, which can detect the target of image features of different sizes simultaneously [348]. The network model of YOLOv5 consists of four main parts: input, backbone, neck, and prediction, the details of which can be found at [364]. YOLOv5 has five versions: YOLOv5n (nano), YOLOv5s (small), YOLOv5m (medium), YOLOv5l (large), and YOLOv5x (extra-large), the depth and width of which are gradually increasing and each one offering different detection accuracy and performance [364]. In this study, YOLOv5s is adopted as it has a smaller size and good accuracy. Moreover, it works well with smaller datasets and can be used with embedded devices [364].

6.3.3. Disease detection model training

NVIDIA GeForce RTX 3090 is used to train all the models in three phases. The classification model developed in stage 1 is implemented in PyTorch (an open-source machine learning framework based on the torch library developed by Meta AI[8]). Using transfer learning (TL) approach, ResNet-50 pre-trained on ImageNet is used [366]. The pre-trained model saves a lot of time as it is already trained on a dataset and hence contains the weights and biases of previous training that represent the features of the dataset it was trained on, which are often transferable to different datasets [350]. Hence, model parameters are initialized using the TL approach and then retrained on a custom dataset prepared in section 3.1.1 with a learning rate of 0.0001, batch size of 64, input size of 224 x 224 x 3, and epochs of 100. The model was tuned using the Adam optimizer. The hyperparameters and their values used for

[8] https://pytorch.org/hub/pytorch_vision_resnet/

the classification model in phase 1 are given in Table 6-2. For the classification model in phase 2, a batch size of 64 is used, and values of the remaining hyperparameters are kept the same, as shown in Table 6-2.

Table 6- 2. Values of hyperparameters used for classification model in phase 1.

Hyperparameters	Values
Weights	ImageNet
Learning rate	0.0001
Batch size	32
Loss function	Categorical CrossEntropy
Epochs	100

For training of the detection models, the dataset is split into 75% train, 20% validation, and 5% test sets. The first model Faster-RCNN is implemented in Detectron2 — Facebook AI Research's next-generation library written in PyTorch that provides state-of-the-art detection and segmentation algorithms. For Faster-RCNN, the annotations format is COCO JSON. The pre-trained model (trained on the COCO dataset) from the model zoo of the Detectron2[9] 'Faster-RCNN with ResNet-101 + FPN' is used, where FPN stands for Feature Pyramid Network [367]. The model is trained for 3000 iterations with an initial learning rate of 0.01 for the first 500 iterations and then 0.001 for the next 2500 iterations.

The second model, YOLOv5s, is implemented in PyTorch. Again, pre-trained version of the algorithm is used to enhance the training process and reduce time. For YOLOv5s, the annotation format is YOLO Darknet TXT but with the addition of a YAML file containing model configuration and class values. The model is trained for 3000. The hyperparameters and their values for the two models are shown in Table 6-3.

[9] https://ai.facebook.com/tools/detectron2/

Table 6- 3. Values of hypermeters used for two objection detection methods.

Hyperparameters	Methods	
	Faster-RCNN	**YOLOv5s**
Input size	600×600	416×416
Batch size	16	16
Learning rate	0.001	lr0=0.01, lrf=0.001
Momentum	0.89	0.937
Gamma value	0.1	fl_gamma = 0.0
Weight decay	0.0001	0.0005
Training time	1.5hrs	50 minutes

6.3.4. Ontology model

The complete development and details of all the concepts and instances of the ontology model 'AquaONT' developed by the authors are available in Chapter 3. AquaONT is a unified ontology model that represents and stores the essential knowledge of an aquaponic 4.0 system. It consists of six concepts: Consumer Product, Ambient Environment, Contextual Data, Production System, Product Quality, and Production Facility. In this study, two classes, 'Consumer Product' and 'Product Quality' are used for knowledge extraction. The 'Consumer Product' class provides an abstract view of the type, growth status, and growth parameters of ready-to-harvest crops in an aquaponic system. Whereas the 'Product Quality' class provides knowledge on crop attributes related to pathology (crop diseases, causes, and the ways and means by which these can be managed or controlled) and morphology (canopy dimensions such as area, length, width, etc.). Four crops: lettuce, basil, parsley, and spinach, are considered in this study. Their growth conditions and morphological and pathological attributes stored as instances of the respective classes are extracted once the crop and disease are classified. Figure 6-6 shows the hierarchical architecture of the 'Consumer Product' and 'Product Quality' classes with their instances for the 'Basil' crop in Protégé (an open-source ontology editor and framework developed at Stanford University) environment.

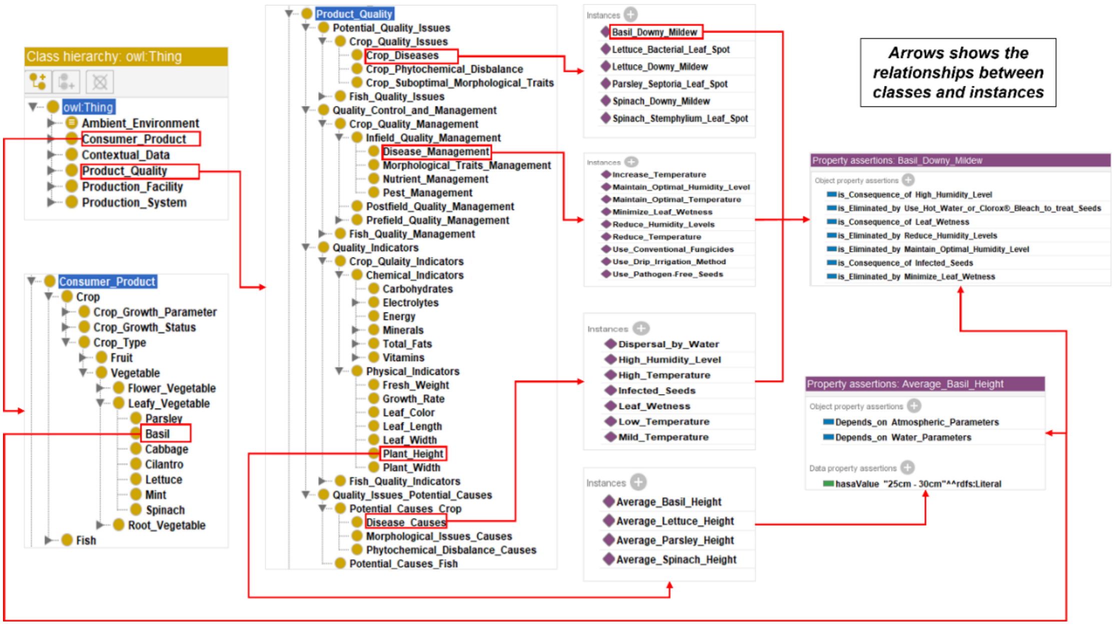


Figure 6- 6. Hierarchical structure of 'Consumer Product' and 'Product Quality' classes and respective instances in relation to Basil Crop.

6.3.5. Cloud-based application

The trained model of the crop disease detection system is then saved and deployed on a cloud-based application built on Streamlit. The ontology model 'AquaONT' is also deployed on the application, and relevant classes are integrated with the final disease detection model through the Owlready2 library. The layout of the application is shown in Figures 6-7. It consists of two user inputs 'Select Model' and 'Upload Image'. 'Select Model' provides an option to select the model as per requirement, which in this study are 'Crop Classification' referring to phase 1, 'Disease or No Disease' referring to phase 2, and 'Disease Type, causes and Treatments' referring to phase 3 of the proposed disease detection system. After model selection, an image is uploaded which is used by all the models. Once the disease is detected and classified, the causes and treatments of the disease are extracted from the ontology model automatically and displayed on the application panel. This kind of information is useful as it will allow agricultural practitioners to determine the causes of diseases and take precautionary steps in the early stages to avoid crop wastage and economic loss.

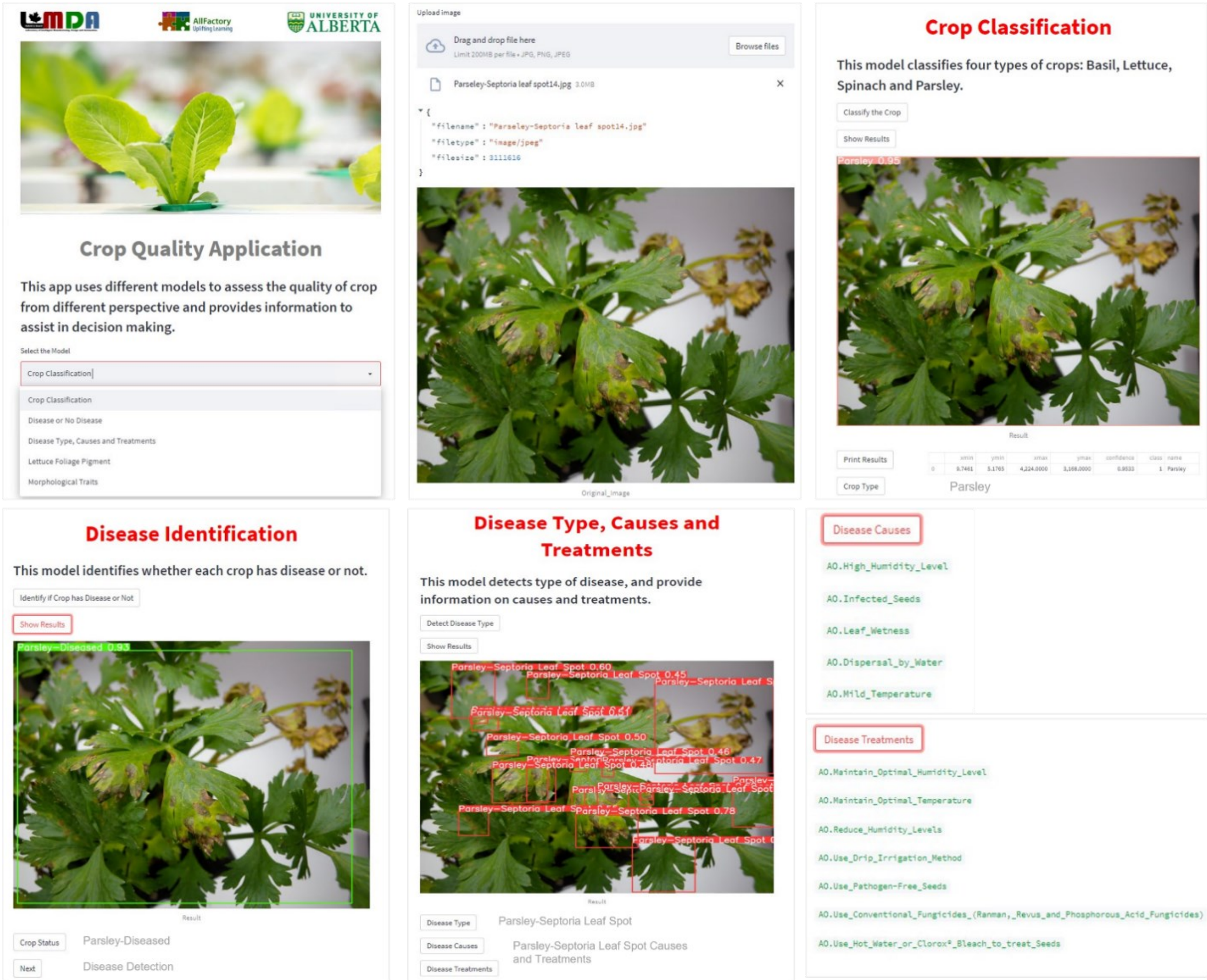


Figure 6- 7. Layout of cloud-based application for disease detection.

6.4. Experimental Results and discussions

This section presents the results of experiments performed in the current research work. First, the performance evaluation of deep learning models in three phases of the disease detection system is discussed. Next, the trained and validated system is tested on new data. In the end, the significance of the complete system is presented.

The performance of the classification model in phase 1 is evaluated using a validation dataset. For this phase, there are four classes to be classified: lettuce, basil, spinach, and parsley. The distribution of labeled images in the validation set for this model is shown in Table 6-4.

Table 6- 4. Dataset distribution of validation set for phase 1.

Class (Health + Diseased)	Number of images
Lettuce	160
Basil	104
Spinach	160
Parsley	104

The performance of the model is presented in the form of a confusion matrix (CM) shown in Figure 6-8.

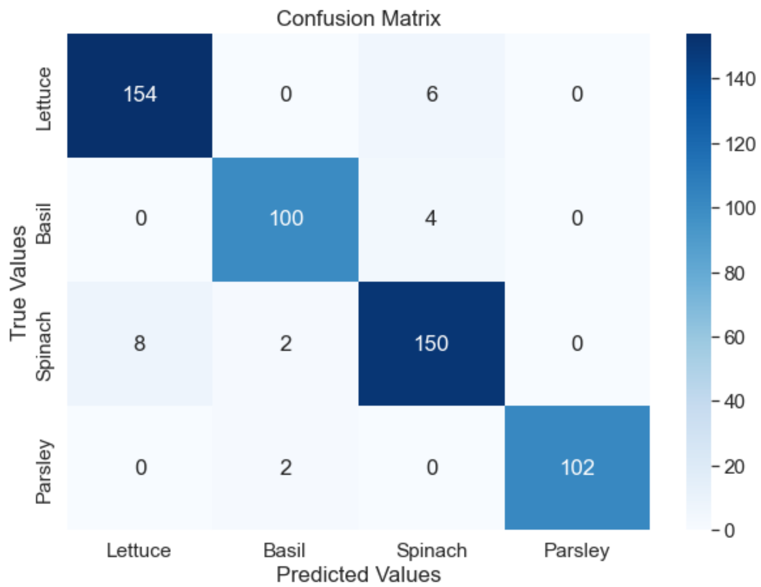


Figure 6- 8. Confusion matrix of classification results in phase 1.

The overall accuracy, precision, recall, and F-measure are computed by using the formulae given below and are summarized in Table 6-5.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Where, TP = True positives, TN=True negatives FP=False positives and FN=False negatives.

Table 6- 5. Class-wise results of classification model in phase1.

Crop	Accuracy	Precision	Recall	F1-Score
Lettuce	0.97	0.95	0.96	0.96
Basil	0.98	0.96	0.96	0.96
Spinach	0.96	0.94	0.94	0.94
Parsley	0.99	1	0.98	0.99
Average	-	96.25%	96%	96.25%
Overall accuracy	95.83%			

The classification model in phase 1 has achieved an overall accuracy of 95.83%, average precision of 96.25%, an average recall of 96%, and an average F1-score of 96.25%. From Table 6-5, it can be seen that the performance metrics of the spinach class are lower than the other classes. This is because eight examples of spinach have been classified as lettuce, and two examples have been classified as basil. This can happen as sometimes spinach leaves might look similar to lettuce or basil, particularly during the initial stages of their growth cycle. Next, the performance of the classification model in phase 2 is evaluated in a similar fashion. For phase 2, there are six classes that model has to classify, which are mentioned in section 6.3.2.2. Table 6-6 shows the distribution of the validation set used for the model in phase 2.

Table 6- 6. Distribution of validation dataset for phase 2.

Class	Number of images
Lettuce-Healthy (LH)	48
Lettuce-Diseased (LD)	112
Basil-Healthy (BH)	48
Basil-Diseased (BD)	56
Spinach-Healthy (SH)	48

Spinach-Diseased (SD)	112
Parsley-Healthy (PH)	48
Parsley-Diseased (PD)	56

The CM for this model is shown in Figure 6-9 and performance metrics are summarized in Table 6-7.

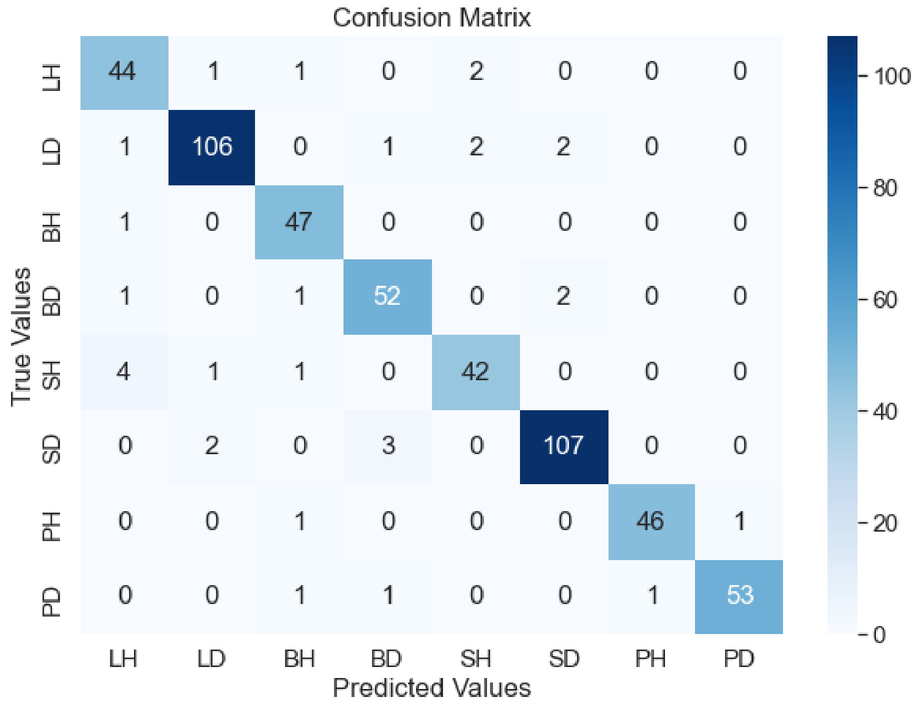


Figure 6- 9. Confusion matrix of classification results in phase 2.

Table 6- 7. Performance metrics of classification model in phase 2.

Class	Accuracy	Precision	Recall	F1-score
LH	0.979	0.86	0.92	0.89
LD	0.981	0.96	0.95	0.95
BH	0.989	0.90	0.98	0.94
BD	0.983	0.91	0.93	0.92
SH	0.981	0.91	0.88	0.89
SD	0.983	0.96	0.96	0.96
PH	0.994	0.98	0.96	0.97
PD	0.992	0.98	0.96	0.97
Average	-	94%	94%	93.6%
Overall accuracy	94.13%			

The classification model in phase 2 has achieved an overall accuracy of 94.13%, average precision of 94%, average recall of 94%, and an average F1-score of 93.6%.

It can be observed from the CM in Figure 6-6 that the model is also prone to confusion in distinguishing between some of the classes. For instance, six examples of LD (Lettuce-Diseased) are classified among LH (1), BD (1), SH (2), and SD (2). This might be due to a lack of clarity in identifying leaf patterns and diseased spots.

The performance of selected models for the detection phase (phase 3) is also evaluated using a validation dataset. For this phase, there are six different diseases that models have to detect in crop leaves. These six diseases and their distribution in the validation dataset are given in Table 6-8.

Table 6- 8. Distribution of validation dataset in phase 3.

Class	Number of images
Lettuce-Bacterial Leaf Spot (LBS)	56
Lettuce-Downy Mildew (LDM)	56
Basil-Downy Mildew (BDM)	56
Parsley-Septoria Leaf Spot (PSS)	56
Spinach-Downy Mildew (SDM)	56
Spinach- Stemphylium Leaf Spot (SSS)	56

In this phase, the metric that is used to evaluate and compare the performance of two models, i-e, Faster-RCNN, and YOLOv5s, is mean Average Precision (mAP). The mAP is the primary evaluation indicator used for the evaluation of object detection models [337]. In particular, $mAP_{@0.5}$ (mean value of mAP at IOU threshold = 0.5) is evaluated. The comparison of the two models against all the classes is presented in Table 6-9.

Table 6- 9. Class-wise comparison of two detection models.

Class	Faster-RCNN	YOLOV5s
Lettuce-Bacterial Leaf Spot (LBS)	77.32	83.86
Lettuce-Downy Mildew (LDM)	73.89	78.63
Basil-Downy Mildew (BDM)	75.47	80.11
Parsley-Septoria Leaf Spot (PSS)	78.63	84.55
Spinach-Downy Mildew (SDM)	74.19	79.87
Spinach-Stemphylium Leaf Spot (SSS)	79.52	85.74
$mAP_{@0.5}$	76.34	82.13

From Table 6-9, it can be seen that YOLOv5s with $mAP_{@0.5}$ of 82.13% have outperformed Faster RCNN. The two models have achieved the best mAP score for

Lettuce-Bacterial Leaf Spot (LBS), Parsley-Septoria Leaf Spot (PSS), and Spinach-Stemphylium Leaf Spot (SSS), whereas a low mAP score is observed for Lettuce-Downy Mildew (LDM), Basil-Downy Mildew (BDM), and Spinach-Downy Mildew (SDM). Downy Mildew causes light green to yellow angular spots on the upper surfaces of leaves and hence looks similar independently of the crop type. This causes confusion for the detector in distinguishing the crop-specific Downy Mildew. But with more data, this issue can easily be resolved. Later in the growth cycle, the plant tissue affected with Downy Mildew turns tan in spinach, purplish brown in basil, and light brown in lettuce, which are correctly identified by the detector.

The performance evaluations of models in three phases have shown that detection models are not as straightforward as classification models. This is because an image consists of many objects which belong to either the same class or different classes. Hence, three things have to be verified during evaluation, including object class, bounding box (object location), and confidence.

In the end, the two detection models are compared in terms of inference time which is an important metric that determines the detection speed. It is observed that the one-stage detector i-e., YOLOv5s with a detection speed of 52.8 FPS (frames per second) is faster than Faster-RCNN with a detection speed of 43.2 FPS. Moreover, it is also observed that YOLOv5s accurately detect objects of varying sizes with little to no overlapping boxes. All the comparisons between the two detection models show that YOLOv5s have a clear advantage in terms of accuracy and run speed. Therefore, in this study, YOLOv5s are used for developing the disease detection system.

After training and validation, the complete system with YOLOv5s is tested using the test set. The system has shown promising results by effectively classifying and detecting the diseases in specified crops, which shows the system's robustness in dealing with a variety of objects having different shapes, patterns, textures, colors, etc. Figure 6-10 shows examples where the system has accurately classified the crop and detected the diseased and healthy spots in crop leaves. Images in the first row of Figure 6-10 are the results from three phases of the disease detection system for the Lettuce crop, which is suffering from Bacterial Leaf Spot disease. Similarly, row 2 and row 3 are the results from three phases of the system showing Spinach and Parsley, respectively, and the diseases they are suffering from, such as Downy Mildew and Septoria leaf spot disease, respectively.

The final crop disease detection system is then deployed on a cloud-based application developed in section 6.3.5. Figure 6-7 shows the layout of the application. The ontology model discussed in section 6.3.4 is also integrated with the final system to build a complete real-time crop diagnostic system. The images are acquired wirelessly from the aquaponic facility through an interface developed on the Google Cloud Platform by the authors in Chapter 5 . The images are stored in a folder to be used by the crop diagnostic system. Once the crop type and its disease are identified, the causes and treatments are automatically extracted from the ontology model and displayed on the application panel. For instance, Figure 6-7 shows an example of a working crop diagnostic system for parsley crops. The disease detected by the system after image uploading is Septoria Leaf Spot. The crop diagnostic system extracts knowledge about potential causes and general treatments of this disease from AquaONT. The primary causes of Septoria Leaf Spot in Parsley could be high humidity levels, infected seeds, leaf wetness, etc. This disease could also be caused due to irregular variations in air temperature. The potential preventive measures and treatments suggested by the system for this disease include: maintaining optimal humidity and temperature levels in accordance with Parsley crop and indoor aquaponics environment throughout the growth cycle, treating seeds before germination with hot water or Clorox bleach, using conventional fungicides if the disease is spread out in multiple plants. Downy Mildew disease is one of the most common diseases observed in different crops [368]. In the greenhouse or indoor farming environment, the potential causes of this disease are the same irrespective of crop type, which includes: high humidity, cool temperatures, infected seeds, and leaf wetness [368]. Therefore, the methods to treat Downy Mildew in lettuce, basil, and spinach are also similar. This means that the classification of Downy Mildew disease with respect to crop type does not impact the results related to disease treatments. Despite this independence, it is still significant to perform the classification of Downy Mildew for each crop individually as its symptoms for three crops, lettuce, basil, and parsley, change later in the growth cycle. This might cause confusion for the detector to distinguish Downy Mildew from other diseases. For instance, the lettuce tissue affected with Downy Mildew eventually turns brown in later stages and these symptoms are similar to the Bacterial Leaf Spot symptom in lettuce, and both diseases have different treatment methods.

Figure 6- 10. Results from proposed disease detection system.

The significance of the proposed system is that it can act as a vital tool for agriculturalists who wants to develop and digitize aquaponics farm. This system will allow them to diagnose diseases at early stages and also assist them in decision-making regarding crop characteristics and treatments of diseases. Moreover, this study will also promote the introduction of new implementations, such as research on the complex relationship between dynamic parameters (environmental and water) and diseases in aquaponics farms and self-adapting farms in case of disease detection. These smart technologies in the aquaponic system will reduce crop wastage and ensure both economic and environmental benefits.

6.5. Conclusions and Future Prospects

This study proposes a crop diagnostic system for leafy green crops grown in an aquaponics environment. Four leafy green crops, lettuce, basil, spinach, and parsley, are considered. The first dataset is developed that contains 2640 healthy and diseased images of these four crops collected from various sources. Next, a system is proposed that can efficiently and effectively identify crops and diseases. The detection system works in three phases. The first phase classifies the crop type, the second phase classifies whether the crop is healthy or diseased, and then in the third phase, the disease is detected if the crop is classified as unhealthy in the previous phase. All the models used in this study are initialized using transfer learning and then trained on a dataset prepared for leafy green crops. The performance of the models is evaluated, and promising results are achieved. For instance, in the detection phase, YOLOv5s with mAP@0.5 of 82.13% and detection speed of 52.8 FPS has outperformed Faster-RCNN. Based on the performance, YOLOv5s is selected as the final model for this study. The ontology model that contains knowledge related to the causes and treatments of diseases is then integrated with the final crop disease detection system. Finally, a cloud-based application is designed where the final crop diagnostic system consisting of a disease detection system and ontology model is deployed. The proposed system proves to be accurate and flexible enough to be used in real scenarios and hence is not limited to being disturbed by potential changing conditions and environments. It can be a helpful tool for agricultural practitioners who want to explore modern farming practices and want to integrate smart techniques into their farms. This system will not only help them in disease diagnosis and quantification but will also assist them in decision-making regarding potential treatments against identified diseases at early stages.

For future work, the system will be extended to include other leafy green crops. Moreover, the dataset will also be extended, and more real-field images will be incorporated. Moreover, a mobile application will be constructed, reducing the latency, and providing data privacy, which normally occurs in cloud-based systems.

Chapter 7 Non-destructive identification of foliage chlorosis in lettuce crop grown in aquaponic facility using image processing

7.1. Introduction

Just like traditional agriculture, crops grown in aquaponics may also face quality issues — causing crop wastage and impacting the overall productivity of the farm. Various quality indicators such as crop morphological attributes (crop height, width, area, volume), biomass production, and foliage color can be used as the measure of crop quality and yield potentiality [21]. One of the quality indicators is foliage color which indicates the chlorophyll content and is used to evaluate the health of the crop. If the color of the foliage is green, it represents that the crop is healthy. Whereas, if it is yellow, it signifies that the crop is suffering from some type of disease or deficiency, which is causing interference in the production of chlorophyll contents [24]. The primary contributors of chlorosis or leaf yellowing in crops grown in aquaponics could be i) inadequate environmental conditions (humidity, temperature, illumination, etc.), ii) incorrect design of the system leading to an irregular supply of nutrient-enriched water which fails to reach the roots of plants, iii) poor water quality (improper pH), iv) root damage or compacted roots, v) insufficient concentrations of required minerals such as N-NO_3, P, K, Ca, and Mg in the effluent, vi) nutrient deficiency in plants, and vii) diseases or pests [24], [369].

The conventional method to identify the plant's health is based on visual observation, requiring certain expertise from agriculture practitioners [370]. Visual detection, however, is a time-consuming and laborious task, and there is a probability of misdiagnosis, especially in the early growth stages [370]. Other methods include chemical analyses and leaf color chart (LCC) matching, which again are costly, time-consuming, and destructive techniques. Chemical methods involve the collection of plant tissue for laboratory analyses of plant leaves. One of the most widely used chemical methods is the Kjeldahl digestion assay [371]. Although this method is accurate, sample preprocessing and delays in laboratory analyses hinder its widespread usage. The standard LCC tool is also available and used as a reference to estimate leaf color and plant health [372]. This technique is widely used in many countries, but it is a manual inspection process and hence, time-consuming.

To overcome these challenges, agriculture methods have been automated for years, and hence a number of non-destructive methods have been proposed to detect plant health. One of the methods is the spectral reflection method which uses the property of chlorophyll that has different reflection intensities at different wavebands to assess the health of the plant. Several portable meters, such as SPAD (Soil Plant Analysis Development), are developed based on this method [373]. The spectral instruments are fast and fairly accurate, but they are very expensive. Hyperspectral imaging and spectral remote sensing also use the spectral reflection principle [374]. Again, hyperspectral instruments are costly and require specific environmental conditions for proper sampling. With the development of technology, some researchers applied computer vision techniques to detect the health of plants based on their nutritional status. Computer vision is low-cost and non-destructive, but it requires a large amount of data for training and achieving the desired performance of the model [375].

7.1.1. Research motivation and contribution

Considering the aforementioned challenges, this study aims to propose a methodology based on an image processing technique to evaluate the health of lettuce crops grown in aquaponics facilities based on their foliage color. To be more certain, the estimation of chlorophyll content or nutrient deficiency is out of the scope of this study. The focus of the study is to determine the plant health by extracting the foliage and its Red (R), Green (G), Blue (B) channel values using 'HSV Space segmentation'. HSV stands for hue, saturation, and value [376]. The foliage color detection model is then developed using mean values of R, G, B channels and a color distance model. The color distance model is to compute the foliage color difference from the threshold values. Numerous color distance models are available for this purpose, such as Euclidean Distance, Color Approximation Distance, CIEXYZ, CIE76, CIE94, CIEDE2000, and CMC l:c [24]. In this study, Euclidean Distance (ED) model is used as it is the simplest method of finding the distance between two colors within an RGB color space [24]. Moreover, it works well when a single color is to be compared to a single color, and the need is to simply know whether a distance is greater or smaller, which is the case with the proposed model in this study. The model is built in Jupyter Notebook and saved in a local directory. A cloud-based application is also developed using Streamlit, which is an open-source app framework for machine learning and data science [377]. The performance of the model is validated, achieving an accuracy of 95%.

7.1.2. Chapter organization

The remainder of the chapter is structured as follows: Section 2 will present the related work, Section 3 will explain the methodology used to develop the system, Section 4 will present results and discussion along with model significance and limitations, and finally, Section 5 will discuss the conclusions and future work.

7.2. Related work

This section presents the recent and relevant image processing-based models that have used different color spaces and techniques to identify leaf chlorosis and assess the quality of crops. Yang et al. proposed a model based on a support vector machine (SVM) and advanced imaging processing techniques such as image binarization, mask, and filling approaches for the extraction of selective color features such as a* (CIELAB color space), G (green from RGB color space), and H (hue from HSV color space) to detect the yellow and rotten lettuce leaves in hydroponics system [378]. The model has achieved an accuracy of 98.33%. Maity et al. proposed a model based on Otsu's method and k-means clustering technique to detect faulty regions in leaves [379]. Wang et al. developed an HSV and decision tree-based method for greenness identification of maize seedling images captured in the outdoor field [380]. Benjamin et al. proposed a methodology based on the color analysis technique to determine the quality of tomato leaves using Otsu's method, SVM, k-NN (k-nearest neighbor), and multi-layer perceptron (MLP) [381]. Their model obtained an accuracy of 86.45% when classifying the healthy tomato leaves from the diseased tomato leaves and an accuracy of 97.39% when classifying the type of disease suffered by a diseased leaf. Sharad et al. developed a system based on LAB (L*: lightness, a*: red/green value, b*: blue/yellow value) space-based color histogram, k-NN and random forest to detect the quality of apple leaves [382]. This approach has achieved an accuracy of 98.63%.

The aforementioned models have made great contributions to literature, but some limitations are observed. For instance, most models have used images belonging to one scenario. Either they are taken indoors in the lab environment or outdoors in open-air fields. Secondly, some models have used non-destructive chemical approaches to collect the preliminary data, particularly while assessing the health of plants based on chlorophyll content, nitrogen level, or nutrient deficiency. Considering that, in this study, a fully automated, low-cost, and non-destructive model is proposed that is built considering a variety of lettuce images from different sources.

7.3. Research Methodology

The block diagram illustrating the five sequential modules of research methodology is shown in Figure 7-1. Each module, along with its elements, is described in the next subsections.

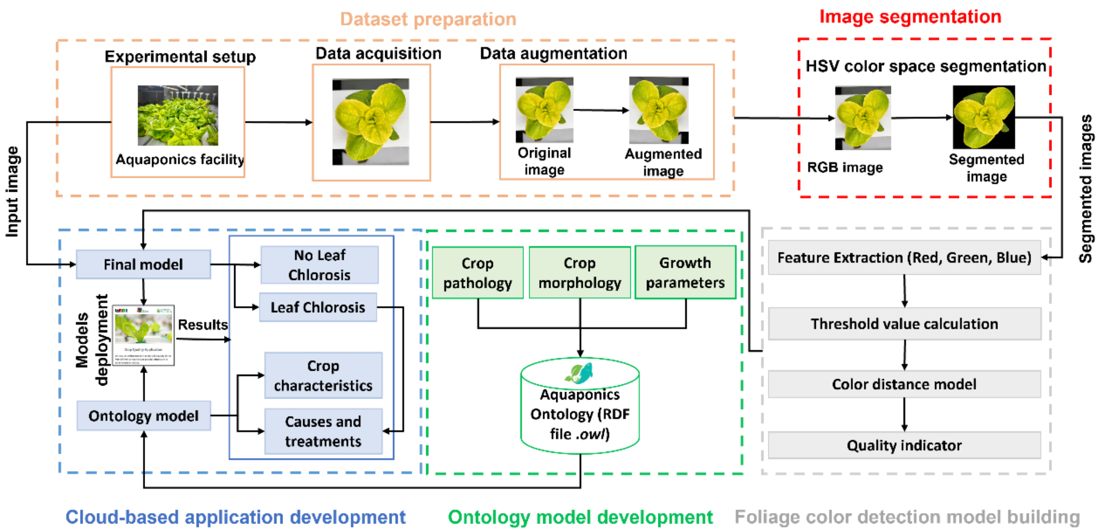


Figure 7- 1. Research methodology outline.

7.3.1. Data preparation

The image dataset is constructed using a variety of little gem romaine lettuce images from diverse sources. This involves top-view images of lettuce grown in the Allfactory 4.0, an NFT-based aquaponic facility at the University of Alberta, Canada, focusing on smart indoor farming. These images are divided into two classes based on the color of foliage: green foliage (no leaf chlorosis), and yellow foliage (leaf chlorosis). To increase the model flexibility to segment lettuce foliage irrespective of background and to ensure it correctly determines the plant's health, the dataset is complemented with more lettuce images obtained from Ecosia, a search engine based in Berlin, Germany. Figure 7-2 shows examples of some of the images.

Figure 7- 2. Image dataset: a and b are acquired from an aquaponic facility and c, d, e, and f are downloaded from Ecosia.org.

Next, the image augmentation process is performed to increase the dataset and reliability of the segmentation process despite the location and orientation of objects in the image by generating new images from existing images. In this study, Albumentations, a Python library, is used for fast and flexible image augmentations [359]. The different augmentation techniques applied are horizontal flip, vertical flip, 90° rotation, and glass noise. In total, 100 images (50 from both classes) were selected randomly for the augmentation, which created 100 new images. The new images were added to their respective classes, increasing the length of the dataset to 500 images. Figure 7-3 shows examples of augmentations.

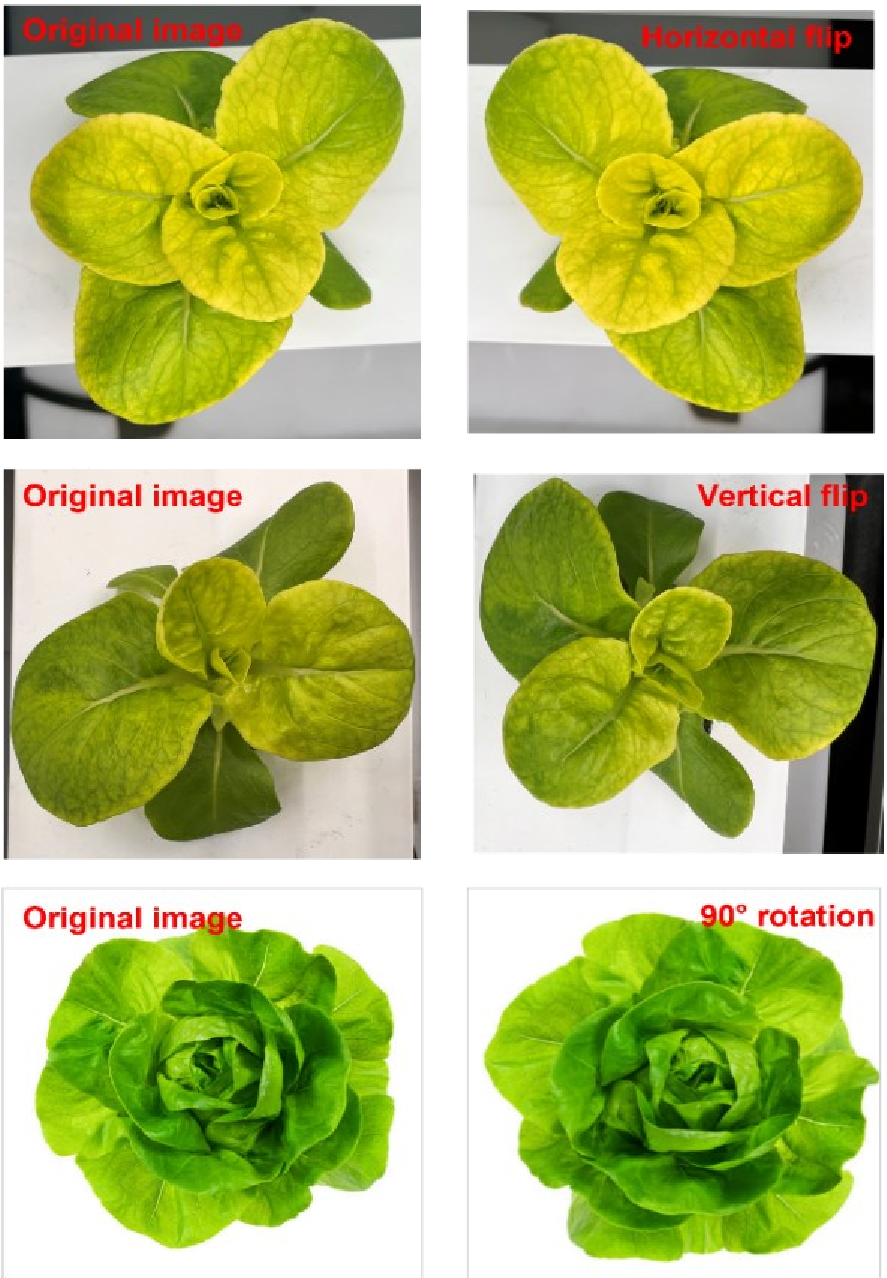


Figure 7- 3. Data augmentation performed on different images.

7.3.2. Image segmentation

Image segmentation was performed to extract the lettuce foliage from the background for further processing. This study uses the HSV segmentation model to segment the image [376]. There are two stages of the image segmentation process, which are detailed in the next two subsections.

