Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE E TECNOLGIE AGRARIE, AMBIENTALI E ALIMENTARI

Ciclo XXXI

Settore Concorsuale: 07/B2

Settore Scientifico Disciplinare: AGR/03

SENSORS, ROBOTICS AND ARTIFICIAL INTELLIGENCE IN PRECISION ORCHARD MANAGEMENT (POM)

Presentata da: **Kushtrim Bresilla**

Coordinatore Dottorato

Prof. Dr. Massimiliano Petracci

Supervisore

Prof. Dr. Luca Corelli Grappadelli

Co-supervisore

Dr. Luigi Manfrini

Esame finale anno 2019

Declaration of Authorship

I, Kushtrim BRESILLA, declare that this thesis titled, "Sensors, Robotics and Artificial Intelligence in Precision Orchard Management (POM)" and the work presented in it are my own. I confirm that:

- This work was done mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“We are what we repeatedly do, therefore excellence is not an act, but a habit.”

Aristotle
Will Durant, "The Story of Philosophy" 1926

“After a time, one may find that 'having' is not as pleasing as 'wanting'.”

Spock
Star Trek, “Amok Time” 1968

UNIVERSITA DI BOLOGNA

Abstract

Faculty of Agriculture

SCIENZE E TECNOLOGIE AGRARIE, AMBIENTALI E ALIMENTARI

Doctor of Philosophy

Sensors, Robotics and Artificial Intelligence in Precision Orchard Management (POM)

by Kushtrim BRESILLA

As agriculture becomes more high tech, a growing number of farmers are using GPS-equipped autonomous systems supported by platforms that collect data on plants, soil, and weather. This emerging field of precision agriculture is based on observing, measuring and responding to crops with smart technology. Robots are taking on many tasks in agriculture nowadays, including planting greenhouse crops, harvesting apples, pruning vineyards and/or plant health monitoring. Precision agriculture refers to the way farmers manage crops to ensure efficiency of inputs such as water and fertilizer, and to maximize productivity, quality, and yield as well minimizing pests, unwanted flooding, and diseases. It begins the cycle by collecting information about their crop yields. This is done through a series of sensors specialized for a purpose, then the data is transferred to cloud for intelligent processing and further data analytics and at the end is either suggested to take actions or is done autonomously.

Image/video processing for fruit detection in the tree using hard-coded feature extraction algorithms has shown high accuracy on fruit detection during recent years. While accurate, these approaches even with high-end hardware are still computationally intensive and too slow for real-time systems. This study details the use of deep convolution neural networks architecture based on single-stage detectors. Using deep-learning techniques eliminates the need for hard-code specific features for specific fruit shapes, color and/or other attributes. This architecture takes the input image and divides into AxA grid, where A is a configurable hyper-parameter that defines the fineness of the grid. To each grid cell an image detection and localization algorithm is applied. Each of those cells is responsible to predict bounding boxes and confidence score for fruit (apple and pear in the case of this study) detected in that cell. We want this confidence score to be high if a fruit exists in a cell, otherwise to be zero, if no fruit is in the cell. More than 100 images of apple and pear trees were taken. Each tree image with approximately 50 fruits, that at the end resulted on more than 5000 images of apple and pear fruits each. Labeling images for training consisted on manually specifying the bounding boxes for fruits, where (x, y) are the center coordinates of the box and (w, h) are width and height. This architecture showed an accuracy of more than 90% fruit detection. Based on correlation between number of visible fruits, detected fruits on one frame and the real number of fruits on one tree, a model was created to accommodate this error rate. Processing speed is higher than 20FPS which is fast enough for any grasping/harvesting robotic arm or other real-time applications.

Cropload estimation is essential for efficient orchard management starting with optimizing chemical thinning, planning the labor force, harvest equipment, and vehicles for transportation of fruit from field to the packing plant. A new approach to determine tree cropload by using a combination of deep neural networks with convolution neural networks (CNN). For this we use a model to detect fruit in a canopy. To accurately determine cropload, images of the canopy were fed to another CNN which correlated tree canopy size and shape to the number of hidden fruit due to occlusion, then the output is feed to an simple deep neural network (DNN) together with detected fruit number to predict/estimate the overall canopy cropload. The model was then compared to an algorithmic method where parameters like cultivar, canopy shape and training system are manually estimated. Results show 60% to 80% accuracy of the algorithmic method while the deep learning model shows accuracy of around 95% prior to harvest, and more than 90% after chemical thinning, when fruit reached 20mm. Based on this, another CNN was developed to estimate the effectiveness of fruit thinning. Using these models, individual tree thinning can be estimated, so that when mechanical thinning would be applied, it would be based on those data.

With the increase of population in the world, the demand for quality food is increasing too. In recent years, increasing demand and environmental factors have heavily influenced the agricultural production. Automation and robotics for fruit and vegetable production/monitoring have become the new standard. An autonomous Unmanned Aerial Vehicle (UAV) able to navigate through rows orchard rows was described. The UAV is comprised of a flight controller (AP stack), a microcontroller for analog reading of different sensors, and an On-Board Computer (OBC). Pictures are taken through a camera and streamed through WiFi to a Ground Control Computer (GCC) running a convolution neural network model. Based on prior training, the model outputs three directions: RIGHT, LEFT and STRAIGHT. A moving average of multiple frames per second is extracted and sent to a build-in Proportional-Integral-Derivative (PID) controller on the UAV. After error correction from this feedback, controller sends the direction to the flight controller using MAVLink protocol's radio channel overrides, thus performing autonomous navigation.

* * *

Contents

List of Figures

List of Tables

List of Abbreviations

ML	Machine Learning
AI	Artificial Intelligence
DL	Deep Learning
CST	Computer Sciences and Technologies
DNN	Deep Neural Network
CNN	Convolution Neural Network
ANN	Artificial Neural Network
SSD	Singe Shot Detector
YOLO	You Only Look Once
R-CNN	Region-based Convolution Neural Networks
CV	Computer Vision
ROS	Robotic Operating System
OBC	On Board Computer
SI	Satellite Imagery
IoT	Internet of Things
RFID	Radio Frequency Identification
PA	Precision Agriculture
POM	Precision Orchard Management
GAP	Good Agricultural Practices
SSCM	Site Specific Crop Monitoring
ACMS	Automated Crop Management Systems
CSPA	Crop Specific Predictive Analytics
GM	Genetically Modified
DA	Data Analytics
DSS	Decision Support Systems

List of Definitions

Precise	Minutely exact, or sharply defined/stated
Sense	Detecting automatically the physical stimulus
Spatial	Occupying, or having the character of space
Temporal	Sequence of time or relative to a particular time

Significance of this study

What is the aim of this study?

The primary goal of this study is to create a platform for fruit growers to estimate the apple/fruit cropload in their orchards. By using Artificial Intelligence to detect fruits and relate it to canopy shape and estimate the overall cropload. The platform needs to be very modular and can be hot-plugged into autonomous tractors, rovers and/or unmanned areal vehicles.

What is already known in this subject?

There are many studies done to detect fruits in the tree and correlate it to the overall cropload. Most of the studies use hard-coded algorithms to extract features like shape, color, reflectance and so on from fruits. Studies on fruit detection date back to 50s and are use extensively for cropload estimation.

What are the new findings and innovations?

Our mani contributions to this field are with:

- A very accurate deep-learning model for fruit detection with accuracy up to 95% prior harvest,
- A model to relate tree canopy and shape to the fruit occlusion and overall cropload,
- A model to estimate the effectiveness of chemical thinning based on fruit detection and cropload models,
- A modular UAV platform for autonomous navigation within orchards that would carry fruit detection systems,
- And a large dataset (images) with labeled and annotated fruits (apples and pears).

Dedicated to family and friends first,
and to anyone else who might find the findings/methodologies
useful and interesting . . .

Chapter 1

Sensors, Robotics and Artificial Intelligence in Agriculture

Abstract

As agriculture becomes more high tech, a growing number of farmers are using GPS-equipped autonomous systems supported by platforms that collect data on plants, soil, and weather. This emerging field of precision agriculture is based on observing, measuring and responding to crops with smart technology. Robots are taking on many tasks in agriculture nowadays, including planting greenhouse crops, harvesting apples, pruning vineyards and/or plant health monitoring. Precision agriculture refers to the way farmers manage crops to ensure efficiency of inputs such as water and fertilizer, and to maximize productivity, quality, and yield as well minimizing pests, unwanted flooding, and diseases. It begins the cycle by collecting information about their crop yields. This is done through a series of sensors specialized for a purpose, then the data is transferred to cloud for intelligent processing and further data analytics. The next step is presented to the farmer as a suggestion to take actions or is done autonomously. In this paper we present a review of this closed system from sensors and data collection, big data and cloud computing until autonomous systems and robotics.

1.1 Introduction

THE pace of technological change — whether through advances in computer sciences & technology (CST), biotechnology, or such emerging fields as nanotechnology — will almost certainly accelerate in the next decade, with synergies across technologies and disciplines generating advances in research and development, production processes, and the nature of products and services [Harrison and Wolyniak, 2015]. In the CST field, for example, advances in microprocessors will support real-time image recognition and detection,while artificial intelligence will make possible the use of more intelligent robotics in manufacturing that will support the ability to quickly reconfigure machines to produce prototypes and new production runs, with implications for manufacturing logistics and inventories. Further technological advances are expected to continue to increase demand for a highly skilled workforce, support higher productivity growth, and change the organization of business and the nature of employment relationships [Lima and Souza, 2017].

Agriculture is a sector with very specific working conditions and constraints. This isn't only due to the dependency on weather conditions, but as well on the labor market. During times of highly intensive agricultural activities (eg. harvest), there are very pronounced peaks in workload which can only be predicted on a short-term basis due to the weather and seasonality. According to [Fröder, 2014] the world's agricultural workforce is expected to decline around 30% between 2017 and 2030. This expected decline will be driven by structural changes within the agri-food industry, but also because the opportunities for employment are expected to be better in other sectors. Rural areas are already facing difficulties in creating attractive jobs in general, pushing towards an ongoing migration towards urban centres. Those structural changes in agriculture are expected to continue with higher investments in technology. For example, investing in precision farming and digital agriculture are expected to significantly increase [Colbert, Yee, and George, 2016]. New technologies are set to impact the farm labour dynamic in many ways [Pierpaoli et al., 2013], but two developments stand out. One, the increasing use of data collection tools, such as sensors, and increasing sophistication of farm hardware and software is increasing demand for higher analytical and technical skill sets [Aubert, Schroeder, and Grimaudo, 2012; Mulla, 2013]. And two, the advancement of automation and autonomy on farm will decrease the reliance on human resources for low-skill and labour-intensive work while increasing autonomous machinery and robotics presence [Bechar and Vigneault, 2016; Stafford, 2007; Yaghoubi et al., 2013].

A new field of agriculture and engineering has emerged, a concept of precision agriculture (PA) that utilizes the advances in computer sciences and electronics for guidance and automatic steering, yield monitoring, variable rate input application, remote sensing, in-field electronic sensors, section and row control on planters, sprayers and fertilizer applicators, and spatial data management systems in agriculture. For a good agricultural practice (GAP), as described by [Zude-Sasse et al., 2016], temporal changes within or between years and spatial patterns of plant growth are to be considered. The most influenced temporal factors is weather conditions and environmental changes while the most influential spatial is the terrain [Manfrini, Taylor, and Grappadelli, 2009]. The fundamental concept of precision agriculture, collecting data and making decisions based on that data, has been around for many years. This was easier to do without technology on small plots. But as the size of farms grew, this no longer was possible. The larger farms require new techniques and tools [Brase, 2006]. These news tools have been developed over time. The Geographic Information System (GIS) was one of the first precision farming tool developed. As noted by [Brase, 2006] all major changes in agricultural technology have often been treated with derision and controversy. The change from horses or oxens to tractors was difficult for many people. Moving from single cross corn seed to hybrid seed was controversial too. New scientific breakthroughs in genome and gene manipulations are facing major backlash all over the world. Most regulations worldwide stipulate that a new genetically modified (GM) crops have to be compared to their closest non-GM counterpart as a corner stone of the pre-market risk assessment [Slot et al., 2018]. Likewise, with precision agriculture we are facing a major shift in technology. As with many of the previously stated technology shifts, even though the techniques and tools changed, and while the processes remained the same, the expected outcome always increased.

This paper presents an overview of new technologies and methodologies used in PA with emphasis in fruit tree orchards or precision orchard management. It aims to illustrate the challenges PA faces within this fast pace changing technological sector. One is that the agricultural sector is extremely low margin, therefore, investments in innovation are difficult [Pierpaoli et al., 2013]. Although the cost of smart farming is still high for any but the largest farms, this doesn't exclude PA for being an innovative sector in research and application. This paper gives and account on those technical advancements in computer sciences, machinery or else that are giving a significant changing push to the core meaning of PA. It focuses especially on the promise that Artificial Intelligence (AI) is making, on the new achievements surpassing even human capacities and capabilities. For many years it has been suggested that PA is in the crossroad where PA applications are catching up with technological advancements [Stafford, 2000b], where as highly-advanced tractors, UAV and other unmanned vehicles and other technologies will autonomously monitor, analyse and manage agricultural farms [Stafford, 2000a].

1.1.1 Precision Agriculture 4.0

The term precision refers to being minutely exact or sharply defined, and in PA, precise and sharply defined in the context of inputs means application of best measurable "thing" at precise "place" and sharply defined "time" - TPT [Pierce, Robert, and Mangold, 1994; Manfrini, Taylor, and Grappadelli, 2009].

The concept of PA has shifted its meaning from when was first conceptualized. When first global positioning systems (GPS) in conjunction with yield monitors and their strict relation were made available for field crops in late 80's, marks a rather short history of PA [Zude-Sasse et al., 2016]. In later years the introduction of analyzing tools like geo-referencing with GIS, shifted the meaning of PA into a more complete system. GIS is used to manage geo-spatial field data, including field boundaries, imagery, soil information, application maps, yield data, and near-earth and remotely sensed data. The affordability of those tools to measure, monitor and analyse spatial variability, tailored the action-taking process into a site-crop-specific [Stafford, 2000b]. PA also included the autosteering systems, variable rate application (VRA) and zone management. The technology needed to accomplish VRA includes an in-cab computer and software with a field zone application map, fertilizer equipment capable of changing rates during operation and GPS [Chen, Feng, and Zhang, 2012]. Autosteering and telemetry is technology that captures data from farm equipment operating in a field and transfers the data to farmer through internet [Pickett, 2004]. According to [Hedley, 2014] GPS and autosteering alone has increased farm efficiency up to 10% thus making PA a more interesting and valuable concept, despite its slow adoption during early years [Pierpaoli et al., 2013].

Later on researchers started using light as means of observing the crop. When crop is observed or its condition assessed without physically contact, it is a form of remote sensing [Huang et al., 2018]. Observing the colors of the leaves or the overall appearances of plants can determine its condition. Remotely sensed images taken from satellites and aircraft provided a means of assessing wide area of fields for a high vantage point and using them for VRA. Even though remote sensing dates back to the 50's [Bastiaanssen, Molden, and Makin, 2000], only recent technological advances have made the benefits of remote sensing accessible to most farmers. Satellite imagery (SI) in conjunction with VRA are often called "Grid farming" or Remote Agriculture.

Due to the prominence of Internet of Things (IoT) and Cloud Computing, PA is transforming into Smart Agriculture (SA). The system of PA is based on a precise and targeted application of specific measures in the exact spatial arrangement for more complex control structures [Mekala and Viswanathan, 2017]. Such structures are systems that process the input information and execute the output action. The input blocks of the control system are sensors that convert the input quantity of physical quantity. Various types of sensors that expand the opportunities for application, regulation, measurement and recording of data are being used in the PA systems. There is a paradigm shift from use of wireless sensors network (WSN) as a major driver of smart agriculture to the use of IoT and Data Analytics (DA) [Elijah et al., 2018]. The IoT integrates several existing technologies such as WSN, radio frequency identification (RFID), cloud computing, middleware systems and end-user applications for delivering Agriculture-as-a-Service (AaaS) [Gill, Chana, and Buyya, 2017; Stočes et al., 2016]. While PA is just taking in-field variability into account, a Smart Agriculture (SA) goes beyond that by basing management tasks not only on location but also on data, enhanced by context- and situation awareness, triggered by real-time events [Wolfert et al., 2017; Pivoto et al., 2018; OGrady and OHare, 2017; Kaloxylos et al., 2012]

Agriculture is highly repetitive, and, many tasks can be automated. Each farming operation requires many resources, for example planting, maintaining, and harvesting crops need money, energy, labour and resources [Brown, 2012]. For those reasons PA relies onto robotics and fully autonomous systems. For a farmer robot to be fully autonomous, it needs to navigate through very diverse and harsh environment without the human supervision [Slaughter, Giles, and Downey, 2008]. Robots nowadays are wirelessly connected to a central operator to both receive updated instructions regarding the mission, and report status and data. However, making an autonomous farm robot requires better controllers, localization, communication and action taking systems. The technology is similar to that of autonomous cars but applied to agitech. Where it differs is that farming robots often need to manipulate their environment, picking vegetables or fruit, applying pesticides in a localized manner, or planting seeds. All these tasks require sensing, manipulation, and processing of their own.

Until recent years, different traditional computer solving approaches have been extensively adopted in the agricultural field, be that in pest and disease detection, stress and water status monitoring or other. In recent years, with the significant increase in computational power, in particular with special purpose processors optimized for matrix-like data processing and large amount of data calculations (eg. Graphical Processing Unit - GPU), a lot of Machine Learning (ML), Deep Learning (DL) models and methodologies have achieved breakthroughs never achieved before [LeCun, Bengio, and Hinton, 2015; Ensmenger, 2011]. ML Algorithms enable one to analyze massive volumes of data, regardless of complexity, quickly and accurately. The union of many components: computer vision (CV) with ML, autonomous decision making based on trained model outcomes have been shown to be promising in solving different problems in agriculture. The predictive potential made possible by DL will cause a disruptive effect in different segments of the traditional industry as well as agriculture [Patricio and Rieder, 2018].

The process of agriculture towards precision since introduction of GPS, georeferencing, VRA has gone from simple Precision Agriculture (PA1.0) into: firstly with remote sensors and satellite and aircraft imagery to Remote Agriculture (PA2.0), then with big data, cloud-computing and IoT to Smart Agriculture (PA3.0), and lastly with Decision Support Systems (DSS), robotics and AI into Intelligent Self-Sustained & Autonomously-Managed Agriculture (PA4.0). PA4.0 it's a fully sustained cyclic system, where outputs of one segment are the input for the next one. It can be thought of three big steps, each including individual and/or independent smaller steps. First step is monitoring and sensing, then comes analyzing and predicting, lastly is managing and action taking. With drones, robots and intelligent monitoring

FIGURE 1.1: PA as cyclic system

systems now successfully being used in research and field trials, artificial intelligence, or machine learning, is set to revolutionize the future of agriculture as the next phase of precision agriculture is on the horizon. Many operations will be done remotely, processes will be automated, risks will by identified and solved before occurring and farmers will be able to make more informed and rapid decisions.

1.1.2 Artificial Intelligence in PA4.0

Scaling up of farm operations to match the exponential increase in consumption will drive the need for automation technologies in the farms. As the farmers are automating their operation, robots and autonomous systems have become an integral part of PA and are assisting farmers to improve yield and product quality while addressing the increasing supply needs. AI has become the backbone of robotics, as it enables a machine to use language processing and deep learning capabilities to take cognitive decisions. With the development of AI technologies, it is easier to track and predict the right time for planting, irrigation, and harvesting. AI helps to predict the likelihood of rain, the outbreak of diseases or attack of pests and the soil health condition. The information gathered from the field using satellite images and sensors on balloons would be juxtaposed with historical weather and other agronomic data to generate customized data for a specific farm and a specific crop. The advanced sensors and technologies, make the entire task of crop and soil management uncomplicated for the farmers and easily automated.