7.3.2.1. HSV color space

The acquired images from the smartphone are RGB images, where the color of any object in these images is represented with the combination values of R, G, and B channels. The main problem with this color representation is that the objects' colors are affected by variations in illumination conditions [177]. With the HSV color segmentation technique, as the name suggests, HSV color space is used, which describes the objects' colors independent of the illumination effect [376]. The difference between various color spaces is usually based on color representation. For instance, the object's color in HSV color space is represented by three different parameters, namely hue (H), saturation (S), and value (V). H represents the color of the object, whereas S and V values represent the illuminance state of the object's color [376]. This kind of description provides the ability to discriminate the color from the illuminance while avoiding the effect of illumination changes on the object's color. Therefore, the first step of segmentation is to convert the image's color space from the RGB into HSV. Generally, the transformation process from RGB into HSV can be performed using the following equations [378].

$$R' = \frac{R}{255}, \qquad G' = \frac{G}{255}, \quad B' = \frac{B}{255}$$

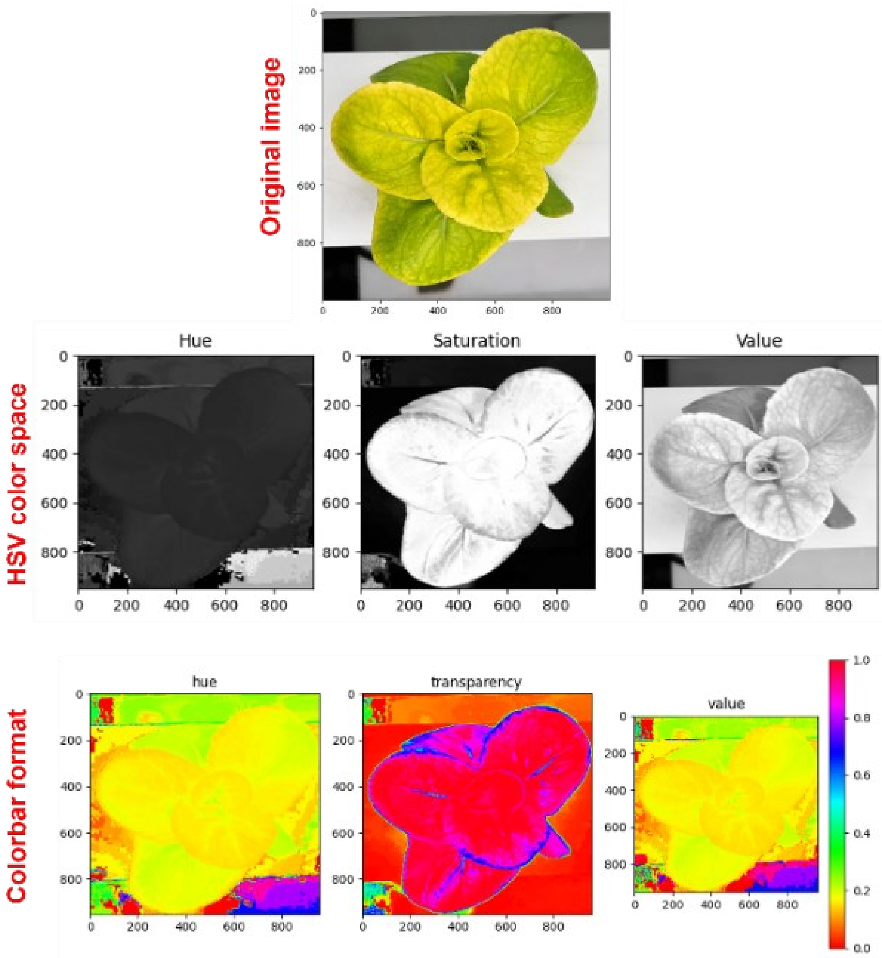
$$M = max(R', G', B'),\ \ m = min(R', G', B'), C = M - m$$

$$H = \begin{cases} 0° & if\ C = 0 \\ 60° \times \left(\frac{G' - B'}{C} mod\ 6\right) & if\ M = R' \\ 60° \times \left(\frac{G' - B'}{C} + 2\right) & if\ M = G' \\ 60° \times \left(\frac{G' - B'}{C} + 4\right) & if\ M = B' \end{cases}$$

$$S = \begin{cases} 0 & if\ M = 0 \\ \frac{C}{M} & if\ M \neq 0 \end{cases}$$

$$V = M$$

After image transformation, a color bar is created, which provides intensity values of H, S, and V channels. These values are used in the next stage for segmenting the image. Figure 7-4 shows an example of the original image, its HSV channel, and color bar format.



F

Figure 7- 4. Illustration of an image, its HSV channel, and color bar format.

7.3.2.2. Image hue thresholding

The second step of image segmentation is to determine the suitable threshold value to distinguish between foreground and background. For this purpose, the hue image obtained in the first step is used as it provides a suitable grayscale image that can be used to classify objects based on the color content. The upper and lower range of the hue channel is obtained from the color bar. This range is used to define an upper and lower threshold value for lettuce foliage in a hue image in the form of a mask. This mask is then applied to the R, G, and B channels of the original image, which are then stacked to get the segmented image. The final segmented image is saved in RGB format. To save time, the segmentation process is automated, and by the end of the process, each segmented image is saved in a common directory.

7.3.3. Foliage color detection model

After image segmentation, we have the lettuce foliage in the RGB channel. The R, G, and B values of each lettuce foliage (foreground) are extracted from segmented images. These images are represented as i and j for two classes g (Green Foliage – No Leaf Discoloration) and y (Yellow Foliage – Leaf Discoloration), respectively. Then the mean value of each color channel red (μ_R), green (μ_G), and blue (μ_B) is computed, which are represented by Equations 1 and 2.

$$\mu_{g,i} = [\mu_{R,i}, \mu_{G,i}, \mu_{B,i}] \qquad (1)$$

$$\mu_{y,j} = [\mu_{R,j}, \mu_{G,j}, \mu_{B,j}] \qquad (2)$$

$\mu_{g,\,i}$ and $\mu_{y,\,j}$ represents the mean values of three color channels of foreground (lettuce foliage). Equations 3 – 8 are used for computing the mean values of channels.

$$R_{i/j} = \sum_{1}^{n_{R,i/j}} R_{n,i/j} \qquad (3)$$

$$G_{i/j} = \sum_{1}^{n_{G,i/j}} G_{n,i/j} \qquad (4)$$

$$B_{i/j} = \sum_{1}^{n_{B,i/j}} B_{n,i/j} \qquad (5)$$

$$\mu_{R,i/j} = \frac{R_{i/j}}{n_{R,i/j}} \qquad (6)$$

$$\mu_{G,i/j} = \frac{G_{i/j}}{n_{G,i/j}} \qquad (7)$$

$$\mu_{B,i/j} = \frac{B_{i/j}}{n_{B,i/j}} \qquad (8)$$

Where $R_{i/j}$, $G_{i/j}$, and $B_{i/j}$ refer to the summation of red, green, and blue values of lettuce foliage in two classes, i/j refers to either image belonging to g class or y class, and $n_{R,i/j}$, $n_{G,\,i/j}$, and $n_{B,\,i/j}$ represents R, G, B count of lettuce foliage respectively.

The background is segmented images is black with [R, G, B] = [0, 0, 0]. Hence, R, G, and B count and values of background are not included while determining the mean value of R, G, and B channels for the foreground. The process of calculating the mean value of R, G, and B channels is again automated to save time. The values for each channel are automatically saved in an excel file. While saving the results, it is ensured that mean values of R, G, and B are saved for their respective image label and class category g and y. Next, reference or threshold values (g_{ref} and y_{ref}) are determined for both g and y classes, shown in equations 9 and 10. To compute g_{ref}, three average values are calculated, which are related to the mean red, mean green, and mean blue values of images saved in the excel file for the g category. The total number of mean values for each channel is m (total values = 3×m). The first average value is obtained by summing all the green channel values and dividing the results by the total number of green values (m). Similarly, the second and third average values are obtained by summing all the green channel values and all blue channel values of all images in the m category and dividing the results by the number of green (m) and blue values (m), respectively. A similar computation is done for y_{ref} while considering the channel values and their count (l) for images in the y category. Equations 11–16 are used to calculate g_{ref} and y_{ref}.

$$g_{ref} = [\bar{x}_{R,m}, \bar{x}_{G,m}, \bar{x}_{B,m}] \qquad (9)$$

$$y_{ref} = [\bar{x}_{R,l}, \bar{x}_{G,l}, \bar{x}_{B,l}] \qquad (10)$$

$$\bar{x}_{R,m} = \frac{\sum_1^m R_m}{m} \qquad (11)$$

$$\bar{x}_{G,m} = \frac{\sum_1^m G_m}{m} \qquad (12)$$

$$\bar{x}_{B,m} = \frac{\sum_1^m B_m}{m} \qquad (13)$$

$$\bar{x}_{R,l} = \frac{\sum_1^l R_l}{l} \qquad (14)$$

$$\bar{x}_{G,l} = \frac{\sum_1^l G_l}{l} \qquad (15)$$

$$\bar{x}_{B,l} = \frac{\sum_1^{l=} B_l}{l} \qquad (16)$$

Where $\bar{x}_{R,m}$, $\bar{x}_{G,m}$, and $\bar{x}_{B,m}$ are averages of m×red, m×green, and m×blue channels values in the g category respectively. R_m, G_m, and B_m are values of m×red, m×green, and m×blue channels in the g category. Where $\bar{x}_{R,l}$, $\bar{x}_{G,l}$, and $\bar{x}_{B,l}$ are averages of l×red, l×green, and l×blue channels values in the y category, respectively. R_l, G_l, and B_l are values of l×red, l×green, and l×blue channels in the y category.

After determining the reference or threshold values, the color distance model is used to compute the foliage color difference from the threshold values. In this study, Euclidean Distance (ED) model is used, and its general equation is presented below [24].

$$d = \sqrt{\Delta R^2 + \Delta G^2 + \Delta B^2}$$

Where

$$\Delta R = R_2 - R_1, \qquad \Delta G = G_2 - G_1, \Delta B = B_2 - B_1$$

Based on the ED model, two distances d_1 and d_2 are computed using two threshold values g_{ref} and y_{ref}, respectively. d_1 determines the distance from the green color threshold, whereas d_2 determines the distance from the yellow color threshold. For single foliage, both d_1 and d_2 are determined. The lower value of d_1 and a higher value of d_2 suggests that color patterns of foliage are closer to g_{ref} or in other words, green tones. Conversely, a lower value of d_2 and a higher value of d_1 suggests that color patterns of foliage are closer to y_{ref} or in other words, yellow tones. The governing equations for d_1 and d_2 are given below.

$$d_1 = \sqrt{(x_R - \bar{x}_{R,m})^2 + (x_G - \bar{x}_{G,m})^2 + (x_B - \bar{x}_{B,m})^2} \quad (17)$$

$$d_2 = \sqrt{(x_R - \bar{x}_{R,l})^2 + (x_G - \bar{x}_{G,l})^2 + (x_B - \bar{x}_{B,l})^2} \quad (18)$$

Where x_R, x_G, and x_B are the mean values of three channels (R,G, B) of the foreground in the segmented image of the validation samples (their mean R, G, B values are not included while developing the model).

Lastly, the quality indicator (Q) is defined as a function of d_1 and d_2 for evaluating the health of plants based on foliage color. In this essence, when green foliage with no leaf depigmentation is detected, the value of Q is equal to 1, which implies that the crop is healthy. On the other hand, when yellow foliage with leaf depigmentation is detected, the value of Q is equal to 0, suggesting that the crop is unhealthy. The Q is represented as under.

$$Q = f(d_1, d_2) = \begin{cases} 1 & if\ d_1 < d_2 \\ 0 & if\ d_2 < d_1 \end{cases} \quad (19)$$

7.3.4. Ontology model

The complete development and details of all concepts and instances of an ontology model 'AquaONT' is available in previous work by the authors (Chapter 3). AquaONT is a unified ontology model that represents and stores the essential knowledge of an aquaponic 4.0 system. It consists of six concepts: Consumer_Product, Ambient_Environment, Contextual_Data, Production_System, Product_Quality, and Production_Facility. In this study, two classes, 'Consumer_Product' and 'Product_Quality' are used for knowledge extraction. The 'Consumer_Product' class provides an abstract view of the type, growth status, and growth parameters of ready-to-harvest crops in an aquaponic system. Whereas the 'Product_Quality' class provides knowledge on crop attributes related to pathology (abiotic and biotic stresses, causes, and the ways and means by which these can be managed or controlled), morphology (canopy dimensions such as area, length, width, etc.) and foliage color. The lettuce crop is considered in this study. The crop growth and quality attributes are defined as instances of respective classes, which are extracted once the crop foliage is detected as yellow (or leaf chlorosis is detected). Figure 7-5 shows the hierarchical architecture of the 'Consumer_Product' and 'Product_Quality' classes with their instances for the Lettuce crop in Protégé7 (an open-source ontology editor and framework developed at Stanford University) environment.

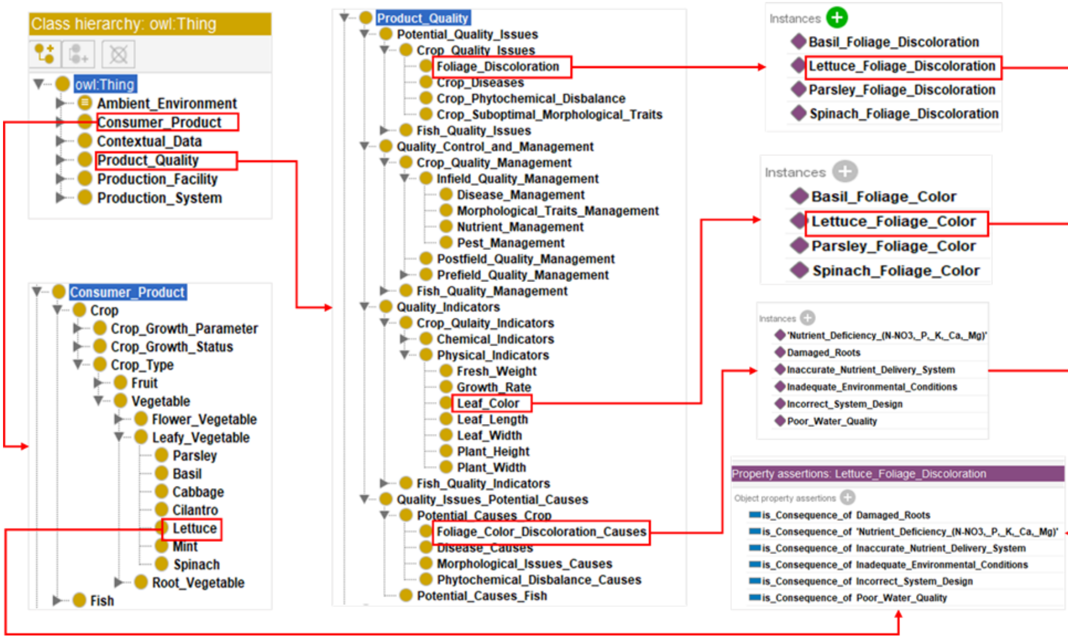


Figure 7- 5. Ontology model showing classes, instances, and relationships between them.

7.3.5. Cloud-based application

The proposed foliage detection model and ontology model are deployed on a cloud-based application built on Streamlit. The layout of the app is shown in Figures 7-6 to Figure 7-9. The app works in six stages. The first and second stages are associated with two user inputs, "Select the Model" and "Upload Image" as shown in Figure 7-6. The first input allows the user to select a relevant quality evaluation model. This app has other quality models integrated into it, which are out of the scope of this study. In this study, the relevant model is "Lettuce Foliage Pigment". After selecting the model, the image is selected using the second input. The third and fourth stages are linked with two widgets, "Preprocess and Segment Image" and "Determine the Crop Status", respectively, shown in Figure. 7-7, that run the sub-processes associated with the model. As the name suggests, the first widget activates the segmentation algorithm, which preprocesses and segments the image selected by the user in the second stage. Likewise, the second widget activates the model developed in the study. The model determines the status of the crop and displays results on the application panel. In the fifth stage, the sensor data from the dashboard is acquired and displayed to monitor the environmental conditions, as shown in Figure 7-8. By clicking 'Sensor Data', the most recent data will be displayed. In the sixth stage, a widget is developed 'Causes and Treatments', which is linked with 'AquaONT'. This widget extracts knowledge from the ontology model related to possible causes of leaf yellowing in the aquaponic facility. Figure 7-9 shows the sixth stage of the app when yellow foliage is detected.

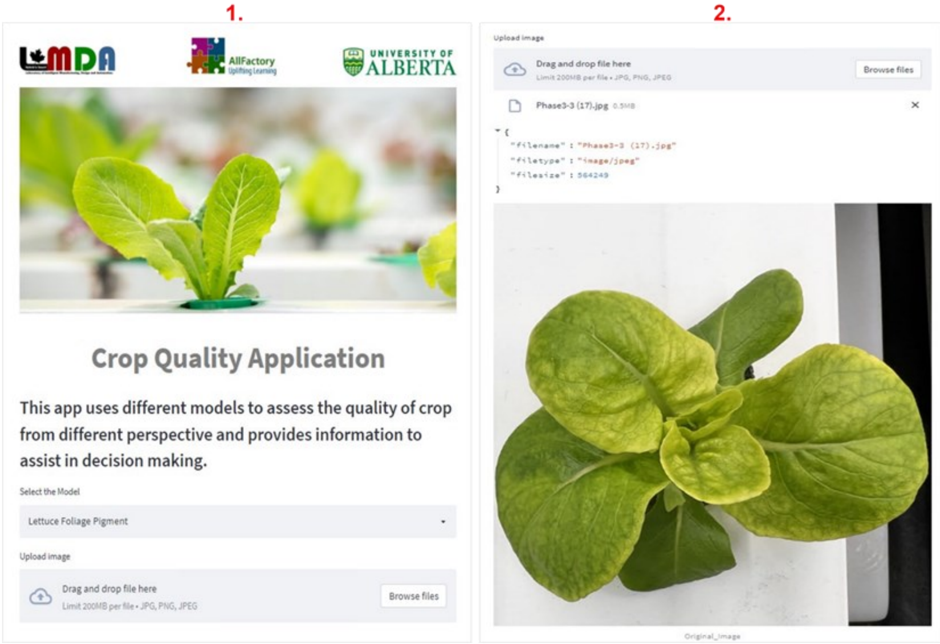


Figure 7- 6. Stage 1 and 2 of cloud -based application.

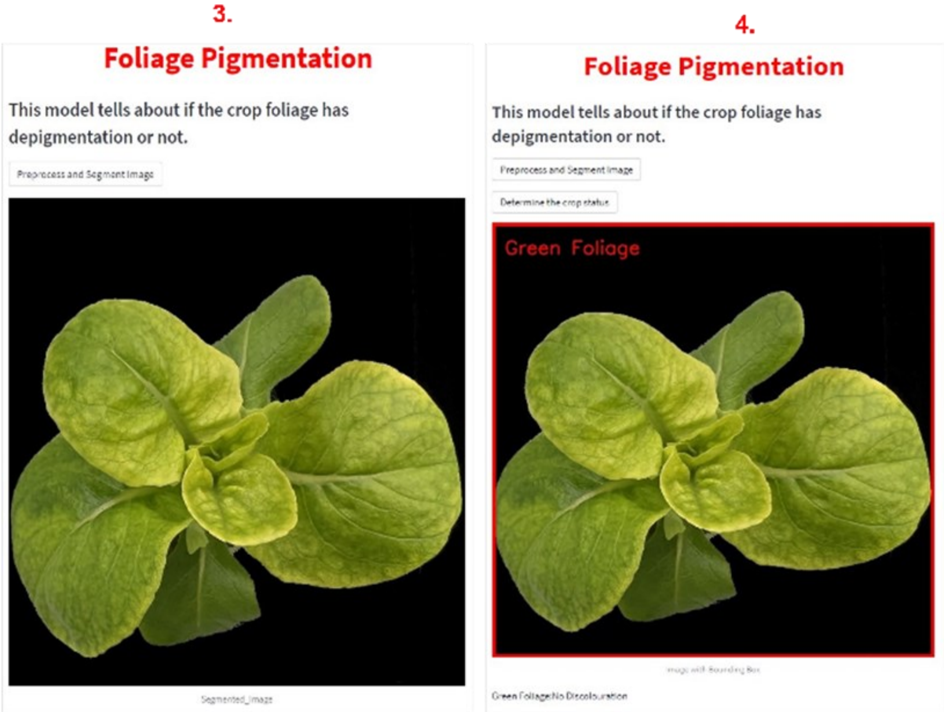
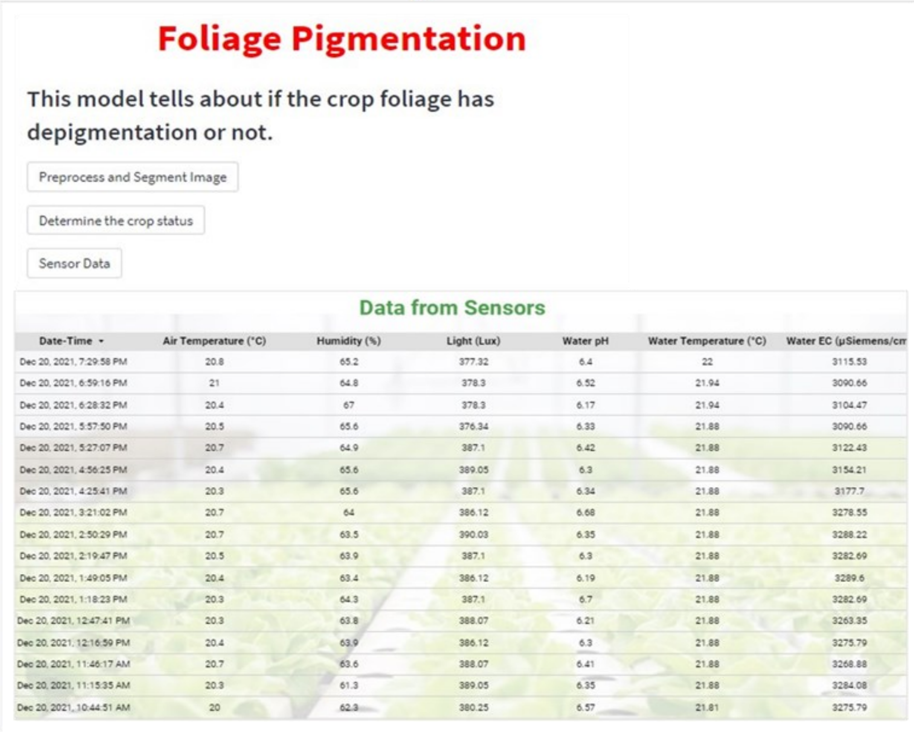


Figure 7- 7. Stages 3 and 4 of a cloud-based application.

5.



Date-Time	Air Temperature (°C)	Humidity (%)	Light (Lux)	Water pH	Water Temperature (°C)	Water EC (µSiemens/cm
Dec 20, 2021, 7:29:58 PM	20.8	65.2	377.32	6.4	22	3115.53
Dec 20, 2021, 6:59:16 PM	21	64.8	378.3	6.52	21.94	3090.66
Dec 20, 2021, 6:28:32 PM	20.4	67	378.3	6.17	21.94	3104.47
Dec 20, 2021, 5:57:50 PM	20.5	65.6	376.34	6.33	21.88	3090.66
Dec 20, 2021, 5:27:07 PM	20.7	64.9	387.1	6.42	21.88	3122.43
Dec 20, 2021, 4:56:25 PM	20.4	65.6	389.05	6.3	21.88	3154.21
Dec 20, 2021, 4:25:41 PM	20.3	65.6	387.1	6.34	21.88	3177.7
Dec 20, 2021, 3:21:02 PM	20.7	64	386.12	6.68	21.88	3278.55
Dec 20, 2021, 2:50:29 PM	20.7	63.5	390.03	6.35	21.88	3288.22
Dec 20, 2021, 2:19:47 PM	20.5	63.9	387.1	6.3	21.88	3282.69
Dec 20, 2021, 1:49:05 PM	20.4	63.4	386.12	6.19	21.88	3289.6
Dec 20, 2021, 1:18:23 PM	20.3	64.3	387.1	6.7	21.88	3282.69
Dec 20, 2021, 12:47:41 PM	20.3	63.8	388.07	6.21	21.88	3263.35
Dec 20, 2021, 12:16:59 PM	20.4	63.9	386.12	6.3	21.88	3275.79
Dec 20, 2021, 11:46:17 AM	20.7	63.6	388.07	6.41	21.88	3268.88
Dec 20, 2021, 11:15:35 AM	20.3	61.3	389.05	6.35	21.88	3284.08
Dec 20, 2021, 10:44:51 AM	20	62.3	380.25	6.57	21.81	3275.79

Figure 7- 8. Stage 5 of a cloud-based application.

6.

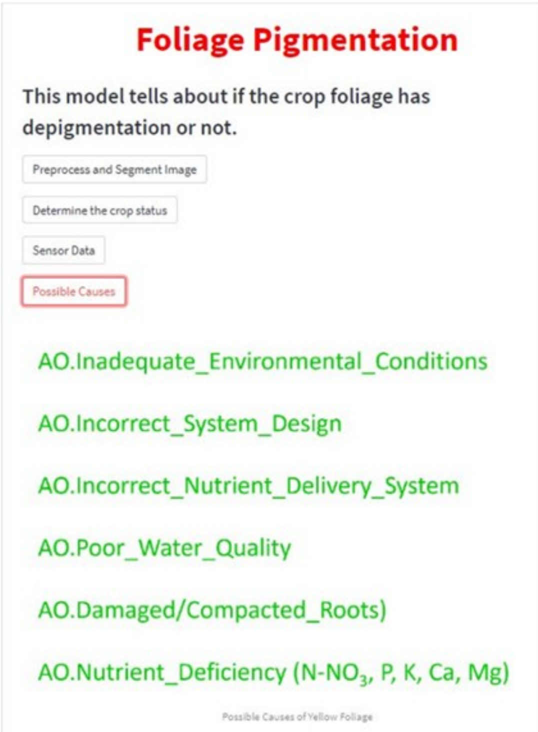


Figure 7- 9. Stages 6 of a cloud-based application

7.4. Results and discussion

This section first presents the validation of the proposed method by a case study. Then, the performance of the proposed method is compared with existing similar methods.

To validate the proposed model, twenty healthy seedings are placed in NFT-based hydroponic systems for five weeks (plantation cycle), after which lettuce is harvested. A 12MP Sony Exmor RS camera sensor is used to capture crop images during this period. Twenty images of 4032×3024 pixels (one image for one lettuce plant) are captured every day at 9:00 am from the top while keeping the distance between camera and channel at a value of 40 centimeters throughout the plantation cycle, i.e., five weeks. In total, 700 images of plants are collected over 5 weeks. During the first three weeks, no significant difference is observed in the color of the foliage. After the third week, foliage chlorosis is observed in eight lettuce plants. Therefore, for further processing, the images captured in the last two weeks of the plantation cycle are considered for model validation. In total, 280 images are divided into two classes based on the color of foliage: Green Foliage – No Leaf Chlorosis (168 images), and Yellow Foliage – Leaf Chlorosis (112 images). The dataset is complemented with more lettuce images with green (32) and yellow (88) foliage, downloaded from Ecosia. The images are added to their respective classes. All the images are resized to 1000 × 1000 pixels and are saved in JPG format. The augmentation process is then performed. In total, 100 images (50 from both classes) are selected randomly for the augmentation, which created 100 new images. The new images were added to their respective classes, increasing the length of the dataset to 500 images. Half of these images belong to the (g) class, and half belongs to the (y) class and hence are saved in two folders named (g) and (y), respectively. Out of 500 images, 100 random images (50 from each folder) were extracted and saved in a separate validation folder to be used for model evaluation. To complement the validation data, 20 images are randomly selected (10 from each class), and their R, G, and B values are altered using Adobe Photoshop in a way that the healthy-looking lettuce appears yellow and the unhealthy lettuce appears green. The validation dataset now has 120 images in total. Figure 7-10 shows an example of new images generated for the validation dataset.

Figure 7- 10. Images generated in Adobe Photoshop.

The segmentation is then performed on all 520 images in the dataset. Figure 7-11 shows an example of segmented images. R, G, and B values and their count are then computed for the foreground (lettuce foliage) of 400 segmented images in two classes g and y. The means values of R, G, and B channels are then computed. Each class has 200 foliage images, so for each class $3 \times 200 = 600$ mean values (3 refers to 3 channels of an image) are obtained which are automatically saved in an excel file.

Figure 7- 11. Example of segmented images.

Out of 600 means values for each class, 200 belong to the red channel, 200 belong to the green channel, and 200 belong to the blue channel. The threshold values g_{ref} and y_{ref} are obtained by dividing the mean values of 3 channels by 200 and are given below.

$$g_{ref} = [\bar{x}_{R,m}, \bar{x}_{G,m}, \bar{x}_{B,m}] = [123.4, 138.2, 19.52]$$

$$y_{ref} = [\bar{x}_{R,l}, \bar{x}_{G,l}, \bar{x}_{B,l}] = [156.6, 155.8, 22.2]$$

Next, the model is validated using a validation dataset consisting of 120 different segmented images belonging to two classes g and y. The mean values of three channels are computed for each image and are inserted in equations 17 and 18 in place of x_R, x_G, and x_B, along with reference values g_{ref} and y_{ref} computed above. The d_1 and d_2 are determined for all 120 images in the validation dataset using equations 17 and 18, respectively. Quality indicator, Q is also determined using equation 19 for 120 images. The performance of the model on the validation dataset is then evaluated by analyzing the ground truth Q value and predicted Q value. In the validation dataset, 60 images have a ground truth Q value of 1, meaning these images contain healthy and green lettuce foliage. 60 images have a ground truth value of 0, meaning these images contain unhealthy and yellow lettuce foliage. The performance is presented in the form of a confusion matrix (CM), shown in Figure 7-12 [177]. The different values of the CM are interpreted as.

- True Positive (TP) = 58; meaning 58 plants were actually healthy, and the model correctly classified them healthy as well.
- True Negative (TN) = 57; meaning 57 plants were actually unhealthy and the model correctly classified them unhealthy as well.
- False Positive (FP) = 3; meaning 3 plants were actually unhealthy, but the model incorrectly classified them as healthy.
- False Negative (FN) = 2; meaning 2 plants were actually healthy, but the model incorrectly classified them as unhealthy.

The performance metrics based on CM are also computed using the formulae given below and are summarized in Table 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Table 7- 1. Performance metrics of foliage color detection model.

Class	N (truth)	N (classified)	Accuracy	Precision	Recall	F1-Score
Q=1	60	61	0.95	0.95	0.97	0.96
Q=0	60	59	0.95	0.97	0.95	0.96
Average	-	-	0.95	0.96	0.96	0.96

In Table 1, N (truth) tells the number of actual cases in a particular class, and N(classified) tells the number of predicted cases belonging to a class. Table 1 shows that the model has achieved an average accuracy of 95%, precision of 96%, recall of

96%, and F1-Score of 96%. The model has correctly classified 115 cases out of a total of 120 cases. Figure 7-13 shows an example of correctly classified cases.

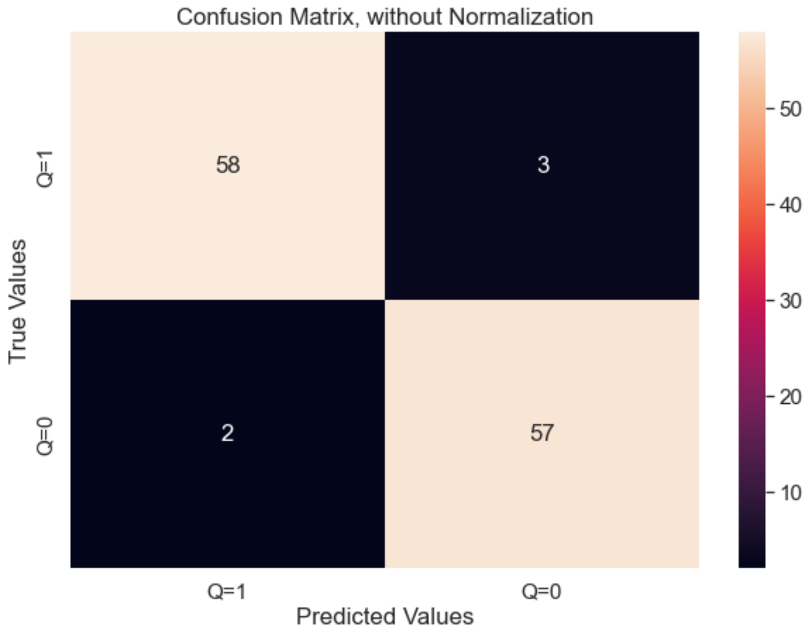


Figure 7- 12. Confusion matrix for quality assessment system.

Figure 7- 13. Example of correctly classified cases.

To further investigate the performance of the proposed methodology, it is compared with existing vision-based methods mentioned in section 7.2. These methods are implemented on the dataset prepared in this study, and their performance is evaluated using the metrics based on CM, which are presented in Table 7-2. The results show that the proposed method has outperformed the similar existing methods, achieving an average accuracy of 96%, precision, recall, and F1-score of 96%. The method proposed by Sharad et al. has shown appreciable performance when implemented on the dataset prepared in this study by achieving an average accuracy of 94%, precision of 94%, recall of 95%, and F1-score of 94.45% [382]. Whereas, with their apple leaf dataset, they have achieved an accuracy of 98.63%.

Table 7- 2. Performance metrics of existing methods.

Methods	Techniques and parameters used	Average Accuracy	Average Precision	Average Recall	Average F1-Score
1 [378]	SVM (support vector machine)	0.91	0.92	0.93	0.925

	and a* (CIELAB color space), G (green from RGB color space), and H (hue from HSV color space)				
2 [379]	Otsu's method and k-means clustering technique	0.92	0.93	0.93	0.93
3 [380]	HSV (hue, saturation, and value) color space and decision tree method	0.89	0.91	0.90	0.905
4 [381]	Otsu' method, SVM, k-NN (k-nearest neighbor) and MLP (multi-layer perceptron)	0.90	0.91	0.91	0.91
5 [382]	L*a*b* color histogram, k-NN, and random forest	0.94	0.95	0.94	0.945
6	Proposed model	0.95	0.96	0.96	0.96

The final model is then deployed in the aquaponic facility through a cloud-based application. This time, instead of manually taking the images, four ELP 1080P webcams (2.8–12 mm HD Varifocal Lens) are installed at a distance of 40 cm from channels for image acquisition. Each camera is programmed through Raspberry Pi 4 (Model B Rev 1) controller to take one image per day at 9:00 am, which along with sensor values from WSM are wirelessly uploaded to 'IoT enabled Aquaponics Dashboard' developed by authors in previous work (Chapter 5). The images and sensor data are available on the cloud as well as locally, and the app developed in this study can access them. The ontology model discussed in section 7.3.4 is also integrated with the proposed model and deployed on a cloud-based application. Once the health status of the lettuce crop is identified as 'Yellow Foliage – Leaf Chlorosis', the potential causes are automatically extracted from the ontology model and displayed on the application panel. Figures from 7-6 to 7-9 show an example of the working of the proposed method and application forlettuce when its foliage is detected to be yellow. The primary causes of lettuce foliage chlorosis could be inadequate environmental conditions (humidity, air temperature) or poor water quality (inadequate pH or EC) or nutrient deficiency, etc. By analyzing sensor data and possible causes of leaf chlorosis, it is possible to reach the specific cause of the problem. For instance, if sensor data shows that all parameters are within their optimal ranges, then the problem could be related to nutrient delivery or the design of the system. A reasonable treatment can be suggested after problem identification.

The proposed model is developed using open-source frameworks, and hence it can easily be expanded or adjusted as per the requirement by adjusting the threshold values. The significance of the model is that it is fully automated and offers a non-destructive, low cost and reliable approach to identifying leaf chlorosis and determining the quality of lettuce plants along with possible causes. In contrast to computer vision and machine learning-based models, the proposed methodology requires less data.

7.5. Conclusions and future work

In this study, the major problem of lettuce foliage chlorosis in the aquaponic system is discussed. The image processing approach ‘HSV Color Segmentation’ is used to segment the lettuce images obtained from various resources. The segmented images are divided into two classes ‘Green Foliage-No Discoloration’ and ‘Yellow Foliage-Leaf Discoloration’. Then, the foliage color detection model is developed, and a quality indicator is defined for evaluating the health of the lettuce crop. The model is validated, and its performance is evaluated. The results show that the overall accuracy of the model is 95%. A cloud-based application is also developed for the aquaponic facility where the proposed model is deployed.

In future work, the current limitations of the model will be addressed by considering other crops as well as multiple objects in an image. Moreover, further work will be done to identify the causes and possible treatments of leaf discoloration.

Chapter 8 Estimation of morphological traits of foliage and effective plant spacing in an NFT-based aquaponic system

8.1. Introduction

Crop quality is evaluated by a number of indicators, as discussed in Chapters 1, 6 and 7. One of the indicators is crop morphological traits such as length, width, area, and perimeter, which are used to assess the health status as well as the market value of the crops. Hence, it is vital to monitor these parameters throughout the plantation cycle [383]. Moreover, plant spacing directly impacts crop quality, defined as the distance between the growing sites of two consecutive plants. In traditional agriculture, crops compete with each other for resources such as solar radiation, nutrients, and moisture uptake to gain energy for their growth. They require reasonable root space and vegetative space [384]. The inadequate plant spacing may lead to a number of problems. For instance, the plants sited closely produce fewer leaves, flowers, and seeds, which causes reductions in final crop yields. Moreover, overcrowded plants are also susceptible to potential diseases, foliage damage as crops mature, and invasion by unwanted pathogens [384]. Many disease agents require a humid environment to develop and in crowded plantations, reduced airflow prevents moisture from evaporating from leaf surfaces, increasing the likelihood of diseases. Similarly, excessive plant spacing can also be a problem, as it hinders the pollination process. Unlike traditional agriculture, the aspect of plant spacing is different in the NFT-based aquaponic system. The crop growing area (hydroponics) in the NFT system is a combination of enclosed channels consisting of circular or squared-shaped pockets known as plant sites where plants reside in small plastic cups allowing their roots to access water and absorb nutrients enriched effluent from aquaculture. Plant spacing in NFT systems refers to either distance between two plants on the same channel or the distance between plants on adjacent channels. The plant spacing attribute on the same channel is normally fixed and is designed considering the full-grown size of the plant before the actual plantation. The plant spacing attribute between channels is also kept fixed in open-air farms but in indoor NFT-based systems, it can be varied depending on the crop type to efficiently utilize the limited space. In NFT systems, all the nutrients are directly delivered to the crops' root system; therefore, there is no competition for resources for the root system. However, vegetative space still requires special attention as crops spend the most energy on vegetation to absorb more light [15]. Plant spacing varies as a function of crop species and their morphological traits such as length, width, area, and perimeter [385]. Therefore, the optimum plant spacing must be maintained according to crop type and its morphological characteristics to achieve maximum yields. Additionally, the crop growth cycle consists of various phases, from seedling to vegetative, and in each phase, the area of crop foliage changes as it grows over time — requiring different plant spacing and illumination conditions [386].

8.1.1. Research motivation and contribution

The aforementioned challenges pose a need for a self-adaptive aquaponic system that determines the crop morphological traits, determines the distance between two plants and adjusts the spacing between channels. Traditionally, manual methods that require a high level of expertise and advanced equipment were used to determine morphological traits. These methods produce accurate results, but they are costly, labor-intensive, and time-consuming [387]. Numerous semi-automatic tools are developed to accelerate plant phenotyping, such as LeafJ, Easy Leaf Area, and TraitEx [388]–[390]. But these tools require pre-processing of the input images for utilizing multiple automation degrees [387]. To overcome the stated bottleneck, this study aims to propose an approach to automatically estimate morphological traits (foliage area, length, and width), the distance between plants, and effective plant spacing between adjacent channels. Based on this information, the aquaponic system can adapt itself by adjusting the position of grow channels. The proposed approach is deployed on a cloud-based application developed in Chapter 7 and integrated with ontology model proposed in Chapter 3.

8.1.2. Chapter organization

The rest of this chapter is organized as follows. Section 2 shows the related work. Section 3 discusses the proposed approach. Section 4 presents the results and findings. Section 5 provides discussion and future work and finally, Section 6 presents the conclusion.

8.2. Related work

With developments in deep learning and computer vision techniques, several methods have been developed and instantly grown in different visual recognition tasks such as the estimation of morphological traits of crops. A review of some of the latest and more relevant methods is presented here. Weaver et al. have proposed a tool 'LeafMachine' based on CNN (convolutional neural network) and SVM (support vector machine) to measure the leaf morphological traits from digitized herbarium specimen images autonomously [338]. Triki et al. used the same dataset and proposed a new and enhanced approach, 'Deep Leaf' based on Mask-RCNN (region-based convolutional neural network), to determine the length, width, area, and perimeter of leaves [313]. Hirigoyen et al. developed a machine learning model using SVM and RF (random forest) techniques to determine the leaf area index in Eucalyptus plantations [391]. Lu et al. proposed a Mask R-CNN-based model to determine the growth rate of lettuce crops as a function of leaf area and time in a hydroponics system. Juyal et al. proposed a method to estimate the length and width of trees for calculating the overall volume using Mask-RCNN [392]. Reyes et al. proposed a methodology to determine the size of crops (height, width, depth, side view area, top view area) for assessing crops' growth rate and fresh weight using Mask-RCNN [177]. Even though the methods above have made a significant contribution to the research community, the analysis shows that none of the studies has focused on using these traits as a key feature to assess plant spacing. Hence, to complement the existing efforts, this study proposed a new methodology to determine the morphological traits of lettuce crops grown in aquaponics facilities and assess the plant spacing between grow channels. Additionally, based on the capability of instance segmentation, Mask-RCNN is used to estimate the morphological attributes in this study [393], [383].

8.3. Research Methodology

An automated system is developed to monitor the morphological traits of the lettuce crop and estimate the plant spacing in aquaponics facilities using deep learning techniques. The overall methodology is divided into three sequential modules: i) Dataset preparation and pre-processing ii) Model development and parameter estimation, and iii) Model training and implementation.

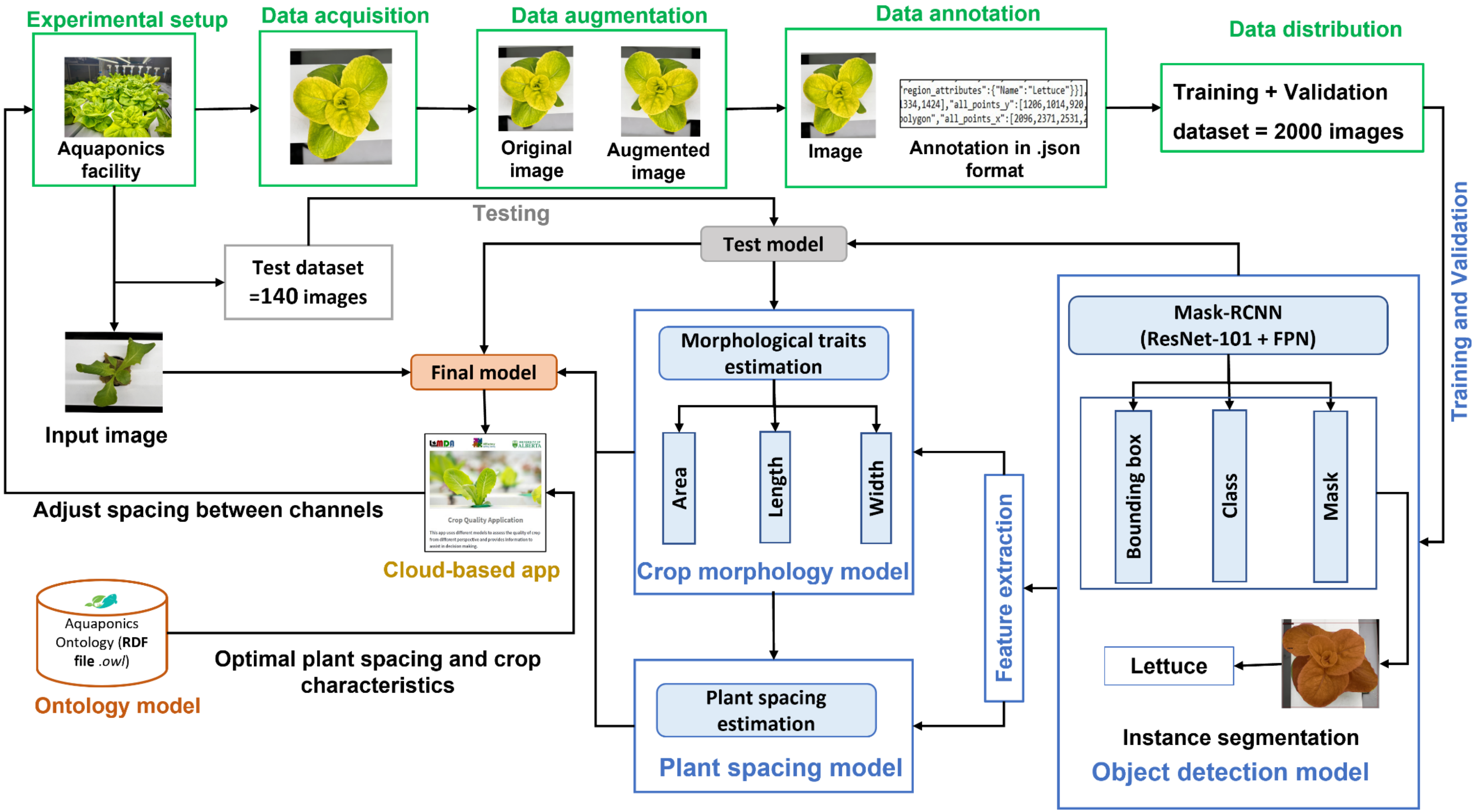


Figure 8- 1. Pipeline for estimating the foliage area for effective plant spacing.

Figure 8-1 shows the complete pipeline of this approach, and details of all stages are explained in the following subsections.

8.3.1. Dataset preparation and preprocessing

8.3.1.1. Data acquisition and manual measurements

For this study, iPhone X12 Pro Max is used to capture images of thirty lettuce plants grown in an NFT-based aquaponic facility situated at the University of Alberta, Canada. Fifty images of $4{,}032 \times 3{,}024$ pixels are captured every day, half (25 images) at 9:00 am and a half at 6:00 pm from the top while keeping the distance between camera and channel at a value of 40cm throughout the plantation cycle, i-e., five weeks. Each image contains two plants planted in adjacent channels. There are six grow channels, each having five plant sites. The channels are horizontally stacked, so the total number of plants in all the rows of channels is six. In total, 1750 images were collected over the period of 5 weeks and are saved in JPG format. Figure 8-2 shows an example of some of the images of lettuce from different growth stages.

Figure 8- 2. Sample images from an aquaponic facility.

The manual measurements of morphological traits such as length, width, and height of 30 plants are also recorded twice a day at 9:00 am and 6:00 pm for five weeks using a caliper. For the ground truth value of the foliage area, the number of pixels is counted manually by selecting the area of interest in Adobe Photoshop, which is then converted to $cm^{2.}$ The area is also recorded twice a day at the given timings for five weeks. As the plants grow, their area, length, and width also increase, reducing the distance between the two plants. This distance needs to be measured throughout the plantation cycle to determine the effective plant spacing and adjust the spacing between plants. The actual distance is recorded for all the plants by taking the manual measurements twice daily at 9:00 am and 6:00 pm for five weeks using a scale. Fifty distance values are calculated daily for 30 plants growing on adjacent channels. Moreover, the plant spacing is also decided by using the recorded distance and initial plant spacing value. All the manual measurements are recorded in a common excel file.

8.3.1.2. Data augmentation

Next, the image augmentation process is performed to increase the dataset, and avoid overfitting. It enhances the reliability of the segmentation process despite the location and orientation of objects in the image by generating new images from existing images. This study uses Albumentations, a Python library, for fast and flexible image augmentations [359]. The different augmentation techniques applied are horizontal flip, vertical flip, 90° rotation, and glass noise. In total, 250 images were selected randomly for the augmentation, which created 250 new images — increasing the length of the dataset to 2000 images. Figure 8-3 shows an example of augmentation.

Figure 8- 3. Example of an augmented image.

To reduce the computation time of the training and testing of the model, the sizes of all the images were scaled to 640×480. A total of 2000 images are used for the model training and parameter optimization of Mask R-CNN, with 80% as the training set and 20% at the validation set. After the training, the model is tested with a new set of images captured from a new batch of plants to evaluate the performance of the trained model.

8.3.1.3. Data annotation

Data annotation is one of the vital steps for the successful development of object detection models. The process is manual and involves labeling the desired objects in an image with a label or tag that refers to a particular class. The labeled data is used during the training of the model. There is a number of open-source annotation tools, but in this study, VGG Image Annotator (VIA) is used [394]. This graphical annotation tool supports a variety of deep learning algorithms. For instance, it generates annotations in JSON and CSV formats. In this study, Mask-RCNN is used, which accepts annotations in JSON format. The relevant regions of the image are labeled, and the remaining region defaults to the background. Figure 8-4 shows an example of ground truth bounding boxes and masks obtained after data annotation.

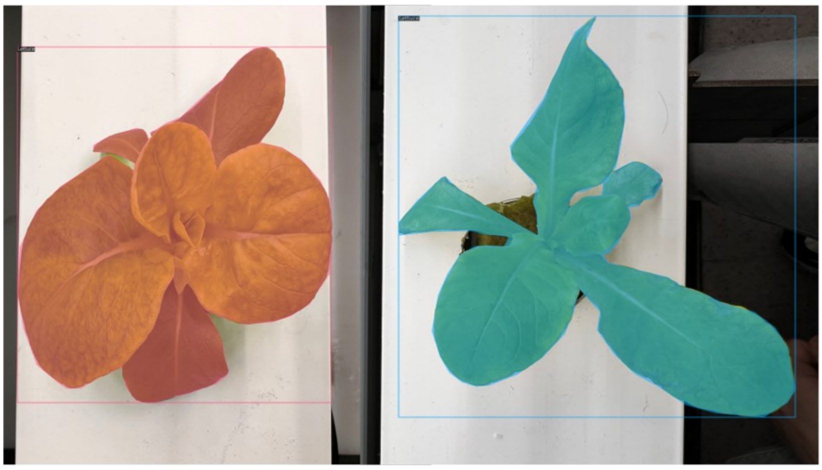


Figure 8- 4. Bounding boxes and masks after data annotation.

8.3.2. Object detection and instance segmentation

After data collection, object detection and instance segmentation are performed on crop images to achieve class, mask, and bounding box values for lettuce foliage. In this study, Mask R-CNN is used, which is the state-of-the-art method in the field of object detection. It is an extension of Faster R-CNN with an extra branch at the end of the model that activates instance segmentation for each output proposal box using a fully connected layer (FC) in parallel with the identification and localization tasks [395]. Instance segmentation is a computer vision task for detecting and localizing an object in an image through the identification of boundaries at a detailed pixel level [396]. The Mask R-CNN framework has three primary steps shown in Figure 8-5. In the first step, feature maps are extracted from input images through a backbone network, which in the second step, are sent to the region proposal network (RPN) to generate regions of interest (RoIs). Finally, in the third step, the generated RoIs are mapped to extract the corresponding target features in the shared feature maps, which are then sent to a multi-branch prediction network consisting of FC, regression layer and FCN (fully convolutional networks) to generate three outputs: classification scores, bounding boxes, and segmentation masks.

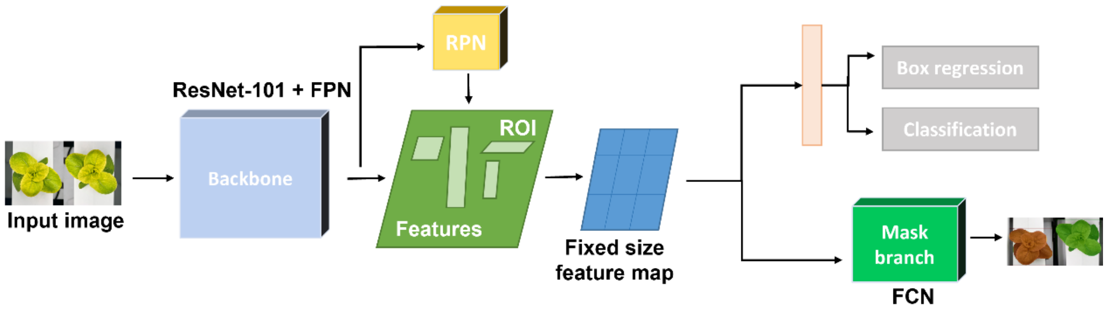


Figure 8- 5. Architecture of Mask-RCNN, adapted from [395] .

Different deep neural networks have been designed over the years, such as AlexNet, ZF, VGG, GoogleNet, Inception, Xception, and ResNet, each having various pros and cons [383]. In this study, ResNet-101 (101 layers deep) is used as the backbone network for feature extraction because it tackles the issues of gradient disappearance and training degradation [362]. The architecture of ResNet is discussed in Chapter 6. Additionally, to better represent the lettuce in images on multiple scales, the backbone network is extended by adding a feature pyramid network (FPN) which is especially effective for the detection of small targets. The convolution feature maps from the backbone network are used as input for the RPN network. To generate the regions of interest (RoIs), nine anchors with different area scales and length-width ratios are used to slide on the feature maps. The corresponding features of each RoI are extracted from the feature maps, and RoIAlign is applied to adjust the dimension of each RoI to meet the input requirements of next layers. The adjusted RoI features are then sent to multi-branch prediction network consisting of three prediction branches: the FC layer for classification prediction, the regression layer for coordinate corrections of bounding boxes, and the fully convolutional network (FCN) for instance segmentation to generate the target mask.

8.3.2.1.Mask-RCNN training

NVIDIA GeForce RTX 3090 is used for training the Mask-RCNN. A total of 2000 images are split into 80% for the training set and 20% for the validation set. The Mask-RCNN is implemented in Detectron2 — Facebook AI, Research's next-generation library, written in PyTorch that provides state-of-the-art detection and segmentation algorithms. Since the dataset is small, the pre-trained version of Mask-CNN (trained on the COCO dataset) from the model zoo of the Detectron2[10] 'Mask-RCNN with ResNet-101 + FPN' is applied using the transfer learning approach. COCO is a huge dataset with 328k images, including 91 categories for object detection and image segmentation. The general features of all categories are extracted from COCO using the pre-trained model [397]. The parameters of the model can be adjusted to a better state based on the pre-trained model, regardless of the size of the dataset. Residual networks, such as ResNet-44/47/50/71/101, are used as a backbone network of Mask R-CNN, which are different from each other in the layer number of convolutional modules [397]. Among all the architectures, ResNet-101 has shown the highest detection accuracy and is, therefore, chosen as the backbone network for Mask-RCNN in this study. The model is trained for 600 iterations with the image input batch size given as 32. The initial learning rate is kept at 0.001 for the first 100 iterations and is then adjusted per 100 iterations with an adjustment factor of 0.95. The category scores, bounding boxes, and masks of lettuce foliage for each input image are then obtained as the outputs of the model. The training process is completed in around 1 hour for 600 iterations, and the model loss function achieved a convergence state.

The total loss of the proposed approach consists of two parts: the loss of classification and regression operations by RPN, and the training loss in the multi-branch predictive network, and can be calculated by using the formula given below [395].

$$L_{final} = L_{RPN} + L_{multi\_branch}$$

Where L_{final} represents the total loss, L_{RPN} represents the training loss of the RPN (anchors classification loss (SoftMax Loss) and bounding box regression loss (SmoothL1 Loss)), and L_{multi_branch} represents the training loss due to the three

[10] https://ai.facebook.com/tools/detectron2/

branch structures (SoftMax Loss, SmoothL1 Loss, and Mask Loss). L_{RPN} and L_{multi_branch} are calculated as follows:

$$L_{RPN} = \frac{1}{N_{cls1}} \sum\nolimits_i L_{cls}(p_i, p_i^*) + \lambda_1 \frac{1}{N_{reg1}} \sum\nolimits_i p_i^* L_{reg}(t_i, t_i^*)$$

$$\begin{aligned} L_{multi\_branch} &= L(p_i, p_i^*, t_i, t_i^*, s_i, s_i^*) \\ &= \frac{1}{N_{cls2}} \sum\nolimits_i L_{cls}(p_i, p_i^*) \\ &+ \lambda_2 \frac{1}{N_{reg2}} \sum\nolimits_i p_i^* L_{reg}(t_i, t_i^*) + \gamma_2 \frac{1}{N_{mask}} \sum\nolimits_i L_{mask}(s_i, s_i^*) \end{aligned}$$

Where constant N_* represents the number of corresponding anchors or bounding boxes. The hyperparameters λ^* and γ^* balance the training losses of the regression and mask branch. Classification loss L_{cls}, regression loss L_{reg}, and mask loss L_{mask} are derived from the following formulas:

$$L_{reg}(t_i^*, t_i) = smooth_{L1}(t_i^* - t_i), smooth_{L1}(x) = \begin{cases} 0.5x^2 \ if \ |x| < 1 \\ |x| - 0.5 \ otherwise \end{cases}$$

$$L_{reg}(p_i^*, p_i) = -logp^* p$$

$$L_{reg}(s_i^*, s_i) = -(s^* \log(s) + (1 - s^*) \log(1 - s))$$

Where p_i represents the classification probability of anchor i, and p_i^* represents the ground-truth label probability of anchor i; The variable t_i represents the difference between the prediction bounding box and the ground-truth label box in four parameter vectors (the horizontal, vertical coordinate value of the center point in the bounding box; the width and height of the bounding box), and t_i^* indicates the difference between the ground-truth label box and the positive anchor; s^* and s represent the mask binary matrices from the prediction and ground-truth label, respectively.

The loss function value and accuracy per iteration of the model for 600 iterations are shown in Figure 8-6. The loss function shows a downward trend during the training process, as seen in Figure 8-6, which means that the prediction loss deviation is gradually decreasing by updating the loss function of the small sample batches during the optimization process. The loss function values for both the training set and validation set are reduced to less than 0.2 and tend to be stable when the number of iterations is more than 550. This indicates that the training of the model runs well, with a detection accuracy of lettuce foliage of more than 0.98.

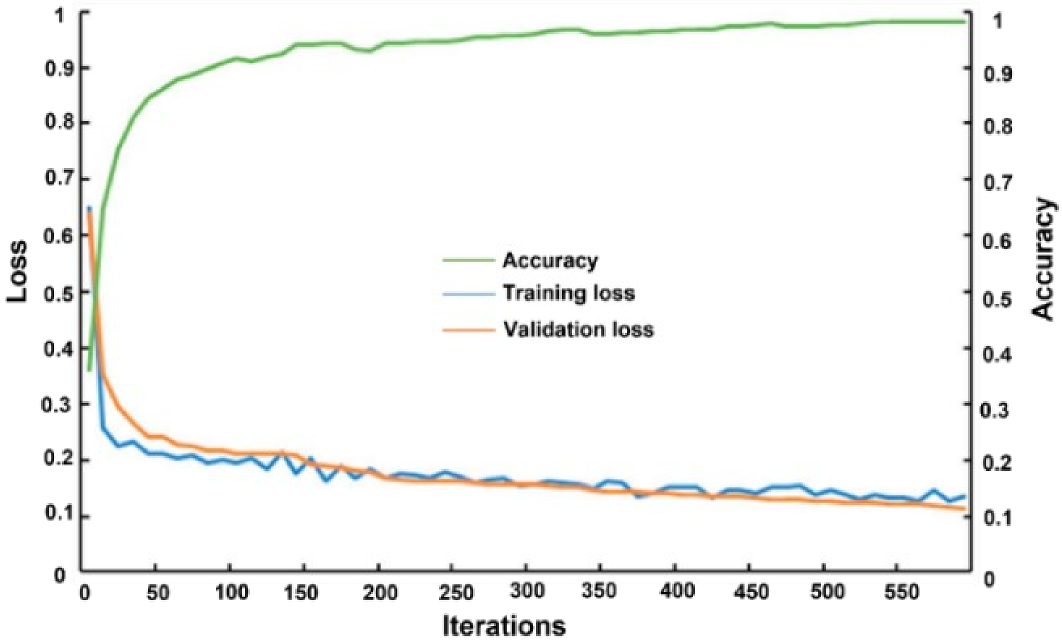


Figure 8- 6. Training losses, validation losses and accuracy per iteration.

8.3.3. Crop morphological traits estimation model

From the instance segmentation process, the predicted mask and bounding box of each instance are retrieved to determine lettuce morphological traits such as foliage area, length, and width. In this case, there are two masks, and two bounding boxes as each image consists of two lettuce foliage. The foliage area of two lettuce plants is calculated by extracting features from their respective predicted masks. The masks are a set of matrices that contain pixels belonging to the area of the segmented object (area of interest), which in this case are two lettuce foliage. These pixels are retrieved from the prediction, and further processing is done to give this data useful meaning. The distance between the camera and the object affects the pixel count of the image. The closer the camera, the greater the number of pixels of an object in an image and vice versa. Hence, in this study, the distance between the camera and channel is kept fixed at a value of 40cm while taking the images and is denoted by (D). As we know, the height of the lettuce increases throughout the plantation cycle and affects the pixel count. . To calculate the foliage area (A_f), the height (h) of the crop is also taken into account. At the end of the plantation cycle, a boxplot is created shown in Figure C.1 of appendix C for the heights of 30 lettuce plants recorded manually and a scatter plot is created shown in Figure C.2 of appendix C using median values of plants' heights to derive a linear regression relationship between height (h) and a number of days (x), which is represented in equation 1.

$$h = 0.2521x + 2.9641 \quad (1)$$

To adequately measure the morphological traits, the relationship between real-world metrics, such as cm (centimeter) and actual pixel values on the image should be identified. Triki et al. used scale bar objects to determine this relationship [387]. In this study, the width of the channel (w_c) which is known to be 10cm is used for this purpose. Equation 2 shows the relationship between (w_c) and (p_c). At a constant distance of D the relationships (k) and (k') between pixel count (p_c) and channel width (w_c) are given in equation 3 and 4.

$$w_c\ (cm) \equiv p_c\ (pixels) \quad (2)$$

$$k = \frac{p_c}{w_c} \quad \left(\frac{pixels}{cm}\right) \quad (3)$$

$$k' = \frac{{p_c}^2}{{w_c}^2} \quad \left(\frac{pixels}{cm^{\ 2}}\right) \quad (4)$$

Let (p_m) is the pixel count of the predicted mask, which is dependent on the height h of the plant. The height of the plant changes throughout the plantation cycle and hence affects the pixel count of the predicted mask. To compute foliage area (A_f), equation 5 is developed.

$$A_f = \frac{p_m \times h}{k' \times D} \quad (5)$$

Next, bounding boxes are retrieved from the model in the form of coordinates of opposite rectangle corners (top left (x_1, y_1) and bottom right (x_2, y_2). These coordinates are used to calculate the approximate width (W_f) and length (L_f) of foliage shown in Figure 8-7, using equations 6 and 7, respectively.

$$W_f = x_2 - x_1 \quad (pixels) \quad (6)$$

$$L_f = y_1 - y_2 \quad (pixels) \quad (7)$$

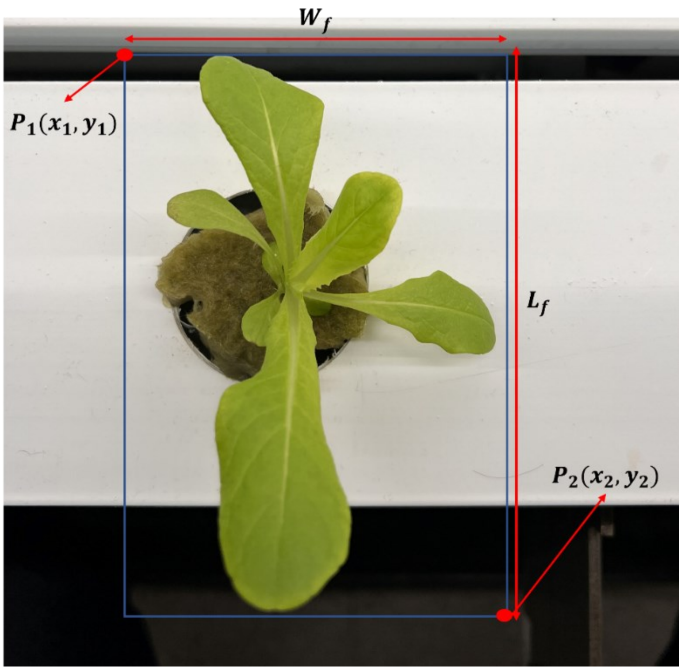


Figure 8- 7. Dimensional characteristics of foliage.

The equations for the width and length of foliage in real-world metric units (centimeter) are given as,

$$W_r = \frac{W_f \times h}{k \times D} \quad (cm) \qquad (8)$$

$$L_r = \frac{L_f \times h}{k \times D} \quad (cm) \qquad (9)$$

The above process is performed for two lettuce foliage in a segmented image.

8.3.4. Mathematical model for plant spacing calculation

Figure 8-8 shows the most common configurations of plant spacing used for NFT channels along with several dimensional characteristics. In configuration 1, the plant sites on two adjacent channels are in-line with each other and in configuration 2, the plant sites are at an angle. In this study, configuration 1 is considered for a case study but the proposed model can also be applicable to configuration 2. The plant spacing (S) is to be maintained on individual grow channels as well as between the channels (S') to avoid foliage occlusions and overlapping which limit crop growth and affect crop quality. Generally, while designing the aquaponic system, the plant spacing on the channel, (S) and between channels (S') is kept constant. The latter parameter can be controlled to make a self-adaptive aquaponics. From Figure 8-8, (S') is the distance from the center of the plant site of one channel to the center of the plant site of the adjacent channel and is given in equation 10.

$$S' = \frac{w_{c1}}{2} + \frac{w_{c2}}{2} + s \qquad (10)$$

Where ($w_{c1} = w_{c2}$), (w_{c1}) is the width of channel 1, (w_{c2}) is the width of channel 2 and (s) is the distance between two adjacent channels. The width of channels is constant whereas (s) is a dynamic parameter. By changing (s), the horizontal channel spacing i-e., the distance between channels can be changed.

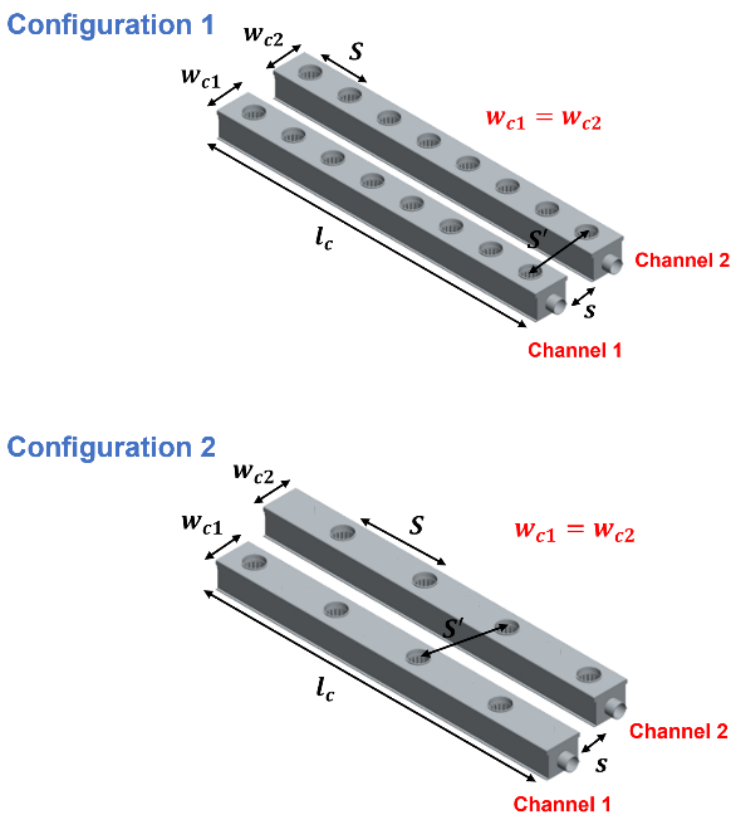


Figure 8- 8. Dimensional characteristics of grow channels with respect to two configurations.

To compute the distance between plants automatically, the initial and new values of (S') and (s) are computed. Let (S'_{n-1}) and (s_{n-1}) refers to initial or previous values for (S') and (s), respectively, and (S'_n) and (s_n) refers to new values for (S') and (s), respectively. Equations 11 and 12 represent the initial and new values of (S').

$$S'_{n-1} = \frac{w_{c1}}{2} + \frac{w_{c2}}{2} + s_{n-1} \qquad (11)$$

$$S'_{n} = \frac{w_{c1}}{2} + \frac{w_{c2}}{2} + s_{n} \qquad (12)$$

To compute (s_n), an incremental parameter (s_i) is defined, which determines the variation in plant spacing due to changes in foliage morphological traits. The incremental parameter (s_i) is calculated using the Euclidean distance (d) between two bounding boxes achieved through the prediction process shown in Figure 8-9 and (S'_{n-1}). The parameter (s_i) will only increment if (d) will be less than (S'_{n-1}) i-e., $if\ d < S'_{n-1}$. The value of (d), on the other hand, is dependent on the foliage area and length, which decreases as the area (A_f) and length (L_f) of foliage increases. Equations 13, 14 and 15 provide the relationship for (d), (s_n) and (s_i), respectively.

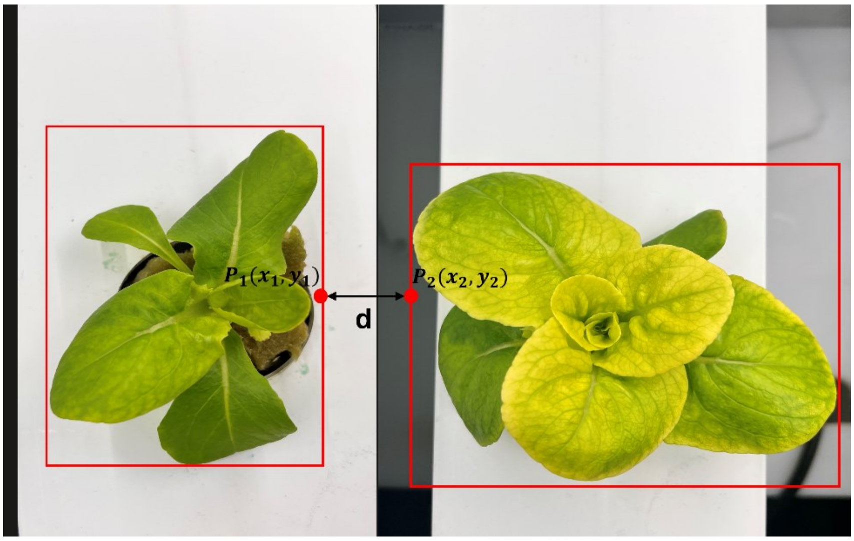


Figure 8- 9. Calculation of distance between plants using bounding boxes.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (13)$$

$$s_n = s_{n-1} + s_i \qquad (14)$$

$$s_i = S'_{n-1} - d \qquad (15)$$

The new values for (S') are then obtained by incrementing the values of (s) using the above process for all plants' pairs in images and converted to metric units (cm) using Equation 2 The updated values of (s) can be used to estimate plant population (PD) using the equation 16 proposed by authors in previous work (Chapter 4).

$$P_D = \frac{N_{PSC} \times N_C}{L \times ((N_C \times \mathrm{W}) + ((N_C - 1) \times H_{CS})} \quad (16)$$

Where (N_{PSC}) is the number of plant sites (circular or squared-shaped pockets) per channel, (N_C) is the total number of channels, L is the length of each channel and is equivalent to (l_c) shown in Figure 7, (W) is the width of each channel and is equivalent to (w_{cj}) shown in Figure 7 (j=1,2,3…..n), and (H_{CS}) is horizontal channel spacing and equivalent to (s).

8.3.5. Ontology model

The complete development and details of all the concepts and instances of the ontology model 'AquaONT' developed by authors in previous work are available in Chapter 3. AquaONT is a unified ontology model that represents and stores the essential knowledge of an aquaponics 4.0 system. It consists of six concepts: Consumer_Product, Ambient_Environment, Contextual_Data, Production_System, Product_Quality, and Production_Facility. In this study, two classes, 'Consumer Product' and 'Production_System' are used for knowledge extraction. The 'Consumer Product' class provides an abstract view of the type, growth status, and growth parameters of ready-to-harvest crops in an aquaponic system. Whereas the 'Production_System' class provides knowledge on design parameters of the grow channels such as width, length and depth of channel and size of plant site on channel [6]. Figure 8-10 shows the hierarchical architecture of the 'Consumer Product' and 'Product Quality' classes with their instances for the 'Lettuce' crop in Protégé[11] (an open-source ontology editor and framework developed at Stanford University) environment. The length and width of the channels can be extracted from ontology to be used to determine the plant spacing as well as the production capacity of the aquaponic facility using equation 16.

[11] https://protege.stanford.edu/products.php#desktop-protege

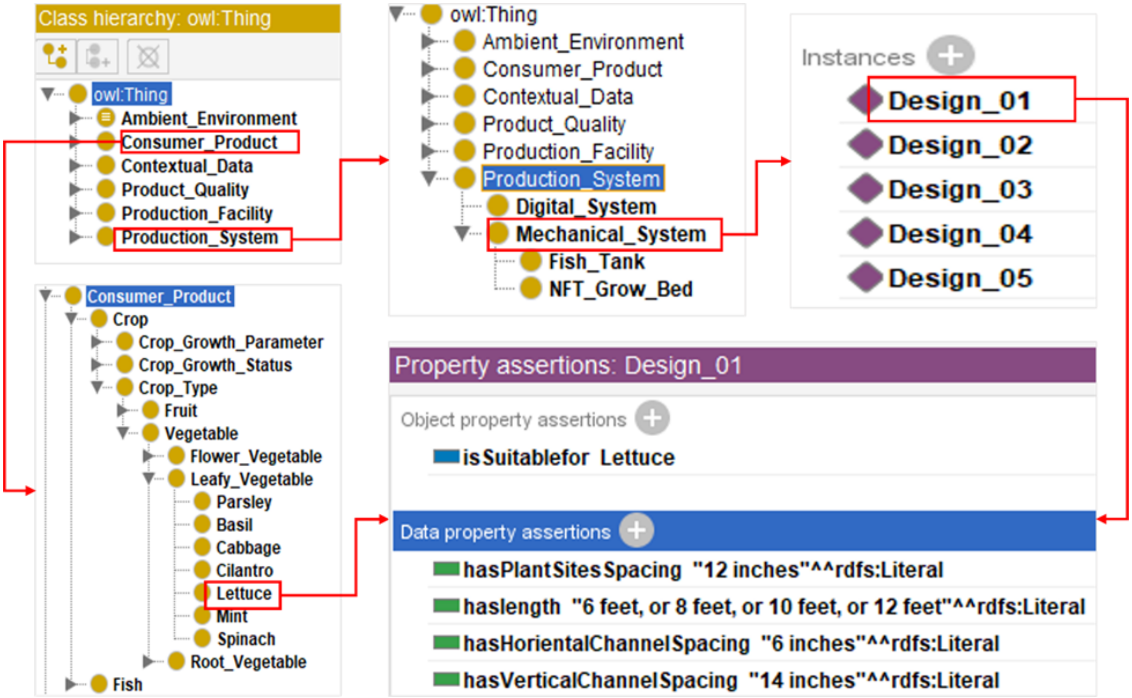


Figure 8- 10. Ontology model showing the instances of 'Consumer_Product' and 'Production_System' classes.

8.3.6. Cloud-based application

A cloud-based application is developed using Streamlit, where the final versions of both models developed in sections 8.3.3 and 8.3.4 are deployed The layout of the application is shown in Figure 8-11. It consists of three tabs: i) Select model, ii) Upload image, and iii) Determine morphological traits and plant spacing. The first and second tabs are user inputs where the model is selected, and the image is uploaded respectively. The third tab activates the crop morphological model and plant spacing model respectively. The ontology model is also deployed on a cloud-based application through the Owlready2 library. Once, the morphological traits are estimated and effective plant spacing is determined, the length of the channel is extracted from the ontology model to determine plant population and overall yield. Moreover, crop quality can also be assessed by comparing the predicted morphological traits of lettuce and standard values stored in the ontology model. This type of application is useful as it provides access to remotely monitor and control the production facility.

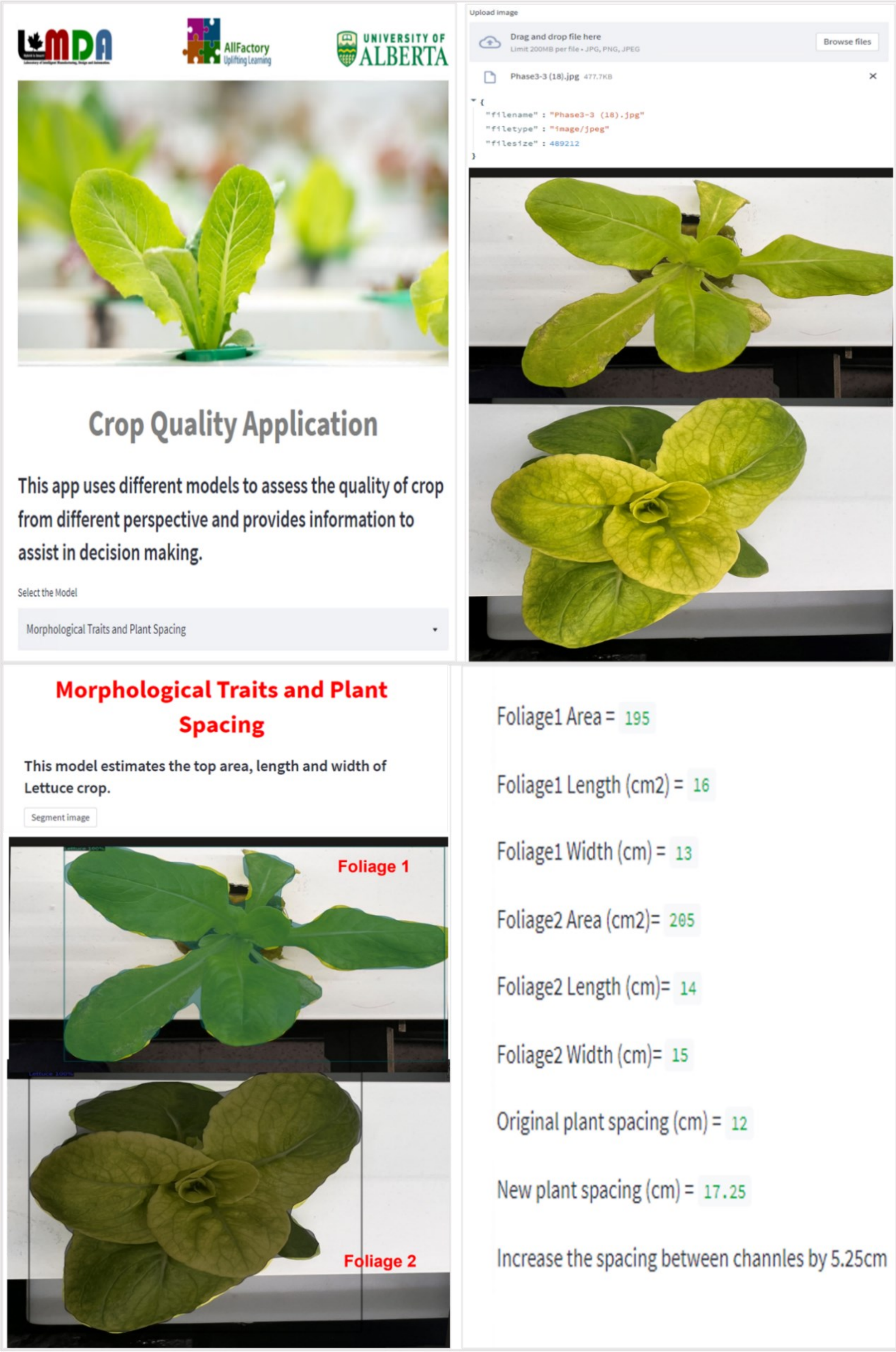


Figure 8- 11. Cloud-based application for estimation of morphological traits of lettuce and plant spacing.

8.4. Results and findings

To validate the research methodology, a new batch of plants are grown. The experimental setup used to capture new dataset and validation results and findings are presented in the following subsections

8.4.1. Experimental setup

The experimental setup is built in the Allfactory 4.0, an NFT-based aquaponic facility situated at the University of Alberta, Canada, which focuses on smart indoor farming [246]. The aquaponic system is divided into five crop growth phases which represent the complete growth cycle of the crop. For this study, only phase 1 is considered, which consists of six horizontally stacked grow channels and each channel has five plant sites. The length and width of each channel are 125 cm and 10 cm respectively. The distance between the center of plant sites (circular pockets) of two consecutive channels is 12 cm. A fresh batch of lettuce crop is grown for which fifty seeds of Little gem lettuce (Lactuca sativa L.) are placed in growth chambers with an ambient temperature of 18°C, relative humidity of 70%, and illumination of a 12-hour (12 hours light / 12 hours dark) photoperiod [333]. Twenty-one days after sowing, 30 healthy lettuce seedlings are transplanted in Rockwool cubes and placed in the six NFT channels in phase 1. The seedings are placed in NFT-based hydroponic systems for a period of five weeks (plantation cycle), after which each lettuce is harvested. A wireless sensing module (WSM) consisting of five sensors (pH, temperature, humidity, water temperature, electroconductivity and light) is installed on the system to monitor the system and gather the sensor data. The complete development and working of WSM are detailed in a previous work by the authors (Chapter 5). Moreover, the images are captured in a similar fashion as discussed in section 3.1.1. But this time camera module consisting of four ELP 1080P webcams (2.8–12 mm HD Varifocal Lens) is used. All the webcams are attached at a distance of 40cm and are scheduled to take one picture per day for five weeks. In total, 8 plant samples are chosen for images and each image contains two plants. These plants are grown on adjacent channels and each day four pictures are taken automatically. At the end of the plantation cycle, there are 140 images which are then used for testing and evaluation of the proposed models. The actual measurements for morphological traits and distance between plants are recorded in a similar manner as discussed in section

8.3.1.1. The actual values of plant spacing (S') are also computed manually using the actual distance between plants and formulas mentioned in section 8.3.4.

8.4.2. Evaluation of trained Mask-RCNN model

To evaluate the detection accuracy of the trained Mask-RCNN model, the intersection over union (IoU) metric is used, which compares the predicted detection with ground truth. IoU is the ratio of the area of overlap between ground truth and predicted bounding boxes or masks divided by the area encompassed by both the predicted bounding box and the ground-truth bounding box [397]. For this purpose, the test dataset consisting of 140 images is used. The IoU is calculated as follows:

$$IoU = \frac{A_T \cap A_D}{A_T \cup A_D}$$

Where A_T and A_D represent the target bounding box of the ground truth and the detected bounding box from the model, respectively. All 140 test images have shown IoU coefficient of 0.9 or above, indicating that there is significant overlap between the two bounding boxes. Figure 8-12 shows an example of an image showing detection with IoU of above 0.9. The green bounding box is the ground truth, whereas the red bounding box is predicted by the model. All the targets in the lettuce image are aimed to be detected and marked with target category scores, bounding boxes, and instance segmentation masks. The detection performance of the Mask-RCNN is shown in Figure 8-13. The final model is then used to evaluate foliage morphological traits and plant spacing.