The use of cognitive technologies in agriculture could also help determine the best crop choice or the best managing practices for a crop mix adapted to various objectives, conditions and better suited for farm's needs. AI can use diverse capabilities to understand how crops react to different soil types, weather forecasts and local conditions. By analyzing and correlating information about weather, type of crops, types of soil or infestations in a certain area, probability of diseases, data about what worked best, year to year outcomes, marketplace trends, prices or consumer needs, farmers can make decisions to maximize return on crops. These artificial intelligence and cognitive systems will save time, increase safety and reduce potential human error while improving effectiveness.

The principle of AI is one where a machine can perceive its environment, and through a certain capacity of flexible rationality, take action to address a specified goal related to that environment. ML is when this same machine, according to a specified set of protocols, improves in its ability to address problems and goals related to the environment as the statistical nature of the data it receives increases. As the system receives an increasing amount of similar sets of data that can be categorized into specified protocols, its ability to rationalize increases, allowing it to better predict on a range of outcomes. Remote sensors, satellites, robots and autonomous systems can gather information 24 hours per day over an entire field thus feeding the system with huge amount of data, and in turn self-improving itself in reaching its goal.

Agricultural technology adoption is increasing each year, and machinery based applications, such as navigation technologies and yield monitors, are leading the way. It appears that adoption is driven in large part by ease of use; the easier a technology is to use, the more likely it is to be adopted on the farm [Pierpaoli et al., 2013].

1.2 Site specific crop monitoring - SSCM

The first step in PA4.0 is data collection and sensor monitoring. Different sensors, communications protocols and transmission protocols are used. The purpose of this step is that through crop monitoring for nutrients, water-stress, disease, insect attack, overall plant health etc. the farmer is aware in real-time of what happens in the field. This information can either be presented in raw-form to the farmer or processed with on-board computers, local computers or cloud.

1.2.1 Sensors

A number of sensing technologies are used in precision agriculture, providing data that helps farmers monitor and optimize crops, as well as adapt to changing environmental factors [Zude-Sasse et al., 2016]. By design most of agricultural sensors can be grouped into:

Location Sensors
use signals from GPS satellites to determine latitude, longitude, and altitude to within 20cm. Three satellites minimum are required to triangulate a position. Precise positioning is the cornerstone of precision agriculture.

Optical Sensors
use light to measure soil and plant properties. The sensors measure different frequencies of light reflectance in near-infrared, mid-infrared, and polarized light spectrums. Sensors can be placed on vehicles or aerial platforms such as drones or even satellites. Soil reflectance and plant color data are just two variables from optical sensors that can be aggregated and processed. In this category there are RGB, multispectral, hyperspectral, thermal, fluorescent cameras and Time of Flight ranging sensors.

Electrochemical Sensors
provide key information required in precision agriculture: moisture, pH and soil nutrient levels and other chemical triggers. Sensor electrodes work by detecting specific ions in the soil/plant. Moisture sensors assess moisture levels by measuring the dielectric constant (an electrical property that changes depending on the amount of moisture present).

Mechanical Sensors
measure force applied to the sensor or "mechanical resistance." Some of those sensors use a probe that penetrates the soil and records resistive forces through use of load cells or strain gauges. A similar form of this technology is used on large tractors to predict pulling requirements for ground engaging equipment. Tensiometers, detect the force used by the roots in water absorption and are very useful for irrigation interventions, Fruit gauges detect fluctuations on fruit growth of the fruit [Morandi et al., 2007] with accuracy of micrometers. Trunk/stem dendrometers, air pressure etc are all mechanical sensors.

Thermal Sensors
basically consists of two different metals such as nickel, copper, tungsten or aluminum etc, that are bonded together to form a Bi-metallic strip. The different linear expansion rates of the two dissimilar metals produces a mechanical bending movement when the strip is subjected to temperature change. The **Thermistor** is another type of temperature sensor, whose name is a combination of the words THERM-ally sensitive res-ISTOR. A thermistor is a special type of resistor which changes its physical resistance when exposed to changes in temperature. Their main advantage over bi-metal types is their speed of response to any changes in temperature, accuracy and repeatability. Most of the sap flow sensors are based on thermal sensors [Scalisi, Bresilla, and Grilo, 2018].

1.2.2 Data transmission

During years with the increase development of IoT technologies and decrease in price of computer raw power, wireless technologies have rapidly emerged [Zhang, Wang, and Wang, 2002]. Different technologies are used for transmitting without wires but likes of: light (infrared and/or laser point-to-point communications) and radio-frequencies (Bluetooth, WiFi, ZigBee, LoRa, SigFox, CDMA and GSM/GPRS) [Mirhosseini, Barani, and Nezamabadi-pour, 2017; Park, Lee, and Yoo, 2015; Carrabs et al., 2015; Ferentinos and Tsiligiridis, 2007].

In agriculture, this technology has been adopted due to the acceptable cost (Vougioukas et al., 2013). While this has still not seen a significant success in farmers orchards, it is one of the most used technology that gathers hight resolution spatial and temporal data about the environment and the specific crop that is being monitored.

Hardware required for Wireless Sensor Network (WSN) and/or Agriculture Wireless Sensor Network (AWSN) [Rehman et al., 2014; Ruiz-Garcia et al., 2009] is composed of: (ranking based on importance)

1. Radio-frequency communication protocol
2. Energy-efficient processor
3. High-resolution analog inputs (sensors)
4. Long-life energy source
5. Hig-speed reaction outputs (actuators)
6. Development platform

Those systems are usually composed of a few sinks and large quantity of small sensors nodes. Each wireless sensor node communicates with a gateway unit which can communicate with other computers via other networks. Communication protocol (CP) consists of the application layer, transport layer, network layer, data link layer, physical layer, power management, mobility management and the task management [Akyildiz et al., 2002; Ruiz-Garcia et al., 2009; Ojha, Misra, and Raghuwanshi, 2015].

Technology	LoRa	Zigbee	Bluetooth	WiFi	RFID
Feature	Mesh	Mesh	Star	Star	P2P
Power	Very Low	Low	Ultra low	Moderate	Very Low
Data rate (up to)	300Kbps	200Kbps	1Mbps	100Mbps	100Kbps
Coverage	20km	500m	300m	50m	3m
Cost	Very low	Low	Low	Medium	Low

TABLE 1.1: Comparison of Radio-Frequencies used in AWSN

CPs built over wireless standards, such as 802.15.4, facilitate the device networking and bridge the gap between the internet-enabled gateways and the end-nodes. Such protocols include ZigBee, ONE-NET, Sigfox, WirelessHART, ISA100.11a, and

6LowPan, Bluetooth Low Energy (BLE), LoRa/LoRaWAN, DASH7 and LoWiFi [Tzounis et al., 2017]. During years different technologies and advances in low power consumption have been developed providing a better means for AWSN. Technologies like Bluetooth were thought the best compromise between datarate, speed and distance [Vuran et al., 2018; Choudhury et al., 2015]. Bluetooth is probably the closest peer to WSNs, but its power consumption has been of secondary importance in its design. Later on, more robust, lower power consumption and better meshed technologies have been applied. [Tab 1.1] [Rehman et al., 2014; Wang and Li, 2013].

As shown in [Fig 1.2], most promising technolog with very high distance transmission rate, low power consumption and reasonable data rate is LoRa. LoRa can be used for free in spectrum bands 434 MHz in Europe and 960 MHz in US. LoRa has many parameters; the most important one is the Spreading Factor (SF). SF is a set of parameters that specify transmit power, subfrequency and air time. LoRa define spreading factors numbered from 6 to 12, where LoRaWAN is using from 7 to 12. The lower is the SF, the higher is the throughput, and the lower is the distance covered. Also, lower SF means lower power consumption [Fig 1.2]. Many algorithms are implemented in order to automatically assign SF among nodes thus making an efficient data-rate/energy-consumption ratio.

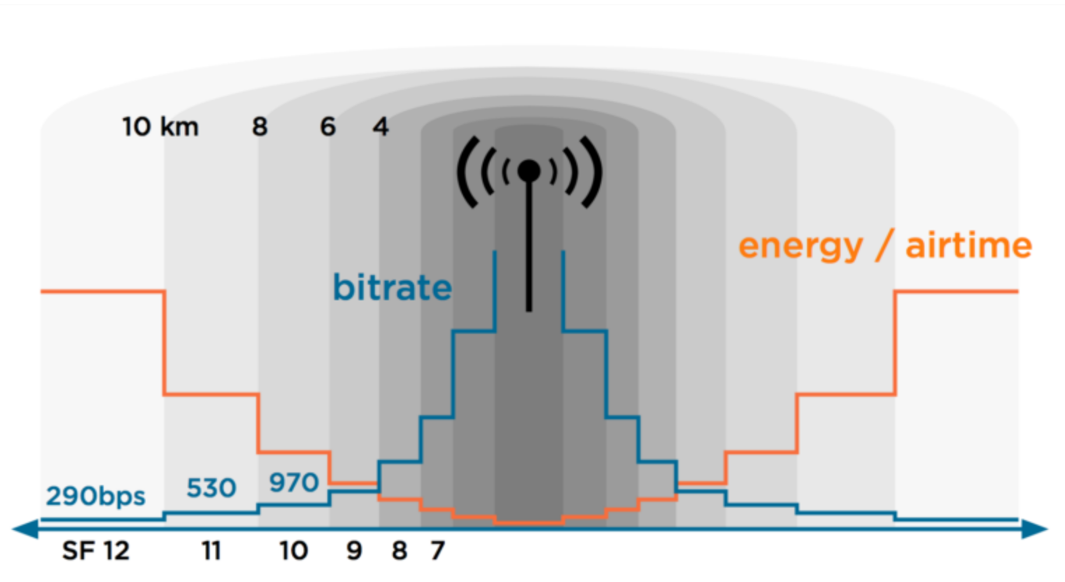


FIGURE 1.2: LoRa

Versatile and convenient form factors, low-cost devices, high-processing AWSN can nowadays be used, on small batteries with or without assist of mounted solar panels, and operate for long periods of time. Those modern embedded devices have sufficient resources to support even more demanding sensors, such as image sensors and on-board image analysis.

Roughly all Agriculture WSNs can be grouped into three categories:

1. **Barebone monitoring without processing**. Most AWSN fall into this category, AWSN in this case are used just as intermediary nodes between sensors and internet gateways. AWSN log the data from the sensors, then through the RF and CP send the data to the server where further processing is done and served to the user/farmer [Akyildiz and Kasimoglu, 2004; Tzounis et al., 2017; Le and Tan, 2015].

2. **Monitoring and simple processing**. A huge number of AWSN is now shifting into smarter self-processing nodes. While they take data from sensor (be that soil, plant or environment), they are able to give early warnings and some barebone predictive analytics about the crop. Usually they are attached to actuators to automate some simplified actions like irrigation valves or other VRA [Hu et al., 2010; Akyildiz and Kasimoglu, 2004; Mainwaring et al., 2002].

3. **Monitoring and advanced processing**. In this category fall the more compute-hungry AWSNs. A lot of early pest detection and insect traps use small but powerfull AWSN that have cameras for image analysis and detection [Chougule, Jha, and Mukhopadhyay, 2016]. Those systems include a much power-hungy processor in form of an embedded computer able to monitor crops in real time [Liqiang et al., 2011; Lin and Liu, 2008; Lv, Shen, and Hu, 2009; Chougule, Jha, and Mukhopadhyay, 2016].

AWSNs are spread across all domains of PA, starting with crop soil monitoring [Wang and Li, 2013; Chen, Feng, and Zhang, 2012; Sun et al., 2009]. According to [Kim, Evans, and Iversen, 2008] spatial and temporal variations of soil moisture can be matched by precision irrigation management, which can increase application efficiencies, reduce environmental impacts and even improve yields. However for better and real-time usage Wireless Underground Sensor Networks (WUSNs) have recently been investigated for unattended soil monitoring. Unlike wired sensor, which need to be deployed and removed frequently during the process of planting, WUSNs are deployed in the ground at a safe depth and do not interfere with agriculture machine operations, such as tillage practices [Akyildiz and Stuntebeck, 2006]. These networks consist of wirelessly-connected underground sensor nodes that communicate through soil [Dong, Vuran, and Irmak, 2013]. Each device contains all necessary parts to make the device self contained WSN: sensors, memory, a processor, a radio, an antenna, and a power source.

The widest adoption of AWSN is in the high-value crops e.g. orchard and/or vineyards [Vougioukas et al., 2013]. A study by [Torres et al., 2017] was developed to acquire a suitable knowledge to manage irrigation and verify the influences of living mulches on the vine by using wireless sensor networks to measure the vapor pressure deficit, soil water potential and soil water content. In another study ann intelligent data acquisition and service system for apple orchard was developed by [Guo et al., 2014] for acquiring apple tree growth information in time and managing orchard production remotely by Portable Digital Assistant (PDA) through the ZigBee Wireless Sensors Network (WSN) deployed in the apple orchard.

Implementing AWSN has many challenges [Tzounis et al., 2017]. Biggest ones are heterogeneity and signal penetration through the vegetation [Vougioukas et al., 2013] and security [Sicari et al., 2015].

1.2.3 Traceability

RFID is an emerging technology that makes use of wireless communication. The protocol was originally developed for short-range product identification, typically covering the 2 mm - 2 m read range, and has been promoted as the replacement technology for the optical bar-code found, with the use of EPC (Electronic Product Code) [Ruiz-Garcia et al., 2009]. RFID has the ability to allow energy to penetrate certain goods and to read a tag that is not visible [Dobkin and Wandinger, 2005]. RFID systems are comprised of three main components: the tag or transponder, the reader or transceiver that reads and writes data to a transponder, and the computer containing database and information management software [Li et al., 2006].

Fruit pack houses attach passive RFID tags to bins holding fruit associated with a number. Every time a bin passes through an RFID portal, its tag is read, and data is collected and sent to a database, thus creating an audit trail or chain-of-custody [Gautam et al., 2017]. The audit trail contains valuable information about those fruit, treatments done and when were they done, when it was picked and so on, which machinery were used, the worker harvesting them (or autonomous machine). With this information readily available, players in the supply chain can quickly and efficiently mitigate a recall. Now, instead of pulling every single product from storage or even store shelves, it is possible to pinpoint and pull just the batch that contains the affected products [Ghaani et al., 2016].

Semi-Active RFID tags equipped with battery-powered sensors allow farmers to collect temperature data from pallets and bins. By reading the RFID tag, farmers can make sure a pallet maintains a certain temperature and no infestation occurs. And, if the pallet has hit a temperature above certain threshold at any point in its life-cycle, the owner can evaluate the specific bins or adjust the expiration data as appropriate [Piramuthu, Farahani, and Grunow, 2013]. Similarly, RFID tags now exist that can monitor humidity, pressure, and event movement, arming users with even more data to ensure food safety.

Tree-based monitoring and tracing was described by [Wu et al., 2013; Ampatzidis and Vougioukas, 2009; Luvisi et al., 2011]. Electronic archives for each apple tree were set up and the whole record of orchard productivity was utilized to inform the management process.And [Ampatzidis et al., 2012] presented a real-time monitoring system that can track and record individual picker efficiency during harvest of tree crops. It integrated a digital weighing scale, RFID reader/writer. As harvested fruit is dumped into a standard collection bin situated on the scale, the system reads simultaneously the picker's ID (RFID tag) and records the incremental weight of fruit.

1.3 Crop specific data analytics - CSDA

This is the second step of PA4.0 and the most important one. After data are collected through sensors in the field, and then transmitted through different protocols to internet-gates, the data is forwarded to specific servers and/or computers for analysis. There are different approaches, softwares and platforms that analyse the data which will be discussed. Those data afterwards are either presented to the farmer so he can check again and decide further actions or can be sent directly to autonomous systems that act immediately based on specified algorithms and/or routines.

1.3.1 Big Data and Cloud Computing

During the significant development of technology of recent decades, the agriculture world has quietly been introduced to data aggregation technology. Companies built data systems into their machinery, farmers started enabling data acquisition devices in their farms, and larger farms started using software to manage their operations. The adoption of those new data sources has been slow [Pierpaoli et al., 2013], and systems are often unwanted because they create lock-in to the software, incompatible with the variety of other tools and/or brands used on the farm. Despite that, the amount ot data once the systems are functional, is very high [Fountas et al., 2006]. The challenge here is not the amount ot data, but what should be done with it, and how can it be processed to help farmers take better decisions [Wolfert et al., 2017; McBratney et al., 2005].

A new phenomenon of Big Data has drawn a huge attention from all domains of research [Wolfert et al., 2017]. According to [Li and Li, 2018] in 2011, the data volume size (copied or created) was around 2 ZetaBytes, and the trend shows this is exponentially increasing. So what is "big data" and what does it mean? Big data is a very abstract concept, and to determine the difference from “massive data” it has some additional features. According to [Ozkose, Ari, and Gencer, 2015], big data should have those three main characteristics: volume, velocity and variety. Some authors would add another characteristic: value, calling it the 4V-s. Big data is an evolving term that defines any large amount of structured, semi-structured and unstructured data that has the potential to be mined for information. Big data is a set of methods and technologies that require new forms of integration to uncover large hidden values from large dataset that are complex, diverse and of a massive scale.

Most agriculture datasets have data related to crop patterns, weather parameters, environmental conditions, soil types, soil nutrients, GIS and GPS data, farmer records, agriculture machinery data, such as yield monitoring and VRA [1.2].

[Nguyen, Nguyen, and Kim, 2017] proposes a a platform for collecting and analyzing agricultural big data by suporting multiple methods of collecting data from various data sources using Flume and MapReduce, by using multiple choices of data storage including HDFS, HBase, and Hive and multiple analysis modules with Spark and Hadoop.

Data stage	Process	Difficulties
Data capture	Sensor logs, UAV and Satellite Imagery, GPS locations, NDVI, Thermal ...	Availability and Quality
Data storage	Cloud platforms, Distributed File Systems, Hybrid Storage Systems, Block-chains, P2P ...	Cost and Safety of data
Data transfer	Internet, Intranet	Integrity and Speed
Dta transform	Data cleaning, Machine Learning and AI, Normalize and other statistics	Automation and Preprocessing
Data analytics	Yield models, Planting instructions, Benchmarking, Decision ontologies ...	Heterogeneity and Scalability
Data Marketing	Visualization and Representation	Ownership and Privacy

TABLE 1.2: Big-Data stages in PA

Another term being extensively used in agriculture is cloud computing of agricultural data [Woodard, 2016; Wolfert et al., 2017; Kaloxylos et al., 2012]. The idea is certainly innovative and new in agriculture but is a very old term in other science domains [Kamilaris, Kartakoullis, and Prenafeta-Boldu, 2017]. The idea of an "intergalactic computer network" was introduced in the 1960s by JCR Licklider, who was responsible for enabling the development of the Advanced Research Projects Agency Network (ARPANET) in 1969. His vision was for everyone on the globe to be interconnected and accessing programs and data at any site, from anywhere.

Cloud adoption is undoubtedly the cornerstone of digital transformation, it is the foundation for rapid, scalable application development and delivery and as such it has an important role in the development of modern agriculture. High-performance computation may allow for faster and more accurate agricultural management, which could improve decision-making quality, reduce information asymmetry, and increase profits [Jayaraman et al., 2016].

A research done by [Xia et al., 2018] evaluates the feasibility of applying cloud computing technology for spectrum-based classification of apple chilling injury, by using frameworks like Spark and support vector machines (SVM) classification models for multivariate classification and analysis of the spectral data sets. The results showed that the efficiency of the cloud computing platform was significantly improved by increasing the spectral data set capacity or number of working nodes.

[Wang et al., 2018] demonstrated a highly-integrated, cloud-based, low-cost and user-friendly portable NIRS system with its key components and main structure. The system was used to predict the maturity level and TSS content of sweet cherry samples. They named the system 'Seefruit' and they foresee the system to become a universal application for fruit quality detection, which offers a fundamental framework for future related research.

According to [Tan, 2016] Cloud computing is particularly beneficial for decision support in PA for specialty crops. Firstly because of its nature on scalability, cloud computing is able to handle large amount of data and scale automatically. Secondly through cloud computing it is possible to change quickly the number of server instances and other resources, based on the demand. Finally, agriculture decision support systems are increasingly hosted on Internet, to take advantage of internet-connected devices IoT and to build an online community. They developed a framework for cloud-based Decision Support for orchards and Automation systems that can acquire data from various sources, synthesize application-specific decisions, and control field devices from the cloud.

Predictive Analytics

One of the most exciting technologies presently being used and widely being transformed and developed has been the use of predictive analytics. Predictive analytics as a whole can be comprised of numerous different statistical abilities from modeling, machine learning, and data mining. Used for agriculture, these methods allow to analyze what has happened in the past on the farm, as well as what currently is happening and is going to happen, to make use of the data to predict the future and make decisions that impact the bottom line and end use of on-farm products [Kamilaris, Kartakoullis, and Prenafeta-Boldu, 2017; Nguyen, Nguyen, and Kim, 2017].

By learning from historical and future data based on measured variables, management and outcomes of decisions can more readily be made that can greatly impact efficiencies and processes. This is no easy task, as decisions and recommendations about the future require true datasets that are well defined from field to field, even different areas within the field. This insight helps producers to make otherwise challenging agronomic decisions that can take time to reach the field every day quicker and easier. It gives them the opportunity to make a fast decision off of digital information, often with the ability to be unbiased to the source, but relied upon the facts. True agronomic knowledge is essential for success and the right outputs for each digital tool. A small decision on the timing of an input application, could mean the difference between profitability or loss for that application. Predictive scores are given to each opportunity to help determine processes and decision making through analyzing datasets and confidence. Predictive analytics can support to discover relationship, and most importantly trends from those data [Zhang et al., 2017].

[Barbouchi et al., 2016], proposed a yield prediction from the input images generated by Radarsat 2. This work has a novel calculation technique from the relationship between the yield obtained from situ and backscatter. The data is collected at the end of the season with the Radarsat2 images. Single data is integrated with the data obtained during acquisition stage and next data during flowering period.

While in anther study [Badr et al., 2016] proposed a platform for geospatial data obtained from the yield from the single farm level up to the continental scale. They identified the coupling of Big-Data approach with the integrated repository and services (PAIRS), for the decision support system for agro-technology transfer (DSSAT) crop model. This foresees the global scale of geospatial analytics, and PAIRS provides the solution for the heterogeneous data to integrate the crop models with the dataset size of hundreds of terabytes.

In another paper [Suwantong et al., 2016], the evaluation takes place by NDVI technique which is calculated from the energy of electromagnetic waves obtained using the target crop. NDVI is calculated by the cosine function which is triply modulated with mean value, first stage of variables, and amplitude. The approach for determining the initial period of one crop at 8-day composite is obtained by moderate resolution imaging spectre radiometer.

By using a center of big data and intelligently storing, screening, calibrating, minning and extracting monitoring data [Zhang et al., 2017] established the crop growth model based on big data, which can predict and forecast the water requirement of crops in different growth periods and make the decision of automatic irrigation and fertilization, finally realizing timely and proper irrigation of crops.

Another paper [Sahu, Chawla, and Khare, 2018] presents research work utilizing a novel algorithm as to foresee the status of crop by monitoring the agriculture land data and advice which area of crop would be suitable for that land. Contingent on different farms, was intended to regularize the dataset fields that will suit for all different crop managements. The algorithm had the functionality of loading the dataset in database and comparing the previous dataset with the current processing dataset and give viable crop prediction scenarios.

[Bendre, Thool, and Thool, 2015] forecasts using a regression model and big data handle by cloud computing platforms which shows a considerable potential of data fusion in field of crop and water management. As per results model predicted the temperature and rainfall in the region by suggesting various decisions to farmers for deciding the crop pattern and water management in the future.

1.3.2 Computer Vision

Another important way of retrieving high volume of data (BigData) in agriculture is through cameras by using machine vision techniques. Computer vision (CV) is the ability of cameras and other visual sensors to capture raw video and process it into useful, actionable information. Information generated by CV is a type of unstructured data, its a collection of very huge amount of data with no structure or labels. The sheer volume of data is not the only challenge, a far bigger challenge is to understand it and act on it in real time.

In PA there are different areas where CV is applied:

- Processing and quality control
- Vegetation indices for yield mapping
- Thermal imagery for stress status
- Pest and disease monitoring

- Production growth and monitoring

One of the most interesting of those is the production growth and monitoring. Especially fruit counting and detection. Agriculture is most of the time repetitive, repetitive work of seeding, weeding, feeding, pruning, picking and harvesting, sorting and so on. Agricultural robots automate those slow, repetitive and dull tasks for farmers, allowing them to focus more on strategic matters and improving overall production yields [Edan, Han, and Kondo, 2009]. One of the most popular robotic application in agriculture are the autonomous harvesting and picking robots. That's because the speed and accuracy has increased significantly in recent years [Tao and Zhou, 2017; Bechar and Vigneault, 2016]. While the robots in addition to harvesting and picking can check at the same time the maturity level and sort based on size [Edan, Han, and Kondo, 2009]. However, there are many challenges for an autonomous robotic system to complete that task. In principle, for the robot to be fully capable to perform harvesting and picking, it needs a sophisticated detection algorithm in order to overcome challenges as naturally occurring changes in illumination, shape, pose, colour, and viewpoint [Barnea, Mairon, and Ben-Shahar, 2016].

The earliest fruit detection systems date since 1968 [Jiménez et al., 1999]. Using different methods and approaches based on photometric information (light reflectance difference from fruit and leaves in visible or infrared spectrum), these detectors were able to differentiate fruit from other parts of the tree. According to the reviews devoted to fruit detection by [Jiménez et al., 1999] and later on by [Kapach et al., 2012], there were many problems related to growth habit that had to be considered. The unstructured and uncontrolled outdoor environment also presents many challenges for computer vision systems in agriculture.

Light conditions have a major influence on fruit detection feasibility: direct sunlight results in saturated spots without color information and in shadows that cause standard segmentation procedures to split the apples surfaces into several fragments. In order to decrease the non-uniform illumination (daytime lighting can be bright, strong, directional and variable), [Payne et al., 2014] described a machine vision techniques to detect fruit based on images acquired during night time using artificial light sources. The results described show 78% fruit detection, 10% errors and suggesting that artificial lighting at night can provide consistent illumination without strong directional shadows.

In a different approach, [Kelman and Linker, 2014] presented an algorithm for localizing spherical fruit that have a smooth surface, such as apples, using only shape analysis and in particular convexity. It is shown that in the images used for the study, more than 40% of the apple profiles were none-convex, more than 85% of apple edges had 15% or more non-convex profiles, and more than 45% of apple edges had 50% or more non-convex profiles. Overall, 94% of the apples were correctly detected and 14% of the detections corresponded to false positives. Despite hight accuracy number, the model is very specific to apples and wound not be extensible to other fruit crops with less spherical shapes. [Kapach et al., 2012] explains colour highlights and spherical attributes, which tend to appear more often on the smoother, more specular, and typically elliptical regions like fruit where the surface normal bisects the angle between illumination and viewing directions.

A method for estimating the number of apple fruit in the orchard using a thermal camera was developed by [Stajnko, Lakota, and Hočevar, 2004]. It shows an algorithm able to threshold, count and report fruit' morphological characteristics from thermal images captured under natural conditions in the orchard. Snd since fruit have bigger volume than leaves, they keep heat for higher amount of time, making them easily to detect while the temperatures start to cool down. However position of foliage on the tree prevent sunshine heating all fruit to the same degree for the same period. fruit inside the canopy are exposed to the sunshine for a shorter time than those outside the canopy, so they could be cooled in almost the same time as leaves, making very difficult to detect the temperature gradient between inner fruit and leaves.

With the development of better sensor cameras and vision techniques in recent years, more sophisticated approaches have been used for apple detection. Range-based devices such as stereo vision cameras, ultrasonic sensors, laser scanners and Time of Flight cameras measure the distance from the sensor to the observed objects, providing accurate range information in real-time and are consistent with varying lighting condition and are used more widely in agricultural machines nowadays. [Si, Liu, and Feng, 2015] describes location of apples in trees using stereoscopic vision. The advantage of the active triangulation method is that the range data may be obtained without much computation and the speed is very high for any robotic harvesting application. While [Jiang, Peng, and Ying, 2008] developed a binocular stereo vision tomato harvester in greenhouse. In this method, a pair of stereo images was obtained by stereo cameras, and transformed to grey-scale images. According to the grey correlation, corresponding points of the stereo images were searched, and a depth image was obtained by calculating distances between tomatoes and stereo cameras based on triangulation principle.

[Barnea, Mairon, and Ben-Shahar, 2016] describes RGB and range data to analyse shape-related features of objects both in the image plane and 3D space. By combining both highlight detection and 3D shape/range data a colour-agnostic fruit detection framework was build. The framework is composed of two steps one following the other. In the first step a high level based feature detector is applied to detect the most probably regions of the frame that can contain fruit. The second step follows immediately after by using a depth-based object classification of the resultant feature vector using a support vector machine (SVM) on those regions passed by step one. In another work [Nguyen et al., 2014] developed a multi-phase algorithm to detect and localize apple fruit by combining an RGB-D camera and point cloud processing techniques. [Tao and Zhou, 2017] developed an automatic apple recognition system based on the fusion of color and 3D features.

For many years, traditional computer vision approaches have been extensively adopted in the agricultural field. More recently, with the significant increase in computational power, in particular with special purpose processors optimized for matrix-like data processing and large amount of data calculations (eg. Graphical Processing Unit - GPU), a lot of deep learning, CNN models and methodologies specifically developed have achieved unprecedented breakthroughs [LeCun, Bengio, and Hinton, 2015].

[Sa et al., 2016], developed a model called Deepfruit, for fruit detection. Adopting a Faster R-CNN model, goal was to build an accurate, fast and reliable fruit detection system. Model after training was able to achieve 0.838 precision and recall in the detection of sweet pepper. In addition they used a multi-modal fusion approach that combines the information from RGB and NIR images. The bottle-neck of the model is that in order to deploy on a real robot system, the processing performance required is a GPU of 8GB or more.

It is well known that all deep leaning models, to have a high accuracy, they need high number of data [Krizhevsky, Sutskever, and Hinton, 2012]. In case of CNN, more pictures of the object of interest, the better the classification/detection performance is. In a model called DeepCount, [Rahnemoonfar and Sheppard, 2017] developed a CNN architecture based on Inception-ResNet for counting fruit. In order to use less training data, [Rahnemoonfar and Sheppard, 2017] used a different approach. They use another model to generate synthetic images/data to feed te main model to train on. Those generated images were simply a brownish and greenish color background with red circles drawn above it to simulate the background and tomato plant with fruit on it. They used twenty-four thousand pictures generated to feed into the model. The model was then tested on real world images and showed an accuracy from 85-80%.

To better understand the amount of data needed for better fruit detection, [Bargoti and Underwood, 2017] used different data augmentation techniques and transfer learning from other fruit. It is shown that transferring weights between different fruit did not have significant performance gains, while data augmentation like flip and scale were found to improve performance resulting in equivalent performance with less than half the number of training images.

1.4 Automated crop management systems - ACMS

The last step in PA4.0 is automation. When enough observations and predictive analytics are made, then farmers are given the choice to take immediate action in order to improve production yield, while reducing resources required. This is done through some chain automation pipelines or autonomous systems (robots). Robots are gradually changing every industry and agriculture isn't an exception. The use of robotics in PA isn't widespread yet. However, it's expected to grow significantly in coming years.

1.4.1 Robotics and Autonomous Systems

Many modern farmers are already high-tech. Digitally-controlled farm implements are regularly in use. There are partially and fully automatic devices for most aspects of agricultural functions from grafting to planting, from harvesting to sorting, packaging and boxing. Farmers use software systems and aerial survey maps and data to guide their field operations. They also use auto-steer systems included in many new tractors which follow GPS and software guidance. Some farmers are already transitioning some of their operations to full autonomy. Thus forward-thinking farm owners today may be able to skip over slow, incremental improvements and jump directly to robotic and autonomous automation.

Fully autonomous vehicles have been studied for many years, with a number of innovations explored as early as the 1920s. The concept of fully autonomous agricultural vehicles is far from new; examples of early driverless tractor prototypes using leader cable guidance systems date back to the 1950s and 1960s [[Basu et al., 2018]. The potential for combining computers with image sensors provided huge opportunities for machine vision based guidance systems.

In Agriculture, autonomous systems can be grouped in three main categories [Emmi et al., 2014]:

1. Big autonomous tractors
2. Small specialized robots
3. Swarm or fleet robotics

Autonomous tractors have been studied and in use in agriculture for many years. However precision agriculture was the one that helped advance vehicle guidance in terms of providing position information that is required for vehicle guidance [Reid et al., 2000]. The key elements of automatic guidance are navigation sensors, a vehicle motion model, a navigation planner, and a steering controller

A tractor usually operates on all terrains, and there are a lot of unpredictable disturbances and noise sources to the signals from the navigation sensors. Therefore, it is necessary to have an effective means for signal conditioning and system state estimation in the sensor fusion modules [Noguchi et al., 2001]. The topography, vegetation landscape, soil composition, texture and structure, air visibility, illumination, light quality and atmospheric conditions change at rates varying from seconds to months and on scales from millimeters to kilometers [Bechar and Vigneault, 2016]. In order to perform well, the next-gen agriculture autonomous robots (Agbots) must be able to recognize and understand the physical properties of each specific object encountered, and also be able to work under both varying field and controlled environment conditions [Eizicovits and Berman, 2014]. Therefore, sensing systems, robotics arms, specialized manipulators, effectors should be able to work under different and unstable environmental conditions [Bechar and Vigneault, 2017].

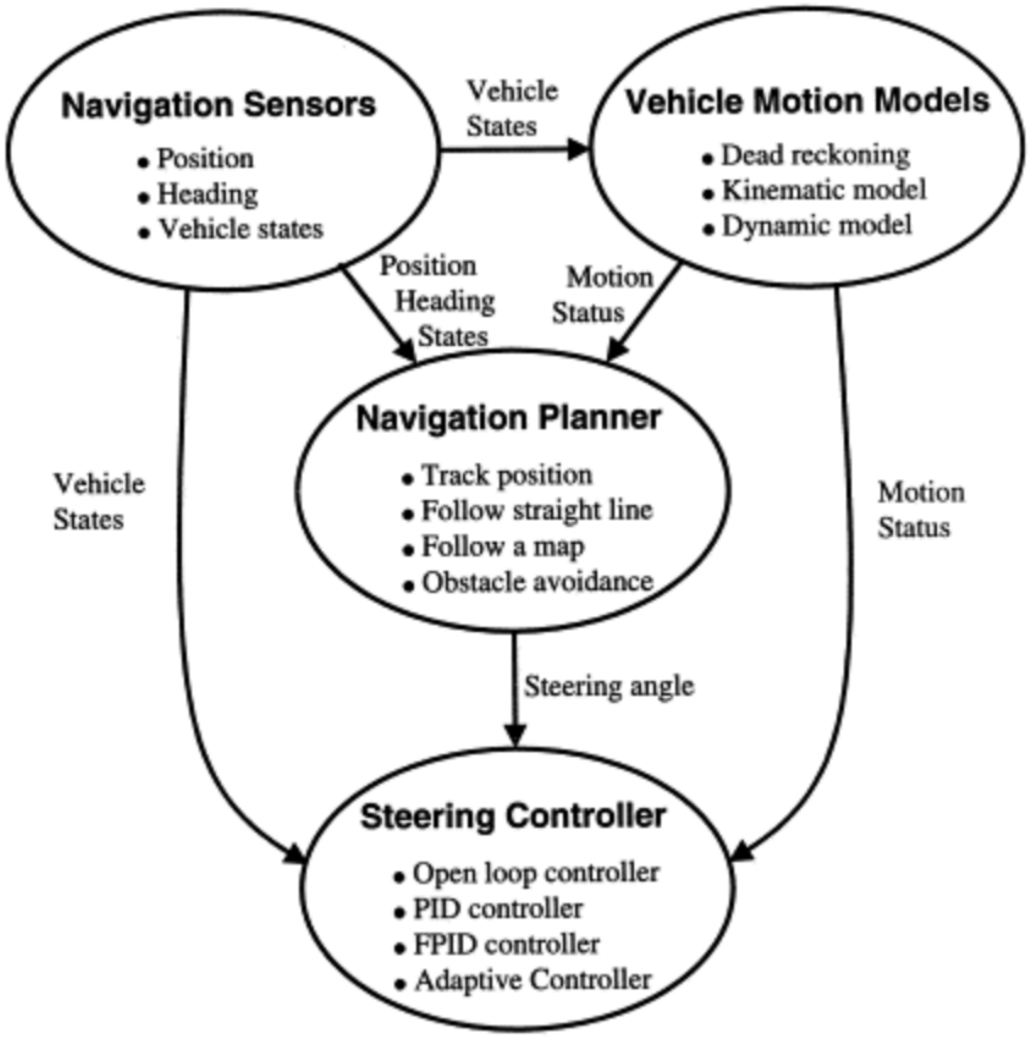


FIGURE 1.3: Basic elements of autonomous tractor

Robots have wide applications in PA, ranging from soil analysis, seedling, weed control, environmental monitoring, harvesting and so on, but in a broader perspective they can be groped in:

- Harvest Management
- Autonomous navigation
- Pest management and spraying
- Weed management and mowing
- Soil Management
- Irrigation Management
- Remote camera sensing UAVs
- Pruning and Thinning
- Sorting and packing
- Seedling and nursery
- Transporting and cleaning
- Other

According to [Bechar, Meyer, and Edan, 2009; Oren, Bechar, and Edan, 2011] a robot, to perform a fully autonomous agricultural action needs to go through four continuous steps: first, the robot senses and acquires raw data from and about the environment, task and/or its state using various sensors; secondly, the robot processes and analyses the data received from its sensors to generate reasoning and a perception of the environment, the task or its state to some level of situation awareness; thirdly, the robot generates an operational plan based on its perception of the environment and state, or the task objectives; and lastly, the robot executes the required actions included in the operational plan.

Stage	Task
Sensing	Sensory inputs like: RGB cameras, LIDAR, Sonar, rotary encoders, potentiometers, resistors...
Analyzing	Landmark detection, Point-cloud analyzing, Kinematics and Inverse kinematics of manipulator...
Planning	Trajectory estimation, object voidance...
Action	Performing the planned action, triggering other actios based on location, time, another action ...