Figure 8- 12. Object detection by Mask-RCNN with IoU.

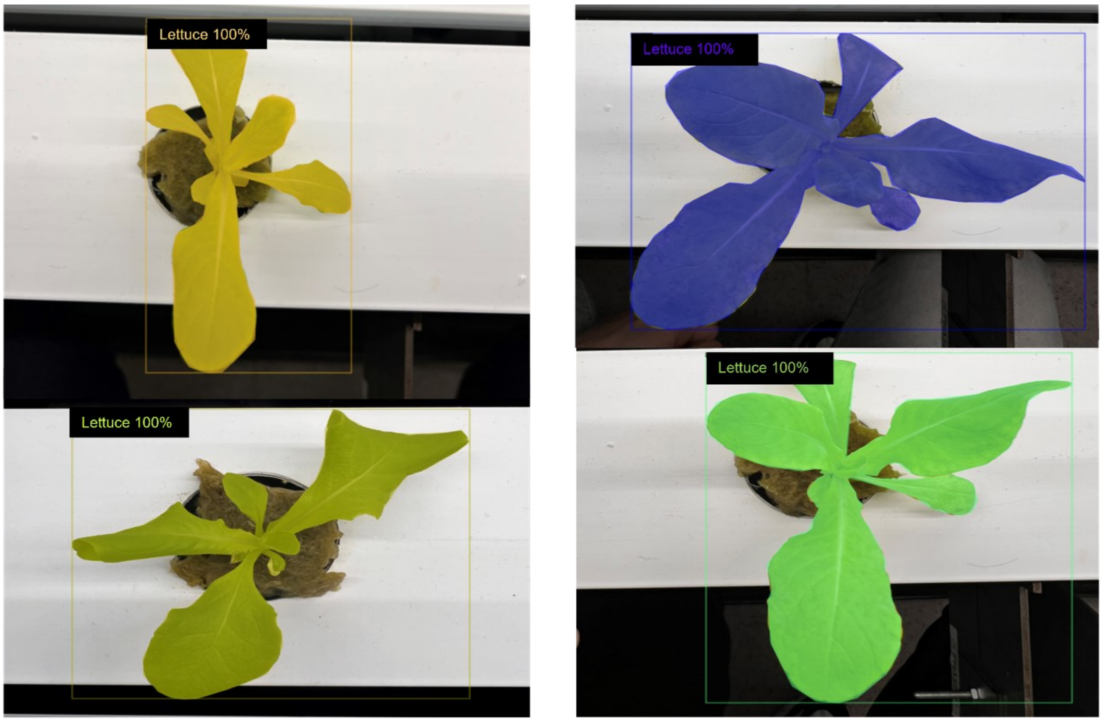


Figure 8- 13. Detection and instance segmentation of lettuce foliage.

8.4.3. Evaluation of morphological model

To evaluate the crop morphological model, the length, width, and area of lettuce foliage are computed from masks predicted by the trained model. The actual and the measured values for each morphological trait are compared for images in the test dataset. The increasing trend in the morphological traits is observed — indicating the growing behavior of the plants (increase in size). The estimation error between manual and masked dimensions for each trait per plan is then measured using root mean squared error (RMSE) using the formula given below.

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=0}^{k}(x_j - y_j)^2}$$

Where x_j is actual value, y_j is predicted value, Table 8-2 lists the RMSE calculated using the above equation for each measurement per plant. The test image set refers to a set of images taken throughout the plantation cycle for one pair of plants (Foliage 1 and Foliage 2). The test dataset in this study consists of four plant pairs. The minimum error is observed, which demonstrates that the trained model is reliable for the calculation of the leaf area, length, and width in real-world scenarios.

Table 8- 1. RMSE of manual and estimated measurements of morphological traits.

Test image set	Instance	Length, L_r(cm)	Width, W_r (cm)	Area, A_f (cm^2)
1	Foliage 1	1.1	1.8	3.35
	Foliage 2	1.3	1.1	2.9
2	Foliage 1	0.95	1.5	3.1
	Foliage 2	1.05	1.3	2.79
3	Foliage 1	1.27	0.99	3.2
	Foliage 2	1.37	1.21	3.12
4	Foliage 1	1.31	1.32	3.23
	Foliage 2	1.18	1.27	2.87

8.4.4. Evaluation of plant spacing model

To evaluate the plant spacing model, first, the distance between bounding boxes of two instances (two plants per image) is computed after instance segmentation is performed on images in the test dataset. The plant spacing is then computed following the process mentioned in section 3.4. The predicted values are then compared with actual measurements, which indicates that plant spacing increases with an increase in foliage area and length. Table 3 presents the RMSE for each test image set. The lower RMSE values indicate the reliability of the proposed model.

Table 8- 2. RMSE of manual and estimated measurements of plant spacing.

Test image set	Instance 1	Instance 2	Plant spacing, S' (cm)
1	Foliage 1	Foliage 2	1.1
2	Foliage 1	Foliage 2	0.95
3	Foliage 1	Foliage 2	1.27
4	Foliage 1	Foliage 2	1.31

8.5. Discussion and future work

The objective of this study is to automatically measure the crop morphological traits from the lettuce images acquired from the aquaponic facility and determine the effective plant spacing between grow channels. The idea is to develop an approach that can lead to a self-adaptive aquaponic system, where based on crop morphological attributes, the grow channels adjust their positions effectively to ensure high crop yield and quality by avoiding occlusions and foliage overlapping. The presented work has used the Mask-RCNN algorithm to enable a fine-grained detection of lettuce foliage within the images and the generation of a pixel-wise segmentation mask for each detected instance. The segmentation masks distinguish foliage from the background and provide a mechanism to identify non-foliage classes, such as channel, Rockwool cubes, etc., by considering their high heterogeneity in form and texture.

The manual measurement of morphological traits is time-consuming and labor-intensive. Even for an experienced agriculturalist, it would take a significant amount of time to measure all morphological traits. The proposed approach can automatically estimate the morphological traits of lettuce foliage that vary in size and shape. Plant spacing is one of the key features that impact crop growth. It is dependent on crop type as well as crop morphological traits. In this study, plant spacing is automatically measured for each segmented foliage using the mathematical model.

The final model is deployed on a cloud-based application and integrated with the ontology model. The ontology model provides information about crop characteristics and grow bed design parameters for a variety of crops grown in NFT-based aquaponics. The application acts as a decision support system, which analyses the results from the models, compares them with the relevant knowledge from the ontology model and suggests the final action by sending control signal to the aquaponic facility for automatically adjusting the grow channels based on the value of plant spacing predicted by the proposed model.

While promising results are achieved from the proposed models, there is still scope of improvement. For instance, only one crop is considered in this study for the estimation of morphological attributes and assessing the plant spacing between channels. Considering that, the potential solutions for estimating the morphological attributes of multiple crops will be investigated in future work. Subsequently, the dataset will be increased with more image variations and other leafy green crops as well as fruits, flowers, and non-flowering plants. Furthermore, the impact of morphological traits on other design parameters of the aquaponic facility will also be studied. A case study using this model is also developed to estimate the wet biomass of lettuce grown in deep water culture aquaponics which is presented in Appendix D.

8.6.Conclusions

In this study, an automatic tool is developed to predict the morphological traits of lettuce crops such as foliage area, length, and width and estimate effective plant spacing for an NFT-based aquaponic facility. The results have shown that the growth of plants is estimated within 2cm of error for both length and width, 4cm for area and 1.5cm for plant spacing. The final model is then deployed on a cloud-based application and an ontology model is integrated with it. The proposed method is accurate and flexible and hence can easily be applied in real scenarios. This contribution has great significance to the research community as it promotes the implementation of a self-adaptive aquaponic system that can be constantly improved using dynamic data. Moreover, the presented methods offer the opportunity to rely on smart technologies for the application of new concepts such as research on complex relationships between optimal parameters, and detection of nutrient deficiency in crops using computer vision which will pave the way for large-scale implementation of aquaponics farming technology.

Chapter 9 Conclusions, discussions, and future work

9.1. General conclusions

Over the past few years, the world has been facing several challenges, such as drastic increases in population, resource scarcity, and climate change. These factors also threaten food security due to increasing demands for agri-food products. In the facet of this upheaval, traditional farming practices seem vulnerable to matching the growing need for nutritious food. This poses a need to explore modern farming practices such as indoor vertical farming methods, namely aquaponics, hydroponics, and aeroponics. These farming methods represent a class of horticulture that involves the cultivation of plants in a soilless-controlled environment by subjecting the bare roots to a nutrient-enriched solution. The benefits of these modern farming methods are evident from sustainable and environmental-friendly urban growth to enhanced crop yields at a faster rate and reduced labor costs.

In this research, the aquaponic system is considered, which integrates hydroponics (soilless growing of plants) with aquaculture (farming of fishes) in a recirculating system, where the wastes and metabolites produced by cultured fish are removed by nitrification and taken up by the plants. Imitating the natural cycle, aquaponics presents a symbiotic relationship between plant cultivation and fish farming and hence acts as a promising resource-efficient and sustainable food production technology that can address food security issues. Although there is a growing interest in aquaponics, the complex relationship between various parameters, limited availability of crop and fish species suitable for this technology, special requirements for system design, nutrient concentrations, and water quality along with disease management are hindering its adoption at commercial scale.

The integration of smart technologies, automation, and control in an aquaponic system enables intelligent data-driven decisions related to crop quality control, design configuration of the system, and autonomous and robust monitoring and control of the system's operations, and hence can address the challenges mentioned above. But realizing such a highly digitized aquaponic system requires system knowledge, efficient data integration, and information flow among different domains. Research has shown that ontologies can be used for this purpose as they store information and knowledge from multiple heterogeneous sources, ensure efficient data integration and infer new knowledge that has not been explicitly incorporated. This knowledge and different machine learning and deep learning techniques can be utilized to develop a data-driven decision support system for system management, crop quality control, and disease prevention.

Therefore, this thesis attempts to unify various findings in the field of a smart aquaponic system which can be divided into seven stages. In the first stage (Chapter 2), a detailed analysis and review of the present literature related to digitization in the agriculture industry are performed to get insights into the status of smart technologies in agriculture. It is found that the pace at which digitization is happening in the agriculture industry is slow because most agricultural operations deal with living subjects, like animals and plants or perishable products, and developing systems is harder than non-living human-made systems. Other reasons include dependence on natural conditions (climate, soil, humidity) and variations in crop species and their growth conditions. Moreover, it is also found that in the case of aquaponics, complex design and management of the system add additional challenges. Further analysis of the literature shows that no unified ontology model is currently available for storing knowledge of an aquaponic system knowledge. Moreover, no automatic crop quality and disease diagnostic system for an aquaponic system has been developed that considers multiple leafy green crops and associated system requirements. Finally, no decision support system is available that integrates an ontology model and crop diagnostic system, which can assist agricultural practitioners in decision-making regarding crop quality, facility layout, and a system's operations.

Considering the research gaps identified from literature analysis in the first stage, an ontology model "AquaONT" is developed using the "Methontology" approach in the second stage (Chapter 3) to model and store relevant knowledge pertaining to different domains of an aquaponic 4.0 system (a digital farm that uses smart technologies to improve the system's design and operations) identified through the literature as well as other resources. The knowledge from AquaONT can be extracted for optimal operation of IoT devices by comparing contextual data coming from a farm with standard/ideal data from experts, taking corrective actions on qualitative issues of crop and fish (disease and quality management), and ensuring correct design configuration of the system based on crop characteristics. This information can assist agricultural practitioners and farmers in clear decision-making regarding IoT devices, sensors, and other components necessary for farm development.

The design of the aquaponic system, such as grow bed configuration (spacing between plants and channels), significantly impact crop growth and overall quality. Therefore, in the third stage (Chapter 4), a framework based on 'AquaONT' is proposed for automated decision-making regarding crop production and respective grow bed design in aquaponics farms. To deploy the framework, a GUI is developed that uses inferred and existing knowledge from AquaONT and a mathematical model to calculate design parameters for grow bed design-based crop characteristics. The GUI is also integrated with CAD software which, upon receiving the calculated design parameters, performs parametric modeling and provides the layout of grow bed for the specified crop. This virtual tool can assist agriculturalists in decision-making regarding which design configuration is suitable for a particular crop choice and space availability.

In the fourth stage (Chapter 5), a cloud-based dashboard is developed for the aquaponic facility established in the Aquaponics 4.0 Learning Factory (AllFactory), University of Alberta. The dashboard displays the sensor data and crop images in a real-time. The data is acquired from an aquaponic facility and uploaded wirelessly to the cloud. The proposed dashboard can be used as a standalone tool for farm and crop monitoring, or it can be integrated with other systems. For instance, this research is integrated with the ontology model developed in stage 2, where it populates the instances of the 'Contextual_Data' class, with sensor data. Moreover, it is also integrated with crop inspection and decision support systems developed in the next stages, demonstrating this dashboard's usefulness.

For a functional and efficient aquaponics farm, a number of challenges must be addressed. One of these challenges is related to crop diseases that impact crop quality and cause crop wastage. Hence, early and correct detection of diseases is crucial to witness the efficiency of aquaponics farms. In the fifth stage of the thesis (Chapter 6), an automatic crop disease detection system is proposed for detecting diseases in four leafy green crops, lettuce, basil, spinach, and parsley, grown in an aquaponic facility. To develop this system, the first dataset comprising 2640 images is constructed. Then the system is developed that works in three phases. The first phase is the crop classification phase which tells about the type of crop. The second phase is the disease identification phase which determines whether the crop has a disease or not. The final phase is the disease detection phase which localizes and detects the diseased and healthy spots in leaves and also tells about the disease category. The third phase only activates if the input coming from the second phase is one of the diseased crops. The proposed approach has shown promising results with accuracy in three phases, reaching 95.83%, 94.13%, and 82.13%, respectively. A cloud-based application is then developed where the final system is deployed along with an ontology model developed in stage two for identifying causes and treatments once the disease is detected. This cloud-based application act as a decision support system (DSS) which can be used by agricultural practitioners for early disease diagnosis and taking precautionary measures.

Following stage five, another tool based on the image processing technique is proposed in stage six (Chapter 7) to evaluate the health of lettuce crops grown in aquaponics facilities based on their foliage color. 'HSV Space segmentation' technique is used to segment images and extract Red (R), Green (G), and Blue (B) channel values. The foliage color detection model is then developed using mean R, G, and B channel values and a color distance model. The performance of the model is evaluated, achieving an accuracy of 95%. After that, the proposed model is deployed on DSS developed in stage five. The model developed in this study can be used for the quality evaluation of other crops by adjusting the threshold values.

Finally, in the last stage (Chapter 8), a model is developed using Mask-RCNN to estimate the crop morphological traits in a particular crop top view area for safe plant spacing, which is one of the key design parameters in aquaponics facilities as mentioned in stage three (Chapter 4). The correct spacing between plants in NFT channels is crucial to ensure healthy crop growth, which otherwise is not possible in case of incorrect spacing. This is because each plant has certain canopy dimensions, and if they grow too close to each other, their foliage can overlap, which can stress the plants and limit their growth potential. Moreover, plants grown very close to each other are prone to diseases. This model is also deployed on DSS developed in stage five.

9.2. Research contribution

The primary contributions of this research are summarized as follows:

- Investigated and analyzed the extent of digital technologies adoption in agriculture in the context of service type, technology readiness level, and farm type. Identified the research gaps, particularly in the area of aquaponics.
- Developed an ontology model to store relevant knowledge of an aquaponic system pertaining to environmental parameters, contextual data, production system, product quality, production facility, and product type. Formalized the relationships among these domains to enable the decision-making process.
- Constructed a virtual decision support tool using ontology model and parametric modeling model to determine to grow bed design configuration based on crop characteristics for healthy growth of crops in aquaponic facility.
- Developed a cloud-based dashboard for data acquisition and monitoring of aquaponic facility.
- Developed a crop disease diagnostic system using deep learning and computer vision techniques for early detection and identification of diseases in leafy green crops grown in an aquaponic facility.
- Developed a novel real-time image processing technique to determine the quality of the crop based on foliage color.
- Designed a system to estimate crop morphological traits for plant site spacing using deep learning and computer vision approaches.
- Integrated all the developed models and systems to ensure efficient data and information flow.
- Developed a DSS in the form of a cloud-based application using the Streamlit, where ontology model and crop quality inspection systems are deployed. This DSS assists in the decision-making process regarding facility layout, crop quality and optimal environmental conditions, etc.

9.3. Research limitations and future research

This research is subjected to the following limitations:

- While conducting the literature analysis using the PRISMA approach, only three online repositories are considered for the literature search (Scopus, IEEE, and Science Direct), and a specified search string is used. In both scenarios, it is highly unlikely that the overall findings would change. But still, there is a possibility that additional databases, keywords, and synonyms might return more studies. For future work, additional research databases and aspects can be considered to provide a holistic overview of the agricultural industry in digitization. Moreover, studies targeting agriculture 5.0, in general, can also be included.

- At the moment, AquaONT can only be applied to an NFT-based aquaponic system that grow leafy green vegetables such as lettuce, spinach, parsley, and basil. Moreover, it does not contain knowledge about microbial inoculum or nitrifying bacteria. Hence, for future work, AquaONT can be expanded to include aspects and knowledge of: i) other aquaponic 4.0 systems such as deep-water culture (DWC) and media-based aquaponics, ii) fruits and other vegetables, and iii) nitrifying bacteria.
- The crop disease diagnostic system is developed considering only four leafy green crops. The dataset is also developed that contains around 2640 healthy and diseased images of these crops. The system can be extended for future work to include other crops such as fruits and cereal crops. Moreover, the dataset can also be expanded to incorporate more real-field images.
- During the development of the model for crop quality evaluation based on color foliage, only the lettuce crop is considered. The model has not taken into account other crops. However, it can be implemented to other crops as well by adjusting some parameters. The second limitation is that model works with images having only one object. The model does not perform the segmentation process efficiently with multiple objects in the background. In future work, the current limitations of the model can be addressed by considering other crops as well as multiple objects in an image.
- The lettuce crop is considered for developing a model to estimate the foliage area for plant site spacing. But this model can easily be applied to other crops as well after training it on new crop data. For future work, the model is expanded to include other crops along with a dataset containing images of other crops.
- Due to certification requirements related to fish, the aquaculture system is not taken into account, which limits the research to the hydroponics component of the process. However, to imitate the aquaponics process, nutrient concentrations in water effluent are kept at similar levels as that of aquaponics. In the future, the scope of current research can be broadened by incorporating aquaculture along

with hydroponics, collectively analyzing the impact of both systems on crop growth and quality and thus facilitating the large-scale commercial implementation of this sustainable food production system.

References

[1] F. Schierhorn and M. Elferink, "Global Demand for Food Is Rising. Can We Meet It?," *Harv. Bus. Rev.*, vol. 7, no. May, p. 2016, 2017, [Online]. Available: https://hbr.org/2016/04/global-demand-for-food-is-rising-can-we-meet-it.

[2] T. Y. Kyaw and A. K. Ng, "Smart Aquaponics System for Urban Farming," in *Energy Procedia*, 2017, vol. 143, pp. 342–347, doi: 10.1016/j.egypro.2017.12.694.

[3] W. K. Mok, Y. X. Tan, and W. N. Chen, "Technology innovations for food security in Singapore: A case study of future food systems for an increasingly natural resource-scarce world," *Trends Food Sci. Technol.*, vol. 102, pp. 155–168, Aug. 2020, doi: 10.1016/j.tifs.2020.06.013.

[4] H. Valin *et al.*, "The future of food demand: Understanding differences in global economic models," *Agric. Econ. (United Kingdom)*, vol. 45, no. 1, pp. 51–67, Jan. 2014, doi: 10.1111/agec.12089.

[5] A. R. Yanes, P. Martinez, and R. Ahmad, "Towards automated aquaponics: A review on monitoring, IoT, and smart systems," *Journal of Cleaner Production*. 2020, doi: 10.1016/j.jclepro.2020.121571.

[6] R. Abbasi, P. Martinez, and R. Ahmad, "An ontology model to support the automated design of aquaponic grow beds," *Procedia CIRP*, vol. 100, pp. 55–60, 2021, doi: 10.1016/j.procir.2021.05.009.

[7] S. A. Gillani, R. Abbasi, P. Martinez, and R. Ahmad, "Ontology-based Interactive Learning Approach for Transdisciplinary Teaching in Learning Factory," *SSRN Electron. J.*, Mar. 2022, doi: 10.2139/SSRN.4071925.

[8] K. Al-Kodmany, "The vertical farm: A review of developments and implications for the vertical city," *Buildings*. 2018, doi: 10.3390/buildings8020024.

[9] Y. Wei, W. Li, D. An, D. Li, Y. Jiao, and Q. Wei, "Equipment and Intelligent Control System in Aquaponics: A Review," *IEEE Access*, vol. 7, pp. 169306–169326, 2019, doi: 10.1109/ACCESS.2019.2953491.

[10] H. Monsees, W. Kloas, and S. Wuertz, "Decoupled systems on trial: Eliminating bottlenecks to improve aquaponic processes," *PLoS One*, vol. 12, no. 9, pp. 1–18, 2017, doi: 10.1371/journal.pone.0183056.

[11] L. Pérez-Urrestarazu, J. Lobillo-Eguíbar, R. Fernández-Cañero, and V. M. Fernández-Cabanás, "Suitability and optimization of FAO's small-scale aquaponics systems for joint production of lettuce (Lactuca sativa) and fish (Carassius auratus)," *Aquac. Eng.*, 2019, doi: 10.1016/j.aquaeng.2019.04.001.

[12] C. Somerville, M. Cohen, E. Pantanella, A. Stankus, and A. Lovatelli, *Small-scale aquaponic food production. Integrated fish and plant farming*. 2014.

[13] R. López-Pozos, G. A. Martínez-Gutiérrez, R. Pérez-Pacheco, and M. Urrestarazu, "The effects of slope and channel nutrient solution gap number on the yield of tomato crops by a nutrient film technique system under a warm climate," *HortScience*, vol. 46, no. 5, pp. 727–729, 2011, doi: 10.21273/hortsci.46.5.727.

[14] E. Loreti and G. G. Striker, "Plant responses to hypoxia: Signaling and adaptation," *Plants*, vol. 9, no. 12, pp. 1–4, 2020, doi: 10.3390/plants9121704.

[15] M. M. Maboko and C. P. Du Plooy, "Effect of plant spacing on growth and yield of lettuce (lactuca sativa L.) in a soilless production system," *South African J. Plant Soil*, 2009, doi: 10.1080/02571862.2009.10639954.

[16] S. Goddek, *Aquaponics Food Production Systems*. 2019.

[17] M. Asmi Romli, S. Daud, R. A. Aliana Raof, Z. Awang Ahmad, and N. Mahrom, “Aquaponic Growbed Water Level Control Using Fog Architecture Related content Aquaponic Growbed Water Level Control Using Fog Architecture,” *J. Phys*, p. 12014, 2018, doi: 10.1088/1742-6596/1018/1/012014.

[18] K. P. Ferentinos, N. Katsoulas, A. Tzounis, T. Bartzanas, and C. Kittas, “Wireless sensor networks for greenhouse climate and plant condition assessment,” *Biosyst. Eng.*, vol. 153, pp. 70–81, Jan. 2017, doi: 10.1016/J.BIOSYSTEMSENG.2016.11.005.

[19] G. Stouvenakers *et al.*, “Plant Pathogens and Control Strategies in Aquaponics,” *Aquaponics Food Prod. Syst.*, pp. 353–378, 2019, doi: 10.1007/978-3-030-15943-6_14.

[20] A. L. Stewart and S. Ahmed, “Effects of climate change on fruit nutrition,” *Fruit Crop. Diagnosis Manag. Nutr. Constraints*, pp. 77–93, doi: 10.1016/B978-0-12-818732-6.00007-1.

[21] R. Abbasi, P. Martinez, and R. Ahmad, “The digitization of agricultural industry – a systematic literature review on agriculture 4.0,” *Smart Agric. Technol.*, vol. 2, p. 100042, Dec. 2022, doi: 10.1016/J.ATECH.2022.100042.

[22] R. Abbasi, P. Martinez, and R. Ahmad, “An ontology model to represent aquaponics 4.0 system’s knowledge,” *Inf. Process. Agric.*, Dec. 2021, doi: 10.1016/J.INPA.2021.12.001.

[23] “Imagining construction’s digital future | McKinsey.” https://www.mckinsey.com/business-functions/operations/our-insights/imagining-constructions-digital-future (accessed Jul. 13, 2021).

[24] T. Haider *et al.*, “A Computer-Vision-Based Approach for Nitrogen Content Estimation in Plant Leaves,” *Agric. 2021, Vol. 11, Page 766*, vol. 11, no. 8, p. 766, Aug. 2021, doi: 10.3390/AGRICULTURE11080766.

[25] M. A. Munnaf, G. Haesaert, M. Van Meirvenne, and A. M. Mouazen, “Site-specific seeding using multi-sensor and data fusion techniques: A review,” *Adv. Agron.*, vol. 161, pp. 241–323, Jan. 2020, doi: 10.1016/BS.AGRON.2019.08.001.

[26] G. Aceto, V. Persico, and A. Pescapé, “A Survey on Information and Communication Technologies for Industry 4.0: State-of-the-Art, Taxonomies, Perspectives, and Challenges,” *IEEE Communications Surveys and Tutorials*. 2019, doi: 10.1109/COMST.2019.2938259.

[27] A. Gacar, H. Aktas, and B. Ozdogan, “Digital agriculture practices in the context of agriculture 4.0,” *Pressacademia*, vol. 4, no. 2, pp. 184–191, 2017, doi: 10.17261/pressacademia.2017.448.

[28] Y. Liu, X. Ma, L. Shu, G. P. Hancke, and A. M. Abu-Mahfouz, “From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges,” *IEEE Trans. Ind. Informatics*, vol. 17, no. 6, pp. 4322–4334, Jun. 2021, doi: 10.1109/TII.2020.3003910.

[29] F. da Silveira, F. H. Lermen, and F. G. Amaral, “An overview of agriculture 4.0 development: Systematic review of descriptions, technologies, barriers, advantages, and disadvantages,” *Comput. Electron. Agric.*, vol. 189, p. 106405, Oct. 2021, doi: 10.1016/J.COMPAG.2021.106405.

[30] M. S. Farooq, S. Riaz, A. Abid, K. Abid, and M. A. Naeem, “A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming,” *IEEE Access*, vol. 7, pp. 156237–156271, 2019, doi: 10.1109/ACCESS.2019.2949703.

[31] J. Miranda, P. Ponce, A. Molina, and P. Wright, “Sensing, smart and sustainable technologies for Agri-Food 4.0,” *Comput. Ind.*, vol. 108, pp. 21–36, Jun. 2019, doi: 10.1016/J.COMPIND.2019.02.002.

[32] M. Lezoche, H. Panetto, J. Kacprzyk, J. E. Hernandez, and M. M. E. Alemany Díaz, “Agri-food 4.0: A survey of the supply chains and technologies for the future agriculture,” *Comput. Ind.*, vol. 117, p. 103187, May 2020, doi: 10.1016/J.COMPIND.2020.103187.

[33] I. Bhakta, S. Phadikar, and K. Majumder, “State-of-the-art technologies in precision

agriculture: a systematic review," 2019, doi: 10.1002/jsfa.9693.

[34] S. O. Araújo, R. S. Peres, J. Barata, F. Lidon, and J. C. Ramalho, "Characterising the Agriculture 4.0 Landscape—Emerging Trends, Challenges and Opportunities," *Agron. 2021, Vol. 11, Page 667*, vol. 11, no. 4, p. 667, Apr. 2021, doi: 10.3390/AGRONOMY11040667.

[35] M. Bacco, P. Barsocchi, E. Ferro, A. Gotta, and M. Ruggeri, "The Digitisation of Agriculture: a Survey of Research Activities on Smart Farming," *Array*, vol. 3–4, no. October, p. 100009, Sep. 2019, doi: 10.1016/j.array.2019.100009.

[36] X. Huang, C. Zanni-Merk, and B. Crémilleux, "Enhancing deep learning with semantics: An application to manufacturing time series analysis," *Procedia Comput. Sci.*, vol. 159, no. 2018, pp. 437–446, 2019, doi: 10.1016/j.procs.2019.09.198.

[37] L. Jiyu *et al.*, "Distribution law of rice pollen in the wind field of small UAV," *Int. J. Agric. Biol. Eng.*, vol. 10, no. 4, pp. 32–40, Jul. 2017, doi: 10.25165/IJABE.V10I4.3103.

[38] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, Mar. 2021, doi: 10.1136/BMJ.N71.

[39] C. Pylianidis, S. Osinga, and I. N. Athanasiadis, "Introducing digital twins to agriculture," *Comput. Electron. Agric.*, vol. 184, p. 105942, May 2021, doi: 10.1016/J.COMPAG.2020.105942.

[40] Aqeel-ur-Rehman, Z. A. Shaikh, N. A. Shaikh, and N. Islam, "An integrated framework to develop context aware sensor grid for agriculture," *Aust. J. Basic Appl. Sci.*, 2010.

[41] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.

[42] A. Tzounis, N. Katsoulas, T. Bartzanas, and C. Kittas, "Internet of Things in agriculture, recent advances and future challenges," *Biosyst. Eng.*, vol. 164, pp. 31–48, Dec. 2017, doi: 10.1016/J.BIOSYSTEMSENG.2017.09.007.

[43] V. P. Kour and S. Arora, "Recent Developments of the Internet of Things in Agriculture: A Survey," *IEEE Access*, vol. 8, pp. 129924–129957, 2020, doi: 10.1109/ACCESS.2020.3009298.

[44] R. Saemaldahr, B. Thapa, K. Maikoo, and E. B. Fernandez, "Reference Architectures for the IoT: A Survey," *Lect. Notes Data Eng. Commun. Technol.*, vol. 72, pp. 635–646, Dec. 2020, doi: 10.1007/978-3-030-70713-2_58.

[45] A. Somov *et al.*, "Pervasive Agriculture: IoT-Enabled Greenhouse for Plant Growth Control," *IEEE Pervasive Comput.*, vol. 17, no. 4, pp. 65–75, 2018, doi: 10.1109/MPRV.2018.2873849.

[46] B. Yimwadsana, P. Chanthapeth, C. Lertthanyaphan, and A. Pornvechamnuay, "An IoT controlled system for plant growth," *Proceeding 2018 7th ICT Int. Student Proj. Conf. ICT-ISPC 2018*, Nov. 2018, doi: 10.1109/ICT-ISPC.2018.8523886.

[47] S. N. Nnadi and F. E. Idachaba, "Design and Implementation of a Sustainable IOT Enabled Greenhouse Prototype," *IEEE 5G World Forum, 5GWF 2018 - Conf. Proc.*, pp. 457–461, Oct. 2018, doi: 10.1109/5GWF.2018.8517006.

[48] J. Yang, M. Liu, J. Lu, Y. Miao, M. A. Hossain, and M. F. Alhamid, "Botanical Internet of Things: Toward Smart Indoor Farming by Connecting People, Plant, Data and Clouds," *Mob. Networks Appl.*, vol. 23, no. 2, pp. 188–202, Apr. 2018, doi: 10.1007/S11036-017-0930-X.

[49] F. Francis, P. L. Vishnu, M. Jha, and B. Rajaram, "IOT-Based Automated Aeroponics System," *Lect. Notes Electr. Eng. Intell. Embed. Syst.*, vol. 492, pp. 337–345, Nov. 2017, doi: 10.1007/978-981-10-8575-8_32.

[50] C. A. Jamhari, W. K. Wibowo, A. R. Annisa, and T. M. Roffi, "Design and Implementation of IoT System for Aeroponic Chamber Temperature Monitoring," *Proceeding - 2020 3rd Int. Conf. Vocat. Educ. Electr. Eng. Strength. Framew. Soc. 5.0 through Innov. Educ. Electr. Eng.*

Informatics Eng. ICVEE 2020, Oct. 2020, doi: 10.1109/ICVEE50212.2020.9243213.

[51] K. C. Chang, P. K. Liu, Z. W. Kuo, and S. H. Liao, "Design of persimmon growing stage monitoring system using image recognition technique," *2016 IEEE Int. Conf. Consum. Electron. ICCE-TW 2016*, Jul. 2016, doi: 10.1109/ICCE-TW.2016.7520978.

[52] T. Namgyel *et al.*, "IoT based hydroponic system with supplementary LED light for smart home farming of lettuce," pp. 221–224, Jan. 2019, doi: 10.1109/ECTICON.2018.8619983.

[53] M. Manju, V. Karthik, S. Hariharan, and B. Sreekar, "Real time monitoring of the environmental parameters of an aquaponic system based on internet of things," *ICONSTEM 2017 - Proc. 3rd IEEE Int. Conf. Sci. Technol. Eng. Manag.*, vol. 2018-January, pp. 943–948, Jun. 2017, doi: 10.1109/ICONSTEM.2017.8261342.

[54] M. S. Mekala and P. Viswanathan, "CLAY-MIST: IoT-cloud enabled CMM index for smart agriculture monitoring system," *Measurement*, vol. 134, pp. 236–244, Feb. 2019, doi: 10.1016/J.MEASUREMENT.2018.10.072.

[55] T. Wiangtong and P. Sirisuk, "IoT-based Versatile Platform for Precision Farming," *Isc. 2018 - 18th Int. Symp. Commun. Inf. Technol.*, pp. 438–441, Dec. 2018, doi: 10.1109/ISCIT.2018.8587989.

[56] M. A. Zamora-Izquierdo, J. Santa, J. A. Martínez, V. Martínez, and A. F. Skarmeta, "Smart farming IoT platform based on edge and cloud computing," *Biosyst. Eng.*, vol. 177, pp. 4–17, Jan. 2019, doi: 10.1016/J.BIOSYSTEMSENG.2018.10.014.

[57] N. K. Jacob, "IoT powered portable aquaponics system," *ACM Int. Conf. Proceeding Ser.*, Mar. 2017, doi: 10.1145/3018896.3018965.

[58] M. Mehra, S. Saxena, S. Sankaranarayanan, R. J. Tom, and M. Veeramanikandan, "IoT based hydroponics system using Deep Neural Networks," *Comput. Electron. Agric.*, vol. 155, pp. 473–486, Dec. 2018, doi: 10.1016/J.COMPAG.2018.10.015.

[59] B. A. A.-Z. Naser, A. L. Saleem, A. Ali, S. Alabassi, and M. Al-Baghdadi, "Design and construction of smart IoT-based aquaponics powered by PV cells," 2019.

[60] M. Odema, I. Adly, A. Wahba, and H. Ragai, "Smart Aquaponics System for Industrial Internet of Things (IIoT)," *Adv. Intell. Syst. Comput.*, vol. 639, pp. 844–854, 2017, doi: 10.1007/978-3-319-64861-3_79.

[61] W. Vernandhes, N. S. Salahuddin, A. Kowanda, and S. P. Sari, "Smart aquaponic with monitoring and control system based on IoT," *Proc. 2nd Int. Conf. Informatics Comput. ICIC 2017*, vol. 2018-Janua, pp. 1–6, Feb. 2018, doi: 10.1109/IAC.2017.8280590.

[62] B. Keswani *et al.*, "Adapting weather conditions based IoT enabled smart irrigation technique in precision agriculture mechanisms," *Neural Comput. Appl. 2018 311*, vol. 31, no. 1, pp. 277–292, Sep. 2018, doi: 10.1007/S00521-018-3737-1.

[63] N. K. Nawandar and V. R. Satpute, "IoT based low cost and intelligent module for smart irrigation system," *Comput. Electron. Agric.*, vol. 162, pp. 979–990, Jul. 2019, doi: 10.1016/J.COMPAG.2019.05.027.

[64] P. M. Gupta, M. Salpekar, and P. K. Tejan, "Agricultural practices Improvement Using IoT Enabled SMART Sensors," *2018 Int. Conf. Smart City Emerg. Technol. ICSCET 2018*, Nov. 2018, doi: 10.1109/ICSCET.2018.8537291.

[65] M. Dholu and K. A. Ghodinde, "Internet of Things (IoT) for Precision Agriculture Application," *Proc. 2nd Int. Conf. Trends Electron. Informatics, ICOEI 2018*, pp. 339–342, Nov. 2018, doi: 10.1109/ICOEI.2018.8553720.

[66] T. A. A. Ali, V. Choksi, and M. B. Potdar, "Precision Agriculture Monitoring System Using Green Internet of Things (G-IoT)," *Proc. 2nd Int. Conf. Trends Electron. Informatics, ICOEI 2018*, pp. 481–487, Nov. 2018, doi: 10.1109/ICOEI.2018.8553866.

[67] A. J. Rau, J. Sankar, A. R. Mohan, D. Das Krishna, and J. Mathew, "IoT based smart irrigation system and nutrient detection with disease analysis," *TENSYMP 2017 - IEEE Int. Symp. Technol. Smart Cities*, Oct. 2017, doi: 10.1109/TENCONSPRING.2017.8070100.

[68] A. Thorat, S. Kumari, and N. D. Valakunde, "An IoT based smart solution for leaf disease detection," *2017 Int. Conf. Big Data, IoT Data Sci. BID 2017*, vol. 2018-January, pp. 193–198, Apr. 2018, doi: 10.1109/BID.2017.8336597.

[69] K. Foughali, K. Fathallah, and A. Frihida, "Using Cloud IOT for disease prevention in precision agriculture," *Procedia Comput. Sci.*, vol. 130, pp. 575–582, Jan. 2018, doi: 10.1016/J.PROCS.2018.04.106.

[70] S. S. Koshy, V. S. Sunnam, P. Rajgarhia, K. Chinnusamy, D. P. Ravulapalli, and S. Chunduri, "Application of the internet of things (IoT) for smart farming: a case study on groundnut and castor pest and disease forewarning," *CSI Trans. ICT 2018 63*, vol. 6, no. 3, pp. 311–318, Oct. 2018, doi: 10.1007/S40012-018-0213-0.

[71] S. S. Patil and S. A. Thorat, "Early detection of grapes diseases using machine learning and IoT," *Proc. - 2016 2nd Int. Conf. Cogn. Comput. Inf. Process. CCIP 2016*, Dec. 2016, doi: 10.1109/CCIP.2016.7802887.

[72] M. I. Pavel, S. M. Kamruzzaman, S. S. Hasan, and S. R. Sabuj, "An IoT based plant health monitoring system implementing image processing," *2019 IEEE 4th Int. Conf. Comput. Commun. Syst. ICCCS 2019*, pp. 299–303, Feb. 2019, doi: 10.1109/CCOMS.2019.8821783.

[73] M. U. Aftab, O. Ashraf, M. Irfan, M. Majid, A. Nisar, and M. A. Habib, "A Review Study of Wireless Sensor Networks and Its Security," *Commun. Netw.*, vol. 7, pp. 172–179, 2015, doi: 10.4236/cn.2015.74016.

[74] X. Yu, P. Wu, W. Han, and Z. Zhang, "A survey on wireless sensor network infrastructure for agriculture," *Comput. Stand. Interfaces*, vol. 1, no. 35, pp. 59–64, Jan. 2013, doi: 10.1016/J.CSI.2012.05.001.

[75] J. Gutierrez, J. F. Villa-Medina, A. Nieto-Garibay, and M. A. Porta-Gandara, "Automated irrigation system using a wireless sensor network and GPRS module," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 1, pp. 166–176, Jan. 2014, doi: 10.1109/TIM.2013.2276487.

[76] H. Navarro-Hellín, R. Torres-Sánchez, F. Soto-Valles, C. Albaladejo-Pérez, J. A. López-Riquelme, and R. Domingo-Miguel, "A wireless sensors architecture for efficient irrigation water management," *Agric. Water Manag.*, vol. 151, pp. 64–74, Mar. 2015, doi: 10.1016/J.AGWAT.2014.10.022.

[77] S. A. Nikolidakis, D. Kandris, D. D. Vergados, and C. Douligeris, "Energy efficient automated control of irrigation in agriculture by using wireless sensor networks," *Comput. Electron. Agric.*, vol. 113, pp. 154–163, Apr. 2015, doi: 10.1016/J.COMPAG.2015.02.004.

[78] I. Mat, M. R. Mohd Kassim, A. N. Harun, and I. Mat Yusoff, "IoT in Precision Agriculture applications using Wireless Moisture Sensor Network," *ICOS 2016 - 2016 IEEE Conf. Open Syst.*, pp. 24–29, Mar. 2017, doi: 10.1109/ICOS.2016.7881983.

[79] J. A. Jiang *et al.*, "Toward a higher yield: a wireless sensor network-based temperature monitoring and fan-circulating system for precision cultivation in plant factories," *Precis. Agric.*, vol. 19, no. 5, pp. 929–956, Oct. 2018, doi: 10.1007/S11119-018-9565-6.