TABLE 1.3: Continuous stages of an Agbot [Bechar and Vigneault, 2016]

In a research [Malavazi et al., 2018] developed an approach for autonomous robot navigation inside crops using LIDAR (Light Detection and Ranging). The research presents a new approach to extract lines from a point cloud with application to agricultural robot autonomous navigation in a GPS denied environment. However they show that to change row, for instance, is quite difficult with only one LiDAR sensor in front of the robot. Indeed, when the robot is at the end of a row, it does not have any information about what is behind it.

A new approach in robotics farming has started to gain attraction due to its flexibility. Swarm robotics is a new type of robotics that allows simple individual robots to work together to perform complex tasks. However, swarm robotics research is still confined into the lab, and no application in the field is currently available.

The associated theoretical foundations for fleets/swarm robotics and applications for agriculture are being researched and developed [Emmi et al., 2014]. There are many advantages using a swarm of robots; using a group of robots cooperating with each other to achieve a well-defined objective is an emerging and necessary concept to achieve agricultural goals. Artificial swarm intelligence has been inspired by biological studies of behavior of ants, bees, wasps and termites and has been the spark for changing the perspective on how robots were understood, and gives a new trend to their functionality; such as solving problems through large population. [Anil et al., 2015]. The individual robots can be regarded as agents with simple and single abilities. Some of them have the ability to evolve themselves when dealing with certain problems to make better compatibility and decisions [Tan and Zheng, 2013].

However, for a robotic agricultural application, considerable information must be processed, and a wide number of actuation signals must be controlled, which may present a number of technical drawbacks. Thus, an important limitation is that the number of total sensors, actuators, and computers/controllers... increases according to the number of swarm units, a failure in one robot component causes the entire swarm to malfunction. This influences swarm reliability, which is of extremely important for the application of automated systems to real tasks in agriculture [Emmi et al., 2013].

Mobile Agricultural Robot Swarms (MARS) aims at the development of small and stream-lined mobile agricultural robot units. The concept addresses looming challenges of to optimize plant specific precision farming, leading to reduced input of seeds, fertilizer and pesticides and to increased yields, to reduce the massive soil compaction as well as energy consumption of heavy machinery and to meet the increasing demand for flexible to use, highly automated and simple to operate systems, anticipating challenges arising from climate change as well as shortage of skilled labour.

Swarm Robotics for Agricultural Applications (SAGA) is using a group of small unmanned aerial vehicles to monitor a sugar beet field and cooperatively map the presence of harmful weed. They aim to determine when to perform weeding and on which parts of the farm land, by multi-rotor UAVs enhanced with on-board camera and vision processing, radio communication systems and suitable protocols to support safe swarm operations. They propose a solution that exploits multiple UAVs that can focus on areas of interest while abandoning those areas of the field that do not require closer inspection.

1.5 Conclusions

We stated the narrative over and over: by the year 2050 the global agricultural community will have to nearly double its output to feed 9 billion people. Efficiency and productivity will increase in the coming years as **Precision Agriculture** becomes bigger and farms become more connected. But while the growing number of connected devices and sensors in the farm represents a big opportunity for farmers, it also adds complexity. The solution lies in making use of cognitive technologies that help understand, learn, adapt, reason, act, interact, and increase efficiency. Key innovations in new sensors, IoT, cloud computing, bg data, artificial intelligence and robotics will assist farmers with answers and recommendations on specific problems.

As information becomes critical for good decision-making as part of the in-field support, it must be collected, stored and interpreted in a timely manner. Because of the overhead cost to maintain local computers, because of the computing power needed needed, and because of expertise needed, many farmers will rely on cloud computing and remote servers for data processing anf predictive analytics. Startups that are aiming to capture, integrate and analyze all those data coming off the farm, whether from sensors, drones, machinery, or imagery, stand to significantly improve the decision-making process for farmers. With more and more companies and start-ups coming up with new and innovative agricultural tools and platforms, interoperability is rapidly becoming a point of concern. The various available tools and technologies often do not follow the same technology standards/platforms as a result of which there is a lack of uniformity in the final analysis referred to farmers.

Robots may one day consist of complete autonomous swarms of relatively small, smart, and cheap units with an optimized allocation of all resources which will manage farms, collect data and perform various tasks. Robots will likely make inroads fastest in areas where the labor is backbreaking, and peak harvest times create a short supply of workers. However most robots built to date are for just one specialized tasks. A fully autonomous system, be that a single unit or multiple units where the whole production chain can be monitored, analysed and managed is still not here. Getting farmers thoroughly acquainted with the concept of smart farming its of utmost importance and proven to be very challenging tasks. In addition to the learning curve there is the motivation and uncertainty to new concepts such as robotics, cloud, IoT and so on. Unfortunately though, the benefits do not become apparent from the very beginning and many farmers still view this use of advanced technology in agriculture as risky and uncertain.

References

Akyildiz, Ian F. and Ismail H. Kasimoglu (2004). "Wireless sensor and actor networks: research challenges". In: *Ad Hoc Networks* 2.4, pp. 351–367. DOI: 10.1016/j.adhoc.2004.04.003. URL: https://doi.org/10.1016%2Fj.adhoc.2004.04.003.

Akyildiz, Ian F. and Erich P. Stuntebeck (2006). "Wireless underground sensor networks: Research challenges". In: *Ad Hoc Networks* 4.6, pp. 669–686. DOI: 10.1016/j.adhoc.2006.04.003. URL: https://doi.org/10.1016%2Fj.adhoc.2006.04.003.

Akyildiz, Ian F et al. (2002). "Wireless sensor networks: a survey". In: *Computer networks* 38.4, pp. 393–422.

Ampatzidis, Y. et al. (2012). "Preliminary testing of a system for evaluating picker efficiency in tree fruit". In: vol. 5, pp. 3726–3737. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84871765509&partnerID=40&md5=3df562f44fe09aa56a93e2705380e32c.

Ampatzidis, Y.G. and S.G. Vougioukas (2009). "Field experiments for evaluating the incorporation of RFID and barcode registration and digital weighing technologies in manual fruit harvesting". In: *Computers and Electronics in Agriculture* 66.2, pp. 166–172. DOI: 10.1016/j.compag.2009.01.008. URL: https://doi.org/10.1016%2Fj.compag.2009.01.008.

Anil, H. et al. (2015). "Revolutionizing Farming Using Swarm Robotics". In: *2015 6th International Conference on Intelligent Systems, Modelling and Simulation*. IEEE. DOI: 10.1109/isms.2015.2. URL: https://doi.org/10.1109%2Fisms.2015.2.

Aubert, Benoit A., Andreas Schroeder, and Jonathan Grimaudo (2012). "IT as enabler of sustainable farming: An empirical analysis of farmers adoption decision of precision agriculture technology". In: *Decision Support Systems* 54.1, pp. 510–520. DOI: 10.1016/j.dss.2012.07.002. URL: https://doi.org/10.1016%2Fj.dss.2012.07.002.

Badr, G. et al. (2016). "Toward large-scale crop production forecasts for global food security". In: *IBM Journal of Research and Development* 60.5/6, 5:1–5:11. DOI: 10.1147/jrd.2016.2591698. URL: https://doi.org/10.1147%2Fjrd.2016.2591698.

Barbouchi, Meriem et al. (2016). "Yield estimation of the winter wheat using Radarsat 2 polarimetric SAR reponse". In: *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. IEEE. DOI: 10.1109/atsip.2016.7523143. URL: https://doi.org/10.1109%2Fatsip.2016.7523143.

Bargoti, Suchet and James Underwood (2017). "Deep fruit detection in orchards". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. DOI: 10.1109/icra.2017.7989417. URL: https://doi.org/10.1109%2Ficra.2017.7989417.

Barnea, Ehud, Rotem Mairon, and Ohad Ben-Shahar (2016). "Colour-agnostic shape-based 3D fruit detection for crop harvesting robots". In: *Biosystems Engineering*

146, pp. 57–70. DOI: 10.1016/j.biosystemseng.2016.01.013. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.01.013.

Bastiaanssen, Wim G.M, David J Molden, and Ian W Makin (2000). “Remote sensing for irrigated agriculture: examples from research and possible applications”. In: *Agricultural Water Management* 46.2, pp. 137–155. DOI: 10.1016/s0378-3774(00)00080-9. URL: https://doi.org/10.1016%2Fs0378-3774%2800%2900080-9.

Basu, Subhajit et al. (2018). “Legal framework for small autonomous agricultural robots”. In: *AI & SOCIETY*. DOI: 10.1007/s00146-018-0846-4. URL: https://doi.org/10.1007%2Fs00146-018-0846-4.

Bechar, A., J. Meyer, and Y. Edan (2009). “An Objective Function to Evaluate Performance of Human–Robot Collaboration in Target Recognition Tasks”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 39.6, pp. 611–620. DOI: 10.1109/tsmcc.2009.2020174. URL: https://doi.org/10.1109%2Ftsmcc.2009.2020174.

Bechar, Avital and Clément Vigneault (2016). “Agricultural robots for field operations: Concepts and components”. In: *Biosystems Engineering* 149, pp. 94–111. DOI: 10.1016/j.biosystemseng.2016.06.014. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.06.014.

– (2017). “Agricultural robots for field operations. Part 2: Operations and systems”. In: *Biosystems Engineering* 153, pp. 110–128. DOI: 10.1016/j.biosystemseng.2016.11.004. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.11.004.

Bendre, M. R., R. C. Thool, and V. R. Thool (2015). “Big data in precision agriculture: Weather forecasting for future farming”. In: *2015 1st International Conference on Next Generation Computing Technologies NGCT*. IEEE. DOI: 10.1109/ngct.2015.7375220. URL: https://doi.org/10.1109%2Fngct.2015.7375220.

Brase, T.A. (2006). *Precision Agriculture*. Thomson/Delmar Learning. ISBN: 9781401881054. URL: https://books.google.co.za/books?id=hEBkQgAACAAJ.

Brown, J.H. (2012). “Robotics in agriculture”. In: vol. 2. cited By 0, pp. 1427–1439. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84878042468&partnerID=40&md5=53d51432d1e740d68e9143bb3fb6b4ad.

Carrabs, Francesco et al. (2015). “Maximizing lifetime in wireless sensor networks with multiple sensor families”. In: *Computers & Operations Research* 60, pp. 121–137. DOI: 10.1016/j.cor.2015.02.013. URL: https://doi.org/10.1016%2Fj.cor.2015.02.013.

Chen, Shu Fa, Cheng Long Feng, and Shi Ping Zhang (2012). “Design and Research on Variable-Rate Fertilization Computer Control System”. In: *Applied Mechanics and Materials* 249-250, pp. 610–615. DOI: 10.4028/www.scientific.net/amm.249-250.610. URL: https://doi.org/10.4028%2Fwww.scientific.net%2Famm.249-250.610.

Choudhury, Sushabhan et al. (2015). “ZigBee and Bluetooth Network based Sensory Data Acquisition System”. In: *Procedia Computer Science* 48, pp. 367–372. DOI: 10.

1016/j.procs.2015.04.195. URL: https://doi.org/10.1016%2Fj.procs.2015.04.195.

Chougule, Archana, Vijay Kumar Jha, and Debajyoti Mukhopadhyay (2016). "Using IoT for integrated pest management". In: *2016 International Conference on Internet of Things and Applications (IOTA)*. IEEE. DOI: 10.1109/iota.2016.7562688. URL: https://doi.org/10.1109%2Fiota.2016.7562688.

Colbert, Amy, Nick Yee, and Gerard George (2016). "The Digital Workforce and the Workplace of the Future". In: *Academy of Management Journal* 59.3, pp. 731–739. DOI: 10.5465/amj.2016.4003. eprint: https://journals.aom.org/doi/pdf/10.5465/amj.2016.4003. URL: https://journals.aom.org/doi/abs/10.5465/amj.2016.4003.

Dobkin, Daniel M and Titus Wandinger (2005). "A radio-oriented introduction to RFID-protocols, tags and applications". In: *High Frequency Electronics* 4.8, pp. 32–46.

Dong, Xin, Mehmet C. Vuran, and Suat Irmak (2013). "Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems". In: *Ad Hoc Networks* 11.7, pp. 1975–1987. DOI: 10.1016/j.adhoc.2012.06.012. URL: https://doi.org/10.1016%2Fj.adhoc.2012.06.012.

Edan, Yael, Shufeng Han, and Naoshi Kondo (2009). "Automation in Agriculture". In: *Springer Handbook of Automation*. Springer Berlin Heidelberg, pp. 1095–1128. DOI: 10.1007/978-3-540-78831-7_63. URL: https://doi.org/10.1007%2F978-3-540-78831-7_63.

Eizicovits, Danny and Sigal Berman (2014). "Efficient sensory-grounded grasp pose quality mapping for gripper design and online grasp planning". In: *Robotics and Autonomous Systems* 62.8, pp. 1208–1219. DOI: 10.1016/j.robot.2014.03.011. URL: https://doi.org/10.1016%2Fj.robot.2014.03.011.

Elijah, Olakunle et al. (2018). "An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges". In: *IEEE Internet of Things Journal*, pp. 1–1. DOI: 10.1109/jiot.2018.2844296. URL: https://doi.org/10.1109%2Fjiot.2018.2844296.

Emmi, Luis et al. (2013). "Fleets of robots for precision agriculture: a simulation environment". In: *Industrial Robot: An International Journal* 40.1, pp. 41–58. DOI: 10.1108/01439911311294246. URL: https://doi.org/10.1108%2F01439911311294246.

Emmi, Luis et al. (2014). "New Trends in Robotics for Agriculture: Integration and Assessment of a Real Fleet of Robots". In: *The Scientific World Journal* 2014, pp. 1–21. DOI: 10.1155/2014/404059. URL: https://doi.org/10.1155%2F2014%2F404059.

Ensmenger, Nathan (2011). "The Quest for Artificial Intelligence: A History of Ideas and Achievements". In: *Isis* 102.3, pp. 588–589. DOI: 10.1086/663065. URL: https://doi.org/10.1086%2F663065.

Ferentinos, Konstantinos P. and Theodore A. Tsiligiridis (2007). “Adaptive design optimization of wireless sensor networks using genetic algorithms”. In: *Computer Networks* 51.4, pp. 1031–1051. DOI: 10.1016/j.comnet.2006.06.013. URL: https://doi.org/10.1016%2Fj.comnet.2006.06.013.

Fountas, S. et al. (2006). “A model of decision-making and information flows for information-intensive agriculture”. In: *Agricultural Systems* 87.2, pp. 192–210. DOI: 10.1016/j.agsy.2004.12.003. URL: https://doi.org/10.1016%2Fj.agsy.2004.12.003.

Fröder, Christoph (2014). *Employment in European Agriculture: Labour Costs, Flexibility and Contractual Aspects*. Institut der deutschen Wirtschaft Köln, Geopa-Copa.

Gautam, Rahul et al. (2017). “Traceability using RFID and its formulation for a kiwifruit supply chain”. In: *Computers & Industrial Engineering* 103, pp. 46–58. DOI: 10.1016/j.cie.2016.09.007. URL: https://doi.org/10.1016%2Fj.cie.2016.09.007.

Ghaani, Masoud et al. (2016). “An overview of the intelligent packaging technologies in the food sector”. In: *Trends in Food Science & Technology* 51, pp. 1–11. DOI: 10.1016/j.tifs.2016.02.008. URL: https://doi.org/10.1016%2Fj.tifs.2016.02.008.

Gill, Sukhpal Singh, Inderveer Chana, and Rajkumar Buyya (2017). “IoT Based Agriculture as a Cloud and Big Data Service”. In: *Journal of Organizational and End User Computing* 29.4, pp. 1–23. DOI: 10.4018/joeuc.2017100101. URL: https://doi.org/10.4018%2Fjoeuc.2017100101.

Guo, X. et al. (2014). “Intelligent data acquisition and cloud services for apple orchard”. In: *International Journal of Agricultural and Biological Engineering* 7.2, pp. 146–153. DOI: 10.3965/j.ijabe.20140702.018. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84899651938&doi=10.3965%2fj.ijabe.20140702.018&partnerID=40&md5=e745230e785b969b600a39e3c57f8611.

Harrison, Peter and Joseph Wolyniak (2015). “The History of ‘Transhumanism’”. In: *Notes and Queries* 62.3, pp. 465–467. DOI: 10.1093/notesj/gjv080. URL: https://doi.org/10.1093%2Fnotesj%2Fgjv080.

Hedley, Carolyn (2014). “The role of precision agriculture for improved nutrient management on farms”. In: *Journal of the Science of Food and Agriculture* 95.1, pp. 12–19. DOI: 10.1002/jsfa.6734. URL: https://doi.org/10.1002%2Fjsfa.6734.

Hu, Jing et al. (2010). “Design and implementation of wireless sensor and actor network for precision agriculture”. In: *2010 IEEE International Conference on Wireless Communications, Networking and Information Security*. IEEE. DOI: 10.1109/wcins.2010.5544151. URL: https://doi.org/10.1109%2Fwcins.2010.5544151.

Huang, Yanbo et al. (2018). “Agricultural remote sensing big data: Management and applications”. In: *Journal of Integrative Agriculture* 17.9, pp. 1915–1931. DOI: 10.1016/s2095-3119(17)61859-8. URL: https://doi.org/10.1016%2Fs2095-3119%2817%2961859-8.

Jayaraman, Prem et al. (2016). "Internet of Things Platform for Smart Farming: Experiences and Lessons Learnt". In: *Sensors* 16.11, p. 1884. DOI: 10.3390/s16111884. URL: https://doi.org/10.3390%2Fs16111884.

Jiang, Huanyu, Yongshi Peng, and Yibin Ying (Jan. 2008). "Measurement of 3-D Locations of Ripe Tomato by Binocular Stereo Vision for Tomato Harvesting". In: *American Society of Agricultural and Biological Engineers Annual International Meeting 2008, ASABE 2008* 11.

Jiménez, A.R. et al. (1999). "Automatic fruit recognition: a survey and new results using Range/Attenuation images". In: *Pattern Recognition* 32.10, pp. 1719–1736. DOI: 10.1016/s0031-3203(98)00170-8. URL: https://doi.org/10.1016/s0031-3203(98)00170-8.

Kaloxylos, Alexandros et al. (2012). "Farm management systems and the Future Internet era". In: *Computers and Electronics in Agriculture* 89, pp. 130–144. DOI: 10.1016/j.compag.2012.09.002. URL: https://doi.org/10.1016%2Fj.compag.2012.09.002.

Kamilaris, Andreas, Andreas Kartakoullis, and Francesc X. Prenafeta-Boldu (2017). "A review on the practice of big data analysis in agriculture". In: *Computers and Electronics in Agriculture* 143, pp. 23–37. DOI: 10.1016/j.compag.2017.09.037. URL: https://doi.org/10.1016%2Fj.compag.2017.09.037.

Kapach, Keren et al. (2012). "Computer vision for fruit harvesting robots - State of the art and challenges ahead". In: *International Journal of Computational Vision and Robotics* 3.1/2, p. 4. DOI: 10.1504/ijcvr.2012.046419. URL: https://doi.org/10.1504%2Fijcvr.2012.046419.

Kelman, Eliyahu (Efim) and Raphael Linker (2014). "Vision-based localisation of mature apples in tree images using convexity". In: *Biosystems Engineering* 118, pp. 174–185. DOI: 10.1016/j.biosystemseng.2013.11.007. URL: https://doi.org/10.1016%2Fj.biosystemseng.2013.11.007.

Kim, Yunseop, R.G. Evans, and W.M. Iversen (2008). "Remote Sensing and Control of an Irrigation System Using a Distributed Wireless Sensor Network". In: *IEEE Transactions on Instrumentation and Measurement* 57.7, pp. 1379–1387. DOI: 10.1109/tim.2008.917198. URL: https://doi.org/10.1109%2Ftim.2008.917198.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. NIPS'12. Lake Tahoe, Nevada: Curran Associates Inc., pp. 1097–1105. URL: http://dl.acm.org/citation.cfm?id=2999134.2999257.

Le, Tuan Dinh and Dat Ho Tan (2015). "Design and deploy a wireless sensor network for precision agriculture". In: *2015 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science (NICS)*. IEEE. DOI: 10.1109/nics.2015.7302210. URL: https://doi.org/10.1109%2Fnics.2015.7302210.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). “Deep learning”. In: *Nature* 521.7553, pp. 436–444. DOI: 10.1038/nature14539. URL: https://doi.org/10.1038/nature14539.

Li, Suhong et al. (2006). “Radio frequency identification technology: applications, technical challenges and strategies”. In: *Sensor Review* 26.3, pp. 193–202.

Li, Xuewei and Xueyan Li (2018). “Big Data and Its Key Technology in the Future”. In: *Computing in Science & Engineering* 20.4, pp. 75–88. DOI: 10.1109/mcse.2018.042781329. URL: https://doi.org/10.1109%2Fmcse.2018.042781329.

Lima, Yuri Oliveira de and Jano Moreira de Souza (2017). “The future of work: Insights for CSCW”. In: *2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE. DOI: 10.1109/cscwd.2017.8066668. URL: https://doi.org/10.1109%2Fcscwd.2017.8066668.

Lin, Jzau-Sheng and Chun-Zu Liu (2008). “A monitoring system based on wireless sensor network and an SoC platform in precision agriculture”. In: *2008 11th IEEE International Conference on Communication Technology*. IEEE. DOI: 10.1109/icct.2008.4716133. URL: https://doi.org/10.1109%2Ficct.2008.4716133.

Liqiang, Zhao et al. (2011). “A Crop Monitoring System Based on Wireless Sensor Network”. In: *Procedia Environmental Sciences* 11, pp. 558–565. DOI: 10.1016/j.proenv.2011.12.088. URL: https://doi.org/10.1016%2Fj.proenv.2011.12.088.

Luvisi, Andrea et al. (2011). “Implanting RFIDs into Prunus to facilitate electronic identification in support of sanitary certification”. In: *Biosystems Engineering* 109.2, pp. 167–173. DOI: 10.1016/j.biosystemseng.2011.03.001. URL: https://doi.org/10.1016%2Fj.biosystemseng.2011.03.001.

Lv, Ruichao, Lianfeng Shen, and Jing Hu (2009). “Design and Implementation of a Wireless Sensor Network Gateway Supporting Multi-Mode Wide Area Access and Video Monitoring”. In: *2009 First International Conference on Information Science and Engineering*. IEEE. DOI: 10.1109/icise.2009.460. URL: https://doi.org/10.1109%2Ficise.2009.460.

Mainwaring, A. et al. (2002). “Wireless sensor networks for habitat monitoring”. In: pp. 88–97. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0036986465&partnerID=40&md5=88de11e091b77d63b4e012d6047d50cc.

Malavazi, Flavio B.P. et al. (2018). “LiDAR-only based navigation algorithm for an autonomous agricultural robot”. In: *Computers and Electronics in Agriculture* 154, pp. 71–79. DOI: 10.1016/j.compag.2018.08.034. URL: https://doi.org/10.1016%2Fj.compag.2018.08.034.

Manfrini, Luigi, James Taylor, and Luca Grappadelli (Apr. 2009). “Spatial Analysis of the Effect of Fruit Thinning on Apple Crop Load”. In: 74, pp. 54–60.

McBratney, Alex et al. (2005). “Future Directions of Precision Agriculture”. In: *Precision Agriculture* 6.1, pp. 7–23. DOI: 10.1007/s11119-005-0681-8. URL: https://doi.org/10.1007%2Fs11119-005-0681-8.

Mekala, Mahammad Shareef and P. Viswanathan (2017). "A Survey: Smart agriculture IoT with cloud computing". In: *2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS)*. IEEE. DOI: 10.1109/icmdcs.2017.8211551. URL: https://doi.org/10.1109%2Ficmdcs.2017.8211551.

Mirhosseini, Mina, Fatemeh Barani, and Hossein Nezamabadi-pour (2017). "Design optimization of wireless sensor networks in precision agriculture using improved BQIGSA". In: *Sustainable Computing: Informatics and Systems* 16, pp. 38–47. DOI: 10.1016/j.suscom.2017.08.006. URL: https://doi.org/10.1016%2Fj.suscom.2017.08.006.

Morandi, Brunella et al. (2007). "A low-cost device for accurate and continuous measurements of fruit diameter". In: *HortScience* 42.6, pp. 1380 –1382. ISSN: 00185345.

Mulla, David J. (2013). "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps". In: *Biosystems Engineering* 114.4, pp. 358–371. DOI: 10.1016/j.biosystemseng.2012.08.009. URL: https://doi.org/10.1016%2Fj.biosystemseng.2012.08.009.

Nguyen, Tien Thanh et al. (2014). "Apple detection algorithm for robotic harvesting using a RGB-D camera". In:

Nguyen, Van-Quyet, Sinh Ngoc Nguyen, and Kyungbaek Kim (2017). "Design of a Platform for Collecting and Analyzing Agricultural Big Data". In: *Journal of Digital Contents Society* 18.1, pp. 149–158. DOI: 10.9728/dcs.2017.18.1.149. URL: https://doi.org/10.9728%2Fdcs.2017.18.1.149.

Noguchi, Noboru et al. (2001). "Autonomous Agricultural Tractor with an Intelligent Navigation System". In: *IFAC Proceedings Volumes* 34.19, pp. 197–202. DOI: 10.1016/s1474-6670(17)33136-1. URL: https://doi.org/10.1016%2Fs1474-6670%2817%2933136-1.

OGrady, Michael J. and Gregory M.P. OHare (2017). "Modelling the smart farm". In: *Information Processing in Agriculture* 4.3, pp. 179–187. DOI: 10.1016/j.inpa.2017.05.001. URL: https://doi.org/10.1016%2Fj.inpa.2017.05.001.

Ojha, Tamoghna, Sudip Misra, and Narendra Singh Raghuwanshi (2015). "Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges". In: *Computers and Electronics in Agriculture* 118, pp. 66–84. DOI: 10.1016/j.compag.2015.08.011. URL: https://doi.org/10.1016%2Fj.compag.2015.08.011.

Oren, Y., A. Bechar, and Y. Edan (2011). "Performance analysis of a human–robot collaborative target recognition system". In: *Robotica* 30.05, pp. 813–826. DOI: 10.1017/s0263574711001020. URL: https://doi.org/10.1017%2Fs0263574711001020.

Ozkose, Hakan, Emin S. Ari, and Cevriye Gencer (2015). "Yesterday, Today and Tomorrow of Big Data". In: *Procedia - Social and Behavioral Sciences* 195, pp. 1042–1050. DOI: 10.1016/j.sbspro.2015.06.147. URL: https://doi.org/10.1016%2Fj.sbspro.2015.06.147.

Park, Junyoung, Sunggu Lee, and Sungjoo Yoo (2015). "Time slot assignment for convergecast in wireless sensor networks". In: *Journal of Parallel and Distributed Computing* 83, pp. 70–82. DOI: 10.1016/j.jpdc.2015.05.004. URL: https://doi.org/10.1016%2Fj.jpdc.2015.05.004.

Patricio, Diego Inacio and Rafael Rieder (2018). "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review". In: *Computers and Electronics in Agriculture* 153, pp. 69–81. DOI: 10.1016/j.compag.2018.08.001. URL: https://doi.org/10.1016%2Fj.compag.2018.08.001.

Payne, A. et al. (2014). "Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging". In: *Computers and Electronics in Agriculture* 100, pp. 160–167. DOI: 10.1016/j.compag.2013.11.011. URL: https://doi.org/10.1016%2Fj.compag.2013.11.011.

Pickett, T.D. (2004). "Future navigation needs in agriculture". In: cited By 1, pp. 1–10. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-8344269315\&partnerID=40\&md5=94805c8c1ed72ae6ed9a5c75916a7c20.

Pierce, F. J., P. C. Robert, and G. Mangold (1994). "Site Specific Management: The Pros, the Cons, and the Realities". In: *Proceedings of the Integrated Crop Management Conference*. Iowa State University, Digital Press. DOI: 10.31274/icm-180809-454. URL: https://doi.org/10.312742Ficm-180809-454.

Pierpaoli, Emanuele et al. (2013). "Drivers of Precision Agriculture Technologies Adoption: A Literature Review". In: *Procedia Technology* 8, pp. 61–69. DOI: 10.1016/j.protcy.2013.11.010. URL: https://doi.org/10.1016%2Fj.protcy.2013.11.010.

Piramuthu, Selwyn, Poorya Farahani, and Martin Grunow (2013). "RFID-generated traceability for contaminated product recall in perishable food supply networks". In: *European Journal of Operational Research* 225.2, pp. 253–262. DOI: 10.1016/j.ejor.2012.09.024. URL: https://doi.org/10.1016%2Fj.ejor.2012.09.024.

Pivoto, Dieisson et al. (2018). "Scientific development of smart farming technologies and their application in Brazil". In: *Information Processing in Agriculture* 5.1, pp. 21–32. DOI: 10.1016/j.inpa.2017.12.002. URL: https://doi.org/10.1016%2Fj.inpa.2017.12.002.

Rahnemoonfar, Maryam and Clay Sheppard (2017). "Deep Count: Fruit Counting Based on Deep Simulated Learning". In: *Sensors* 17.4, p. 905. DOI: 10.3390/s17040905. URL: https://doi.org/10.3390%2Fs17040905.

Rehman, Aqeel ur et al. (2014). "A review of wireless sensors and networks applications in agriculture". In: *Computer Standards & Interfaces* 36.2, pp. 263–270. DOI: 10.1016/j.csi.2011.03.004. URL: https://doi.org/10.1016\%2Fj.csi.2011.03.004.

Reid, John F. et al. (2000). "Agricultural automatic guidance research in North America". In: *Computers and Electronics in Agriculture* 25.1-2, pp. 155–167. DOI: 10.1016/

s0168-1699(99)00061-7. URL: https://doi.org/10.1016%2Fs0168-1699%2899%2900061-7.

Ruiz-Garcia, Luis et al. (2009). "A Review of Wireless Sensor Technologies and Applications in Agriculture and Food Industry: State of the Art and Current Trends". In: *Sensors* 9.6, pp. 4728–4750. DOI: 10.3390/s90604728. URL: https://doi.org/10.3390\%2Fs90604728.

Sa, Inkyu et al. (2016). "DeepFruits: A Fruit Detection System Using Deep Neural Networks". In: *Sensors* 16.8, p. 1222. DOI: 10.3390/s16081222. URL: https://doi.org/10.3390%2Fs16081222.

Sahu, Shriya, Meenu Chawla, and Nilay Khare (2018). "Viable Crop Prediction Scenario in BigData Using a Novel Approach". In: *Advances in Intelligent Systems and Computing*. Springer Singapore, pp. 165–177. DOI: 10.1007/978-981-13-1498-8_15. URL: https://doi.org/10.1007%2F978-981-13-1498-8_15.

Scalisi, Alessio, Kushtrim Bresilla, and Filipa Simões Grilo (2018). "The Quest for Artificial Intelligence: A History of Ideas and Achievements". In: *Italus Hortus* 24, 39–50. ISSN: 1127-3496. DOI: 10.26353/j.itahort/2017.2.3950. URL: http://doi.org/10.26353/j.itahort/2017.2.3950.

Si, Yongsheng, Gang Liu, and Juan Feng (2015). "Location of apples in trees using stereoscopic vision". In: *Computers and Electronics in Agriculture* 112, pp. 68–74. DOI: 10.1016/j.compag.2015.01.010. URL: https://doi.org/10.1016%2Fj.compag.2015.01.010.

Sicari, S. et al. (2015). "Security, privacy and trust in Internet of Things: The road ahead". In: *Computer Networks* 76, pp. 146–164. DOI: 10.1016/j.comnet.2014.11.008. URL: https://doi.org/10.1016%2Fj.comnet.2014.11.008.

Slaughter, D.C., D.K. Giles, and D. Downey (2008). "Autonomous robotic weed control systems: A review". In: *Computers and Electronics in Agriculture* 61.1, pp. 63–78. DOI: 10.1016/j.compag.2007.05.008. URL: https://doi.org/10.1016%2Fj.compag.2007.05.008.

Slot, M. M. et al. (2018). "The assessment of field trials in GMO research around the world and their possible integration in field trials for variety registration". In: *Transgenic Research* 27.4, pp. 321–329. DOI: 10.1007/s11248-018-0076-z. URL: https://doi.org/10.1007%2Fs11248-018-0076-z.

Stafford, John V. (2000a). In: *Precision Agriculture* 2.3, pp. 229–229. DOI: 10.1023/a:1011894914362. URL: https://doi.org/10.1023%2Fa%3A1011894914362.

– (2000b). "Implementing Precision Agriculture in the 21st Century". In: *Journal of Agricultural Engineering Research* 76.3, pp. 267–275. DOI: 10.1006/jaer.2000.0577. URL: https://doi.org/10.1006%2Fjaer.2000.0577.

Stafford, J.V., ed. (2007). *Precision agriculture '07*. Wageningen Academic Publishers. DOI: 10.3920/978-90-8686-603-8. URL: https://doi.org/10.3920%2F978-90-8686-603-8.

Stajnko, D., M. Lakota, and M. Hočevar (2004). “Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging”. In: *Computers and Electronics in Agriculture* 42.1, pp. 31–42. DOI: 10.1016/s0168-1699(03)00086-3. URL: https://doi.org/10.1016/s0168-1699(03)00086-3.

Stočes, Michal et al. (2016). “Internet of Things (IoT) in Agriculture - Selected Aspects”. In: *Agris on-line Papers in Economics and Informatics* VIII.1, pp. 83–88. DOI: 10.7160/aol.2016.080108. URL: https://doi.org/10.7160%2Faol.2016.080108.

Sun, Y. et al. (2009). “A solar-powered wireless cell for dynamically monitoring soil water content”. In: *Computers and Electronics in Agriculture* 69.1, pp. 19–23. DOI: 10.1016/j.compag.2009.06.009. URL: https://doi.org/10.1016%2Fj.compag.2009.06.009.

Suwantong, Rata et al. (2016). “Moving Horizon Estimator with Pre-Estimation for crop start date estimation in tropical area”. In: *2016 American Control Conference (ACC)*. IEEE. DOI: 10.1109/acc.2016.7525476. URL: https://doi.org/10.1109%2Facc.2016.7525476.

Tan, Li (2016). “Cloud-based Decision Support and Automation for Precision Agriculture in Orchards”. In: *IFAC-PapersOnLine* 49.16, pp. 330–335. DOI: 10.1016/j.ifacol.2016.10.061. URL: https://doi.org/10.1016%2Fj.ifacol.2016.10.061.

Tan, Ying and Zhong yang Zheng (2013). “Research Advance in Swarm Robotics”. In: *Defence Technology* 9.1, pp. 18–39. DOI: 10.1016/j.dt.2013.03.001. URL: https://doi.org/10.1016%2Fj.dt.2013.03.001.

Tao, Yongting and Jun Zhou (2017). “Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking”. In: *Computers and Electronics in Agriculture* 142, pp. 388–396. DOI: 10.1016/j.compag.2017.09.019. URL: https://doi.org/10.1016%2Fj.compag.2017.09.019.

Torres, Roque et al. (2017). “Effects of soil and climate in a table grape vineyard with cover crops. Irrigation management using sensors networks”. In: *Ciencia e Tecnica Vitivinicola* 32.1, pp. 72–81. DOI: 10.1051/ctv/20173201072. URL: https://doi.org/10.1051%2Fctv%2F20173201072.

Tzounis, Antonis et al. (2017). “Internet of Things in agriculture, recent advances and future challenges”. In: *Biosystems Engineering* 164, pp. 31–48. DOI: 10.1016/j.biosystemseng.2017.09.007. URL: https://doi.org/10.1016%2Fj.biosystemseng.2017.09.007.

Vougioukas, S. et al. (2013). “Influence of foliage on radio path losses (PLs) for wireless sensor network (WSN) planning in orchards”. In: *Biosystems Engineering* 114.4, pp. 454–465. DOI: 10.1016/j.biosystemseng.2012.08.011. URL: https://doi.org/10.1016%2Fj.biosystemseng.2012.08.011.

Vuran, Mehmet C. et al. (2018). “Internet of underground things in precision agriculture: Architecture and technology aspects”. In: *Ad Hoc Networks* 81, pp. 160–173.

DOI: 10.1016/j.adhoc.2018.07.017. URL: https://doi.org/10.1016%2Fj.adhoc.2018.07.017.

Wang, N. and Z. Li (2013). "Wireless sensor networks (WSNs) in the agricultural and food industries". In: *Robotics and Automation in the Food Industry*. Elsevier, pp. 171–199. DOI: 10.1533/9780857095763.1.171. URL: https://doi.org/10.1533%2F9780857095763.1.171.

Wang, Tao et al. (2018). "SeeFruits: Design and evaluation of a cloud-based ultra-portable NIRS system for sweet cherry quality detection". In: *Computers and Electronics in Agriculture* 152, pp. 302–313. DOI: 10.1016/j.compag.2018.07.017. URL: https://doi.org/10.1016%2Fj.compag.2018.07.017.

Wolfert, Sjaak et al. (2017). "Big Data in Smart Farming – A review". In: *Agricultural Systems* 153, pp. 69–80. DOI: 10.1016/j.agsy.2017.01.023. URL: https://doi.org/10.1016%2Fj.agsy.2017.01.023.

Woodard, Joshua (2016). "Big data and Ag-Analytics". In: *Agricultural Finance Review* 76.1, pp. 15–26. DOI: 10.1108/afr-03-2016-0018. URL: https://doi.org/10.1108%2Fafr-03-2016-0018.

Wu, Chenxing et al. (2013). "An Orchard Management Systemwith RFID-Based Apple Tree Identify Detection". In: *2013 Fourth International Conference on Digital Manufacturing & Automation*. IEEE. DOI: 10.1109/icdma.2013.42. URL: https://doi.org/10.1109%2Ficdma.2013.42.

Xia, JiAn et al. (2018). "Visible-near infrared spectrum-based classification of apple chilling injury on cloud computing platform". In: *Computers and Electronics in Agriculture* 145, pp. 27–34. DOI: 10.1016/j.compag.2017.12.012. URL: https://doi.org/10.1016%2Fj.compag.2017.12.012.

Yaghoubi, S et al. (Jan. 2013). "Autonomous robots for agricultural tasks and farm assignment and future trends in agro robots". In: 13, pp. 1–6.

Zhang, Naiqian, Maohua Wang, and Ning Wang (2002). "Precision agriculture—a worldwide overview". In: *Computers and Electronics in Agriculture* 36.2-3, pp. 113–132. DOI: 10.1016/s0168-1699(02)00096-0. URL: https://doi.org/10.1016/s0168-1699(02)00096-0.

Zhang, Peng et al. (2017). "The Construction of the Integration of Water and Fertilizer Smart Water Saving Irrigation System Based on Big Data". In: *22017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*. IEEE. DOI: 10.1109/cse-euc.2017.258. URL: https://doi.org/10.1109%2Fcse-euc.2017.258.