[80] M. Srbinovska, C. Gavrovski, V. Dimcev, A. Krkoleva, and V. Borozan, "Environmental parameters monitoring in precision agriculture using wireless sensor networks," *J. Clean. Prod.*, vol. 88, pp. 297–307, Feb. 2015, doi: 10.1016/J.JCLEPRO.2014.04.036.

[81] P. C. Menon, "IoT enabled Aquaponics with wireless sensor smart monitoring," *Proc. 4th Int. Conf. IoT Soc. Mobile, Anal. Cloud, ISMAC 2020*, pp. 171–176, Oct. 2020, doi: 10.1109/I-SMAC49090.2020.9243368.

[82] T. Cao-Hoang and C. N. Duy, "Environment monitoring system for agricultural application

based on wireless sensor network," *7th Int. Conf. Inf. Sci. Technol. ICIST 2017 - Proc.*, pp. 99–102, May 2017, doi: 10.1109/ICIST.2017.7926499.

[83] O. N. Samijayani, R. Darwis, S. Rahmatia, A. Mujadin, and D. Astharini, "Hybrid zigbee and wifi wireless sensor networks for hydroponic monitoring," 2020, Accessed: Aug. 23, 2021. [Online]. Available: http://www.icecce.com/wp-content/uploads/2020/06/Program_ICECCE.pdf.

[84] P. M. Mell and T. Grance, "The NIST definition of cloud computing," 2011, doi: 10.6028/NIST.SP.800-145.

[85] T. Alwada'n, "CLOUD COMPUTING AND MULTI-AGENT SYSTEM : MONITORING AND SERVICES," 2018.

[86] X. Shi *et al.*, "State-of-the-art internet of things in protected agriculture," *Sensors (Switzerland)*, vol. 19, no. 8. Multidisciplinary Digital Publishing Institute, p. 1833, Apr. 17, 2019, doi: 10.3390/s19081833.

[87] J. Wang, H. Yue, and Z. Zhou, "An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network," *Food Control*, vol. 79, pp. 363–370, Sep. 2017, doi: 10.1016/J.FOODCONT.2017.04.013.

[88] A. Kaloxylos *et al.*, "A cloud-based farm management system: Architecture and implementation," *Comput. Electron. Agric.*, vol. 100, pp. 168–179, 2014, doi: 10.1016/J.COMPAG.2013.11.014.

[89] F. Yang, K. Wang, Y. Han, and Z. Qiao, "A Cloud-Based Digital Farm Management System for Vegetable Production Process Management and Quality Traceability," *Sustain. 2018, Vol. 10, Page 4007*, vol. 10, no. 11, p. 4007, Nov. 2018, doi: 10.3390/SU10114007.

[90] D. S. Paraforos *et al.*, "A Farm Management Information System Using Future Internet Technologies," *IFAC-PapersOnLine*, vol. 49, no. 16, pp. 324–329, 2016, doi: 10.1016/J.IFACOL.2016.10.060.

[91] Y. Ampatzidis, L. Tan, R. Haley, and M. D. Whiting, "Cloud-based harvest management information system for hand-harvested specialty crops," *Comput. Electron. Agric.*, vol. 122, pp. 161–167, Mar. 2016, doi: 10.1016/J.COMPAG.2016.01.032.

[92] S. Fountas *et al.*, "Farm management information systems: Current situation and future perspectives," *Comput. Electron. Agric.*, vol. 115, pp. 40–50, Jul. 2015, doi: 10.1016/J.COMPAG.2015.05.011.

[93] H. Y. Chang, J. J. Wang, C. Y. Lin, and C. H. Chen, "An agricultural data gathering platform based on internet of things and big data," *Proc. - 2018 Int. Symp. Comput. Consum. Control. IS3C 2018*, pp. 302–305, Feb. 2019, doi: 10.1109/IS3C.2018.00083.

[94] F. J. Ferrández-Pastor, J. M. García-Chamizo, M. Nieto-Hidalgo, J. Mora-Pascual, and J. Mora-Martínez, "Developing Ubiquitous Sensor Network Platform Using Internet of Things: Application in Precision Agriculture," *Sensors (Basel).*, vol. 16, no. 7, Jul. 2016, doi: 10.3390/S16071141.

[95] M. R *et al.*, "mySense: A comprehensive data management environment to improve precision agriculture practices," *Comput. Electron. Agric.*, vol. 162, pp. 882–894, Jul. 2019, doi: 10.1016/J.COMPAG.2019.05.028.

[96] S. T. Oliver, A. González-Pérez, and J. H. Guijarro, "An IoT proposal for monitoring vineyards called senviro for agriculture," *ACM Int. Conf. Proceeding Ser.*, Oct. 2018, doi: 10.1145/3277593.3277625.

[97] D. H. Fan and S. Gao, "IOP Conference Series: Earth and Environmental Science The application of mobile edge computing in agricultural water monitoring system The application of mobile edge computing in agricultural water monitoring system," vol. 191, p. 12015, 2018, doi: 10.1088/1755-1315/191/1/012015.

[98] L. G, R. C, and G. P, “An automated low cost IoT based Fertilizer Intimation System for smart agriculture,” *Sustain. Comput. Informatics Syst.*, vol. 28, p. 100300, Dec. 2020, doi: 10.1016/J.SUSCOM.2019.01.002.

[99] R. Rahmadian and M. Widyartono, “Autonomous Robotic in Agriculture: A Review,” *Proceeding - 2020 3rd Int. Conf. Vocat. Educ. Electr. Eng. Strength. Framew. Soc. 5.0 through Innov. Educ. Electr. Eng. Informatics Eng. ICVEE 2020*, Oct. 2020, doi: 10.1109/ICVEE50212.2020.9243253.

[100] A. Bechar and C. Vigneault, “Agricultural robots for field operations: Concepts and components,” *Biosyst. Eng.*, vol. 149, pp. 94–111, Sep. 2016, doi: 10.1016/J.BIOSYSTEMSENG.2016.06.014.

[101] P. Gonzalez-De-Santos, R. Fernández, D. Sepúlveda, E. Navas, and M. Armada, “Unmanned Ground Vehicles for Smart Farms,” *Agron. - Clim. Chang. Food Secur.*, Feb. 2020, doi: 10.5772/INTECHOPEN.90683.

[102] O. Bawden *et al.*, “Robot for weed species plant-specific management,” *J. F. Robot.*, vol. 34, no. 6, pp. 1179–1199, Sep. 2017, doi: 10.1002/ROB.21727.

[103] J. Gai, L. Tang, and B. L. Steward, “Automated crop plant detection based on the fusion of color and depth images for robotic weed control,” *J. F. Robot.*, vol. 37, no. 1, pp. 35–52, Jan. 2020, doi: 10.1002/ROB.21897.

[104] G. Adamides *et al.*, “Design and development of a semi-autonomous agricultural vineyard sprayer: Human–robot interaction aspects,” *J. F. Robot.*, vol. 34, no. 8, pp. 1407–1426, Dec. 2017, doi: 10.1002/ROB.21721.

[105] R. Berenstein and Y. Edan, “Automatic Adjustable Spraying Device for Site-Specific Agricultural Application,” *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 2, pp. 641–650, Apr. 2018, doi: 10.1109/TASE.2017.2656143.

[106] J. Underwood *et al.*, “Real-time target detection and steerable spray for vegetable crops,” 2015.

[107] N. Srinivasan *et al.*, “Design of an autonomous seed planting robot,” *IEEE Reg. 10 Humanit. Technol. Conf. 2016, R10-HTC 2016 - Proc.*, Apr. 2017, doi: 10.1109/R10-HTC.2016.7906789.

[108] M. U. Hassan, M. Ullah, and J. Iqbal, “Towards autonomy in agriculture: Design and prototyping of a robotic vehicle with seed selector,” *2016 2nd Int. Conf. Robot. Artif. Intell. ICRAI 2016*, pp. 37–44, Dec. 2016, doi: 10.1109/ICRAI.2016.7791225.

[109] M. Nejati, H. S. Ahn, and B. MacDonald, “Design of a sensing module for a kiwifruit flower pollinator robot,” *Australas. Conf. Robot. Autom. ACRA*, vol. 2019-December, Jun. 2020, Accessed: Aug. 23, 2021. [Online]. Available: https://arxiv.org/abs/2006.08045v1.

[110] Y. Ge, Y. Xiong, and P. J. From, “Symmetry-based 3D shape completion for fruit localisation for harvesting robots,” *Biosyst. Eng.*, vol. 197, pp. 188–202, Sep. 2020, doi: 10.1016/J.BIOSYSTEMSENG.2020.07.003.

[111] S. Birrell, J. Hughes, J. Y. Cai, and F. Iida, “A field-tested robotic harvesting system for iceberg lettuce,” *J. F. Robot.*, vol. 37, no. 2, pp. 225–245, Mar. 2020, doi: 10.1002/ROB.21888.

[112] J. del Cerro, C. C. Ulloa, A. Barrientos, and J. de L. Rivas, “Unmanned Aerial Vehicles in Agriculture: A Survey,” *Agron. 2021, Vol. 11, Page 203*, vol. 11, no. 2, p. 203, Jan. 2021, doi: 10.3390/AGRONOMY11020203.

[113] P. N. Patel, M. Patel, R. M. Faldu, and Y. R. Dave, “Quadcopter for Agricultural Surveillance,” 2013.

[114] G. Sylvester, Food and Agriculture Organization of the United Nations., and International Telecommunication Union., “E-agriculture in action : drones for agriculture,” p. 112.

[115] H. Aasen, A. Burkart, A. Bolten, and G. Bareth, “Generating 3D hyperspectral information

with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance," *ISPRS J. Photogramm. Remote Sens.*, vol. 108, pp. 245–259, Oct. 2015, doi: 10.1016/J.ISPRSJPRS.2015.08.002.

[116] J. Bendig *et al.*, "Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 39, pp. 79–87, Jul. 2015, doi: 10.1016/J.JAG.2015.02.012.

[117] M. Du and N. Noguchi, "Monitoring of wheat growth status and mapping of wheat yield's within-field spatial variations using color images acquired from UAV-camera System," *Remote Sens.*, vol. 9, no. 3, Mar. 2017, doi: 10.3390/RS9030289.

[118] M. L. Guillen-Climent, P. J. Zarco-Tejada, J. A. J. Berni, P. R. J. North, and F. J. Villalobos, "Mapping radiation interception in row-structured orchards using 3D simulation and high-resolution airborne imagery acquired from a UAV," *Precis. Agric. 2012 134*, vol. 13, no. 4, pp. 473–500, Mar. 2012, doi: 10.1007/S11119-012-9263-8.

[119] H. Xiang and L. Tian, "Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV)," *Biosyst. Eng.*, vol. 108, no. 2, pp. 174–190, Feb. 2011, doi: 10.1016/J.BIOSYSTEMSENG.2010.11.010.

[120] C. A. Rokhmana, "The Potential of UAV-based Remote Sensing for Supporting Precision Agriculture in Indonesia," *Procedia Environ. Sci.*, vol. 24, pp. 245–253, Jan. 2015, doi: 10.1016/J.PROENV.2015.03.032.

[121] Z. Pan *et al.*, "Effects of citrus tree-shape and spraying height of small unmanned aerial vehicle on droplet distribution," *Int. J. Agric. Biol. Eng.*, vol. 9, no. 4, pp. 45–52, Jul. 2016, doi: 10.25165/IJABE.V9I4.2178.

[122] S. Meivel, K. Dinakaran, N. Gandhiraj, and M. Srinivasan, "Remote sensing for UREA Spraying Agricultural (UAV) system," *ICACCS 2016 - 3rd Int. Conf. Adv. Comput. Commun. Syst. Bringing to Table, Futur. Technol. from Arround Globe*, Oct. 2016, doi: 10.1109/ICACCS.2016.7586367.

[123] B. S. Faiçal *et al.*, "The use of unmanned aerial vehicles and wireless sensor networks for spraying pesticides," *J. Syst. Archit.*, vol. 60, no. 4, pp. 393–404, Apr. 2014, doi: 10.1016/J.SYSARC.2014.01.004.

[124] D. K. Giles and R. C. Billing, "Deployment and performance of a uav for crop spraying," *Chem. Eng. Trans.*, vol. 44, pp. 307–312, May 2015, doi: 10.3303/CET1544052.

[125] X. Xue, Y. Lan, Z. Sun, C. Chang, and W. C. Hoffmann, "Develop an unmanned aerial vehicle based automatic aerial spraying system," *Comput. Electron. Agric.*, vol. 128, pp. 58–66, Oct. 2016, doi: 10.1016/J.COMPAG.2016.07.022.

[126] J. Torres-Sánchez, J. M. Peña, A. I. de Castro, and F. López-Granados, "Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV," *Comput. Electron. Agric.*, vol. 103, pp. 104–113, Apr. 2014, doi: 10.1016/J.COMPAG.2014.02.009.

[127] J. M. Peña, J. Torres-Sánchez, A. I. de Castro, M. Kelly, and F. López-Granados, "Weed Mapping in Early-Season Maize Fields Using Object-Based Analysis of Unmanned Aerial Vehicle (UAV) Images," *PLoS One*, vol. 8, no. 10, p. e77151, Oct. 2013, doi: 10.1371/JOURNAL.PONE.0077151.

[128] J. Baluja *et al.*, "Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV)," *Irrig. Sci.*, vol. 30, no. 6, pp. 511–522, Nov. 2012, doi: 10.1007/S00271-012-0382-9.

[129] P. J. Zarco-Tejada, V. González-Dugo, and J. A. J. Berni, "Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera," *Remote Sens. Environ.*, vol. 117, pp. 322–337, Feb. 2012, doi: 10.1016/J.RSE.2011.10.007.

[130] H. Hoffmann, R. Jensen, A. Thomsen, H. Nieto, J. Rasmussen, and T. Friborg, "Crop water stress maps for an entire growing season from visible and thermal UAV imagery," *Biogeosciences*, vol. 13, no. 24, pp. 6545–6563, Dec. 2016, doi: 10.5194/BG-13-6545-2016.

[131] C. Romero-Trigueros *et al.*, "Effects of saline reclaimed waters and deficit irrigation on Citrus physiology assessed by UAV remote sensing," *Agric. Water Manag.*, vol. 183, pp. 60–69, Mar. 2017, doi: 10.1016/J.AGWAT.2016.09.014.

[132] M. Romero, Y. Luo, B. Su, and S. Fuentes, "Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management," *Comput. Electron. Agric.*, vol. 147, pp. 109–117, Apr. 2018, doi: 10.1016/J.COMPAG.2018.02.013.

[133] F. Vanegas, D. Bratanov, K. Powell, J. Weiss, and F. Gonzalez, "A Novel Methodology for Improving Plant Pest Surveillance in Vineyards and Crops Using UAV-Based Hyperspectral and Spatial Data," *Sensors 2018, Vol. 18, Page 260*, vol. 18, no. 1, p. 260, Jan. 2018, doi: 10.3390/S18010260.

[134] F. Garcia-Ruiz, S. Sankaran, J. M. Maja, W. S. Lee, J. Rasmussen, and R. Ehsani, "Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees," *Comput. Electron. Agric.*, vol. 91, pp. 106–115, 2013, doi: 10.1016/J.COMPAG.2012.12.002.

[135] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of Big Data challenges and analytical methods," *J. Bus. Res.*, vol. 70, pp. 263–286, Jan. 2017, doi: 10.1016/J.JBUSRES.2016.08.001.

[136] M. Chi, A. Plaza, J. A. Benediktsson, Z. Sun, J. Shen, and Y. Zhu, "Big Data for Remote Sensing: Challenges and Opportunities," *Proc. IEEE*, vol. 104, no. 11, pp. 2207–2219, Nov. 2016, doi: 10.1109/JPROC.2016.2598228.

[137] K. Tesfaye *et al.*, "Targeting drought-tolerant maize varieties in southern Africa: a geospatial crop modeling approach using big data," *Int. Food Agribus. Manag. Rev.*, vol. 19, 2016, Accessed: Aug. 24, 2021. [Online]. Available: https://cgspace.cgiar.org/handle/10568/76332.

[138] B. Vandana and S. S. Kumar, "A novel approach using big data analytics to improve the crop yield in precision agriculture," *2018 3rd IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. RTEICT 2018 - Proc.*, pp. 824–827, May 2018, doi: 10.1109/RTEICT42901.2018.9012549.

[139] S. Sharma, G. Rathee, and H. Saini, "Big data analytics for crop prediction mode using optimization technique," *PDGC 2018 - 2018 5th Int. Conf. Parallel, Distrib. Grid Comput.*, pp. 760–764, Dec. 2018, doi: 10.1109/PDGC.2018.8746001.

[140] A. Ani and P. Gopalakirishnan, "Automated Hydroponic Drip Irrigation Using Big Data," *Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020*, pp. 370–375, Jul. 2020, doi: 10.1109/ICIRCA48905.2020.9182908.

[141] P. Zhang, Q. Zhang, F. Liu, J. Li, N. Cao, and C. Song, "The Construction of the Integration of Water and Fertilizer Smart Water Saving Irrigation System Based on Big Data," *Proc. - 2017 IEEE Int. Conf. Comput. Sci. Eng. IEEE/IFIP Int. Conf. Embed. Ubiquitous Comput. CSE EUC 2017*, vol. 2, pp. 392–397, Aug. 2017, doi: 10.1109/CSE-EUC.2017.258.

[142] R. Sharma, S. S. Kamble, A. Gunasekaran, V. Kumar, and A. Kumar, "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance," *Comput. Oper. Res.*, vol. 119, p. 104926, Jul. 2020, doi: 10.1016/J.COR.2020.104926.

[143] T. Talaviya, D. Shah, N. Patel, H. Yagnik, and M. Shah, "Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides," *Artif. Intell. Agric.*, vol. 4, pp. 58–73, Jan. 2020, doi: 10.1016/J.AIIA.2020.04.002.

[144] M. Mohri, A. Rostamizadeh, and A. Talwalkar, "Foundations in Machine learning," *SpringerBriefs Comput. Sci.*, vol. 0, no. 9783319056050, pp. 39–44, 2014.

[145] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," *Sensors 2018, Vol. 18, Page 2674*, vol. 18, no. 8, p. 2674, Aug. 2018, doi: 10.3390/S18082674.

[146] G. Xu, H. Li, S. Liu, K. Yang, and X. Lin, "VerifyNet: Secure and Verifiable Federated Learning," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, pp. 911–926, 2020, doi: 10.1109/TIFS.2019.2929409.

[147] E. Kamir, F. Waldner, Z. Hochman, E. Kamir, F. Waldner, and Z. Hochman, "Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods," *JPRS*, vol. 160, pp. 124–135, Feb. 2020, doi: 10.1016/J.ISPRSJPRS.2019.11.008.

[148] H. Aghighi, M. Azadbakht, D. Ashourloo, H. S. Shahrabi, and S. Radiom, "Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 12, pp. 4563–4577, Dec. 2018, doi: 10.1109/JSTARS.2018.2823361.

[149] S. Kulkarni, S. N. Mandal, G. S. Sharma, M. R. Mundada, and Meeradevi, "Predictive Analysis to Improve Crop Yield using a Neural Network Model," *2018 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2018*, pp. 74–79, Nov. 2018, doi: 10.1109/ICACCI.2018.8554851.

[150] P. Feng *et al.*, "Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique," *Agric. For. Meteorol.*, vol. 285–286, p. 107922, May 2020, doi: 10.1016/J.AGRFORMET.2020.107922.

[151] Y. Cakir, M. Kirci, and E. O. Gunes, "Yield prediction of wheat in south-east region of Turkey by using artificial neural networks," *2014 3rd Int. Conf. Agro-Geoinformatics, Agro-Geoinformatics 2014*, Sep. 2014, doi: 10.1109/AGRO-GEOINFORMATICS.2014.6910609.

[152] X. E. Pantazi, D. Moshou, T. Alexandridis, R. L. Whetton, and A. M. Mouazen, "Wheat yield prediction using machine learning and advanced sensing techniques," *Comput. Electron. Agric.*, vol. 121, pp. 57–65, Feb. 2016, doi: 10.1016/J.COMPAG.2015.11.018.

[153] Y. Everingham, J. Sexton, D. Skocaj, and G. Inman-Bamber, "Accurate prediction of sugarcane yield using a random forest algorithm," *Agron. Sustain. Dev. 2016 362*, vol. 36, no. 2, pp. 1–9, Apr. 2016, doi: 10.1007/S13593-016-0364-Z.

[154] I. Ahmad *et al.*, "Yield Forecasting of Spring Maize Using Remote Sensing and Crop Modeling in Faisalabad-Punjab Pakistan," *J. Indian Soc. Remote Sens. 2018 4610*, vol. 46, no. 10, pp. 1701–1711, Aug. 2018, doi: 10.1007/S12524-018-0825-8.

[155] M. S. Verma and S. D. Gawade, "A machine learning approach for prediction system and analysis of nutrients uptake for better crop growth in the Hydroponics system," *Proc. - Int. Conf. Artif. Intell. Smart Syst. ICAIS 2021*, pp. 150–156, Mar. 2021, doi: 10.1109/ICAIS50930.2021.9395956.

[156] M. A. Ebrahimi, M. H. Khoshtaghaza, S. Minaei, and B. Jamshidi, "Vision-based pest detection based on SVM classification method," *Comput. Electron. Agric.*, vol. 137, pp. 52–58, May 2017, doi: 10.1016/J.COMPAG.2017.03.016.

[157] D. Moshou, X. E. Pantazi, D. Kateris, and I. Gravalos, "Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier," *Biosyst. Eng.*, vol. 117, no. 1, pp. 15–22, Jan. 2014, doi: 10.1016/J.BIOSYSTEMSENG.2013.07.008.

[158] R. Barosa, S. I. S. Hassen, and L. Nagowah, "Smart Aquaponics with Disease Detection," *2nd Int. Conf. Next Gener. Comput. Appl. 2019, NextComp 2019 - Proc.*, Sep. 2019, doi: 10.1109/NEXTCOMP.2019.8883437.

[159] A. Etienne and D. Saraswat, "Machine learning approaches to automate weed detection by UAV based sensors," *SPIE*, vol. 11008, p. 110080R, May 2019, doi: 10.1117/12.2520536.

[160] A. Bakhshipour and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Comput. Electron. Agric.*, vol. 145, pp. 153–160, Feb. 2018, doi: 10.1016/J.COMPAG.2017.12.032.

[161] A. Morellos *et al.*, "Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy," *Biosyst. Eng.*, vol. 152, pp. 104–116, Dec. 2016, doi: 10.1016/J.BIOSYSTEMSENG.2016.04.018.

[162] E. Acar, M. S. Ozerdem, and B. B. Ustundag, "Machine learning based regression model for prediction of soil surface humidity over moderately vegetated fields," *2019 8th Int. Conf. Agro-Geoinformatics, Agro-Geoinformatics 2019*, Jul. 2019, doi: 10.1109/AGRO-GEOINFORMATICS.2019.8820461.

[163] W. M. Ridwan, M. Sapitang, A. Aziz, K. F. Kushiar, A. N. Ahmed, and A. El-Shafie, "Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia," *Ain Shams Eng. J.*, vol. 12, no. 2, pp. 1651–1663, Jun. 2021, doi: 10.1016/J.ASEJ.2020.09.011.

[164] M. Taki, S. Abdanan Mehdizadeh, A. Rohani, M. Rahnama, and M. Rahmati-Joneidabad, "Applied machine learning in greenhouse simulation; new application and analysis," *Inf. Process. Agric.*, vol. 5, no. 2, pp. 253–268, Jun. 2018, doi: 10.1016/J.INPA.2018.01.003.

[165] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*, vol. 61, pp. 85–117, Apr. 2014, doi: 10.1016/j.neunet.2014.09.003.

[166] A. Canziani, A. Paszke, and E. Culurciello, "An Analysis of Deep Neural Network Models for Practical Applications," May 2016, Accessed: Aug. 24, 2021. [Online]. Available: https://arxiv.org/abs/1605.07678v4.

[167] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *Proc. 2017 Int. Conf. Eng. Technol. ICET 2017*, vol. 2018-January, pp. 1–6, Mar. 2018, doi: 10.1109/ICENGTECHNOL.2017.8308186.

[168] A. Kamilaris and F. X. Prenafeta-Boldu, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, Jul. 2018, doi: 10.1016/j.compag.2018.02.016.

[169] V. Kakani, V. H. Nguyen, B. P. Kumar, H. Kim, and V. R. Pasupuleti, "A critical review on computer vision and artificial intelligence in food industry," *J. Agric. Food Res.*, vol. 2, Dec. 2020, doi: 10.1016/J.JAFR.2020.100033.

[170] R. A. Schwalbert, T. Amado, G. Corassa, L. Pierre Pott, Pvv. Prasad, and I. A. Ciampitti, "Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil," 2019, doi: 10.1016/j.agrformet.2019.107886.

[171] Z. Chu and J. Yu, "An end-to-end model for rice yield prediction using deep learning fusion," *Comput. Electron. Agric.*, vol. 174, Jul. 2020, doi: 10.1016/J.COMPAG.2020.105471.

[172] D. Tedesco-Oliveira, R. Pereira da Silva, W. Maldonado, and C. Zerbato, "Convolutional neural networks in predicting cotton yield from images of commercial fields," *Comput. Electron. Agric.*, vol. 171, p. 105307, Apr. 2020, doi: 10.1016/J.COMPAG.2020.105307.

[173] Y. Chen *et al.*, "Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages," *Remote Sens. 2019, Vol. 11, Page 1584*, vol. 11, no. 13, p. 1584, Jul. 2019, doi: 10.3390/RS11131584.

[174] L. Gong, M. Yu, S. Jiang, V. Cutsuridis, and S. Pearson, "Deep Learning Based Prediction on Greenhouse Crop Yield Combined TCN and RNN," *Sensors 2021, Vol. 21, Page 4537*, vol. 21, no. 13, p. 4537, Jul. 2021, doi: 10.3390/S21134537.

[175] L. Zhang, Z. Xu, D. Xu, J. Ma, Y. Chen, and Z. Fu, "Growth monitoring of greenhouse lettuce based on a convolutional neural network," *Hortic. Res.*, vol. 7, no. 1, Dec. 2020, doi: 10.1038/S41438-020-00345-6.

[176] X. Song, G. Zhang, F. Liu, D. Li, Y. Zhao, and J. Yang, "Modeling spatio-temporal

distribution of soil moisture by deep learning-based cellular automata model," *J. Arid Land*, vol. 8, no. 5, pp. 734–748, Oct. 2016, doi: 10.1007/S40333-016-0049-0.

[177] A. Reyes-Yanes, P. Martinez, and R. Ahmad, "Real-time growth rate and fresh weight estimation for little gem romaine lettuce in aquaponic grow beds," *Comput. Electron. Agric.*, vol. 179, p. 105827, Dec. 2020, doi: 10.1016/j.compag.2020.105827.

[178] M. Dyrmann, H. Karstoft, and H. S. Midtiby, "Plant species classification using deep convolutional neural network," *Biosyst. Eng.*, vol. 151, pp. 72–80, Nov. 2016, doi: 10.1016/J.BIOSYSTEMSENG.2016.08.024.

[179] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Comput. Intell. Neurosci.*, vol. 2016, 2016, doi: 10.1155/2016/3289801.

[180] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Front. Plant Sci.*, vol. 0, no. September, p. 1419, Sep. 2016, doi: 10.3389/FPLS.2016.01419.

[181] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition," *Sensors 2017, Vol. 17, Page 2022*, vol. 17, no. 9, p. 2022, Sep. 2017, doi: 10.3390/S17092022.

[182] A. Anagnostis, G. Asiminari, E. Papageorgiou, and D. Bochtis, "A Convolutional Neural Networks Based Method for Anthracnose Infected Walnut Tree Leaves Identification," *Appl. Sci. 2020, Vol. 10, Page 469*, vol. 10, no. 2, p. 469, Jan. 2020, doi: 10.3390/APP10020469.

[183] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, and C. McCool, "DeepFruits: A Fruit Detection System Using Deep Neural Networks," *Sensors 2016, Vol. 16, Page 1222*, vol. 16, no. 8, p. 1222, Aug. 2016, doi: 10.3390/S16081222.

[184] S. W. Chen *et al.*, "Counting Apples and Oranges with Deep Learning: A Data-Driven Approach," *IEEE Robot. Autom. Lett.*, vol. 2, no. 2, pp. 781–788, Apr. 2017, doi: 10.1109/LRA.2017.2651944.

[185] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, "Apple detection during different growth stages in orchards using the improved YOLO-V3 model," *Comput. Electron. Agric.*, vol. 157, pp. 417–426, Feb. 2019, doi: 10.1016/J.COMPAG.2019.01.012.

[186] F. Terribile *et al.*, "A Web-based spatial decision supporting system for land management and soil conservation," *Solid Earth*, vol. 6, no. 3, pp. 903–928, Jul. 2015, doi: 10.5194/SE-6-903-2015.

[187] A. Felsberger, B. Oberegger, and G. Reiner, "A Review of Decision Support Systems for Manufacturing Systems," *undefined*, 2016.

[188] P. Taechatanasat and L. Armstrong, "Decision Support System Data for Farmer Decision Making," *ECU Publ. Post 2013*, Jan. 2014, Accessed: Aug. 24, 2021. [Online]. Available: https://ro.ecu.edu.au/ecuworkspost2013/855.

[189] H. Navarro-Hellín, J. Martínez-del-Rincon, R. Domingo-Miguel, F. Soto-Valles, and R. Torres-Sánchez, "A decision support system for managing irrigation in agriculture," *Comput. Electron. Agric.*, vol. 124, pp. 121–131, Jun. 2016, doi: 10.1016/J.COMPAG.2016.04.003.

[190] E. Giusti and S. Marsili-Libelli, "A Fuzzy Decision Support System for irrigation and water conservation in agriculture," *Environ. Model. Softw.*, vol. 63, pp. 73–86, Jan. 2015, doi: 10.1016/J.ENVSOFT.2014.09.020.

[191] M. D. M. Kadiyala, S. Nedumaran, P. Singh, C. S., M. A. Irshad, and M. C. S. Bantilan, "An integrated crop model and GIS decision support system for assisting agronomic decision making under climate change," *Sci. Total Environ.*, vol. 521–522, pp. 123–134, Jul. 2015, doi: 10.1016/J.SCITOTENV.2015.03.097.

[192] K. O. Wenkel, M. Berg, W. Mirschel, R. Wieland, C. Nendel, and B. Köstner, "LandCaRe DSS

- An interactive decision support system for climate change impact assessment and the analysis of potential agricultural land use adaptation strategies,” *J. Environ. Manage.*, vol. 127, Sep. 2013, doi: 10.1016/J.JENVMAN.2013.02.051.

[193] D. D. Bochtis, C. G. Sørensen, and O. Green, “A DSS for planning of soil-sensitive field operations,” *Decis. Support Syst.*, vol. 53, no. 1, pp. 66–75, Apr. 2012, doi: 10.1016/J.DSS.2011.12.005.

[194] A. Ghandar, A. Ahmed, S. Zulfiqar, Z. Hua, M. Hanai, and G. Theodoropoulos, “A decision support system for urban agriculture using digital twin: A case study with aquaponics,” *IEEE Access*, vol. 9, pp. 35691–35708, 2021, doi: 10.1109/ACCESS.2021.3061722.

[195] J. A. Sánchez-Molina, N. Pérez, F. Rodríguez, J. L. Guzmán, and J. C. López, “Support system for decision making in the management of the greenhouse environmental based on growth model for sweet pepper,” *Agric. Syst.*, vol. 139, pp. 144–152, Oct. 2015, doi: 10.1016/J.AGSY.2015.06.009.

[196] D. Nestel *et al.*, “An Integrated Decision Support System for Environmentally-Friendly Management of the Ethiopian Fruit Fly in Greenhouse Crops,” *Agron. 2019, Vol. 9, Page 459*, vol. 9, no. 8, p. 459, Aug. 2019, doi: 10.3390/AGRONOMY9080459.

[197] G. Aiello, I. Giovino, M. Vallone, P. Catania, and A. Argento, “A decision support system based on multisensor data fusion for sustainable greenhouse management,” *J. Clean. Prod.*, vol. 172, pp. 4057–4065, Jan. 2018, doi: 10.1016/J.JCLEPRO.2017.02.197.

[198] J. Cañadas, J. A. Sánchez-Molina, F. Rodríguez, and I. M. del Águila, “Improving automatic climate control with decision support techniques to minimize disease effects in greenhouse tomatoes,” *Inf. Process. Agric.*, vol. 4, no. 1, pp. 50–63, Mar. 2017, doi: 10.1016/J.INPA.2016.12.002.

[199] R. M. Sampurno, K. B. Seminar, and Y. Suharnoto, “Weed control decision support system based on precision agriculture approach,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 12, no. 2, pp. 475–484, 2014, doi: 10.12928/TELKOMNIKA.V12I2.1982.

[200] L. Wang, M. Törngren, and M. Onori, “Current status and advancement of cyber-physical systems in manufacturing,” *J. Manuf. Syst.*, vol. 37, pp. 517–527, Oct. 2015, doi: 10.1016/J.JMSY.2015.04.008.

[201] D. G. S. Pivoto, L. F. F. de Almeida, R. da Rosa Righi, J. J. P. C. Rodrigues, A. B. Lugli, and A. M. Alberti, “Cyber-physical systems architectures for industrial internet of things applications in Industry 4.0: A literature review,” *J. Manuf. Syst.*, vol. 58, pp. 176–192, Jan. 2021, doi: 10.1016/J.JMSY.2020.11.017.

[202] A. F. Jimenez, P. F. Cardenas, F. Jimenez, A. Canales, and A. López, “A cyber-physical intelligent agent for irrigation scheduling in horticultural crops,” *Comput. Electron. Agric.*, vol. 178, p. 105777, Nov. 2020, doi: 10.1016/J.COMPAG.2020.105777.

[203] B. Bagheri, S. Yang, H. A. Kao, and J. Lee, “Cyber-physical systems architecture for self-aware machines in industry 4.0 environment,” *IFAC-PapersOnLine*, vol. 28, no. 3, pp. 1622–1627, May 2015, doi: 10.1016/J.IFACOL.2015.06.318.

[204] A. Selmani *et al.*, “Agricultural cyber-physical system enabled for remote management of solar-powered precision irrigation,” *Biosyst. Eng.*, vol. 177, pp. 18–30, Jan. 2019, doi: 10.1016/J.BIOSYSTEMSENG.2018.06.007.

[205] A. Nayak, R. R. Levalle, S. Lee, and S. Y. Nof, “Resource sharing in cyber-physical systems: modelling framework and case studies,” *http://dx.doi.org/10.1080/00207543.2016.1146419*, vol. 54, no. 23, pp. 6969–6983, Dec. 2016, doi: 10.1080/00207543.2016.1146419.

[206] N. Chen, X. Zhang, and C. Wang, “Integrated open geospatial web service enabled cyber-physical information infrastructure for precision agriculture monitoring,” *Comput. Electron. Agric.*, vol. 111, pp. 78–91, Feb. 2015, doi: 10.1016/J.COMPAG.2014.12.009.

[207] D. V. S. Srikar, K. C. Sairam, T. Srikanth, G. Narayanan, K. Vrinda, and D. G. Kurup, "Implementation and Testing of Cyber Physical System in Laboratory for Precision Agriculture," *2018 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2018*, pp. 1906–1908, Nov. 2018, doi: 10.1109/ICACCI.2018.8554601.

[208] I. Ahmad and K. Pothuganti, "Smart Field Monitoring using ToxTrac: A Cyber-Physical System Approach in Agriculture," *Proc. - Int. Conf. Smart Electron. Commun. ICOSEC 2020*, pp. 723–727, Sep. 2020, doi: 10.1109/ICOSEC49089.2020.9215282.

[209] P. Guo, P. O. Dusadeerungsikul, and S. Y. Nof, "Agricultural cyber physical system collaboration for greenhouse stress management," *Comput. Electron. Agric.*, vol. 150, pp. 439–454, Jul. 2018, doi: 10.1016/J.COMPAG.2018.05.022.

[210] B. Stark, S. Rider, and Y. Q. Chen, "Optimal pest management by networked unmanned cropdusters in precision agriculture: A cyber-physical system approach," *IFAC Proc. Vol.*, vol. 46, no. 30, pp. 296–302, Jan. 2013, doi: 10.3182/20131120-3-FR-4045.00019.

[211] C.-R. Rad, O. Hancu, I.-A. Takacs, and G. Olteanu, "Smart Monitoring of Potato Crop: A Cyber-Physical System Architecture Model in the Field of Precision Agriculture," *Agric. Agric. Sci. Procedia*, vol. 6, pp. 73–79, Jan. 2015, doi: 10.1016/J.AASPRO.2015.08.041.

[212] K. Antonopoulos, C. Panagiotou, C. P. Antonopoulos, and N. S. Voros, "A-FARM Precision Farming CPS Platform," *10th Int. Conf. Information, Intell. Syst. Appl. IISA 2019*, Jul. 2019, doi: 10.1109/IISA.2019.8900717.

[213] D. Cimino, A. Ferrero, L. Queirolo, F. Bellotti, R. Berta, and A. De Gloria, "A low-cost, open-source cyber physical system for automated, remotely controlled precision agriculture," *Lect. Notes Electr. Eng.*, vol. 409, pp. 191–203, 2017, doi: 10.1007/978-3-319-47913-2_23.

[214] C. Verdouw, B. Tekinerdogan, A. Beulens, and S. Wolfert, "Digital twins in smart farming," *Agric. Syst.*, vol. 189, p. 103046, Apr. 2021, doi: 10.1016/J.AGSY.2020.103046.

[215] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the Digital Twin: A systematic literature review," *CIRP J. Manuf. Sci. Technol.*, vol. 29, pp. 36–52, May 2020, doi: 10.1016/J.CIRPJ.2020.02.002.

[216] A. Ahmed, S. Zulfiqar, A. Ghandar, Y. Chen, M. Hanai, and G. Theodoropoulos, "Digital Twin Technology for Aquaponics: Towards Optimizing Food Production with Dynamic Data Driven Application Systems," *Commun. Comput. Inf. Sci.*, vol. 1094, no. October, pp. 3–14, Oct. 2019, doi: 10.1007/978-981-15-1078-6_1.

[217] A. Kampker, V. Stich, P. Jussen, B. Moser, and J. Kuntz, "Business models for industrial smart services - the example of a digital twin for a product-service-system for potato harvesting," *Procedia CIRP*, vol. 83, pp. 534–540, 2019, doi: 10.1016/J.PROCIR.2019.04.114.

[218] N. Tsolakis, D. Bechtsis, and D. Bochtis, "AgROS: A Robot Operating System Based Emulation Tool for Agricultural Robotics," *Agron. 2019, Vol. 9, Page 403*, vol. 9, no. 7, p. 403, Jul. 2019, doi: 10.3390/AGRONOMY9070403.

[219] T. Machl, A. Donaubauer, and T. H. Kolbe, "Planning agricultural core road networks based on a digital twin of the cultivated landscape," *J. Digit. Landsc. Archit.*, vol. 2019, no. 4, pp. 316–327, 2019, doi: 10.14627/537663034.

[220] R. G. Alves *et al.*, "A digital twin for smart farming," *2019 IEEE Glob. Humanit. Technol. Conf. GHTC 2019*, Oct. 2019, doi: 10.1109/GHTC46095.2019.9033075.

[221] J. Monteiro, J. Barata, M. Veloso, L. Veloso, and J. Nunes, "Towards sustainable digital twins for vertical farming," *2018 13th Int. Conf. Digit. Inf. Manag. ICDIM 2018*, pp. 234–239, Sep. 2018, doi: 10.1109/ICDIM.2018.8847169.

[222] C. Verdouw and J. Kruize, "Digital twins in farm management : illustrations from the FIWARE accelerators SmartAgriFood and Fractals," *undefined*, 2017.

[223] P. Moghadam, T. Lowe, and E. J. Edwards, "Digital Twin for the Future of Orchard Production

Systems," *Proc. 2019, Vol. 36, Page 92*, vol. 36, no. 1, p. 92, Feb. 2020, doi: 10.3390/PROCEEDINGS2019036092.

[224] S. Aydin and M. N. Aydin, "Semantic and syntactic interoperability for agricultural open-data platforms in the context of IoT using crop-specific trait ontologies," *Appl. Sci.*, vol. 10, no. 13, 2020, doi: 10.3390/app10134460.

[225] Y. He, J. Guo, and X. Zheng, "From Surveillance to Digital Twin: Challenges and Recent Advances of Signal Processing for Industrial Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 120–129, Sep. 2018, doi: 10.1109/MSP.2018.2842228.

[226] A. Villa-Henriksen, G. T. C. Edwards, L. A. Pesonen, O. Green, and C. A. G. Sørensen, "Internet of Things in arable farming: Implementation, applications, challenges and potential," *Biosyst. Eng.*, vol. 191, pp. 60–84, Mar. 2020, doi: 10.1016/J.BIOSYSTEMSENG.2019.12.013.