Zude-Sasse, M. et al. (2016). "Applications of precision agriculture in horticultural crops". In: *European Journal of Horticultural Science* 81.2, pp. 78–90. DOI: 10.17660/ejhs.2016/81.2.2. URL: https://doi.org/10.17660%2Fejhs.2016%2F81.2.2.

Chapter 2

Single-shot convolution neural networks for real-time fruit detection within the tree

Abstract

Image/video processing for fruit detection in the tree using hard-coded feature extraction algorithms has shown high accuracy on fruit detection during recent years. While accurate, these approaches even with high-end hardware are still computationally intensive and too slow for real-time systems. This paper details the use of deep convolution neural networks architecture based on single-stage detectors. Using deep-learning techniques eliminates the need for hard-code specific features for specific fruit shapes, color and/or other attributes. This architecture takes the input image and divides into AxA grid, where A is a configurable hyper-parameter that defines the fineness of the grid. To each grid cell an image detection and localization algorithm is applied. Each of those cells is responsible to predict bounding boxes and confidence score for fruit (apple and pear in the case of this study) detected in that cell. We want this confidence score to be high if a fruit exists in a cell, otherwise to be zero, if no fruit is in the cell. More than 100 images of apple and pear trees were taken. Each tree image with approximately 50 fruits, that at the end resulted on more than 5000 images of apple and pear fruits each. Labeling images for training consisted on manually specifying the bounding boxes for fruits, where (x, y) are the center coordinates of the box and (w, h) are width and height. This architecture showed an accuracy of more than 90% fruit detection. Based on correlation between number of visible fruits, detected fruits on one frame and the real number of fruits on one tree, a model was created to accommodate this error rate. Processing speed is higher than 20FPS which is fast enough for any grasping/harvesting robotic arm or other real-time applications.

2.1 Introduction

AGRICULTURE is a sector with very specific working conditions and constraints. This isn't only due to the dependency on the weather conditions, but as well on the labor market. During times of highly intensive agricultural activities (eg. harvest), there are very pronounced peaks in workload which can only be predicted on a short-term basis due to the weather conditions and seasonality. According to [Fröder, 2014] the world's agricultural workforce is expected to decline around 30% between 2017 and 2030. This expected decline will be driven by structural changes within the agri-food industry, but also because the opportunities for employment are expected to be better in other sectors. Rural areas are already facing difficulties in creating attractive jobs in general, pushing towards an ongoing migration towards urban centres. Those structural changes in agriculture are expected to continue with higher investments in technology. For example, investing in precision farming and digital agriculture are expected to significantly increase [Colbert, Yee, and George, 2016]. New technologies are set to impact the farm labour dynamic in many ways [Pierpaoli et al., 2013], but two developments stand out. One, the increasing use of data collection tools, such as sensors, and increasing sophistication of farm hardware and software is increasing demand for higher analytical and technical skill sets [Aubert, Schroeder, and Grimaudo, 2012; Mulla, 2013]. And two, the advancement of automation and autonomy on farm will decrease the reliance on human resources for low-skill and labour-intensive work while increasing autonomous machinery and robotics presence [Bechar and Vigneault, 2016; Stafford, 2007].

Along with many other emerging concepts of precision agriculture, the agricultural robot has evolved as one of the most promising of all [Tao and Zhou, 2017]. Agriculture is most of the time repetitive, repetitive work of seeding, weeding, feeding, pruning, picking and harvesting, sorting and so on. Agricultural robots automate those slow, repetitive and dull tasks for farmers, allowing them to focus more on strategic matters and improving overall production yields [Edan, Han, and Kondo, 2009]. One of the most popular robotic application in agriculture are the autonomous harvesting and picking robots. That's because the speed and accuracy has increased significantly in recent years [Tao and Zhou, 2017; Bechar and Vigneault, 2016]. While the robots in addition to harvesting and picking can check at the same time the maturity level and sort based on size [Edan, Han, and Kondo, 2009]. However, there are many challenges for an autonomous robotic system to complete that task. In principle, for the robot to be fully capable to perform harvesting and picking, it needs a sophisticated detection algorithm in order to overcome challenges as naturally occurring changes in illumination, shape, pose, colour, and viewpoint [Barnea, Mairon, and Ben-Shahar, 2016].

Over the years, different approaches and techniques have been developed to tackle fruit detection and localization [Jiménez et al., 1999; Song et al., 2014]. All the techniques until recently relied on feature extraction, be that, color, shape, reflectance, etc. [Rahnemoonfar and Sheppard, 2017; Song et al., 2014]. Besides the aforementioned techniques, a new one which recently are gaining momentum and higher accuracy are deep learning techniques [Rahnemoonfar and Sheppard, 2017; Sa et al., 2016].

Many of the state-of-the-art object detector CNN are divided into two main groups. In one side, are models that reach higher accuracy but are slower: the two-stage detectors such as Faster R-CNN (Region-based Convolutional Neural Networks), and/or Mask R-CNN, that use a Region Proposal Network to generate regions of interests in the first stage and send the region proposals down the pipeline for object classification and bounding-box regression. In other side are models that reach lower accuracy but are faster: the single-stage detectors, such as You Only Look Once (YOLO) and Single Shot MultiBox Detector (SSD), that treat object detection as a simple regression problem which takes an input image and learns the class probabilities and bounding box coordinates.

In this paper it is presented a CNN model for fast and accurate fruit detection based on YOLO model [Redmon et al., 2015]. By using those DL techniques, the need for hard-code specific features like specific fruit shapes, color and/or other attributes was eliminated. The network consists of several convolution and pooling layers, tweaked and changed from the standard model. Those modifications made to the model, make it more accurate to detect objects of the same class on close proximity (eg. only apple fruits, or only pear fruits). Even though the model was trained only on apple images (training data), it shows high accuracy on other fruits with similar attributes (green apples and green pears).

2.2 Background and Related Work

The positions of the fruits in the tree are widely distributed, highly depending on the tree size, form, and growth. Furthermore, in addition to their position, fruits vary in size, shape, and reflectance due to the natural variation that exists in nature. Currently, no growth models can predict where fruit will occur. The shape of the fruit, one of the most distinctive features, varies between species and even cultivars (e.g., apples, oranges, etc., are cylindrical, but the width/height ratio are not constant with other fruits like pears) [Bac et al., 2014]. Reflectance (mostly color and near-infrared) of fruit is a visual cue often used to distinguish fruit from other plant parts and still it varies strongly [Tao and Zhou, 2017]. Color and texture are the fundamental character of natural images and plays an important role in visual perception. Color is often a distinctive and indicative cue for the presence of fruit. Most fruits when ripe have a distinctive color: red (apples, strawberries, peaches, etc...), orange (oranges, etc...), or yellow (pears, lemons, peaches, bananas). This makes them stand out from the green foliage when they are ready to pick [Barnea, Mairon, and Ben-Shahar, 2016; Edan, Han, and Kondo, 2009]. However, some fruits even after ripening are still green (apple cv Granny Smith even after ripening does not change color) making them indistinguishable from the foliage on the basis of color alone [Edan, Han, and Kondo, 2009].

The earliest fruit detection systems date since 1968 [Jiménez et al., 1999]. Using different methods and approaches based on photometric information (light reflectance difference from fruits and leaves in visible or infrared spectrum), these detectors were able to differentiate fruits from other part of the tree. According to the reviews devoted to fruit detection by [Jiménez et al., 1999] and later on by [Kapach et al., 2012], there were many problems related to growth habit that had to be considered. The unstructured and uncontrolled outdoor environment also presents many challenges for computer vision systems in agriculture.

Light conditions have a major influence on fruit detection: direct sunlight results in saturated spots without color information and in shadows that cause standard segmentation procedures to split the apple surfaces into several fragments. In order to decrease the non-uniform illumination (daytime lighting can be bright, strong, directional and variable), [Payne et al., 2014] described a machine vision technique to detect fruit based on images acquired during night time using artificial light sources.

They reported 78% fruit detection, 10% errors and suggesting that artificial lighting at night can provide consistent illumination without strong directional shadows.

In a different approach, [Kelman and Linker, 2014] presented an algorithm for localizing spherical fruits that have a smooth surface, such as apples, using only shape analysis and in particular convexity. It is shown that in the images used for the study, more than 40% of the apple profiles were none-convex, more than 85% of apple edges had 15% or more non-convex profiles, and more than 45% of apple edges had 50% or more non-convex profiles. Overall, 94% of the apples were correctly detected and 14% of the detection corresponded to false positives. Despite high accuracy number, the model is very specific to apples and would not be extensible to other fruit crops with less spherical shapes. [Kapach et al., 2012] explains color highlights and spherical attributes, which tend to appear more often on the smoother, more secular, and typically elliptical regions like fruits where the surface normal bisects the angle between illumination and viewing directions. While a method for estimating the number of apple fruits in the orchard using thermal camera was developed by [Stajnko, Lakota, and Hočevar, 2004]. [Si, Liu, and Feng, 2015] describes location of apples in trees using stereoscopic vision. The advantage of the active triangulation method is that the range data may be obtained without much computation and the speed is very high for any robotic harvesting application. [Jiang, Peng, and Ying, 2008] developed a binocular stereo vision tomato harvesting in greenhouse. In this method, a pair of stereo images was obtained by stereo cameras and transformed to gray-scale images. According to the grey correlation, corresponding points of the stereo images were searched, and a depth image was obtained by calculating distances between tomatoes and stereo cameras based on triangulation principle. A similar method was described by [Barnea, Mairon, and Ben-Shahar, 2016] using RGB and range data to analyse shape-related features of objects both in the image plane and 3D space. In another work [Nguyen et al., 2014] developed a multi-phase algorithm to detect and localize apple fruits by combining an RGB-D camera and point cloud processing techniques. [Tao and Zhou, 2017] developed an automatic apple recognition system based on the fusion of color and 3D features.

Until recent years, traditional computer vision approaches have been extensively adopted in the agricultural field. In recent years, with the significant increase in computational power, in particular with special purpose processors optimized for matrix-like data processing and large amount of data calculations (eg. Graphical Processing Unit – GPU), a lot of deep learning, CNN models and methodologies specifically have achieved breakthroughs never achieved before [LeCun, Bengio, and Hinton, 2015].

In 2016, [Sa et al., 2016] developed a model called DeepFruits, for fruit detection. Adopting a Faster R-CNN model, goal was to build an accurate, fast and reliable fruit detection system. The model after training was able to achieve 0.838 precision and recall in the detection of sweet pepper. In addition, they used a multi-modal fusion approach that combines the information from RGB and NIR images. The bottleneck of the model is that in order to deploy on a real robot system, the processing performance required is a GPU of 8GB or more.

To understand better the amount of data needed for better fruit detection, [Bargoti and Underwood, 2017] used different data augmentation techniques and transfer learning from other fruits. It is shown that transferring weights between different fruits did not have significant performance gains, while data augmentation like flip and scale were found to improve performance resulting in equivalent performance with less than half the number of training images.

2.3 Materials and Methodology

2.3.1 Convolution Neural Networks

Convolutional neural networks are a specialized type of artificial neural networks used for image analysis. Since computers sees the image as a matrix of numbers that represent each pixels, it is important that the relation between the pixels (values) remains even after the image is processed through the network. To save this spatial relation between pixels, convolution neural networks are used that have different mathematical operations stacked on top of each-other to create layers of the network.

In this first part of this section we will briefly explain those mathematical operation that serve as building blocks for our model architecture.

Convolution

Convolution is a mathematical operation that takes two functions ($K_{i,j}^{(l)}$ and $Y_j^{(l-1)}$) to produce a third one ($Y_i^{(l)}$). It is the process of adding each pixel of the image to its local neighbors, weighted by the kernel/filter of $m \times n$ size. In each layer, there is a certain amount of filters. The number of filters p that is applied in one stage is equivalent to the depth of the volume of output feature maps. Each filter detects a particular feature at every location on the input. The output $Y_i^{(l)}$ of layer l consists of $p^{(l)}$ feature maps of size $m^{(l)} \times n^{(l)}$. Thus the i^{th} feature map is computed as:

$$Y_i^{(l)} = B_i^{(l)} + \sum_{i=0}^{p^{l-1}} K_{i,j}^{(l)} \times Y_j^{(l-1)}$$

where $B_i^{(l)}$ is bias matrix and $K_{i,j}^{(l)}$ is the kernel connecting the feature map j of previous layer $(l-1)$ with i^{th} feature map in current layer l.

Actvation

It is a function that takes the feature map $Y_j^{(l-1)}$ generated by the convolution and creates the activation map as output $Y_j^{(l)}$. It serves as a gate, to let certain part of the map elements pass while others not. This is strictly element-wise operation:

$$Y_i^{(l)} = f(Y_i^{(l-1)})$$

where f is the function that we multiply with.

In the case of our model used, the activation layer is a special implementation of activation functions of non-linearity and rectification called PReLU:

$$Y_i^{(l)} = max(\alpha \times Y_i^{(l-1)}, Y_i^{(l-1)})$$

so when $Y_j^{(l)} < 0$, it will have a small positive slope of α.

Pooling

Pooling layer is used to reduce the computational requirements through the network and minimize overlapping by reducing the spatial size of the activation map. Pooling has two key components: spatial grouping $F^{(l)}$ and spatial shift $S^{(l)}$. It takes a input volume of $p^{(l-1)} \times m^{(l-1)} \times n^{(l-1)}$ and provides a reduced volume output of size $p^{(l)} \times m^{(l)} \times n^{(l)}$ where:

$$p^{(l)} = p^{(l-1)} \qquad m^{(l)} = (m^{(l-1)} - F^{(l)})/S^{(l)} + 1 \qquad n^{(l)} = (n^{(l-1)} - F^{(l)})/S^{(l)} + 1$$

Most of the classical techniques for object detection used different filters/kernels to extract features and then programmatically apply detection. Those kernels, usually of 3×3 size, were manually given to the program to perform convolution with the image fed to it and then detect objects of interests. Depending on the kernel type, those would extract features like edges, sharpening, color filtering, and many others others [Lee and Rhodes, 1990]. Convolutional neural network, use techniques of loss function optimization and back propagation to automatically generate those kernels (in CNN called weights of the model). During this process the weights are updated with each iteration (epochs), until the best possible version is reached.

2.3.2 Model Architecture

YOLO

YOLO is a state-of-the art convolutional network for detection and localization. There are different versions of YOLO, and in this paper we modified and used YOLO900 (also known as YOLOv2), and as such, in the remaining part of the paper, we refer to YOLO900 as YOLO. Compared to other state-of-the art methods that treat detection, classification and region extraction as different problems, YOLO does all in one pass (hence the name You Only Look Once). To achieve that, YOLO in one hand losses in accuracy but in other hand gains speed. YOLO takes as input an image of max size 608×608 and divides into $S \times S$. Each grid cell is responsible for the bounding box whose center is at the location of the grid cell, and predicts B bounding boxes as well as confidence level and class probability. In a dataset with C class labels, the output tensor is $S \times S \times (C + B \times 5)$. In our modified model, the class C is equal to 1 since we train the network each time with one class (just apples, just pears, etc...) which will be further discussed in section below. If the cell is offset from the top left corner of image by (x_c, y_c) and the bounding box ground truth is of size (g_w, g_y) then the prediction goes as:

$$box_x = \sigma(x_i) + x_c$$
$$box_y = \sigma(y_i) + y_c$$
$$box_w = g_w e^{w_i}$$
$$box_h = g_h e^{h_i}$$

The model assigns five anchor boxes to each cell. For each anchor box we need confidence score, four coordinates and the class number. x_i and y_i are the location of the centre of the anchor box, w_i and h_i are the width and height of the anchor box C_i confidence score of whether there is an object or not, and $p_i(c)$ is the classification loss.

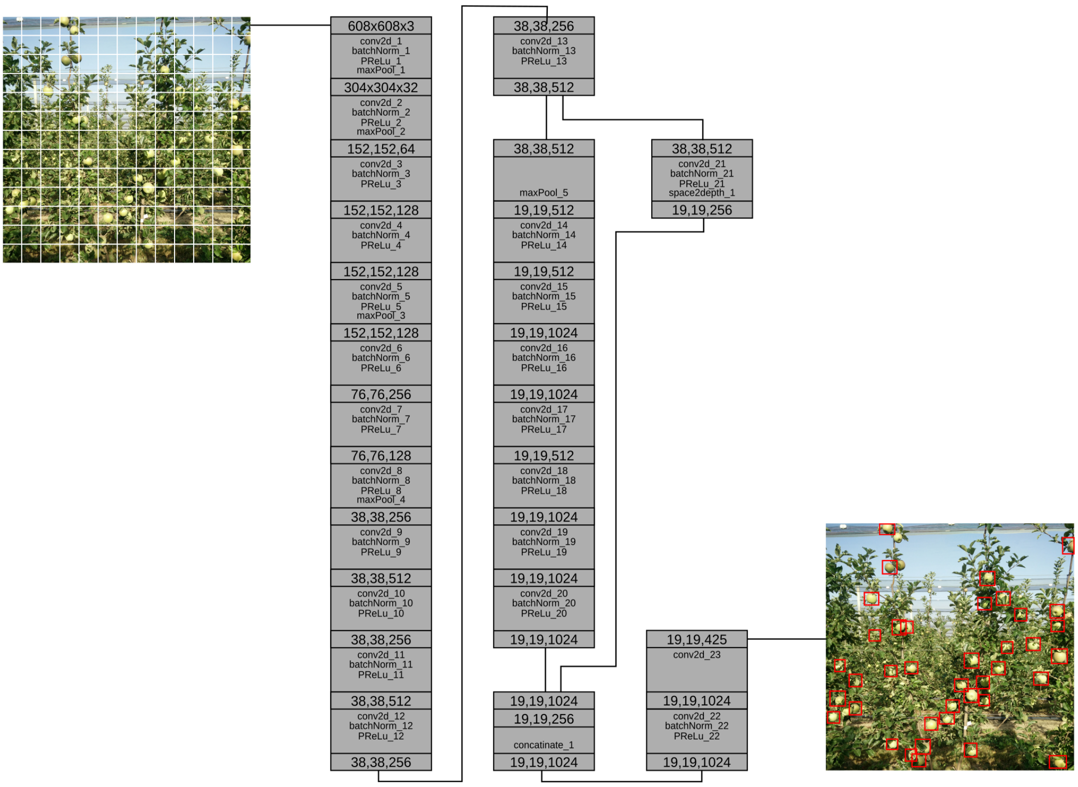


FIGURE 2.1: YOLO model with 24 Layers

The loss function is divided into three parts: coordinate loss (the penalization for bounding box parameters), confidence loss (for whether the grid cell is responsible for the ground truth) and class loss or classification error:

$$\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B}1_{ij}^{obj}[(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2]$$
$$+\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B}1_{ij}^{obj}[(\sqrt{w_i}-\sqrt{\hat{x}_i})^2+(\sqrt{h_i}-\sqrt{\hat{h}_i})^2]$$
$$+\lambda_{obj}\sum_{i=0}^{S^2}\sum_{j=0}^{B}1_{ij}^{obj}(C_i-\hat{C}_i)^2+\lambda_{noobj}\sum_{i=0}^{S^2}\sum_{j=0}^{B}1_{ij}^{noobj}(C_i-\hat{C}_i)^2$$
$$+\lambda_{cls}\sum_{i=0}^{S^2}1_{i}^{obj}\sum_{c\in classes}(p_i(c)-p_i\hat{(c)})^2$$

where λ_{coord}, λ_{obj}, λ_{noobj}, and λ_{cls} are independent parameters. Similarly, 1_i^{obj} denotes if an object i appears in cell i and 1_{ij}^{obj} denotes that the j^{th} bounding box predictor in cell i is responsible for that ij prediction.