[227] H. M. Jawad, R. Nordin, S. K. Gharghan, A. M. Jawad, and M. Ismail, "Energy-efficient wireless sensor networks for precision agriculture: A review," *Sensors (Switzerland)*, vol. 17, no. 8, p. 1781, Aug. 2017, doi: 10.3390/s17081781.

[228] L. Sigrist, N. Stricker, D. Bernath, J. Beutel, and L. Thiele, "Thermoelectric Energy Harvesting from Gradients in the Earth Surface," *IEEE Trans. Ind. Electron.*, vol. 67, no. 11, pp. 9460–9470, Nov. 2020, doi: 10.1109/TIE.2019.2952796.

[229] N. Brinis and L. A. Saidane, "Context Aware Wireless Sensor Network Suitable for Precision Agriculture," *Wirel. Sens. Netw.*, 2016, doi: 10.4236/wsn.2016.81001.

[230] G. Idoje, T. Dagiuklas, and M. Iqbal, "Survey for smart farming technologies: Challenges and issues," *Comput. Electr. Eng.*, vol. 92, p. 107104, Jun. 2021, doi: 10.1016/J.COMPELECENG.2021.107104.

[231] M. Zimmerling, L. Mottola, and S. Santini, "Synchronous Transmissions in Low-Power Wireless: A Survey of Communication Protocols and Network Services," *ACM Comput. Surv.*, vol. 53, no. 6, Feb. 2021, doi: 10.1145/3410159.

[232] F. Tonolini and F. Adib, "Networking across boundaries: Enabling wireless communication through the water-air interface," *SIGCOMM 2018 - Proc. 2018 Conf. ACM Spec. Interes. Gr. Data Commun.*, pp. 117–131, Aug. 2018, doi: 10.1145/3230543.3230580.

[233] L. Chen *et al.*, "Robustness, Security and Privacy in Location-Based Services for Future IoT: A Survey," *IEEE Access*, vol. 5, pp. 8956–8977, 2017, doi: 10.1109/ACCESS.2017.2695525.

[234] Y. Njah and M. Cheriet, "Parallel Route Optimization and Service Assurance in Energy-Efficient Software-Defined Industrial IoT Networks," *IEEE Access*, vol. 9, pp. 24682–24696, 2021, doi: 10.1109/ACCESS.2021.3056931.

[235] A. Rajput and V. B. Kumaravelu, "Scalable and sustainable wireless sensor networks for agricultural application of Internet of things using fuzzy c-means algorithm," *Sustain. Comput. Informatics Syst.*, vol. 22, pp. 62–74, Jun. 2019, doi: 10.1016/J.SUSCOM.2019.02.003.

[236] B. B. Sinha and R. Dhanalakshmi, "Recent advancements and challenges of Internet of Things in smart agriculture: A survey," *Futur. Gener. Comput. Syst.*, vol. 126, pp. 169–184, Jan. 2022, doi: 10.1016/J.FUTURE.2021.08.006.

[237] F. Caffaro and E. Cavallo, "The effects of individual variables, farming system characteristics and perceived barriers on actual use of smart farming technologies: Evidence from the piedmont region, northwestern Italy," *Agric.*, vol. 9, no. 5, May 2019, doi: 10.3390/AGRICULTURE9050111.

[238] JainMohit, KumarPratyush, BhansaliIshita, L. Vera, TruongKhai, and PatelShwetak, "FarmChat," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–22, Dec. 2018, doi: 10.1145/3287048.

[239] B. Mclaughlan, J. Brandli, and F. Smith, "Toward Sustainable High-Yield Agriculture via

Intelligent Control Systems," no. May, 2015.

[240] R. K. Kodali, S. Soratkal, and L. Boppana, "IOT based control of appliances," *Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2016*, pp. 1293–1297, Jan. 2017, doi: 10.1109/CCAA.2016.7813918.

[241] R. Abbasi, A. Reyes, E. Martinez, and R. Ahmad, "Real-time implementation of digital twin for robot based production line," pp. 4–6.

[242] O. Bermeo-Almeida, M. Cardenas-Rodriguez, T. Samaniego-Cobo, E. Ferruzola-Gómez, R. Cabezas-Cabezas, and W. Bazán-Vera, "Blockchain in Agriculture: A Systematic Literature Review," *Commun. Comput. Inf. Sci.*, vol. 883, pp. 44–56, Nov. 2018, doi: 10.1007/978-3-030-00940-3_4.

[243] V. Saiz-Rubio and F. Rovira-Más, "From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management," *Agron. 2020, Vol. 10, Page 207*, vol. 10, no. 2, p. 207, Feb. 2020, doi: 10.3390/AGRONOMY10020207.

[244] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, "Industry 4.0 and Industry 5.0—Inception, conception and perception," *J. Manuf. Syst.*, vol. 61, pp. 530–535, Oct. 2021, doi: 10.1016/J.JMSY.2021.10.006.

[245] P. K. R. Maddikunta *et al.*, "Industry 5.0: A survey on enabling technologies and potential applications," *J. Ind. Inf. Integr.*, p. 100257, Aug. 2021, doi: 10.1016/J.JII.2021.100257.

[246] P. Martinez and R. Ahmad, "AllFactory : an aquaponics 4 . 0 transdisciplinary educational and applied research learning factory at the University of Alberta," pp. 5–7.

[247] P. Cooper, "Data, information, knowledge and wisdom," *Anaesth. Intensive Care Med.*, vol. 18, no. 1, pp. 55–56, Jan. 2017, doi: 10.1016/J.MPAIC.2016.10.006.

[248] A. Rhayem, M. B. A. Mhiri, and F. Gargouri, "Semantic Web Technologies for the Internet of Things: Systematic Literature Review," *Internet of Things*, vol. 11, p. 100206, 2020, doi: 10.1016/j.iot.2020.100206.

[249] B. Drury, R. Fernandes, M. F. Moura, and A. de Andrade Lopes, "A survey of semantic web technology for agriculture," *Inf. Process. Agric.*, vol. 6, no. 4, pp. 487–501, 2019, doi: 10.1016/j.inpa.2019.02.001.

[250] K. Munir and M. Sheraz Anjum, "The use of ontologies for effective knowledge modelling and information retrieval," *Applied Computing and Informatics*. 2018, doi: 10.1016/j.aci.2017.07.003.

[251] C. Martinez-Cruz, I. J. Blanco, and M. A. Vila, "Ontologies versus relational databases: Are they so different? A comparison," *Artificial Intelligence Review*. 2012, doi: 10.1007/s10462-011-9251-9.

[252] R. Tyson, "Sustainable Aquaponic Vegetable and Fish," *Orange Cty. Ext. fact sheet 1-13*, 2005.

[253] N. Hari Kumar, S. Baskaran, S. Hariraj, and V. Krishnan, "An autonomous aquaponics system using 6LoWPAN based WSN," 2016, doi: 10.1109/W-FiCloud.2016.37.

[254] J. Hwang and H. Yoe, "Study on the context-aware middleware for ubiquitous greenhouses using wireless sensor networks," *Sensors*, 2011, doi: 10.3390/s110504539.

[255] T. Davenport, D. De Long, and M. Beers, "Successful Knowledge Management Projects," *Sloan Manage. Rev.*, 1998.

[256] V. Devedzic, "Knowledge modeling - State of the art," *Integrated Computer-Aided Engineering*. 2001, doi: 10.3233/ica-2001-8307.

[257] A. T. Bimba *et al.*, "Towards knowledge modeling and manipulation technologies: A survey," *Int. J. Inf. Manage.*, 2016, doi: 10.1016/j.ijinfomgt.2016.05.022.

[258] T. R. Gruber, “A translation approach to portable ontology specifications,” *Knowl. Acquis.*, 1993, doi: 10.1006/knac.1993.1008.

[259] F. Zhang, Z. M. Ma, and W. Li, “Storing OWL ontologies in object-oriented databases,” *Knowledge-Based Syst.*, 2015, doi: 10.1016/j.knosys.2014.12.020.

[260] L. Ramos, “Semantic Web for manufacturing, trends and open issues: Toward a state of the art,” *Computers and Industrial Engineering*. 2015, doi: 10.1016/j.cie.2015.10.013.

[261] A. Maalel, L. Mejri, H. H. Mabrouk, and H. Ben Ghezela, “Towards an ontology of help to the modeling of accident scenarii: Application on railroad transport,” 2012, doi: 10.1109/SETIT.2012.6481880.

[262] C. X. Zhang, C. G. Cao, F. Gu, and J. X. Si, “Domain-specific formal ontology of archaeology and its application in knowledge acquisition and analysis,” *J. Comput. Sci. Technol.*, 2004, doi: 10.1007/BF02944899.

[263] C. Zhang, C. Cao, Y. Sui, and X. Wu, “A Chinese time ontology for the Semantic Web,” *Knowledge-Based Syst.*, 2011, doi: 10.1016/j.knosys.2011.04.021.

[264] E. Miller, “An introduction to the Resource Description Framework,” *D-Lib Magazine*. 1998, doi: 10.1045/may98-miller.

[265] World Wide Web Consortium, “Semantic Web - W3C,” *W3.org*, 2012. .

[266] D. Allemang and J. Hendler, *Semantic Web for the Working Ontologist: Effective modeling in RDFS and OWL*. 2011.

[267] G. Antoniou, P. Groth, F. van Harmelen, and R. Hoekstra, *A Semantic Web Primer (Third Edition)*. 2012.

[268] I. Horrocks, P. F. Patel-Schneider, S. Bechhofer, and D. Tsarkov, “OWL rules: A proposal and prototype implementation,” *Web Semant.*, 2005, doi: 10.1016/j.websem.2005.05.003.

[269] M. A. Musen, “The protégé project,” *AI Matters*, 2015, doi: 10.1145/2757001.2757003.

[270] M. Kim, H. Kang, S. Kwon, Y. Lee, K. Kim, and C. S. Pyo, “Augmented ontology by handshaking with machine learning,” *Int. Conf. Adv. Commun. Technol. ICACT*, pp. 740–743, 2017, doi: 10.23919/ICACT.2017.7890191.

[271] A. Sheth and C. Ramakrishnan, “Semantic (Web) Technology In Action: Ontology Driven Information Systems for Search, Integration and Analysis,” *IEEE Data Eng. Bull. Spec. issue Mak. Semant. Web Real*, vol. 26, no. 4, pp. 40–8, 2003, [Online]. Available: http://knoesis.wright.edu/library/resource.php?id=00149.

[272] T. Hagedorn, M. Bone, B. Kruse, I. Grosse, and M. Blackburn, “Knowledge Representation with Ontologies and Semantic Web Technologies to Promote Augmented and Artificial Intelligence in Systems Engineering,” *Insight*, vol. 23, no. 1, pp. 15–20, 2020, doi: 10.1002/inst.12279.

[273] H. Zheng, Y. Wang, C. Han, F. Le, R. He, and J. Lu, “Learning and Applying Ontology for Machine Learning in Cyber Attack Detection,” *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 1309–1315, 2018, doi: 10.1109/TrustCom/BigDataSE.2018.00180.

[274] V. Swathi Lakshmi, V. Nithya, K. Sripriya, C. Preethi, and K. Logeshwari, “Prediction of diabetes patient stage using ontology based machine learning system,” *2019 IEEE Int. Conf. Syst. Comput. Autom. Networking, ICSCAN 2019*, pp. 2019–2022, 2019, doi: 10.1109/ICSCAN.2019.8878831.

[275] C. Steinmetz, A. Rettberg, F. G. C. Ribeiro, G. Schroeder, and C. E. Pereira, “Internet of things ontology for digital twin in cyber physical systems,” *Brazilian Symp. Comput. Syst. Eng. SBESC*, vol. 2018-Novem, pp. 154–159, 2018, doi: 10.1109/SBESC.2018.00030.

[276] J. Rogushina, A. Gladun, and S. Pryima, “Use of ontologies for metadata records analysis in big data,” *CEUR Workshop Proc.*, vol. 2318, pp. 46–63, 2018.

[277] M. Alkahtani, A. Choudhary, A. De, and J. A. Harding, “A decision support system based on ontology and data mining to improve design using warranty data,” *Comput. Ind. Eng.*, vol. 128, pp. 1027–1039, 2019, doi: 10.1016/j.cie.2018.04.033.

[278] R. Santodomingo, S. Rohjans, M. Uslar, J. A. Rodríguez-Mondéjar, and M. A. Sanz-Bobi, “Ontology matching system for future energy smart grids,” *Eng. Appl. Artif. Intell.*, vol. 32, pp. 242–257, 2014, doi: 10.1016/j.engappai.2014.02.005.

[279] N. Moran, S. Nieland, G. Tintrup gen. Suntrup, and B. Kleinschmit, “Combining machine learning and ontological data handling for multi-source classification of nature conservation areas,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 54, pp. 124–133, 2017, doi: 10.1016/j.jag.2016.09.009.

[280] P. Delgoshaei, M. Heidarinejad, and M. A. Austin, “COMBINED ONTOLOGY-DRIVEN AND MACHINE LEARNING APPROACH TO MONITORING OF BUILDING ENERGY CONSUMPTION 1 National Institute of Standards and Technology (NIST), Gaithersburg , MD , USA 2 Illinois Institute of Technology , Chicago , IL , USA 3 University of ,” no. i, pp. 667–674, 2018.

[281] S. S. L. Chukkapalli, A. Piplai, S. Mittal, M. Gupta, and A. Joshi, “A Smart-Farming Ontology for Attribute Based Access Control,” 2020, doi: 10.1109/BigDataSecurity-HPSC-IDS49724.2020.00017.

[282] “A smart ontology-based IoT framework for remote patient monitoring _ Elsevier Enhanced Reader.pdf.” .

[283] S. S. L. Chukkapalli *et al.*, “Ontologies and artificial intelligence systems for the cooperative smart farming ecosystem,” *IEEE Access*, vol. 8, pp. 164045–164064, 2020, doi: 10.1109/ACCESS.2020.3022763.

[284] S. Dutta, S. S. L. Chukkapalli, M. Sulgekar, S. Krithivasan, P. K. Das, and A. Joshi, “Context Sensitive Access Control in Smart Home Environments,” *Proc. - 2020 IEEE 6th Intl Conf. Big Data Secur. Cloud, BigDataSecurity 2020, 2020 IEEE Intl Conf. High Perform. Smart Comput. HPSC 2020 2020 IEEE Intl Conf. Intell. Data Secur. IDS 2020*, pp. 35–41, 2020, doi: 10.1109/BigDataSecurity-HPSC-IDS49724.2020.00018.

[285] S. Sree and L. Chukkapalli, *Ontology driven AI and Access Control Systems for Smart Fisheries Ontology driven AI and Access Control Systems for Smart Fisheries*, vol. 1, no. 1. Association for Computing Machinery, 2021.

[286] S. Raaijmakers and C. Brewster, “Exploiting ontologies for deep learning: A case for sentiment mining,” *CEUR Workshop Proc.*, vol. 2198, pp. 1–4, 2018.

[287] “light weight dynamic ontology for IOT.pdf.” .

[288] D. L. McGuinness and F. van Harmelen, “OWL Web Ontology Language Overview,” *W3C recommendation*. 2004.

[289] “Homepage | AGROVOC.” http://www.fao.org/agrovoc/ (accessed Sep. 01, 2021).

[290] E. Arnaud *et al.*, “Towards a reference plant trait ontology for modeling knowledge of plant traits and phenotypes,” 2012, doi: 10.5220/0004138302200225.

[291] A. Thunkijjanukij, A. Kawtrakul, S. Panichsakpatana, and U. Veesommai, “Ontology development: A case study for thai rice,” *Kasetsart J. - Nat. Sci.*, 2009.

[292] M. Sodanil, P. Phonarin, and N. Porrawatpreyakorn, “An ontology-based query expansion for an agricultural expert retrieval system,” 2013, doi: 10.1145/2539150.2539220.

[293] P. Phonarin, S. Nitsuwat, and C. Haruechaiyasak, “AGRIX: An ontology based agricultural expertise retrieval framework,” 2012, doi: 10.4028/www.scientific.net/AMR.403-408.3714.

[294] S. Pokharel, M. A. Sherif, and J. Lehmann, "Ontology based data access and integration for improving the effectiveness of farming in Nepal," 2014, doi: 10.1109/WI-IAT.2014.114.

[295] H. Afzal and M. K. Kasi, "Ontology-based knowledge modeling for rice crop production," 2019, doi: 10.1109/FiCloud.2019.00057.

[296] E. Abrahao and A. R. Hirakawa, "Task ontology modeling for technical knowledge representation in agriculture field operations domain," 2017, doi: 10.1109/ICISE.2017.18.

[297] S. Aydin and M. N. Aydin, "Ontology-based data acquisition model development for agricultural open data platforms and implementation of OWL2MVC tool," *Comput. Electron. Agric.*, 2020, doi: 10.1016/j.compag.2020.105589.

[298] S. Sivamani, N. Bae, and Y. Cho, "A smart service model based on ubiquitous sensor networks using vertical farm ontology," *Int. J. Distrib. Sens. Networks*, vol. 2013, 2013, doi: 10.1155/2013/161495.

[299] S. Sivamani, H. Kim, M. Lee, J. Park, C. Shin, and Y. Cho, "An Ontology Model for Smart Service in Vertical Farms – An OWL-S Approach," *Int. J. u- e- Serv. Sci. Technol.*, 2016, doi: 10.14257/ijunesst.2016.9.1.18.

[300] T. Kim, N. Bae, M. Lee, C. Shin, Cho Yongyun, and J. Park, "A Study of an Agricultural Ontology Model for an Intelligent Service in a Vertical Farm," *Int. J. Smart Home*, vol. 7, no. 4, pp. 117–126, 2013, [Online]. Available: http://www.sersc.org/journals/IJSH/vol7_no4_2013/12.pdf.

[301] N. Husna, M. Hasbullah, and Z. Zakaria, "The Development of STIF in Ontology Based on The Methontology Approach," *UTM Comput. Proc. Innov. Comput. Technol. Appl.*, 2017.

[302] M. Fernandez, A. Gómez-Pérez, and N. Juristo, "Methontology: from ontological art towards ontological engineering," 1997.

[303] R. Ahmad, C. Masse, S. Jituri, J. Doucette, and P. Mertiny, "Alberta Learning Factory for training reconfigurable assembly process value stream mapping," 2018, doi: 10.1016/j.promfg.2018.04.023.

[304] N. Scattini and S. P. MAJ, "Aquaponics – A Process Control Approach," *Mod. Appl. Sci.*, vol. 11, no. 11, p. 43, 2017, doi: 10.5539/mas.v11n11p43.

[305] M. K. M. Zambri, M. S. M. Aras, A. Khamis, H. N. M. Shah, M. H. Harun, and A. A. Yusof, "Design and development of intelligent aquaponics system," *Int. J. Eng. Technol.*, 2018.

[306] T. Gu, X. H. Wang, H. K. Pung, and D. Q. Zhang, "An ontology-based context model in intelligent environments," *arXiv*. 2020.

[307] T. Brosnan and D. W. Sun, "Improving quality inspection of food products by computer vision - A review," *Journal of Food Engineering*. 2004, doi: 10.1016/S0260-8774(03)00183-3.

[308] A. Estellés and A. Alcina, "A model for formalizing characteristics in Protégé-OWL," 2009.

[309] Health Canada, "Nutrient Value of Some Common Foods," *Can. Nutr. File*, 2005.

[310] S. Abburu, "A Survey on Ontology Reasoners and Comparison," *Int. J. Comput. Appl.*, 2012.

[311] N. F. Noy, R. W. Fergerson, and M. A. Musen, "The knowledge model of protégé-2000: Combining interoperability and flexibility," 2000, doi: 10.1007/3-540-39967-4_2.

[312] C. Zhang, G. Zhou, and Q. Lu, "Decision support oriented ontological modeling of product knowledge," 2018, doi: 10.1109/ITNEC.2017.8284816.

[313] N. Sharma, S. Acharya, K. Kumar, N. Singh, and O. P. Chaurasia, "Hydroponics as an advanced technique for vegetable production: An overview," *J. Soil Water Conserv.*, 2018, doi: 10.5958/2455-7145.2018.00056.5.

[314] R. I. Danner, U. Mankasingh, K. Anamthawat-Jonsson, and R. I. Thorarinsdottir, "Designing

aquaponic production systems towards integration into greenhouse farming," *Water (Switzerland)*, 2019, doi: 10.3390/w11102123.

[315] M. Varga, B. Csukas, and B. Kucska, "Implementation of an easily reconfigurable dynamic simulator for recirculating aquaculture systems," *Aquac. Eng.*, vol. 90, no. December 2019, p. 102073, 2020, doi: 10.1016/j.aquaeng.2020.102073.

[316] D. Karimanzira, K. J. Keesman, W. Kloas, D. Baganz, and T. Rauschenbach, "Dynamic modeling of the INAPRO aquaponic system," *Aquac. Eng.*, vol. 75, pp. 29–45, 2016, doi: 10.1016/j.aquaeng.2016.10.004.

[317] A. Survey, V. M. Kushala, and M. C. Supriya, "Deep Learning Ontology : Dimensions in the Field of," *Int. J. Latest Technol. Eng. Manag. Appl. Sci.*, vol. VII, no. I, pp. 17–21, 2018.

[318] I. Naja, M. Markovic, P. Edwards, and C. Cottrill, "A Semantic Framework to Support AI System Accountability and Audit," pp. 1–16.

[319] G. J. Gooley and F. M. Gavine, *Integrated systems Agri-Aquaculture Systems A resouce handbook for australian industry development.* 2003.

[320] W. Lennard and S. Goddek, "Aquaponics: The Basics," in *Aquaponics Food Production Systems*, 2019.

[321] B. C. Mohapatra, N. K. Chandan, S. K. Panda, D. Majhi, and B. R. Pillai, "Design and development of a portable and streamlined nutrient film technique (NFT) aquaponic system," *Aquac. Eng.*, 2020, doi: 10.1016/j.aquaeng.2020.102100.

[322] S. El Kadiri *et al.*, "Current trends on ICT technologies for enterprise information systems," *Comput. Ind.*, 2016, doi: 10.1016/j.compind.2015.06.008.

[323] R. Ahmad, S. Tichadou, and J. Y. Hascoet, "A knowledge-based intelligent decision system for production planning," *Int. J. Adv. Manuf. Technol.*, 2017, doi: 10.1007/s00170-016-9214-z.

[324] E. M. Sanfilippo, "Ontological foundations for feature-based modeling," 2018, doi: 10.1016/j.procir.2018.03.002.

[325] P. Martinez, R. Ahmad, and M. Al-Hussein, "Automatic Selection Tool of Quality Control Specifications for Off-site Construction Manufacturing Products: A BIM-based Ontology Model Approach," *Modul. Offsite Constr. Summit Proc.*, no. July, pp. 141–148, 2019, doi: 10.29173/mocs87.

[326] R. Bulgari, A. Baldi, A. Ferrante, and A. Lenzi, "Yield and quality of basil, Swiss chard, and rocket microgreens grown in a hydroponic system," *New Zeal. J. Crop Hortic. Sci.*, 2017, doi: 10.1080/01140671.2016.1259642.

[327] Haryanto, M. Ulum, A. F. Ibadillah, R. Alfita, K. Aji, and R. Rizkyandi, "Smart aquaponic system based Internet of Things (IoT)," *J. Phys. Conf. Ser.*, vol. 1211, no. 1, p. 012047, Apr. 2019, doi: 10.1088/1742-6596/1211/1/012047.

[328] A. K. Pasha, E. Mulyana, C. Hidayat, M. A. Ramdhani, O. T. Kurahman, and M. Adhipradana, "System Design of Controlling and Monitoring on Aquaponic Based on Internet of Things," *Proceeding 2018 4th Int. Conf. Wirel. Telemat. ICWT 2018*, Nov. 2018, doi: 10.1109/ICWT.2018.8527802.

[329] A. M. Nagayo, C. Mendoza, E. Vega, R. K. S. Al Izki, and R. S. Jamisola, "An automated solar-powered aquaponics system towards agricultural sustainability in the Sultanate of Oman," *2017 IEEE Int. Conf. Smart Grid Smart Cities, ICSGSC 2017*, pp. 42–49, Sep. 2017, doi: 10.1109/ICSGSC.2017.8038547.

[330] D. Wang, J. Zhao, L. Huang, and D. Xu, "Design of A Smart Monitoring and Control System for Aquaponics Based on OpenWrt," *Proc. 5th Int. Conf. Inf. Eng. Mech. Mater.*, vol. 21, pp. 937–942, Jul. 2015, doi: 10.2991/ICIMM-15.2015.171.

[331] A. Reyes-Yanes, S. Gelio, P. Martinez, and R. Ahmad, "Wireless Sensing Module for IoT

Aquaponics Database Construction," *Int. J. Electron. Electr. Eng.*, vol. 9, no. 2, pp. 43–47, Jun. 2021, doi: 10.18178/IJEEE.9.2.43-47.

[332] M. Abbasi, M. H. Yaghmaee, and F. Rahnama, "Internet of Things in agriculture: A survey," *Proc. 3rd Int. Conf. Internet Things Appl. IoT 2019*, Apr. 2019, doi: 10.1109/IICITA.2019.8808839.

[333] S. A. Gillani, R. Abbasi, P. Martinez, and R. Ahmad, "Review on Energy Efficient Artificial Illumination in Aquaponics," *Clean. Circ. Bioeconomy*, vol. 2, p. 100015, Aug. 2022, doi: 10.1016/J.CLCB.2022.100015.

[334] S. B. Dhal, M. Bagavathiannan, U. Braga-Neto, and S. Kalafatis, "Nutrient optimization for plant growth in Aquaponic irrigation using Machine Learning for small training datasets," *Artif. Intell. Agric.*, vol. 6, pp. 68–76, Jan. 2022, doi: 10.1016/J.AIIA.2022.05.001.

[335] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," *Proc. - 1st Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2015*, pp. 768–771, Jul. 2015, doi: 10.1109/ICCUBEA.2015.153.

[336] M. Dutot, L. M. Nelson, and R. C. Tyson, "Predicting the spread of postharvest disease in stored fruit, with application to apples," *Postharvest Biol. Technol.*, vol. 85, pp. 45–56, Nov. 2013, doi: 10.1016/J.POSTHARVBIO.2013.04.003.

[337] A. I. Khan, S. M. K. Quadri, S. Banday, and J. Latief Shah, "Deep diagnosis: A real-time apple leaf disease detection system based on deep learning," *Comput. Electron. Agric.*, vol. 198, p. 107093, Jul. 2022, doi: 10.1016/J.COMPAG.2022.107093.

[338] W. N. Weaver, J. Ng, and R. G. Laport, "LeafMachine: Using machine learning to automate leaf trait extraction from digitized herbarium specimens," *Appl. Plant Sci.*, vol. 8, no. 6, Jun. 2020, doi: 10.1002/APS3.11367.

[339] P. Bedi and P. Gole, "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network," *Artif. Intell. Agric.*, vol. 5, pp. 90–101, Jan. 2021, doi: 10.1016/J.AIIA.2021.05.002.

[340] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A Dataset for Visual Plant Disease Detection," *ACM Int. Conf. Proceeding Ser.*, pp. 249–253, Nov. 2019, doi: 10.1145/3371158.3371196.

[341] M. Á. Rodríguez-García, F. García-Sánchez, and R. Valencia-García, "Knowledge-Based System for Crop Pests and Diseases Recognition," *Electron. 2021, Vol. 10, Page 905*, vol. 10, no. 8, p. 905, Apr. 2021, doi: 10.3390/ELECTRONICS10080905.

[342] R. Studer, V. R. Benjamins, and D. Fensel, "Knowledge Engineering: Principles and methods," *Data Knowl. Eng.*, 1998, doi: 10.1016/S0169-023X(97)00056-6.

[343] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, "Detection of rice plant diseases based on deep transfer learning," *J. Sci. Food Agric.*, vol. 100, no. 7, pp. 3246–3256, May 2020, doi: 10.1002/JSFA.10365.

[344] J. Ma, K. Du, F. Zheng, L. Zhang, and Z. Sun, "Disease recognition system for greenhouse cucumbers based on deep convolutional neural network," *Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric. Eng.*, vol. 34, no. 12, pp. 186–192, Jun. 2018, doi: 10.11975/J.ISSN.1002-6819.2018.12.022.

[345] C. R. Rahman *et al.*, "Identification and recognition of rice diseases and pests using convolutional neural networks," *Biosyst. Eng.*, vol. 194, pp. 112–120, Jun. 2020, doi: 10.1016/J.BIOSYSTEMSENG.2020.03.020.

[346] D. Oppenheim, G. Shani, O. Erlich, and L. Tsror, "Using deep learning for image-based potato tuber disease detection," *Phytopathology*, vol. 109, no. 6, pp. 1083–1087, 2019, doi: 10.1094/PHYTO-08-18-0288-R.

[347] X. Fan, J. Zhou, Y. Xu, and X. Peng, "Corn Disease Recognition under Complicated

Background Based on Improved Convolutional Neural Network," *Nongye Jixie Xuebao/Transactions Chinese Soc. Agric. Mach.*, vol. 52, no. 3, pp. 210–217, Mar. 2021, doi: 10.6041/J.ISSN.1000-1298.2021.03.023.

[348] J. Qi *et al.*, "An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease," *Comput. Electron. Agric.*, vol. 194, p. 106780, Mar. 2022, doi: 10.1016/J.COMPAG.2022.106780.

[349] M. Nandhini, K. U. Kala, M. Thangadarshini, and S. Madhusudhana Verma, "Deep Learning model of sequential image classifier for crop disease detection in plantain tree cultivation," *Comput. Electron. Agric.*, vol. 197, p. 106915, Jun. 2022, doi: 10.1016/J.COMPAG.2022.106915.

[350] A. Abbas, S. Jain, M. Gour, and S. Vankudothu, "Tomato plant disease detection using transfer learning with C-GAN synthetic images," *Comput. Electron. Agric.*, vol. 187, p. 106279, Aug. 2021, doi: 10.1016/J.COMPAG.2021.106279.

[351] C. Liu, H. Zhu, W. Guo, X. Han, C. Chen, and H. Wu, "EFDet: An efficient detection method for cucumber disease under natural complex environments," *Comput. Electron. Agric.*, vol. 189, p. 106378, Oct. 2021, doi: 10.1016/J.COMPAG.2021.106378.

[352] A. Musa, M. Hamada, F. M. Aliyu, and M. Hassan, "An Intelligent Plant Dissease Detection System for Smart Hydroponic Using Convolutional Neural Network," *Proc. - 2021 IEEE 14th Int. Symp. Embed. Multicore/Many-Core Syst. MCSoC 2021*, pp. 345–351, 2021, doi: 10.1109/MCSOC51149.2021.00058.

[353] K. Lisha Kamala and S. Anna Alex, "Apple Fruit Disease Detection for Hydroponic plants using Leading edge Technology Machine Learning and Image Processing," *Proc. - 2nd Int. Conf. Smart Electron. Commun. ICOSEC 2021*, pp. 820–825, 2021, doi: 10.1109/ICOSEC51865.2021.9591903.

[354] I. Yudha Pratama, A. Wahab, and M. Alaydrus, "Deep learning for assessing unhealthy lettuce hydroponic using convolutional neural network based on faster R-CNN with Inception V2," *2020 5th Int. Conf. Informatics Comput. ICIC 2020*, Nov. 2020, doi: 10.1109/ICIC50835.2020.9288554.

[355] M. A. Noyan, "Uncovering bias in the PlantVillage dataset," Jun. 2022, doi: 10.48550/arxiv.2206.04374.

[356] Y. Y. Zheng, J. L. Kong, X. B. Jin, X. Y. Wang, T. L. Su, and M. Zuo, "CropDeep: The Crop Vision Dataset for Deep-Learning-Based Classification and Detection in Precision Agriculture," *Sensors 2019, Vol. 19, Page 1058*, vol. 19, no. 5, p. 1058, Mar. 2019, doi: 10.3390/S19051058.

[357] W. Jearanaiwongkul, C. Anutariya, T. Racharak, and F. Andres, "An Ontology-Based Expert System for Rice Disease Identification and Control Recommendation," *Appl. Sci. 2021, Vol. 11, Page 10450*, vol. 11, no. 21, p. 10450, Nov. 2021, doi: 10.3390/APP112110450.

[358] "Vegetable Pathology – Long Island Horticultural Research & Extension Center." https://blogs.cornell.edu/livegpath/ (accessed Oct. 19, 2022).

[359] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Inf.*, vol. 11, no. 2, Feb. 2020, doi: 10.3390/INFO11020125.

[360] "Illustrated: 10 CNN Architectures | by Raimi Karim | Towards Data Science." https://towardsdatascience.com/illustrated-10-cnn-architectures-95d78ace614d (accessed Jul. 07, 2022).

[361] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 770–778, Dec. 2015, doi: 10.48550/arxiv.1512.03385.

[362] D. Theckedath and R. R. Sedamkar, "Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks," *SN Comput. Sci.*, vol. 1, no. 2, pp. 1–7, 2020, doi: 10.1007/s42979-020-0114-9.

[363] N. D. Nguyen, T. Do, T. D. Ngo, and D. D. Le, "An Evaluation of Deep Learning Methods for Small Object Detection," *J. Electr. Comput. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/3189691.

[364] Glenn, "ultralytics/yolov5." https://github.com/ultralytics/yolov5 (accessed Jul. 08, 2022).

[365] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2015, doi: 10.48550/arxiv.1506.01497.

[366] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/S11263-015-0816-Y.

[367] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," Dec. 2016, doi: 10.48550/arxiv.1612.03144.

[368] Margaret Tuttle McGrath, "Pest management," *Cornell University*, 2021. https://www.vegetables.cornell.edu/pest-management/ (accessed Aug. 03, 2022).

[369] M. F. Taha *et al.*, "Using Deep Convolutional Neural Network for Image-Based Diagnosis of Nutrient Deficiencies in Plants Grown in Aquaponics," *Chemosens. 2022, Vol. 10, Page 45*, vol. 10, no. 2, p. 45, Jan. 2022, doi: 10.3390/CHEMOSENSORS10020045.

[370] S. E. Cook and R. G. V. Bramley, "Coping with variability in agricultural production -implications for soil testing and fertiliser management," *https://doi.org/10.1080/00103620009370524*, vol. 31, no. 11–14, pp. 1531–1551, 2008, doi: 10.1080/00103620009370524.

[371] J. Kjeldahl, "Neue Methode zur Bestimmung des Stickstoffs in organischen Körpern," *Zeitschrift für Anal. Chemie 1883 221*, vol. 22, no. 1, pp. 366–382, Dec. 1883, doi: 10.1007/BF01338151.

[372] W. H. Yang, S. Peng, J. Huang, A. L. Sanico, R. J. Buresh, and C. Witt, "Using Leaf Color Charts to Estimate Leaf Nitrogen Status of Rice," *Agron. J.*, vol. 95, no. 1, pp. 212–217, Jan. 2003, doi: 10.2134/AGRONJ2003.2120.

[373] J. Markwell, J. C. Osterman, and J. L. Mitchell, "Calibration of the Minolta SPAD-502 leaf chlorophyll meter," *Photosynth. Res. 1995 463*, vol. 46, no. 3, pp. 467–472, Jan. 1995, doi: 10.1007/BF00032301.

[374] H. Zheng *et al.*, "Evaluation of RGB, Color-Infrared and Multispectral Images Acquired from Unmanned Aerial Systems for the Estimation of Nitrogen Accumulation in Rice," *Remote Sens. 2018, Vol. 10, Page 824*, vol. 10, no. 6, p. 824, May 2018, doi: 10.3390/RS10060824.

[375] M. Tao *et al.*, "Smartphone-based detection of leaf color levels in rice plants," *Comput. Electron. Agric.*, vol. 173, p. 105431, Jun. 2020, doi: 10.1016/J.COMPAG.2020.105431.

[376] D. D. Burdescu, M. Brezovan, E. Ganea, and L. Stanescu, "A new method for segmentation of images represented in a HSV color space," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5807 LNCS, pp. 606–617, 2009, doi: 10.1007/978-3-642-04697-1_57/COVER/.

[377] "Streamlit • The fastest way to build and share data apps." https://streamlit.io/ (accessed Jul. 06, 2022).

[378] R. Yang *et al.*, "Detection of abnormal hydroponic lettuce leaves based on image processing and machine learning," *Inf. Process. Agric.*, Nov. 2021, doi: 10.1016/J.INPA.2021.11.001.

[379] S. Maity *et al.*, "Fault Area Detection in Leaf Diseases Using K-Means Clustering," *Proc. 2nd Int. Conf. Trends Electron. Informatics, ICOEI 2018*, pp. 1538–1542, Nov. 2018, doi:

10.1109/ICOEI.2018.8553913.

[380] W. Yang, S. Wang, X. Zhao, J. Zhang, and J. Feng, "Greenness identification based on HSV decision tree," *Inf. Process. Agric.*, vol. 2, no. 3–4, pp. 149–160, Oct. 2015, doi: 10.1016/J.INPA.2015.07.003.

[381] B. Luna-Benoso, J. C. Martínez-Perales, J. Cortés-Galicia, R. Flores-Carapia, and V. M. Silva-García, "Detection of Diseases in Tomato Leaves by Color Analysis," *Electron. 2021, Vol. 10, Page 1055*, vol. 10, no. 9, p. 1055, Apr. 2021, doi: 10.3390/ELECTRONICS10091055.

[382] S. Hasan, S. Jahan, and M. I. Islam, "Disease detection of apple leaf with combination of color segmentation and modified DWT," *J. King Saud Univ. - Comput. Inf. Sci.*, Jul. 2022, doi: 10.1016/J.JKSUCI.2022.07.004.

[383] J. Y. Lu, C. L. Chang, and Y. F. Kuo, "Monitoring Growth Rate of Lettuce Using Deep Convolutional Neural Networks," *2019 ASABE Annu. Int. Meet.*, pp. 1-, 2019, doi: 10.13031/AIM.201900341.

[384] I. Zaman *et al.*, "Effect of Plant Spacings on Growth, Physiology, Yield and Fiber Quality Attributes of Cotton Genotypes under Nitrogen Fertilization," *Agron. 2021, Vol. 11, Page 2589*, vol. 11, no. 12, p. 2589, Dec. 2021, doi: 10.3390/AGRONOMY11122589.

[385] S. Singh *et al.*, "Plant spacing and cultivar on quality attributes in sprouting broccoli," *South African J. Bot.*, vol. 148, pp. 737–741, Aug. 2022, doi: 10.1016/J.SAJB.2022.04.049.

[386] A. D. Nakarmi and L. Tang, "Automatic inter-plant spacing sensing at early growth stages using a 3D vision sensor," *Comput. Electron. Agric.*, vol. 82, pp. 23–31, Mar. 2012, doi: 10.1016/J.COMPAG.2011.12.011.

[387] A. Triki, B. Bouaziz, J. Gaikwad, and W. Mahdi, "Deep leaf: Mask R-CNN based leaf detection and segmentation from digitized herbarium specimen images," *Pattern Recognit. Lett.*, vol. 150, pp. 76–83, Oct. 2021, doi: 10.1016/J.PATREC.2021.07.003.

[388] J. Maloof, K. Nozue, … M. M.-J. (Journal of, and undefined 2013, "LeafJ: an ImageJ plugin for semi-automated leaf shape measurement," *jove.com*, Accessed: Jul. 14, 2022. [Online]. Available: https://www.jove.com/t/50028/leafj-an-imagej-plugin-for-semi-automated-leaf-shape-measurement.

[389] H. M. Easlon and A. J. Bloom, "Easy Leaf Area: Automated digital image analysis for rapid and accurate measurement of leaf area," *Appl. Plant Sci.*, vol. 2, no. 7, p. 1400033, Jul. 2014, doi: 10.3732/APPS.1400033.

[390] J. Gaikwad, A. Triki, … B. B.-I. S. and, and undefined 2019, "Measuring morphological functional leaf traits from digitized herbarium specimens using TraitEx Software," *search.proquest.com*, Accessed: Jul. 14, 2022. [Online]. Available: https://search.proquest.com/openview/8c7980e7e2dcbe9723923204a730976f/1?pq-origsite=gscholar&cbl=2049297.

[391] A. Hirigoyen *et al.*, "A machine learning approach to model leaf area index in Eucalyptus plantations using high-resolution satellite imagery and airborne laser scanner data," *Ann. For. Res*, vol. 64, no. 2, pp. 165–183, 2021, doi: 10.15287/afr.2021.2073.

[392] P. Juyal and S. Sharma, "Estimation of Tree Volume Using Mask R-CNN based Deep Learning," *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, Jul. 2020, doi: 10.1109/ICCCNT49239.2020.9225509.

[393] B. R. Kang, H. Lee, K. Park, H. Ryu, and H. Y. Kim, "BshapeNet: Object detection and instance segmentation with bounding shape masks," *Pattern Recognit. Lett.*, vol. 131, pp. 449–455, Mar. 2020, doi: 10.1016/J.PATREC.2020.01.024.

[394] A. Dutta, A. Z.-P. of the 27th A. international, and undefined 2019, "The VIA annotation software for images, audio and video," *dl.acm.org*, pp. 2276–2279, Oct. 2019, doi: 10.1145/3343031.3350535.