Our model

Despite being one of the most popular state-of-the art model, YOLO has problems detecting small objects [Redmon et al., 2015]. The main problem with YOLO is that, the model can detect only one object class per cell, making it very difficult to detect two apples at the same cell. And since in apple tree, we are dealing with small fruits in relation to the canopy, we made different changes to the model, resulting in different accuracy scores.

First method (noted as model M1) we used is to scale up the grid. The standard YOLO takes input image and divides into 13×13. In our case we scaled up to 26×26. This improved the detection, as at this division of input image, we approximately have the grid cell size approximately like the apple size. However, this almost doubled the training time, and decreased the overall detection speed. More so, the input image is still of size 608×608, it is just divided in more finer grid.

Second method (noted as model M2) we took the M1 model and removed some layers, making the model shallower. We saw that, removing pooling layers and some other convolutional layers, we increase in speed while not loosing in accuracy. In this model, we tried to make the model run on higher frames per second, in embedded devices like NVIDIA Jetson TX2. This network consists of the grid 26×26, while the network shallowness is reduced by half, instead of having 23 layers, in this model we have 11 layers. We removed pooling layers, and used instead higher strides and kernel/filter size in convolution layers.

Third method (noted as model M3) we used M2 model as base, and then we added two other blocks (one in entrance and one in the end). The entrance block, noted as "splitter" and the end one noted as "joiner". Splitter, takes the image and separates it into 4 individual images. The resolution before split is 1216×1216 and then it gets splits into exactly four 608×608 images. The joiner at the end, is responsible to put the four pieces together and output the results in 1216×1216 single image with detection. Ths technique of adding blocks to split and join the image introduced by [Chen, Zhang, and Ouyang, 2018], is very effective and accurate, however it decreases the speed of network, as this in essence in feeding four images instead of one.

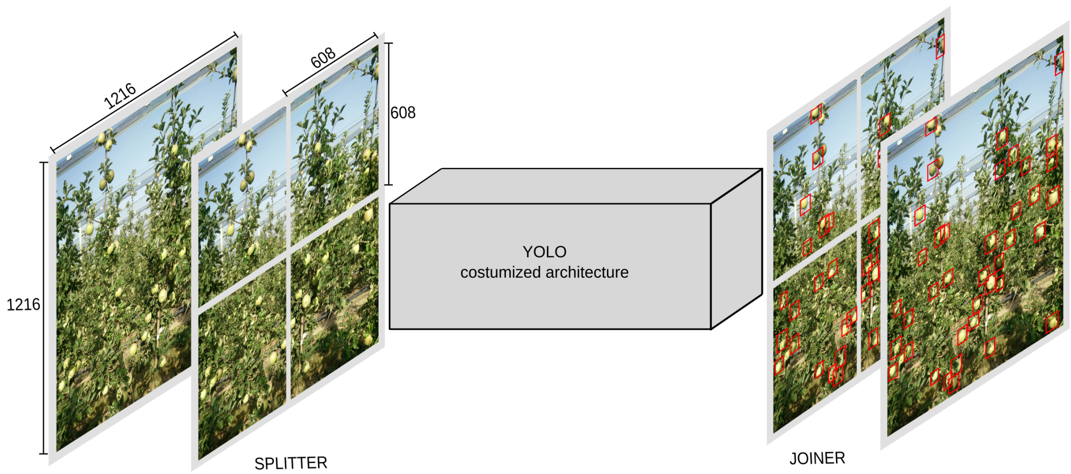


FIGURE 2.2: Model M3, with splitter and joiner blocks

2.3.3 Dataset

Images were collected in the experimental orchards of University of Bologna - Department of Agriculture (Italy), and in privately owned production orchards in Ferrara (Italy). There have been taken multiple images of same tree with different angles, multiple fruit tree species and cultivars, multiple sources (webcam, DSLR camera, smartphone) and during different time of the year with different weather and light conditions.

In total 100 images of apples and 50 of pears are taken. Each image containing approximately 50 fruits, resulting in more than 5000 images to train the models.

Images were taken before the coloring of apples occur, thus all apples were still green.

CAMERA	NUMBER	SIZE	RESIZE	AUGMENTED	SYNTHETIC
Webcam	20	1280x720	608x608	80	0
Smartphone	100 + 30	2340x4608	1216x1216	200	100
DSLR	30	5664x8512	1216x1216	120	0

TABLE 2.1: Image number, sources and augmentation of apple fruits

For pears, the number is comparably lower than apples, but in this case we wanted to observe some transfer learning techniques from apples, thus this amount was enough to carry the study:

CAMERA	NUMBER	SIZE	RESIZE	AUGMENTED	SYNTHETIC
Smartphone	50	2340x4608	1216x1216	60	0

TABLE 2.2: Image number, sources and augmentation of pear fruits

Data preparation

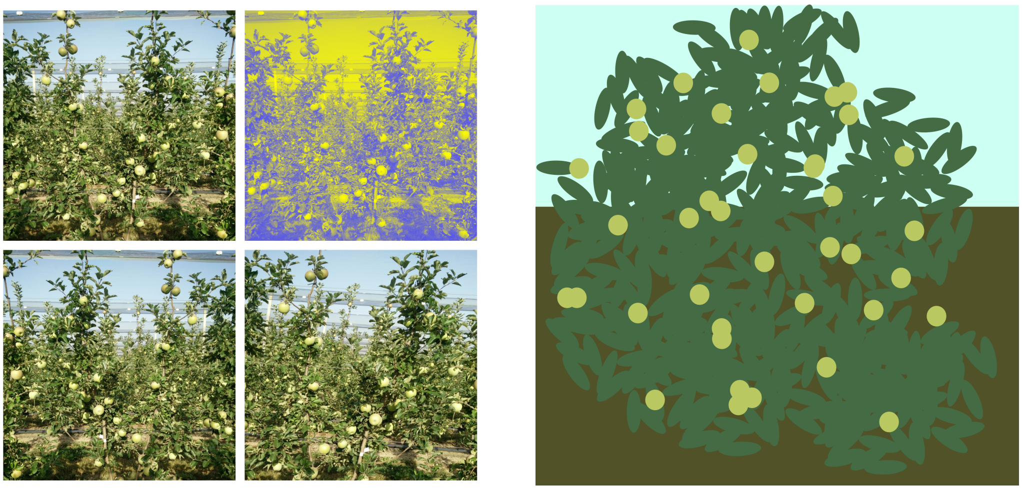
Data preparation and pre-analysis is the most important step in building machine learning algorithms. After images were collected together, we started the exploratory data analysis. Each image was shot in portrait mode, that the tree was fully captured. However, 1/5 of image from top, and 1/5 from the bottom does not contain any apple fruits. In some cases, the bottom part even had fallen apples, that the model will try to detect, and later on results, would be considered as false positive and decrease the F_1 score. To avoid that, and better fit the image in a square dimensions, all the images were cropped. In addition to cropping, the images had to be resized after to be suitable for the model. Depending on the model we used, the images were resized into two different sizes: 608×608 to be used by model M1 and M2 and 1216×1216 to be used by model M3. Everything was done with a script written in Python, that would take a batch of images and output the desired scaling and cropped shape.

Images are divided into Training set and Testing set. For Testing we used only images from smartphone camera. Those images have go through all procedures of other images, until training the model. Those images are never shown to the training, as it is important for testing the accuracy of the models into images not seen before. This way we can see if the model is over-fit only on those training images, or is still able to generalize and detect fruits on images never seen before.

For labeling, we used a free and open-source labeling tool called BBoxLabelTool where each image went through process of labeling. Most of problems during labeling come because of occlusion and overcrowded image. This happens when one object is either partially or completely occluded by the other, and when a large amount of objects are close or attached to each other. Due to the nature of the fruit trees, this is present on every image taken of that tree. In each image, every apple fruit visible was labeled with a bounding box representing the location of apple fruit. This is done manually and very carefully to avoid mislabelling or occlusion. However, this tool annotates the data into PASCAL Visual Object Classes format, and in our case we needed the DARKNET format. All labeling data were automatically converted for each bounding box with a Python script

Data augmentation

Fetching the right amount and type of data which matches the use-case of our research is a difficult task. In addition, the data should have good diversity as the object of interest needs to be present in varying sizes, lighting conditions and poses in order to expect that our network generalizes well during the training phase. To overcome this problem of limited quantity and limited diversity of data, we generate our own data with the existing data which we have. This methodology of generating our own data is known as data augmentation. There are numerous approaches to augment training data, in terms of quantity, it can be either by expanding the dataset with copies of augmented versions, or by randomly augmenting some data from dataset. In terms of diversity, there are techniques like color manipulation, light and contrast, resizing, scaling, flipping, rotation, perspective transformation, adding noise and so on. In this paper, we augmented all images in the dataset by randomly choosing the augmentation technique. For each image, three of those random augmentation were generated. In total 400 new images were created through augmentation.



(A) Augmenting from an image to other four (B) Synthetic generated image

FIGURE 2.3: Image augmentation and synthetic generated image

Synthetic data

As described by [Rahnemoonfar and Sheppard, 2017], deep-learning models require a lot of images to be collected and annotated then fed to the network. This is a very time-consuming and difficult task. In order for us to use less real images, we generated synthetic images for the network to train on. This is done automatically, through a Python script, where a image canvas of 608×608 was used. The upper part of background was colored with blueish color, representing the sky, while the down part was colored with mixture of brown. Above that image, random elliptic dark-green shapes were generated representing leaves while random light-green and light-red circles were generated to represent fruits. In total 100 synthetic images were generated to observe if those data improve the detection of the models.

2.4 Results and Discussion

The results will be will be evaluated using the data (images with labels) from the testing and validation set. The metrics evaluating the accuracy will be according to the well known criteria based on Pascal Visual Object Classes (VOC) that many research on this field use.

2.4.1 Metrics

Pixel-wise accuracy was measured by comparison of ground truth and predicted information. Two metrics of accuracy are used: Confusion Matrix (precision, recall and F_1 score) and Intersect over Union (IoU). While for measuring speed, frames per second (FPS) is used.

		PREDICTED	
		TRUE	FALSE
GROUND TRUTH	TRUE	True Positive (TP)	False Negative (FN)
	FALSE	False Positive (FP)	True Negative (TN)

TABLE 2.3: Confusion Matrix

Precision evaluates the fraction of true positives (TP) detected bounding boxes in the pool of all true positives predictions (TP) and false positive predictions (FP) while recall evaluates the fraction of true positives (TP) detected bounding boxes in the pool of all true positives predictions (TP) and false negatives predictions (FN):

$$precision = \frac{TP}{TP + FP} \qquad recall = \frac{TP}{TP + FN}$$

Precision and recall are tightly related, thus we can use only the F_1 score, which takes in consideration both precision and recall to compute the score and how well the prediction fit to the ground truth:

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall}$$

However in order to compute the correctness of detection, we use intersect over union. IoU is defined by calculating the overlapping area of prediction and ground truth:

$$IoU = \frac{(p_w \times p_h) \cap (g_w \times g_h)}{(p_w \times p_h) \cup (g_w \times g_h)}$$

where p_w and p_h is the prediction bounding box width and height, and g_w and g_h is the ground truth bounding box width and height. A threshold above 0.5 IoU is considered as positive, while under is considered as poor detection.

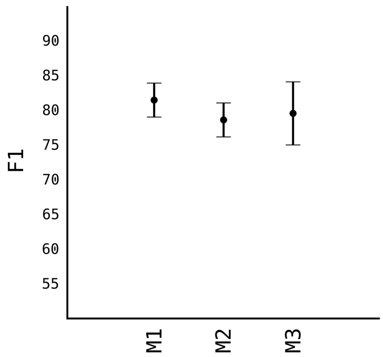
Another metric we used to measure speed is frames per second (FPS). In this case, we tested how fast the model runs in a NVIDIA Jetson TX2 with 300 CUDA cores and NVIDIA GeForce 960M with 960 CUDA cores. FPS were calculated by dividing a second with the time in millisecond of processing a single image.

2.4.2 Results

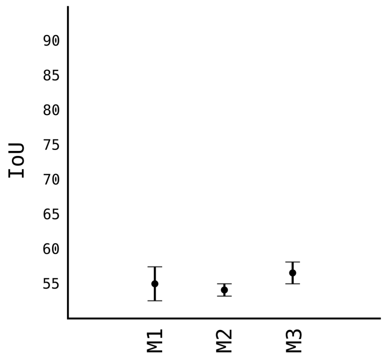
In this section we present the result obtained by comparing different models used, their detection speed after training and the number of images the models are trained and transfer learning techniques.

Model accuracy

Calculating accuracy and the speed of models, before choosing one to proceed, we compared all three models which were trained on 100 images of apples taken from sources mentioned in Table 2.1. As previously explained, stock YOLO is not very accurate in detecting small objects with its standard 13×13 grid. However first model M1 that uses 26×26 grid cell, is vey accurate at detecting objects. As shown in Figure 2.4a, the F_1 score of the M1 is 0.81 while having a relative high IoU score, as shown in Figure 2.4b. IoU in general tends to penalize single instances of bad classification more than the F_1 score quantitatively even when they both are referring to the same bad detection instance. Since the image is resized 608×608 pixels, more wrongly pixels are classified as fruits, thus IoU penalizes them more than F_1 score. With the grid 26×26 of model M1, speed suffers quiet a lot, see Figure 2.5. Average speed on Nvidia Jetson TX2 was 5FPS and on Nvidia GeForce 960M was 8FPS, which is very slow for any real time application use.



(A) F_1 accuracy score

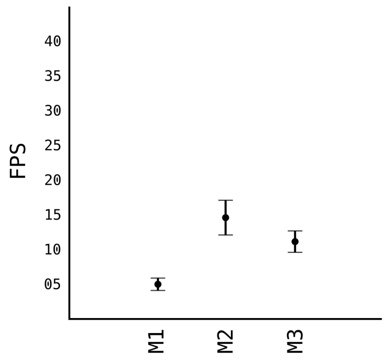


(B) IoU accuracy score

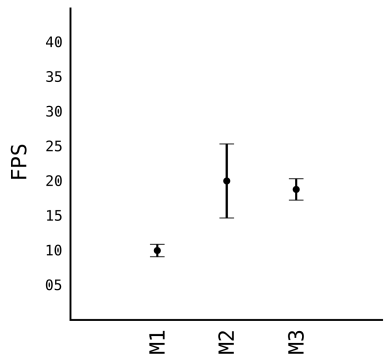
FIGURE 2.4: Comparison of accuracy of different models

When we moved to M2, where we removed some layers from the model, in order to make it shallower, the model lost in F_1 and IoU score. Respectively the F_1 is 0.77

while IoU is 0.53. And since the model is half less deeper than M1, the processing speed from Figure 2.5, is significantly higher, with average FPS of 15 in TX2 and 20 in 960M.



(A) Nvidia Jetson TX2



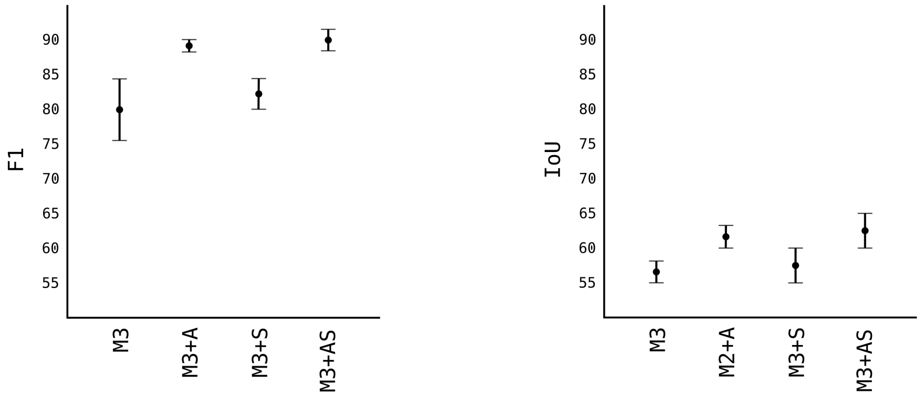
(B) Nvidia GeForce 960M

FIGURE 2.5: Running speed of models in different platforms

When the "splitter" and "joiner" blocks are introduced to the M3 model, they improve on accuracy of M2, both in terms of F_1 score and IoU score. F_1 being 0.79 and IoU 0.58. However we loose some FPS. This is due fact, that every image feed to the network, essentially is a 4 image split. Input size of the image is 1216×1216, and each of those 4 images of 608×608 pixels goes through network. This makes the model almost as accurate as model M1 while increasing the speed comparably from 8FPS to 20FPS (in Nvidia GeForce 960M).

For next results, we choose the M3 model to work with, and used other techniques to improve on upon it, as showed in results below.

Number of images

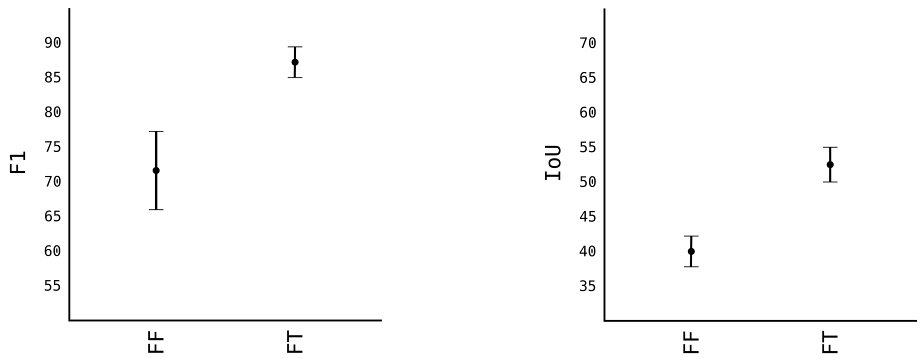


(A) F_1 accuracy score (B) IoU accuracy score

FIGURE 2.6: Image augmentation and synthetic images results

Number if data is one of the most important factors in any deep learning model. Since the models generalize and learns patterns from labeled images, it is of outmost importance the quantity and quality of those data. In the Figure 2.4, we used 100 images with approximately 50 fruits each, reaching F_1 score of 0.79 which not very high. In Figure 2.6 we increased the number of images by augmentation and synthetic generated images and observed the model's score. As expected, the F_1 score of model M3 with 400 more augmented images (noted as M3+A) jumped from 0.79 to 0.89 and IoU from 0.57 to 0.62. When to the model M3 we added 100 synthetic generated images (noted as M3+S), the F_1 score improved very slightly from 0.79 to 0.81, while IoU improved from 0.57 to 0.60. This shows that synthetic generated images helped the model better localize the pixels of the detected apple. When both 400 images from augmented part and 100 synthetic generated were added to M3 (together noted as M3+AS) the model showed the highest score observed, with F_1 of 0.9 and 0.64 IoU.

Transfer learning



(A) F_1 accuracy score (B) IoU accuracy score

FIGURE 2.7: Transfer learning from M3+AS to pears images

For pears, we only had 50 images. From which 40 of them were used for training, and we added another 60 from augmentation, resulting in around 5000 fruits, 15000 less than apples. This amount of images, as showed in Figure 2.4 results not more than F_1 of 0.8. To overcome that, we used transfer-learning techniques. Firstly, we used the weights of model M3+AS and tested in pears without any further training. Noted as FF (Fixed Feature extractor) from Figure 2.7, the results were surprisingly high with F_1 score of 0.74. This showed that model M3+AS generalizes very good even in different fruit tree species even though never trained on images of that specie. When the wights were further trained in very few additional images of pears, with just 5 more epochs the model reached a F_1 score of 0.87 and IoU of 0.54.