[395] K. He, G. Gkioxari, P. Dollar, and R. Girshick, *Mask R-CNN*, vol. 2017-Octob. 2017, pp. 2980–2988.

[396] W. Abdulla, “Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow,” *GitHub Repos.*

[397] Y. Yu, K. Zhang, L. Yang, and D. Zhang, “Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN,” *Comput. Electron. Agric.*, vol. 163, p. 104846, Aug. 2019, doi: 10.1016/J.COMPAG.2019.06.001.

[398] A. Ayadi, A. Samet, F. D. B. De Beuvron, and C. Zanni-Merk, “Ontology population with deep learning-based NLP: A case study on the Biomolecular Network Ontology,” *Procedia Comput. Sci.*, vol. 159, pp. 572–581, 2019, doi: 10.1016/J.PROCS.2019.09.212.

Appendices

Appendix A: Supplementary Information on AquaONT (Chapter 3)

A1: Summary of different ontology-driven smart systems (Table A.1)

Table A- 1. Review on ontology-driven smart systems (Integration of ontology and AI techniques).

Sr. No.	Model	Domain	Purpose	Significance	Approach	Citations
1.	A semantic framework consisting of "RAInS" and "SAO" ontologies is developed for AI systems.	AI	AI system design and implementation	Provide a visual interface for designing accountability plans and managing accountability records.	Modeled the accountability information relevant to the design stage of the AI system.	[318]
2.	A deep learning-based NLP ontology population system is developed.	Biology	Systems and processes analysis	Automatic population of the Biomolecular Network Ontology (BNO).	Used the popular deep learning algorithm "Word2vec" to learn word embeddings using a shallow neural network	[398]
2	An ontology model "OntoSenticNet" is proposed consisting of a composite deep learning classifier.	Text analysis	Sentiment mining	Deep learning-based text mining with ontology-based information.	Used semantic information in ontologies for explaining deep text mining, using neural attention and word embeddings.	[286]
3	A lightweight dynamic ontology "LiO-IoT" using machine learning technique is proposed for IoT systems.	IoT	Fulfilling the requirements of IoT systems.	Ensure semantic interoperability among IoT devices and reducing complexity and introducing dynamicity in the ontology.	Used machine learning technique (clustering) to provide dynamic semantics automatically for including additional concepts.	[287]
4	A smart ontology-based IoT	Healthcare	Covid-19 detection	Control the spread of the	An ontology-based biosensor	[282]

	framework is proposed for the early detection of COVID-19 in patients.		and remote patient monitoring .	coronavirus.	is developed using sensory 1D biomedical Signals (ECG, PPG, temperature, and accelerometer) and machine learning techniques such as feature extraction and classification.	
5	A novel approach is proposed for generating ontology using a deep learning model.	Cyber security	Ontology learning and application for cyber-attack detection.	Improve intelligent intrusion detection for cybersecurity.	Used deep neural network and supervised classifier gradient boosting decision tree (GB) for ontology learning and application.	[273]
6	A decision support system is developed using ontology and machine learning techniques to predict disease and suggest solutions.	Healthca re	Prediction of diabetes stage in patients	Improve the healthcare diagnostic system.	Used ontology containing data such as disease symptoms, causes and treatments, Naive Bayes algorithm, and decision tree for identifying disease stage.	[274]
7	An ontology model is proposed for cyber-security of self-recovering smart grids.	Smart Grids	Monitoring and restoration of the operability of the Smart Grid power systems.	Ensure sustainability and functioning of power systems in the event of destabilization	Used conceptual bases of self-recovery of perspective energy systems in the conditions of information confrontation are employed to develop an ontology model.	[281]
8	An ontology model is proposed to represent	Cyber-physical system	Analysis of DT through	Highlights the main concepts involved in the	Used concepts already existing in the literature,	[275]

	Digital Twin (DT) in the context of a cyber-physical system (CPS).		knowledge representation.	development of DT.	such as the IoT-Lite ontology.	
9	A smart fisheries ontology is developed, and an Attribute-Based Access Control System (ABAC) is implemented to evaluate access requests in farms.	Aquaculture	Protect the internet-connected sensors from potential cyber-attacks and propose AI applications to aid the owners to effectively manage their fish farms.	Secure and efficient smart fisheries ecosystem.	Used an ecosystem consisting of virtual and digital entities and their interactions to develop an ontology model, which is then used to implement ABAC.	[285]
10	A method based on natural language analysis and big data ontology is proposed for metadata analysis,	Big data	Analysis of metadata records in big data.	Improve the efficiency of all stages of big data processing	Used data created by human experts and generated by electronic devices to populate the instances of the ontological big data model.	[276]
11	A decision support system based on ontology-based text mining, self-organizing maps (machine learning algorithm), reliability and cost optimization is developed for design improvement using warranty data.	Design	Identification of manufacturing faults and optimization of design parameters	Enhanced design reliability and reduced manufacturing costs.	Used ontology-based text mining approach to extract hidden knowledge in warranty database, data mining approach SOM to link information with manufacturing data, and statistical analysis for re-evaluation of parameters for cost analysis and design	[277]

					changes.	
12	A system “PALS (Privacy via Anomaly-detection System)” is developed based on ABAC and ontology model to execute access control decisions for smart homes.	Smart home	Detect the intrusions in the smart home.	Enable users to control devices at home by providing them with the ability to change the access control policy.	Used context-sensitive policies that were created based on online privacy policies available from cloud service providers like Google Nest suite of products.	[284]
13	Innovative ontology matching system is developed that finds complex correspondences by processing expert knowledge from external domain ontologies and by using novel matching methods.	Smart grids	Identify and solve interoperability issues within smart grids.	Improve performance of smart grids.	Used innovative graph-based matching methods and innovative mapping algorithms.	[278]
14	A novel ontology-based neural network model “OntoLSTM” is proposed for manufacturing time series classification.	Smart factory	Enhancing deep learning model by using semantic.	Improve products, processes, and decisions.	Used a core manufacturing process ontology to design deep neural networks.	[284]
15	A methodology for the combination of ontology-based knowledge management and machine learning for the classification of multiple spatial data sources.	Nature conservation	Determination of the grassland indicators wetness and alkalinity.	Improve interoperability, reproducibility and exchange of data.	Used a decision tree classifier (DT) approach in combination with ontological formalism to generate classification procedures and results.	[279]
16	A roadmap is presented combining ontology and machine learning approaches for intelligent	Energy consumption	Intelligent monitoring of building energy consumption	Enhanced energy utilization.	Used nearest neighbour algorithm to predict the electricity consumption based on raw	[280]

building monitoring.		data along with knowledge data related to occupancy.

A2: Ontology specification document for AquaONT (Table A.2)

Table A- 2. Ontology specification document for AquaONT describing the scope and purpose.

Ontology Requirement Specification Document	
1	**Purpose**
	The purpose of AquaONT is to structure, model and store the aquaponics 4.0 system's knowledge, and use it to enable data-driven decision making for farmers by developing a functional decision support system.
2	**Domain**
	Agriculture – Smart Indoor Soilless Farming – Aquaponics 4.0
3	**Scope**
	The ontology has to focus on the aquaponics 4.0 system – which is a digital farm based on a smart farming concept. The level of granularity is directly linked with the indoor vertical farming terminologies.
4	**Implementation Language**
	The ontology has to be implemented in OWL 2 web ontology language.
5	**Intended End-Users**
	• User 1. A farmer who wants to build a new indoor aquaponics 4.0 farm and wants to get information on the growth parameters of crops and fish. • User 2. A person who wants to understand the impact of crops on the design configuration of aquaponics grow beds before building a physical farm for his/her startup. • User 3. Researchers who want to build further applications such as decision support systems, and expert systems. • User 4. Industrial practitioners who want to find out the causes of poor crop quality in their aquaponic farm.
6	**Intended Uses**

	• Use 1. Develop a decision support system. The ontology model will be used to develop a system for crop quality monitoring and control. • Use 2. Search for aquaponics 4.0 system information. A farmer will look for general information about the management of the aquaponics 4.0 system. • Use 3. Search for optimal ranges of environmental and water parameters for the sustainability of aquaponics 4.0 farm. • Use 4. Build an interface to enable parametric design automation. The ontology will be used to develop an application to visualize the impact of crop type on the design configuration of grow bed.
7	**Ontology Requirements**
	a. Non-Functional Requirements 1. The ontology must be based on the international and Canadian food standards in existence or under development. **b. Functional requirements** 1. Optimal ranges of environmental parameters: light intensity, humidity, temperature, etc. 2. Optimal ranges of water parameters: pH, DO, TDS, water level, flow rate, nitrogen level etc. 3. Types of crops: leafy green vegetables such as lettuce, basil, spinach, mint etc. 4. Optimal growth conditions for crop survival: environmental plus water parameters. 5. The growth cycle of each crop: seedling, sprouting, vegetative and plant. 6. Species of fish suitable for the crops considered: Tilapia, catfish, trout etc. 7. Optimal growth conditions for fish survival: environmental plus water parameters. 8. The growth cycle of each fish: fingerlings and mature fish. 9. NFT grow channel configuration: standard widths, lengths, and heights available in the commercial market. 10. Specifications of fish tanks, mechanical filters, and biological filters. 11. Specifications of sensors and electronic devices installed at the farm for monitoring and controlling. 12. Crop and fish quality specifications and control standards determined by international and Canadian food agencies.

A3: A formal model of AquaONT showing classes and subclasses (Figure A.1)

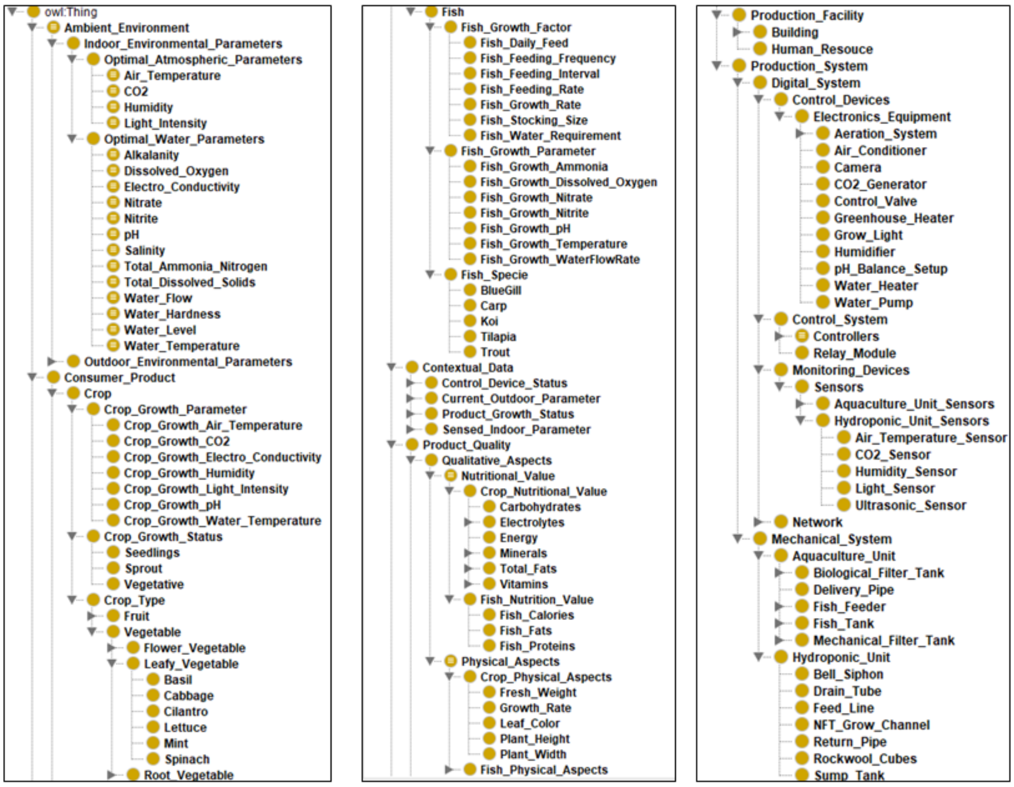


Figure A. 1. AquaONT formal model implemented on Protégé showing classes and subclasses

A4: AquaONT instances for different classes and subclasses (Figure A.2)

Concept	Class	Subclass	Subclass	Instances
Contextual_Data	Sensed_Indoor_Parameters	AO:Sensed_Indoor_Parameter	AO:Sensed_pH	AO:6.7
		AO:Sensed_Indoor_Parameter	AO:Sensed_AT	AO:24°C
		AO:Sensed_Indoor_Parameter	AO:Sensed_WT	AO:25°C
		AO:Sensed_Indoor_Parameter	AO:Sensed_LDR	AO:478Lumens
		AO:Sensed_Indoor_Parameter	AO:Sensed_EC	AO:1000µS/cm
		AO:Sensed_Indoor_Parameter	AO:Sensed_RH	AO:80
Consumer_Product	Crop	AO:Crop_Growth_Parameter	AO:Crop_Growth_CO2	AO:Lettuce_Growth_CO2
		AO:Crop_Growth_Parameter	AO:Crop_Growth_Water_Temperature	AO:Lettuce_Growth_WaterTemperature
		AO:Crop_Growth_Parameter	AO:Crop_Growth_Electro_Conductivity	AO:Lettuce_Growth_Electro_Conductivity
		AO:Crop_Growth_Parameter	AO:Crop_Growth_pH	AO:Lettuce_Growth_pH
		AO:Crop_Growth_Parameter	AO:Crop_Growth_Humidity	AO:Lettuce_Growth_Humidity
		AO:Crop_Growth_Parameter	AO:Crop_Growth_Light_Intensity	AO:Lettuce_Growth_Light_Intensity
		AO:Crop_Growth_Parameter	AO:Crop_Growth_Air_Temperature	AO:Lettuce_Growth_AirTemperature
	Fish	AO:Fish_Growth_Parameter	AO:Fish_Growth_Dissolved_Oxygen	AO:Tilapia_Growth_DO
		AO:Fish_Growth_Parameter	AO:Fish_Growth_WaterFlowRate	AO:Tilapia_Growth_WFR
		AO:Fish_Growth_Parameter	AO:Fish_Growth_pH	AO:Tilapia_Growth_pH
		AO:Fish_Growth_Parameter	AO:Fish_Growth_Nitrate	AO:Tilapia_Growth_Nitrate
		AO:Fish_Growth_Parameter	AO:Fish_Growth_Temperature	AO:Tilapia_Growth_Temperature
		AO:Fish_Growth_Parameter	AO:Fish_Growth_Nitrite	AO:Tilapia_Growth_Nitrite
		AO:Fish_Growth_Parameter	AO:Fish_Growth_Ammonia	AO:Tilapia_Growth_Ammonia
Production_System	Mechanical_System	AO:Hydroponic_Unit	AO:NFT_Grow_Channel	AO:NFT_Grow_Channel_01
		AO:Hydroponic_Unit	AO:Sump_Tank	AO:Sump_Tank_01
		AO:Hydroponic_Unit	AO:Drain_Tube	AO:Drain_Tube_01
		AO:Hydroponic_Unit	AO:Return_Pipe	AO:Return_Pipe_01
	Digital_System	AO:Hydroponic_Unit_Sensors	AO:Humidity_Sensor	AO:Humidity_Sensor_01
		AO:Hydroponic_Unit_Sensors	AO:CO2_Sensor	AO:CO2_Sensor_01
		AO:Hydroponic_Unit_Sensors	AO:Ultrasonic_Sensor	AO:Ultrasonic_Sensor_01
		AO:Hydroponic_Unit_Sensors	AO:Air_Temperature_Sensor	AO:Air_Temperature_Sensor_01
Product_Quality	Crop_Nutritional_Value	AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Vitamin_C_Content
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Carbohydrates_Content
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Calcium_Content
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Vitamin_A_Content
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Total_Fats_Content
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Protein_Content
	Crop_Physical _Aspects	AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Fresh_Weight
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Leaf_Color
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Plant_Width
		AO:Leafy_Vegetable	AO:Lettuce	AO:Lettuce_Plant_Height

Figure A. 2. AquaONT instances for different classes and subclasses.

A5: Object properties of AquaONT (Figure A.3)

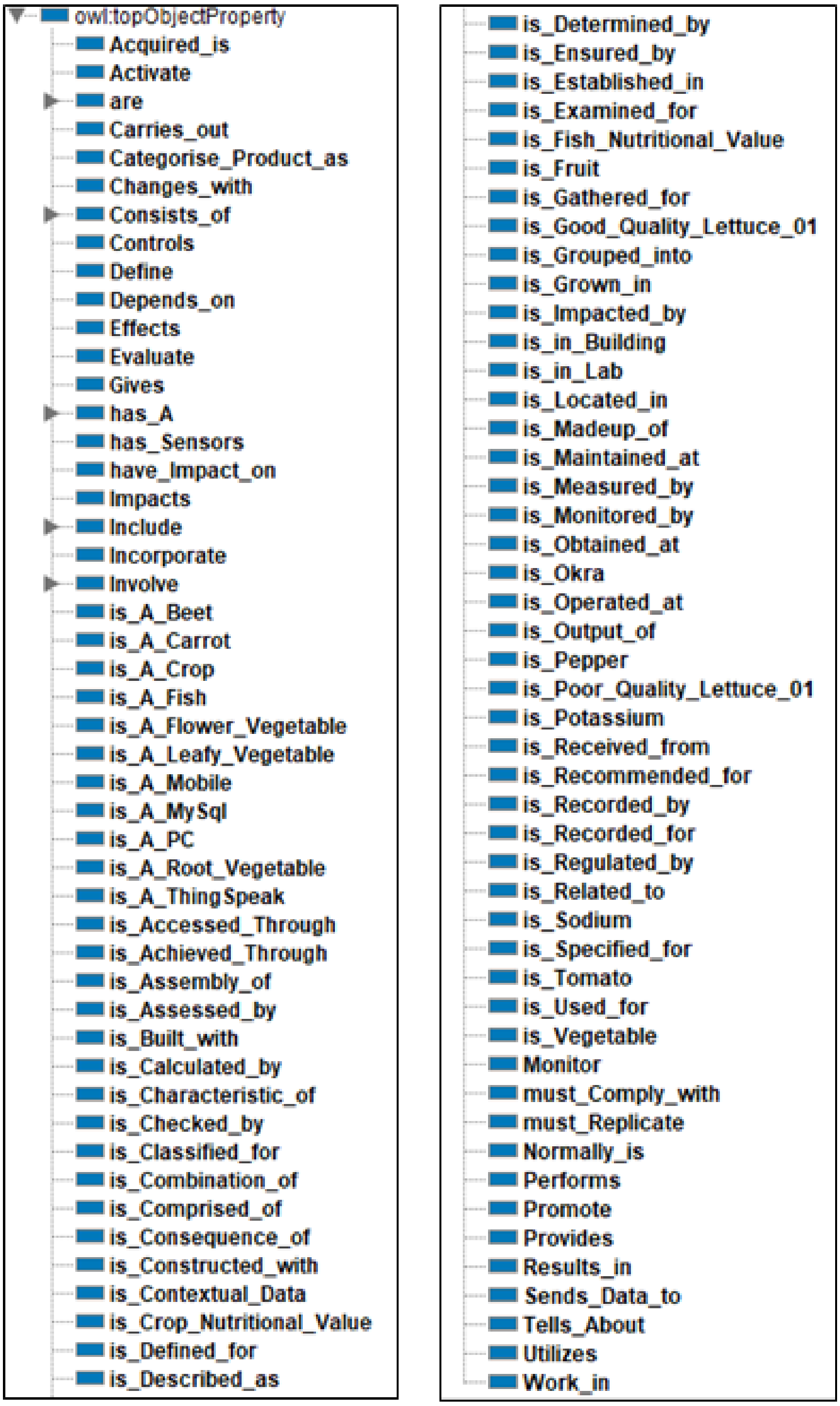


Figure A. 3. Object properties of AquaONT formal model.

A6: Object properties showing relationships between classes and subclasses (Figure A.4)

Domain	ObjectProperty	Range
AO:Contextual_Data	AO:Acquired_is	AO:Sensed_Indoor_Parameter
AO:Arduino	AO:Collects_Data_From	AO:Sensors
AO:Controllers	AO:Commonly_Used_are	AO:Arduino
AO:Sensors	AO:Generally_are	AO:Aquaculture_Unit_Sensors
AO:Sensors	AO:Generally_are	AO:Hydroponic_Unit_Sensors
AO:Humidity_Sensor	AO:Gives	AO:Sensed_RH
AO:Humidity	AO:Impacts	AO:Humidity_Sensor
AO:Fish_Growth_Factor	AO:Incorporate	AO:Fish_Daily_Feed
AO:Fish_Growth_Factor	AO:Incorporate	AO:Fish_Feeding_Frequency
AO:Fish_Growth_Factor	AO:Incorporate	AO:Fish_Feeding_Interval
AO:Control_System	AO:Controls	AO:Control_Devices
AO:Product_Inspection	AO:Define	AO:Nutritional_Inspection
AO:Product_Inspection	AO:Define	AO:Physical_Inspection
AO:Qualitative_Value	AO:Depends_on	AO:Indoor_Environmental_Parameters
AO:Crop_Growth_Parameter	AO:Normally_Include	AO:Crop_Growth_Air_Temperature
AO:Crop_Growth_Parameter	AO:Normally_Include	AO:Crop_Growth_CO2
AO:Crop_Growth_Parameter	AO:Normally_Include	AO:Crop_Growth_Electro_Conductivity
AO:Crop_Growth_Parameter	AO:Normally_Include	AO:Crop_Growth_Humidity
AO:Crop_Growth_Parameter	AO:Normally_Include	AO:Crop_Growth_Light_Intensit
AO:Optimal_Water_Parameters	AO:Mainly_Include	AO:Total_Dissolved_Solids
AO:Optimal_Water_Parameters	AO:Mainly_Include	AO:Water_Flow
AO:Optimal_Water_Parameters	AO:Mainly_Include	AO:Water_Hardness
AO:Optimal_Water_Parameters	AO:Mainly_Include	AO:Water_Level
AO:Optimal_Water_Parameters	AO:Mainly_Include	AO:Water_Temperature
AO:Optimal_Water_Parameters	AO:Mainly_Include	AO:pH
AO:Optimal_Atmospheric_Parameters	AO:Mainly_Involve	AO:Humidity
AO:Optimal_Atmospheric_Parameters	AO:Mainly_Involve	AO:Air_Temperature
AO:Optimal_Atmospheric_Parameters	AO:Mainly_Involve	AO:CO2
AO:Crop_Physical_Aspects	AO:are_Evaluated_by	AO:Fresh_Weight
AO:Crop_Physical_Aspects	AO:are_Evaluated_by	AO:Growth_Rate
AO:Crop_Physical_Aspects	AO:are_Evaluated_by	AO:Leaf_Color
AO:Crop_Physical_Aspects	AO:are_Evaluated_by	AO:Plant_Height
AO:Crop_Nutritional_Value	AO:is_Assessed_by	AO:Minerals
AO:Crop_Nutritional_Value	AO:is_Assessed_by	AO:Vitamins
AO:Crop_Nutritional_Value	AO:is_Assessed_by	AO:Carbohydrates
AO:Crop_Nutritional_Value	AO:is_Assessed_by	AO:Electrolytes
AO:Crop_Nutritional_Value	AO:is_Assessed_by	AO:Energy
AO:Crop_Nutritional_Value	AO:is_Assessed_by	AO:Protein
AO:Aquaculture_Unit_Sensors	AO:Considered_are	AO:Alkalinity_Sensor
AO:Aquaculture_Unit_Sensors	AO:Considered_are	AO:Dissolved_Oxygen_Sensor
AO:Aquaculture_Unit_Sensors	AO:Considered_are	AO:Electrical_Conductivity_Sensor
AO:Aquaculture_Unit_Sensors	AO:Considered_are	AO:Flow_Sensor
AO:Aquaculture_Unit_Sensors	AO:Considered_are	AO:Nitrite_Sensor
AO:Aquaculture_Unit_Sensors	AO:Considered_are	AO:Total_Ammonia_Nitrogen_Sensor
AO:Hydroponic_Unit	AO:is_Built_with	AO:Drain_Tube
AO:Hydroponic_Unit	AO:is_Built_with	AO:Feed_Line
AO:Hydroponic_Unit	AO:is_Built_with	AO:NFT_Grow_Channel
AO:Hydroponic_Unit	AO:is_Built_with	AO:Return_Pipe
AO:Hydroponic_Unit	AO:is_Built_with	AO:Rockwool_Cubes
AO:Hydroponic_Unit	AO:is_Built_with	AO:Sump_Tank
AO:Aquaculture_Unit	AO:is_Constructed_with	AO:Mechanical_Filter_Tank
AO:Aquaculture_Unit	AO:is_Constructed_with	AO:Biological_Filter_Tank
AO:Aquaculture_Unit	AO:is_Constructed_with	AO:Fish_Tank
AO:Aquaculture_Unit	AO:is_Constructed_with	AO:Delivery_Pipe
AO:Aquaculture_Unit	AO:is_Constructed_with	AO:Fish_Feeder

Figure A. 4. Object properties showing relationships between classes and subclasses.

A7: Datatype properties of AquaONT (Figure A.5)

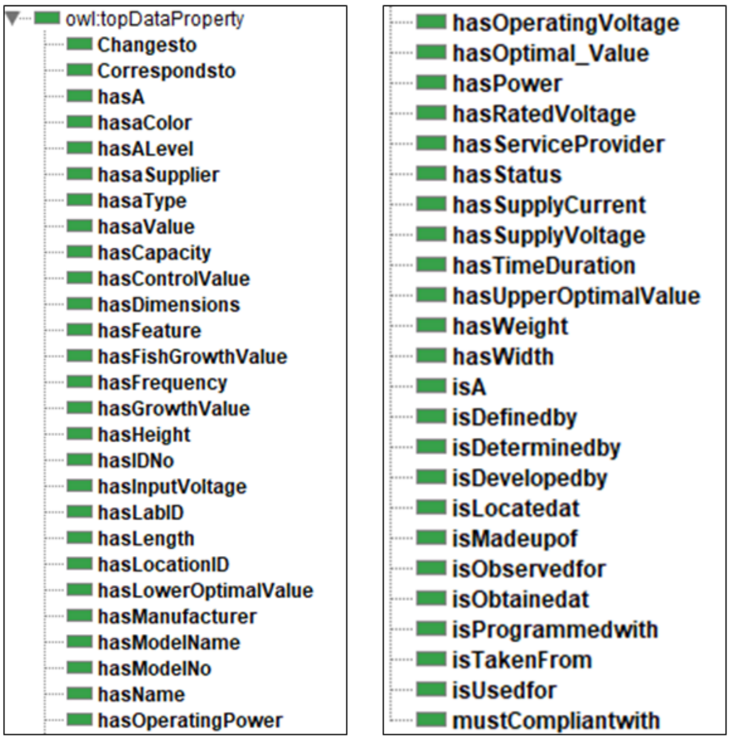


Figure A. 5. Datatype properties of AquaONT.

A8: Datatype properties showing relationships between instances and attributes (Figure A.6)

Instance	DatatypeProperty	Attribute
AO:Lettuce_Growth_CO2	AO:hasGrowthValue	"1000 ppm"^^
AO:Lettuce_Growth_LightIntensity	AO:hasGrowthValue	"15-17 mols/m^2/day of PAR or PPFD 650-800nm wavelength"^^
AO:Optimal_Relative_Humidity	AO:hasLowerOptimalValue	"60.0"^^xsd:float
AO:Optimal_Relative_Humidity	AO:hasUpperOptimalValue	"80.0"^^xsd:float
AO:Light_Sensor_01	AO:hasModelNo	"SI1147-M01-GM"
AO:Light_Sensor_01	AO:hasaSupplier	"Digi Key"
AO:Light_Sensor_01	AO:hasaType	"Ambient, IR, Ultraviolet (UV)"
AO:Light_Sensor_01	AO:hasManufacturer	"Silicon Labs"
AO:Light_Sensor_01	AO:hasIDNo	"LS01"
AO:Light_Sensor_01	AO:hasSupplyVoltage	"1.71V ~ 3.6V"
AO:NFT_Grow_Channel_01	AO:isMadeupof	"PVC"
AO:NFT_Grow_Channel_01	AO:isUsedfor	"Lettuce Growing"
AO:NFT_Grow_Channel_01	AO:hasaSupplier	"Crop King"
AO:NFT_Grow_Channel_01	AO:hasCapacity	"135 Plants"
AO:NFT_Grow_Channel_01	AO:hasIDNo	"NGC01"
AO:NFT_Grow_Channel_01	AO:hasA	"1 inch Square Holes"

Figure A. 6. Datatype properties showing relationships between instances and attribute.

Appendix B : Layout of dashboard for IoT enabled aquaponics system (Chapter 5)

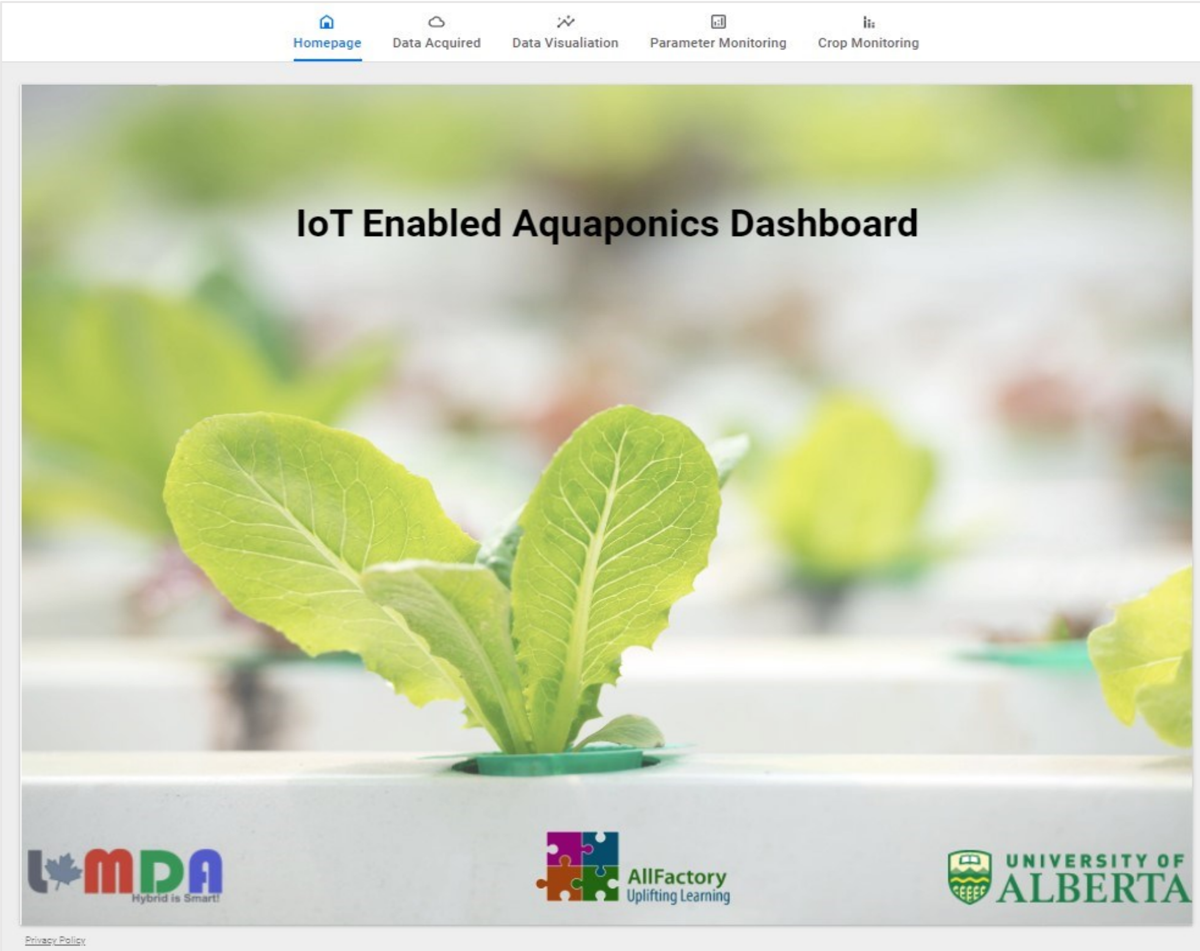
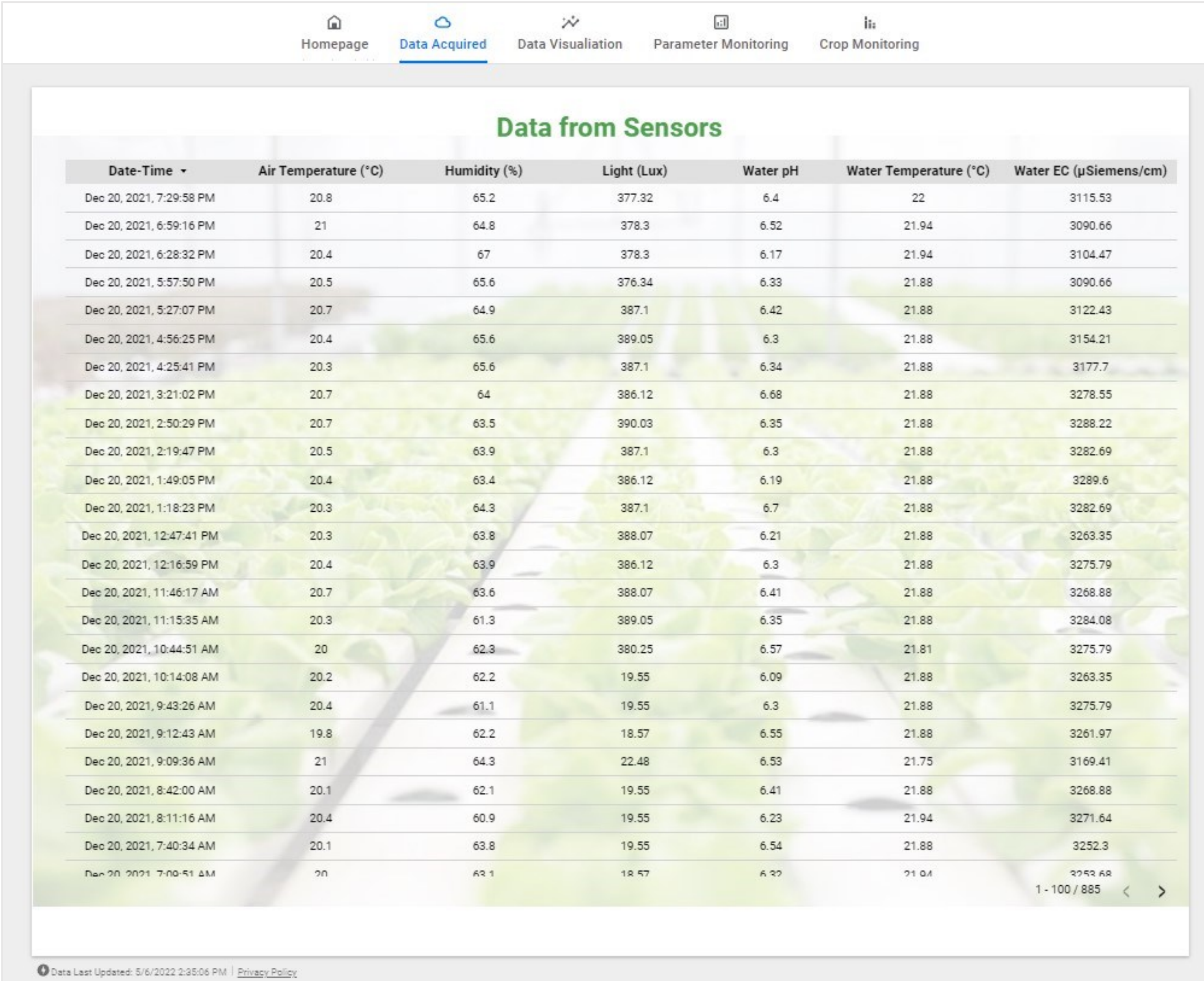


Figure B. 1. Home page of dashboard.



Date-Time	Air Temperature (°C)	Humidity (%)	Light (Lux)	Water pH	Water Temperature (°C)	Water EC (μSiemens/cm)
Dec 20, 2021, 7:29:58 PM	20.8	65.2	377.32	6.4	22	3115.53
Dec 20, 2021, 6:59:16 PM	21	64.8	378.3	6.52	21.94	3090.66
Dec 20, 2021, 6:28:32 PM	20.4	67	378.3	6.17	21.94	3104.47
Dec 20, 2021, 5:57:50 PM	20.5	65.6	376.34	6.33	21.88	3090.66
Dec 20, 2021, 5:27:07 PM	20.7	64.9	387.1	6.42	21.88	3122.43
Dec 20, 2021, 4:56:25 PM	20.4	65.6	389.05	6.3	21.88	3154.21
Dec 20, 2021, 4:25:41 PM	20.3	65.6	387.1	6.34	21.88	3177.7
Dec 20, 2021, 3:21:02 PM	20.7	64	386.12	6.68	21.88	3278.55
Dec 20, 2021, 2:50:29 PM	20.7	63.5	390.03	6.35	21.88	3288.22
Dec 20, 2021, 2:19:47 PM	20.5	63.9	387.1	6.3	21.88	3282.69
Dec 20, 2021, 1:49:05 PM	20.4	63.4	386.12	6.19	21.88	3289.6
Dec 20, 2021, 1:18:23 PM	20.3	64.3	387.1	6.7	21.88	3282.69
Dec 20, 2021, 12:47:41 PM	20.3	63.8	388.07	6.21	21.88	3263.35
Dec 20, 2021, 12:16:59 PM	20.4	63.9	386.12	6.3	21.88	3275.79
Dec 20, 2021, 11:46:17 AM	20.7	63.6	388.07	6.41	21.88	3268.88
Dec 20, 2021, 11:15:35 AM	20.3	61.3	389.05	6.35	21.88	3284.08
Dec 20, 2021, 10:44:51 AM	20	62.3	380.25	6.57	21.81	3275.79
Dec 20, 2021, 10:14:08 AM	20.2	62.2	19.55	6.09	21.88	3263.35
Dec 20, 2021, 9:43:26 AM	20.4	61.1	19.55	6.3	21.88	3275.79
Dec 20, 2021, 9:12:43 AM	19.8	62.2	18.57	6.55	21.88	3261.97
Dec 20, 2021, 9:09:36 AM	21	64.3	22.48	6.53	21.75	3169.41
Dec 20, 2021, 8:42:00 AM	20.1	62.1	19.55	6.41	21.88	3268.88
Dec 20, 2021, 8:11:16 AM	20.4	60.9	19.55	6.23	21.94	3271.64
Dec 20, 2021, 7:40:34 AM	20.1	63.8	19.55	6.54	21.88	3252.3
Dec 20, 2021, 7:09:51 AM	20	63.1	18.57	6.32	21.94	3253.68

Figure B. 2. Data acquisition panel.

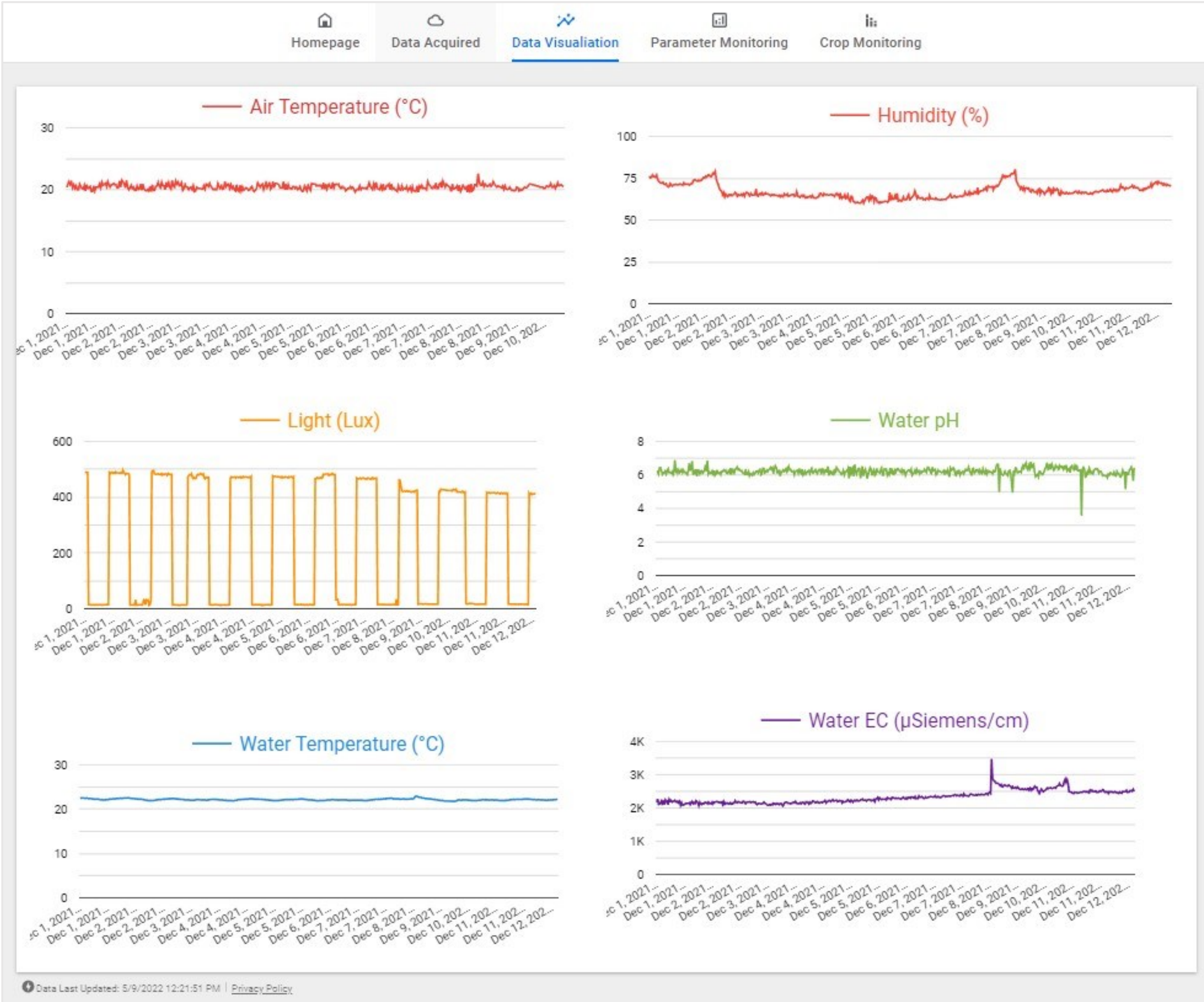


Figure B. 3. Data visualization panel.

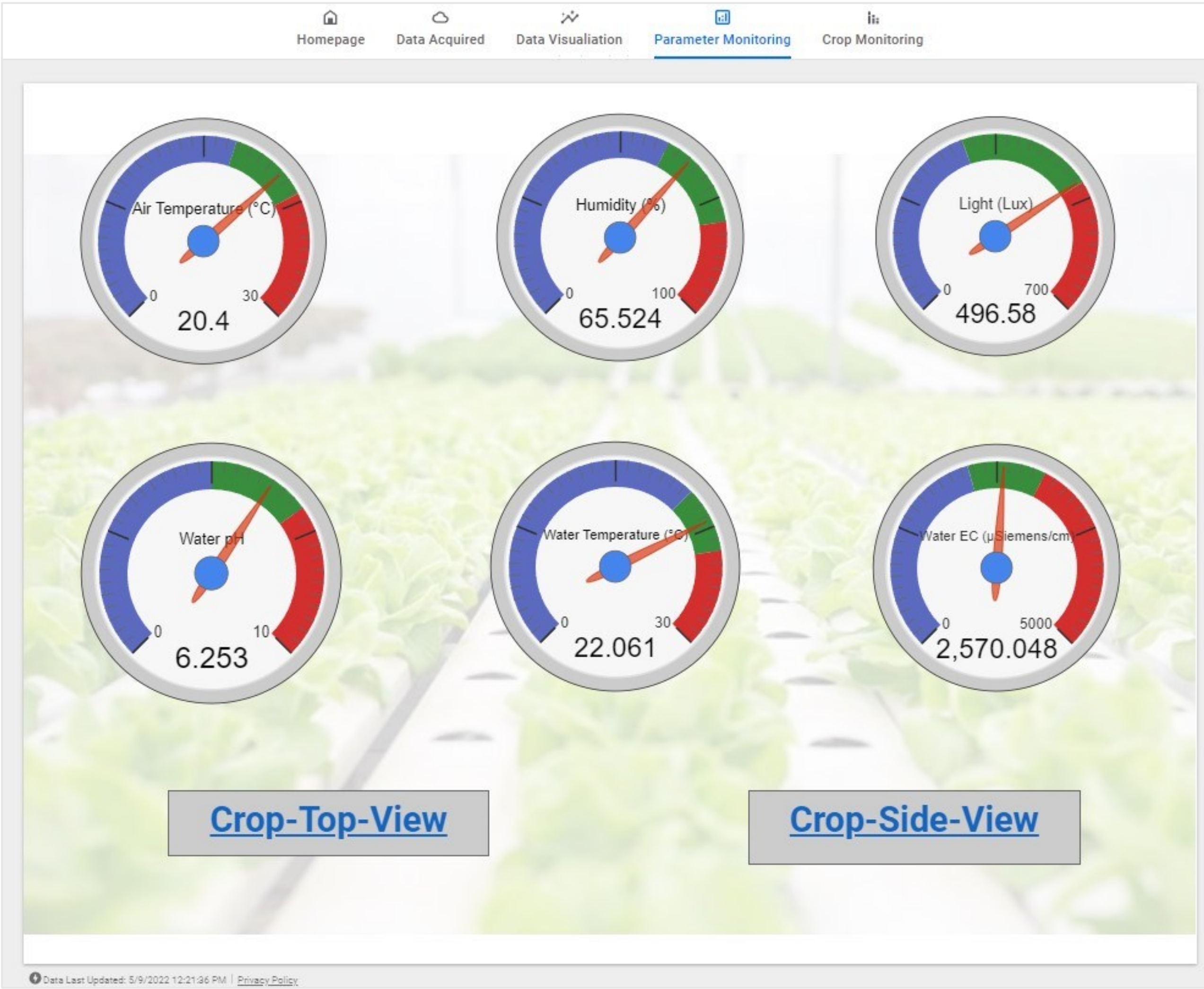


Figure B. 4. Parameter monitoring panel.

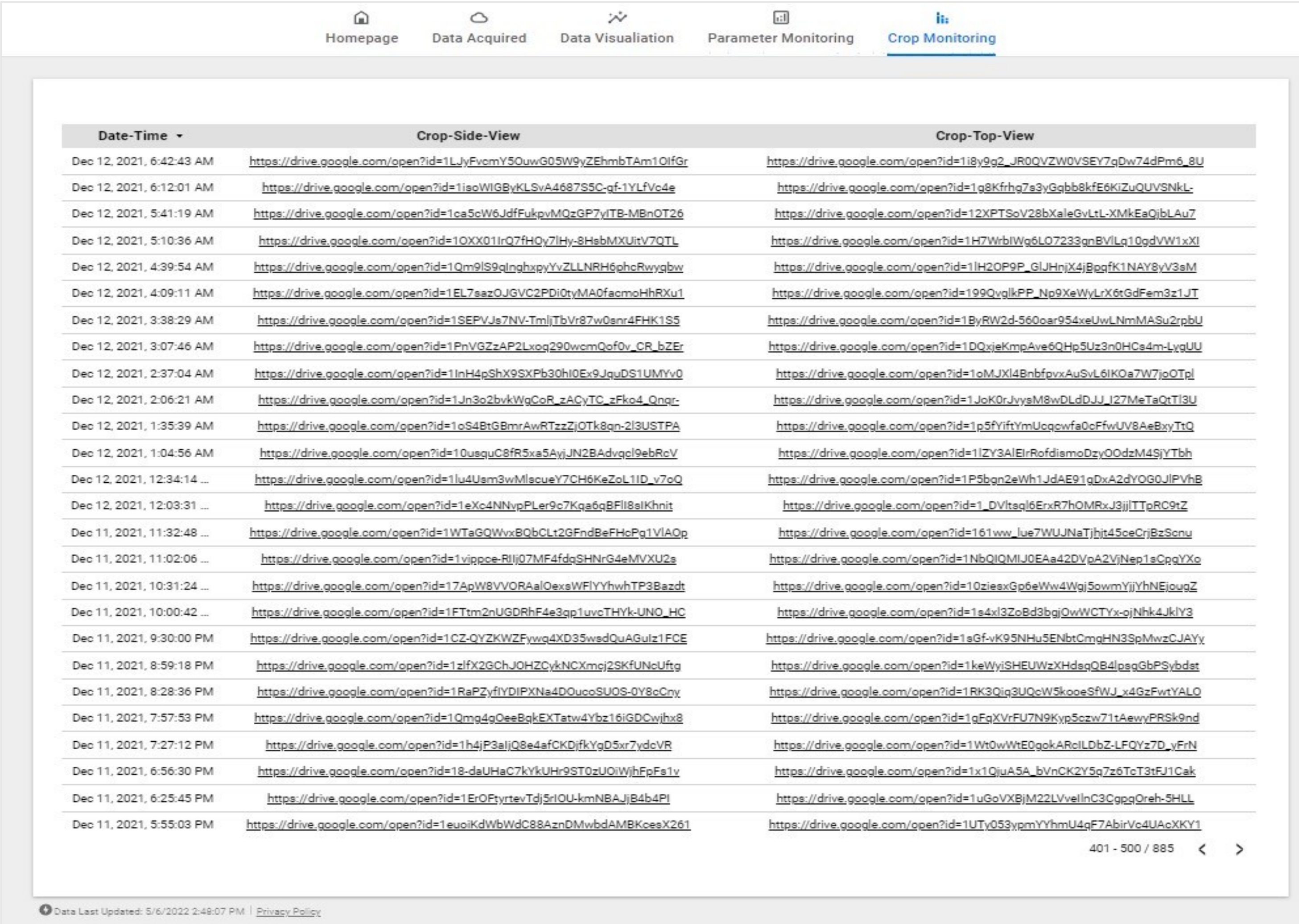


Figure B. 5. Crop monitoring panel.

Appendix C: Variations in heights of 30 plants (Chapter 8)

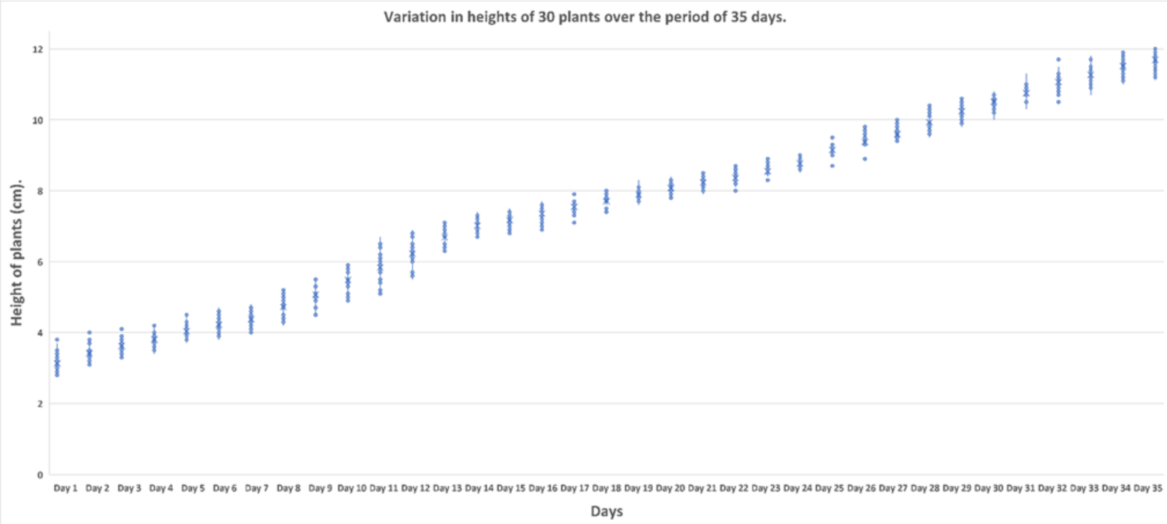


Figure C. 1. Boxplot showing variations in the heights of thirty plants over the period of thirty days.

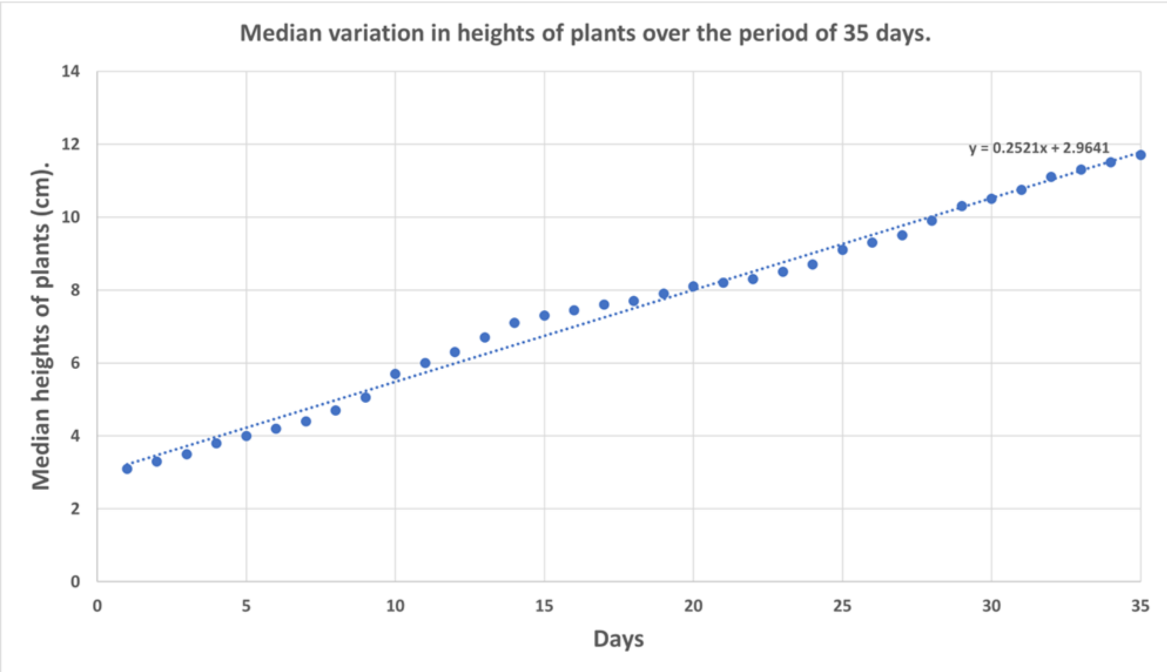


Figure C. 2. Variations in median heights of plants.

Appendix D: Case Study- Biomass prediction from canopy measurements in deep water culture-based aquaponics system. (Related to Chapter 8)

Objective

This case study is conducted in collaboration with Dr. Lisa Stein's team from Biological Sciences Lab. Crop biomass refers to the mass of the crop composed of live cells [1]. It is one of the primary indicators used to assess the plant growth and quality in plant phenotyping and also used an estimator of yield in agriculture [2]. The objective of this case study is to develop a system for the estimation of aboveground crop biomass suitable for in situ deployment. For this purpose, the canopy area is measured by implementing an instance segmentation process on crop images. Using Mask-RCNN, lettuce mask is extracted from the images and its features are computed and correlated with biomass values recorded manually using multi-layer perceptron (MLP) neural network. The significance of this study is that it paves a way to the ultimate goal of equipping the researchers and farm practitioners with a tool for rapid, non-destructive, reliable, and affordable assessment of their crops.

Experimental setup and description

In this aquaponics experiment, lettuce and goldfish are grown in four tanks namely A1, A2, B1 and B2 under different conditions. The pH in Tanks A1 and A2 was maintained at 7.6, whereas in Tanks B1 and B2 pH was maintained at 6. In this case study, Tank B1 and B2 are considered to develop a prediction model for lettuce canopy area and biomass. The primary difference between Tank B1 and B2 was that in the Tank B1 the inoculum or bacteria received from commercial supplier 'NutraPonics' were used for the biofiltration process, while for Tank B2 the desired community of bacteria developed by researchers in the Biology Lab was used. The wet biomass in grams was calculated for all 14 plants in each tank at the end of each harvesting round and recorded in an excel file shown in Table D-1. In total, there were three harvesting rounds with the first and second rounds 30 days, and the third round 45 days. Figure D.1 shows the comparison of fish yield in the four systems. It can be observed that fish growth in the acidic pH was slower as compared to the basic pH. However, in Tank B2 (where a developed bacterial community is used) fish grew better than in Tank B1 (with the original bacterial community). Likewise, Figure D.2 shows the plant biomass of 3 rounds of harvesting; unlike fish, lettuce was growing better at acidic pH. In the first round, the higher biomass in Tank B2 was obtained compared to B1, but in rounds, 2 and 3 (when you set up your cameras) no meaningful difference between Tanks B1 and B2 is observed.

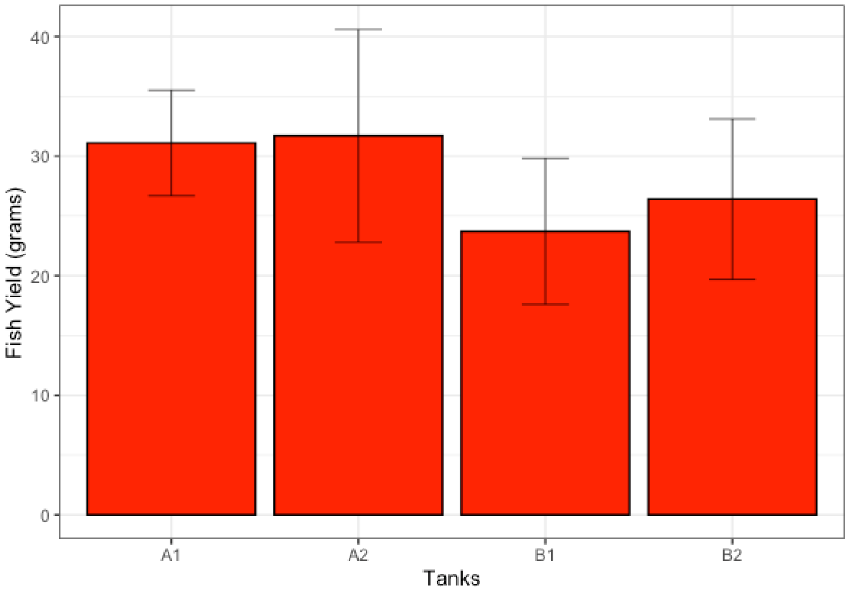


Figure D. 1. Fish yield in four tanks for three harvesting rounds (Biological sciences lab data).

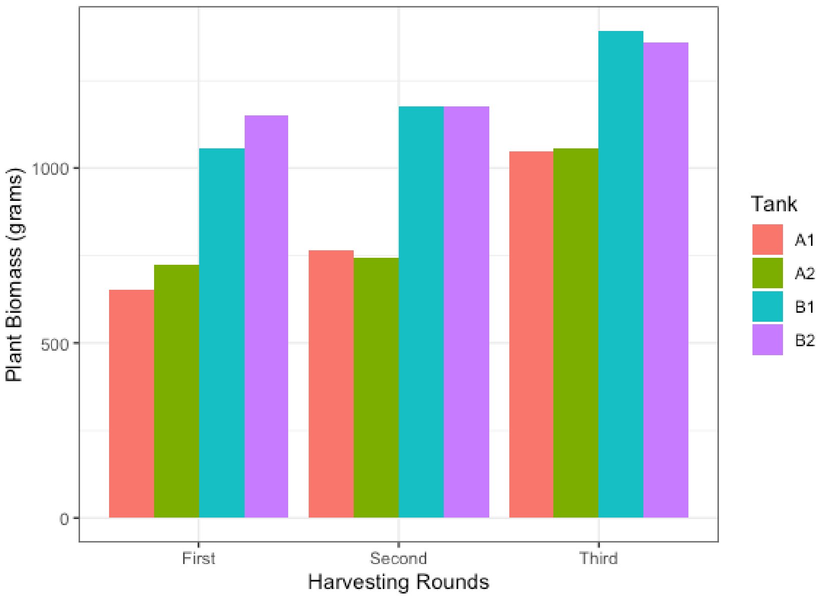


Figure D. 2. Lettuce yield in four tanks for three harvesting rounds (Biological sciences lab data).

Table D- 1. Biomass data for fourteen plants in four tanks for three harvesting rounds (Biological sciences lab data).

Tank	**Round**	**Biomass**
A1	A1.First	36.2
A1	A1.First	52.5
A1	A1.First	51
A1	A1.First	31.1
A1	A1.First	42.5
A1	A1.First	50.3
A1	A1.First	15.3
A1	A1.First	51.3
A1	A1.First	18.5
A1	A1.First	42.5
A1	A1.First	23.7
A1	A1.First	96.3
A1	A1.First	75.4
A1	A1.First	64.4
A1	A1.Second	66.1
A1	A1.Second	49.3
A1	A1.Second	55.6
A1	A1.Second	73
A1	A1.Second	51.1
A1	A1.Second	39.6
A1	A1.Second	44.8
A1	A1.Second	61.4
A1	A1.Second	38.4
A1	A1.Second	42.3
A1	A1.Second	63
A1	A1.Second	71.4
A1	A1.Second	66.9
A1	A1.Second	42.5
A1	A1.Third	46.5
A1	A1.Third	42.9
A1	A1.Third	78.9
A1	A1.Third	57.2
A1	A1.Third	102.1
A1	A1.Third	75.7
A1	A1.Third	12.3
A1	A1.Third	96.9
A1	A1.Third	110.2

A1	A1.Third	54
A1	A1.Third	77.2
A1	A1.Third	112.3
A1	A1.Third	123.1
A1	A1.Third	58.9
A2	A2.First	66.3
A2	A2.First	64.6
A2	A2.First	64
A2	A2.First	54.5
A2	A2.First	40.5
A2	A2.First	44.3
A2	A2.First	73.8
A2	A2.First	65.9
A2	A2.First	57.5
A2	A2.First	37.3
A2	A2.First	34.7
A2	A2.First	50.9
A2	A2.First	24.2
A2	A2.First	43.5
A2	A2.Second	53.6
A2	A2.Second	55.2
A2	A2.Second	71.3
A2	A2.Second	51.3
A2	A2.Second	39.8
A2	A2.Second	58.7
A2	A2.Second	66.1
A2	A2.Second	45.7
A2	A2.Second	81
A2	A2.Second	42.2
A2	A2.Second	54.6
A2	A2.Second	39.6
A2	A2.Second	44.1
A2	A2.Second	40.2
A2	A2.Third	94.3
A2	A2.Third	112.7
A2	A2.Third	34.8
A2	A2.Third	59.9
A2	A2.Third	67.3
A2	A2.Third	91.1
A2	A2.Third	66.1
A2	A2.Third	76.7
A2	A2.Third	105

A2	A2.Third	57
A2	A2.Third	80.3
A2	A2.Third	50.7
A2	A2.Third	64.2
A2	A2.Third	93.6
B1	B1.First	71.6
B1	B1.First	85.4
B1	B1.First	68.4
B1	B1.First	75.4
B1	B1.First	66.7
B1	B1.First	37.9
B1	B1.First	53.6
B1	B1.First	56.1
B1	B1.First	95.9
B1	B1.First	63.1
B1	B1.First	114.2
B1	B1.First	95.4
B1	B1.First	109.1
B1	B1.First	60.9
B1	B1.Second	61.9
B1	B1.Second	82.8
B1	B1.Second	83.4
B1	B1.Second	66.4
B1	B1.Second	70.9
B1	B1.Second	54.2
B1	B1.Second	60.7
B1	B1.Second	66.4
B1	B1.Second	88.3
B1	B1.Second	118
B1	B1.Second	89.5
B1	B1.Second	109.8
B1	B1.Second	114
B1	B1.Second	111.7
B1	B1.Third	43.7
B1	B1.Third	97.8
B1	B1.Third	92.7
B1	B1.Third	124.5
B1	B1.Third	123.3
B1	B1.Third	63.2
B1	B1.Third	104.7
B1	B1.Third	116.1
B1	B1.Third	72.6

B1	B1.Third	81.7
B1	B1.Third	128.6
B1	B1.Third	95.3
B1	B1.Third	103.8
B1	B1.Third	143.8
B2	B2.First	107.2
B2	B2.First	55.5
B2	B2.First	86
B2	B2.First	58.7
B2	B2.First	74.9
B2	B2.First	61.9
B2	B2.First	92.1
B2	B2.First	57.1
B2	B2.First	64.3
B2	B2.First	82.8
B2	B2.First	75.5
B2	B2.First	124.7
B2	B2.First	114.2
B2	B2.First	95.6
B2	B2.Second	88.9
B2	B2.Second	113.6
B2	B2.Second	77
B2	B2.Second	66
B2	B2.Second	63.2
B2	B2.Second	88.2
B2	B2.Second	68.4
B2	B2.Second	61.1
B2	B2.Second	91
B2	B2.Second	68.7
B2	B2.Second	66.4
B2	B2.Second	108.5
B2	B2.Second	101
B2	B2.Second	115.9
B2	B2.Third	60.1
B2	B2.Third	66.1
B2	B2.Third	72.6
B2	B2.Third	79.2
B2	B2.Third	69.7
B2	B2.Third	105.8
B2	B2.Third	114.1
B2	B2.Third	106.3
B2	B2.Third	98

B2	B2.Third	86.8
B2	B2.Third	133.1
B2	B2.Third	72.6
B2	B2.Third	152.4
B2	B2.Third	140.3

Image acquisition and preprocessing

For this case study, four ELP 1080P webcams (2.8–12 mm HD Varifocal Lens) were installed on Tank B1 and B2 (two cameras for each tank for top and side view) at a distance of 40cm from crops grow area to capture images during the second harvesting round. Each camera is programmed through Raspberry Pi 4 (Model B Rev 1) controller to take one image per day at 9:00 am, which were wirelessly uploaded to ‘IoT enabled Aquaponics Dashboard’ developed by authors in previous work (Chapter 5). To develop a model for biomass prediction based on canopy area, the lettuce images captured on the harvesting day are used. A data augmentation process is then performed to supplement and enrich the dataset. This helps increase the model's generalizability and overcome the problem of overfitting. Moreover, it also allows the model to learn as many relevant features as possible. This study uses Albumentations, a Python library, for fast and flexible image augmentations. The different augmentation techniques applied are flip, rotation, noise, blur, and brightness. Figure D.3 shows the sample image.

Figure D. 3. Sample image from the aquaponics system.

Prediction model development and training

The methodology used to estimate biomass of lettuce from the canopy area (top view) is shown in Figure D.4. The Mask-RCNN based morphological model developed in chapter 8 is used to extract features such as pixel count from the predicted mask, which is then used to estimate canopy dimensions or morphological traits. The artificial neural network (ANN) algorithm is used to optimize the canopy area, length and width which act as independent variables with respect to the dependent variable crop biomass.

The ANN works in two phases: i) forward propagation, and ii) backpropagation. In forward propagation, first feature values are multiplied with weights, then bias is added, and lastly, an activation function is applied to each neuron in the neural network. An activation function introduces non-linearity to the data which helps to identify the underlying complex patterns. In backpropagation, the optimal values of the parameters are determined for the model by iteratively updating parameters by partially differentiating gradients of the loss function with respect to the parameters. An optimization function is applied to perform backpropagation which is used to find the optimal value for parameters. The architectural layout of ANN used for the proposed model is shown in Figure D.5.

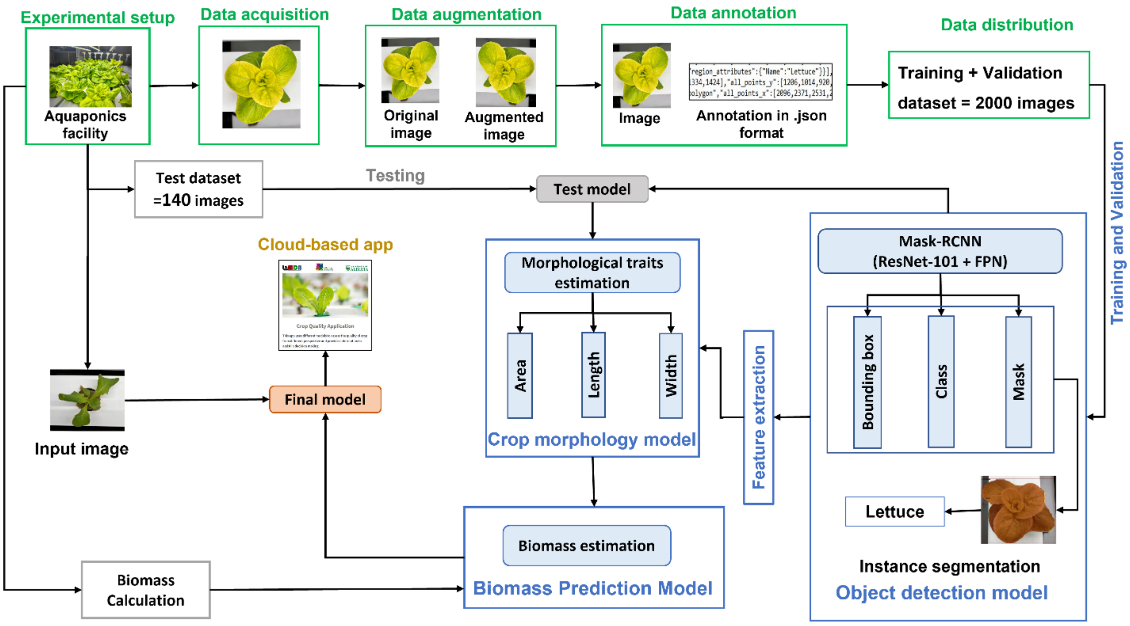


Figure D. 4. Research methodology.

It consists of seven layers including one input layer, one output layer, and one hidden layer. The first layer is input layer with three neurons, each representing one of the lettuce crop biomass yield prediction features x_i. The second layer is hidden layer with seven neurons (units) along with the bias neuron, which allows control of the behaviour of the layer without changing a value. Bias also helps the model to fit in the best way possible. This layer uses Sigmoid activation function. The last layer is output layer that has only one neuron. The i^{th} neuron of the input layer connects with the j^{th} neuron of the hidden layer by weight W_{ij}, and weight between the j^{th} neuron of the hidden layer and the k^{th} neuron of output layer is W_{jk}. In hidden layer, all the inputs are multiplied by their weights, W_{ij}. Weight is the gradient or coefficient of each variable. It shows the strength of the particular input. After assigning the weights, a bias variable is added.

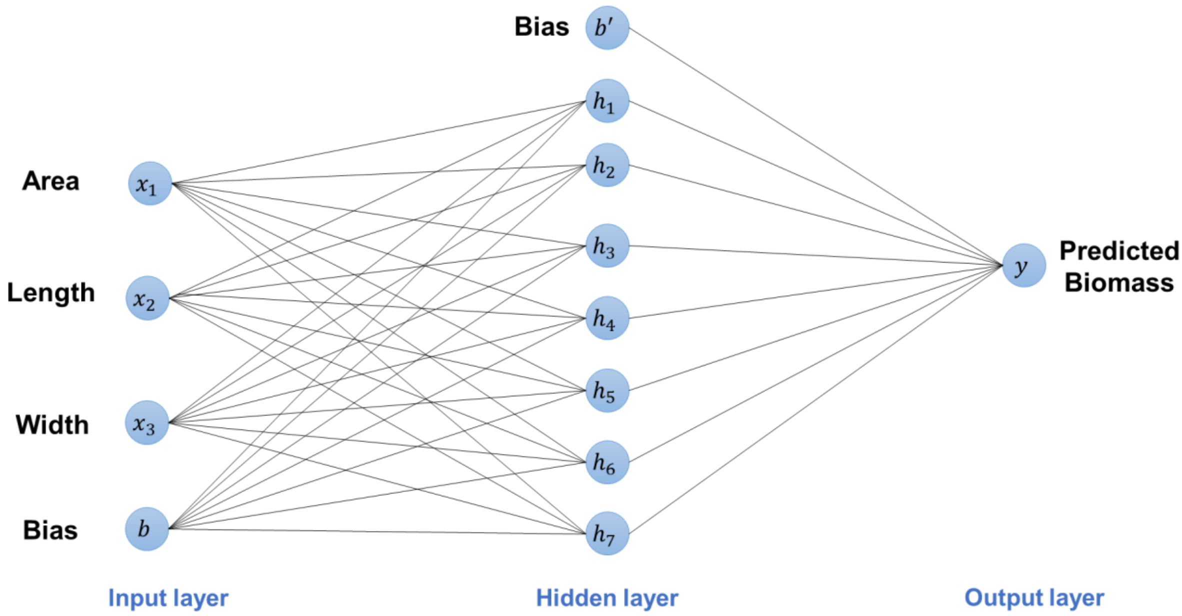


Figure D. 5. The architecture of ANN model used to estimate the crop biomass.

The equation used to perform computations in hidden layers are given below.

$$\hat{h}_j = \sum_{i=1}^{n} (x_i \times W_{ij}) + b$$

In the second step, the activation function (non-linear) is applied to the linear equation h_j. For this model sigmoid function is used, which is given below.

$$h_j = \sigma(\hat{h}_j) = \frac{1}{1 + e^{-\hat{h}_j}}$$

Likewise, for output layer the following equations are used for biomass prediction.

$$\hat{y}_k = \sum_{j=1}^{n} (h_j \times W_{jk}) + b'$$

$$y_k = \sigma(\hat{y}_k) = \frac{1}{1 + e^{-\hat{y}_k}}$$

For model to learn and perform the prediction process, backpropagation is performed. For this purpose, first error is computed with actual and predicted output using loss function such as mean squared error (MSE). It calculates the difference between the truth and prediction and square it. For multiple predictions and ground truths, the formula for MSE is given below.

$$MSE = \frac{1}{N}\sum_{k=1}^{N}(y_k - y'_k)^2$$

Where y_k is predicted output and y'_k is ground truth value. N denotes the total number of output samples. In this study, there is only one output node. The goal with backpropagation is to update each of the weights in the network so to bring the actual output closer to the ground truth output, thereby minimizing the error for each output neuron and the network as a whole. For this purpose, gradient decent optimization function is used by applying chain rule for partial differentiation of loss i-e, MSE with respect to weights. The equation is given as under.

$$\frac{\partial MSE}{\partial W_i} = \frac{\partial y_k}{\partial W_i} \times \frac{\partial \hat{y}_k}{\partial y_k} \times \frac{\partial MSE}{\partial \hat{y}_k}$$

To decrease the error, the above value is subtracted from current weight and multiplied by learning rate, ϵ. This is called weight update rule.

$$W_i' = W_i - (\epsilon \times \frac{\partial MSE}{\partial W_i})$$

Results and Discussion

The proposed model is implemented in TensorFlow and to tarin the model the values of hyperparameters are shown in Table D-2. The constructed ANN model is trained with a ground truth dataset of 100 data points where 10 % of arbitrarily selected data was used for validation. The model was optimized using the Adam optimizer algorithm. It is an extension of Stochastic gradient descent which can be applied in place of conventional stochastic gradient descent to update network weights more competently.

Table D- 2. Hyperparameters for training.

Hyperparameter	**Values**
Learning rate	0.05
Momentum	0.5
Batch size	18
Epochs	120
Activation	Sigmoid

Optimizer	Adam

The ANN model with 120 iterations resulted in an optimum accuracy with respect to minimized training and validation error in 100 iterations. The model stops when training loss minimizes to 2.4305 % and validation loss equals 0.4260 %. Figure D.6 shows the loss minimization of the model for 100 iterations. The model is then tested with another input and output dataset of known values.

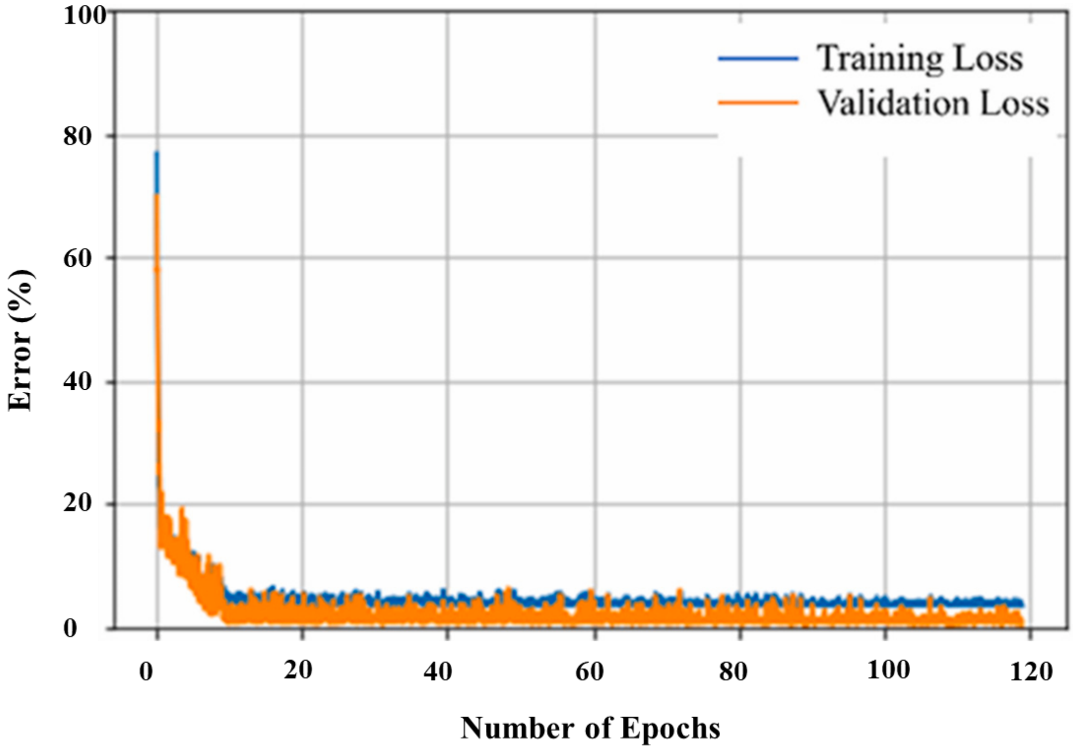


Figure D. 6. Training and validation minimization loss.

The proposed model is evaluated and tested with a prearranged dataset of inputs and output data that is unknown to the model. The predicted biomass by the model from the input variables of the testing dataset is compared with the actual biomass of the same input variables. Figure D.7 shows the model fit for ANN predicted values. The proposed ANN model has shown high accuracy in predicting biomass from the morphological parameters.

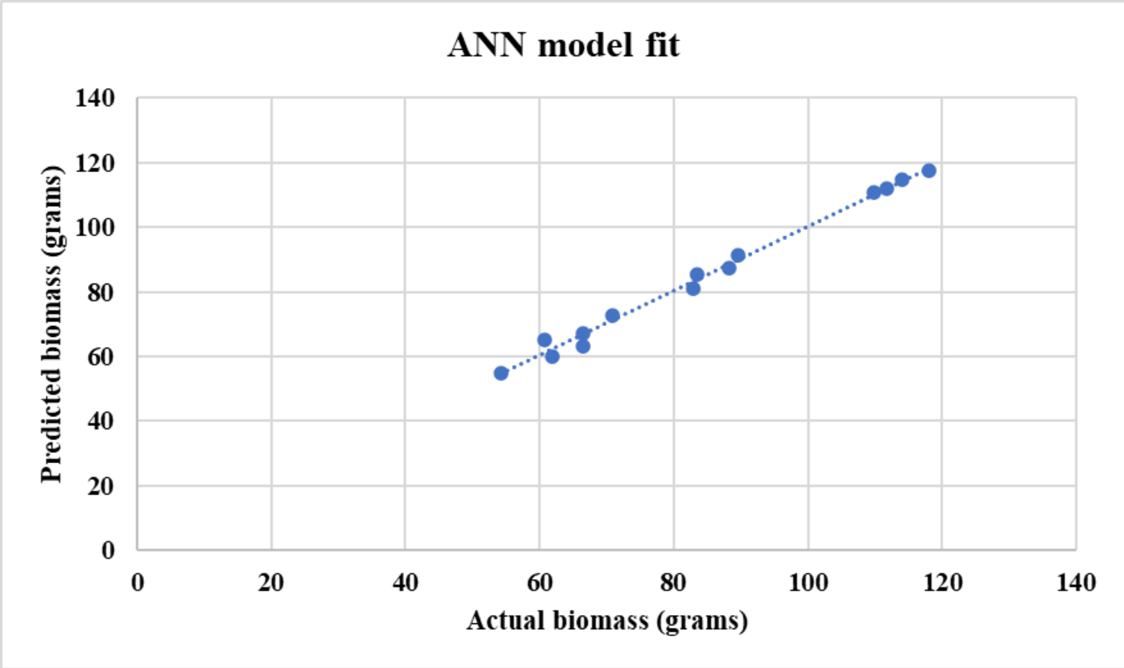


Figure D. 7. Model fit for ANN.

References

1. Buelvas, R.M., Adamchuk, V.I., Leksono, E., Tikasz, P., Lefsrud, M. and Holoszkiewicz, J., 2019. Biomass estimation from canopy measurements for leafy vegetables based on ultrasonic and laser sensors. Computers and Electronics in Agriculture, 164, p.104896.
2. Khan, M.S., Semwal, M., Sharma, A. and Verma, R.K., 2020. An artificial neural network model for estimating Mentha crop biomass yield using Landsat 8 OLI. Precision Agriculture, 21(1), pp.18-33.
3. Basir, M.S., Chowdhury, M., Islam, M.N. and Ashik-E-Rabbani, M., 2021. Artificial neural network model in predicting yield of mechanically transplanted rice from transplanting parameters in Bangladesh. Journal of Agriculture and Food Research, 5, p.100186.