2.4.3 Cropload

Estimating fruit number based just on detected fruit from the model is not very accurate, despite the model being more than 90%accurate. This not due to the model detection pipeline or architecture, but because of the nature of the tree itself. When an image is taken of the tree, there are still apples that are not visible even by human looking at the image. Even looking at the tree from two meeters far (the same distance the images were taken) is difficult to see fruits inside canopy. To solve this, we correlated the number of visible apples and hidden ones. We counted every fruit in a tree outside in the orchard where the images were taken, and then we counted again each fruit in the image (manually). We found out that, depending on the training system this number varies from 70% visible to 85% visible. Thus we calculate the tree cropload as below:

$$cropload = d_i + (d_i \times (1 - F_1)) + (t_i \times (d_i + (d_i \times (1 - F_1))))$$

where d_i is detection number by the model, F_1 accuracy score, and t_i is percentage of hidden in that training system, in our case was 0.05 in one type of training system and 0.1 in another one.

2.5 Conclusions

In this paper we presented an approach for fruit detection based on state-of-the art deep neural networks techniques using single shot detectors (YOLO) as an convolutional neural network to detect fruits of apple and pears in the tree canopy. This study demonstrates that modifications like the input grid on the standard model of YOLO yields better results. Furthermore, removing some layers of the model, we loose in accuracy but we gain in processing speed, this due to less compute resources needed to drive the model. In order to accommodate both speed and accuracy, we created another model, based on YOLO, with just 11 layers, with double grid size and introduced two new blocks to it. Those two blocks are completely independent and can be used with any other convolution neural network model. By splitting the image into smaller pieces and feeding separately to the model, the images retains higher resolution and its clearer. In addition, the objects in each individual block are bigger and easier to detect from the model.

By increasing the number of images we increase the F_1 and IoU score of the model. With 5000 images of apple fruits, the accuracy of the detection was F_1 0.79. With techniques like augmenting, we increased this number by four, into 20000 apple fruits, thus the F_1 score reaches 0.9. In case when we added synthetic images, the F_1 score remained the same, however the IoU improved slightly. This shows that synthetic images can be an easily approach to fast generate images to improve the localization of pixels of detected object. Transfer-learning proved to be a very interesting tool in fruit detection pipeline. Using the model trained solely in apple fruits, and later testing unchanged in pears, we observed F_2 of 0.72. Using weights from apple models, and training few epochs further in very small amount of pear images, the F_2 accuracy reached 0.87.

Current limitation of our platform is the compute power needed for the system to run. Because most neural networks have many layers, especially convolutional neural networks, the most suitable to run the models are the CUDA capable devices. Our continuation of this work will include more images of different fruit species, at different growth stages, with different training systems, and using a mobile platform for capturing those images. The earlier and smaller the fruits is, more difficult is to detect. But is very important for different orchard management procedures, to have known the cropload in order to proceed with more precision for the specific task.

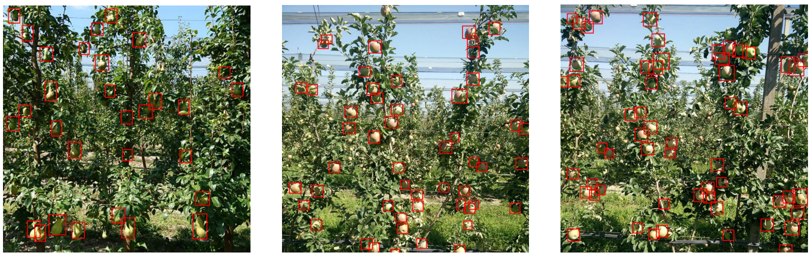


FIGURE 2.8: Detection of apples and pears samples

References

Aubert, Benoit A., Andreas Schroeder, and Jonathan Grimaudo (2012). "IT as enabler of sustainable farming: An empirical analysis of farmers adoption decision of precision agriculture technology". In: *Decision Support Systems* 54.1, pp. 510–520. DOI: 10.1016/j.dss.2012.07.002. URL: https://doi.org/10.1016%2Fj.dss.2012.07.002.

Bac, C. Wouter et al. (2014). "Harvesting Robots for High-value Crops: State-of-the-art Review and Challenges Ahead". In: *Journal of Field Robotics* 31.6, pp. 888–911. DOI: 10.1002/rob.21525. URL: https://doi.org/10.1002%2Frob.21525.

Bargoti, Suchet and James Underwood (2017). "Deep fruit detection in orchards". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. DOI: 10.1109/icra.2017.7989417. URL: https://doi.org/10.1109%2Ficra.2017.7989417.

Barnea, Ehud, Rotem Mairon, and Ohad Ben-Shahar (2016). "Colour-agnostic shape-based 3D fruit detection for crop harvesting robots". In: *Biosystems Engineering* 146, pp. 57–70. DOI: 10.1016/j.biosystemseng.2016.01.013. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.01.013.

Bechar, Avital and Clément Vigneault (2016). "Agricultural robots for field operations: Concepts and components". In: *Biosystems Engineering* 149, pp. 94–111. DOI: 10.1016/j.biosystemseng.2016.06.014. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.06.014.

Chen, Zhong, Ting Zhang, and Chao Ouyang (2018). "End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images". In: *Remote Sensing* 10.1, p. 139. DOI: 10.3390/rs10010139. URL: https://doi.org/10.3390%2Frs10010139.

Colbert, Amy, Nick Yee, and Gerard George (2016). "The Digital Workforce and the Workplace of the Future". In: *Academy of Management Journal* 59.3, pp. 731–739. DOI: 10.5465/amj.2016.4003. eprint: https://journals.aom.org/doi/pdf/10.5465/amj.2016.4003. URL: https://journals.aom.org/doi/abs/10.5465/amj.2016.4003.

Edan, Yael, Shufeng Han, and Naoshi Kondo (2009). "Automation in Agriculture". In: *Springer Handbook of Automation*. Springer Berlin Heidelberg, pp. 1095–1128. DOI: 10.1007/978-3-540-78831-7_63. URL: https://doi.org/10.1007%2F978-3-540-78831-7_63.

Fröder, Christoph (2014). *Employment in European Agriculture: Labour Costs, Flexibility and Contractual Aspects*. Institut der deutschen Wirtschaft Köln, Geopa-Copa.

Jiang, Huanyu, Yongshi Peng, and Yibin Ying (Jan. 2008). "Measurement of 3-D Locations of Ripe Tomato by Binocular Stereo Vision for Tomato Harvesting". In: *American Society of Agricultural and Biological Engineers Annual International Meeting 2008, ASABE 2008* 11.

Jiménez, A.R. et al. (1999). "Automatic fruit recognition: a survey and new results using Range/Attenuation images". In: *Pattern Recognition* 32.10, pp. 1719–1736. DOI: 10.1016/s0031-3203(98)00170-8. URL: https://doi.org/10.1016/s0031-3203(98)00170-8.

Kapach, Keren et al. (2012). "Computer vision for fruit harvesting robots - State of the art and challenges ahead". In: *International Journal of Computational Vision and Robotics* 3.1/2, p. 4. DOI: 10.1504/ijcvr.2012.046419. URL: https://doi.org/10.1504%2Fijcvr.2012.046419.

Kelman, Eliyahu (Efim) and Raphael Linker (2014). "Vision-based localisation of mature apples in tree images using convexity". In: *Biosystems Engineering* 118, pp. 174–185. DOI: 10.1016/j.biosystemseng.2013.11.007. URL: https://doi.org/10.1016%2Fj.biosystemseng.2013.11.007.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). "Deep learning". In: *Nature* 521.7553, pp. 436–444. DOI: 10.1038/nature14539. URL: https://doi.org/10.1038/nature14539.

Lee, Yim-Kul and William T. Rhodes (1990). "Nonlinear image processing by a rotating kernel transformation". In: *Optics Letters* 15.23, p. 1383. DOI: 10.1364/ol.15.001383. URL: https://doi.org/10.1364%2Fol.15.001383.

Mulla, David J. (2013). "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps". In: *Biosystems Engineering* 114.4, pp. 358–371. DOI: 10.1016/j.biosystemseng.2012.08.009. URL: https://doi.org/10.1016%2Fj.biosystemseng.2012.08.009.

Nguyen, Tien Thanh et al. (2014). "Apple detection algorithm for robotic harvesting using a RGB-D camera". In:

Payne, A. et al. (2014). "Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging". In: *Computers and Electronics in Agriculture* 100, pp. 160–167. DOI: 10.1016/j.compag.2013.11.011. URL: https://doi.org/10.1016%2Fj.compag.2013.11.011.

Pierpaoli, Emanuele et al. (2013). "Drivers of Precision Agriculture Technologies Adoption: A Literature Review". In: *Procedia Technology* 8, pp. 61–69. DOI: 10.1016/j.protcy.2013.11.010. URL: https://doi.org/10.1016%2Fj.protcy.2013.11.010.

Rahnemoonfar, Maryam and Clay Sheppard (2017). "Deep Count: Fruit Counting Based on Deep Simulated Learning". In: *Sensors* 17.4, p. 905. DOI: 10.3390/s17040905. URL: https://doi.org/10.3390%2Fs17040905.

Redmon, Joseph et al. (2015). "You Only Look Once: Unified, Real-Time Object Detection". In: *CoRR* abs/1506.02640. arXiv: 1506.02640. URL: http://arxiv.org/abs/1506.02640.

Sa, Inkyu et al. (2016). "DeepFruits: A Fruit Detection System Using Deep Neural Networks". In: *Sensors* 16.8, p. 1222. DOI: 10.3390/s16081222. URL: https://doi.org/10.3390%2Fs16081222.

Si, Yongsheng, Gang Liu, and Juan Feng (2015). "Location of apples in trees using stereoscopic vision". In: *Computers and Electronics in Agriculture* 112, pp. 68–74. DOI: 10.1016/j.compag.2015.01.010. URL: https://doi.org/10.1016%2Fj.compag.2015.01.010.

Song, Y. et al. (2014). "Automatic fruit recognition and counting from multiple images". In: *Biosystems Engineering* 118, pp. 203–215. DOI: 10.1016/j.biosystemseng.2013.12.008. URL: https://doi.org/10.1016%2Fj.biosystemseng.2013.12.008.

Stafford, J.V., ed. (2007). *Precision agriculture '07*. Wageningen Academic Publishers. DOI: 10.3920/978-90-8686-603-8. URL: https://doi.org/10.3920%2F978-90-8686-603-8.

Stajnko, D., M. Lakota, and M. Hočevar (2004). "Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging". In: *Computers and Electronics in Agriculture* 42.1, pp. 31–42. DOI: 10.1016/s0168-1699(03)00086-3. URL: https://doi.org/10.1016/s0168-1699(03)00086-3.

Tao, Yongting and Jun Zhou (2017). "Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking". In: *Computers and Electronics* *in Agriculture* 142, pp. 388–396. DOI: 10.1016/j.compag.2017.09.019. URL: https://doi.org/10.1016%2Fj.compag.2017.09.019.

Chapter 3

Estimating fruit tree cropload and effectiveness of chemical thinning based on fruit detection

Abstract

Cropload estimation is essential for efficient orchard management starting with optimizing chemical thinning, planning the labor force, harvest equipment, and vehicles for transportation of fruit from field to the packing plant. In this paper is presented a new approach to determine tree cropload by using a combination of deep neural networks with convolution neural networks (CNN). For this we use a model to detect fruit in a canopy. To accurately determine cropload, images of the canopy were fed to another CNN which correlated tree canopy size and shape to the number of hidden fruit due to occlusion, then the output is feed to an simple deep neural network (DNN) together with detected fruit number to predict/estimate the overall canopy cropload. The model was then compared to an algorithmic method where parameters like cultivar, canopy shape and training system are manually estimated. Results show 60% to 80% accuracy of the algorithmic method while the deep learning model shows accuracy of around 95% prior to harvest, and more than 90% after chemical thinning, when fruit reached 20mm. Based on this, another CNN was developed to estimate the effectiveness of fruit thinning. Using these models, individual tree thinning can be estimated, so that when mechanical thinning would be applied, it would be based on those data

3.1 Introduction

CROPLOAD estimation is one of the most important managing factors pre-harvest assessments and it is essential for efficient and effective management of orchards [Gongal et al., 2016; Cheng et al., 2017]. For growers, flowering intensity and yield forecast is very valuable information. Very early in season, this information helps them to properly insure their crops so that they can be compensated in the event of financial loss due to weather [Gongal et al., 2016; Cohen, Linker, and Naor, 2011]. Later in the season, cropload estimation helps to optimize chemical fruit thinning [Volz, 1988; Manfrini, Taylor, and Grappadelli, 2009], which is key to improve fruit quality and maximize marketable yield. According to [Mizushima and Lu, 2011] when growers have this information, they can develop more robust and efficient crop load management and harvest strategies that can maximize their profits. Further, cropload estimation prior to harvest is essential for efficient management of the labor force, harvest equipment, and vehicles for transportation of fruit from field to the packing plant [Wulfsohn et al., 2011]. Knowing the amount of production is beneficiary for pack-houses to optimize post-harvest handling and storage [Gautam et al., 2017; Ampatzidis and Vougioukas, 2009].

Crop load adjustment through fruit thinning is a routine practice adopted by fruit growers to obtain a marketable product [Pla and Juste, 1995]. Management of crop load is a balancing act between reducing crop load sufficiently to achieve optimum fruit size and adequate return bloom and not reducing yield excessively. Reduction of crop load can be accomplished through hand thinning or chemically induced thinning, although chemical thinning typically is the preferred method. Manual thinning, although effective, is a labor-intense and expensive operation, accounting for over 30% of the production costs [Gebbers et al., 2013; Layne and Bassi, 2008]; chemical thinning, on the other hand, provides inconsistent results and, though relatively inexpensive, is an unreliable practice. Despite its shortcomings, for the past 50 years chemical thinning using plant-growth-regulating chemicals applied shortly after bloom has been the preferred method apple growers use to thin fruit [Addicott and Lynch, 1955].

There are mny different approaches on orchard design and tree management among farmers. However there is an obvious correlation between the tree management and fruit quality and/or quantity. Trunk cross-sectional area (TCSA) and the production potential of a tree are affected by different tree management and pruning regimes [Westwood and Roberts, 1970]. According to [Forshey, 1986; Wright et al., 2006; Koike et al., 2003] fruit thinning improves leaf:fruit ratio by increasing the leaf area available to each of the remaining fruit. While a study by [Wright, Nichols, and Embree, 2006] reported that the contour method of measurement for apple tree canopy volumes and shape (TCVS) accounted for greater variation in fruit sizes compared to TCSA measurements.

Most of the yield forecasts are based on manual counting of the number of apples on selected trees during the season. Classical yield estimation methods, which consist on manual collection and weighing of the crop in a limited number of plants prior to harvest is tedious and insufficient to obtain representative yield data [Aquino et al., 2018; Manfrini, Taylor, and Grappadelli, 2009]. While field surveying is considered the most reliable method of forecasting crop yields and has been used as a benchmark for validating yield forecast models based on remote sensing when true yields were not available, it is still extremely time-consuming and the small number of trees that can be inspected is insufficient to describe the high variability of yields that exists in apple orchards. To address these issues, machine vision systems involving different types of sensors and image processing techniques are considered as potentially best solutios[Aggelopoulou et al., 2010; Gongal et al., 2016].

This study we present a new approach utilizing Artificial Intelligence models based on Convolutional Neural Networks to estimate tree cropload at different times in the season and to estimate effectiveness chemical thinning. This represents an extension of application of the fruit detection model presented in the previous chapter of this dissertation.

Two approaches were used to estimate overall tree cropload based on the number of apples detected by the model mentioned above. One based on a simple relation between detection and cropload; the second one is based on CNNs. In order to take into account tree canopy volume and shape and to estimate hidden fruit into canopy (behind foliage) or fruit occlusion, a CNN called GoogLeNet was used. In this case CNN is a very good choice as it tries to correlate tree canopy volume and shape (TCVS) to the real number of fruit (hidden and detected).

To determine the effectiveness of chemical thinning, and to determine if and how much manual/mechanical thinning should be applied, two different methods were used and compared. One, based on a simple relation between the cropload and the desired amount of thinning. However, this might not be a good approach as there i a lot of variability, a lot of trees can be small in size or have less cropload than the desired cropload. Second, we use the same CNN (GoogLeNet) to correlate the TCVS to the desired thinning amount.

3.2 Background and Related Work

Nowadays, since there are not accurate flowering and yield forecast methods, growers tend to perform insufficient thinning, preferring to be on the safe side rather than causing irreversibly low yields. This has significant repercussions later in the season when hand thinning is required to compensate for the insufficient chemical thinning. Furthermore, only after the process of natural fruit drop has ended and the fruit have reached a size where growers can estimate yield by visual inspection, then hand thinning is performed. According to [Cohen, Linker, and Naor, 2011] by doing this, two consequences follow:

- it results in smaller fruit at harvest, which, in addition to having a lower market value, require more time for harvesting.
- high crop load before the late thinning reduces flower bud initiation, which lowers the potential bloom in the subsequent season.

As machine vision systems have gotten better during the years, and more open-sourced and free image processing techniques together with AI have received more attention in last years [Gongal et al., 2016; Rahnemoonfar and Sheppard, 2017]. Using imagery for yield estimation is not new in agriculture, with areial image acquisition being reported for annual crops such as wheat [Pan, Li, and Sun, 2007] and rice [Swain et al., 2008] as well as in perennial crops such as blueberries [Zaman et al., 2008]. The prediction of crop yield by remote sensing technology has been one of the most important subjects in precision agriculture [Rodriguez et al., 2004; Yang et al., 2004]. However, most of the high altitude imagery and satellite based sensing still has higher costs compared to ground and/or near ground imagery [Aggelopoulou et al., 2010; Swain et al., 2008]. To understand better the cropload many studies have suggested that the best practice is for individual fruit identification first [**Gongal2016**; Linker, Cohen, and Naor, 2012], followed by a model approximation second [Jiménez et al., 1999; Stajnko, Lakota, and Hočevar, 2004].

Different techniques and methods have been used over the years to identify fruit and correlate between those algorithms prediction and manually counted fruit in an image [Cheng et al., 2017; Dorj, Lee, and Yun, 2017]. Furthermore, different sensors have been used like color camera [Linker and Kelman, 2015; Volz, 1988; Linker, Cohen, and Naor, 2012], thermal camera [Stajnko, Lakota, and Hočevar, 2004], and multispectral and hyperspectral cameras [Kim, Reid, and Zhang, 2008; Mehl et al., 2004]. Multispectral images have been used to estimate the number of oranges [Okamoto and Lee, 2009] and mandarins [Ye et al., 2007] on trees close to harvest and to correlate them to the actual yield.

Other studies [Swain et al., 2008; Escolà et al., 2011; Llorens et al., 2011] used ultrasonic sensor-based automatic yield monitoring system as well as LIDAR sensors for canopy mapping [Cheein et al., 2015; Pfeiffer et al., 2018; Sanz et al., 2018], while [Wei and Salyani, 2004] developed a laser scanning system and corresponding algorithms for measurement of tree canopy height, width, and volume and relate those data to yields. In almost all studies, detection was limited by clustering and occlusions of fruit, and variable lighting conditions in orchards [Ji et al., 2012]

Apart from sensors, estimating fruit yield using digital images has been based on a variety of features associated with field crop and various types of image processing techniques have been investigated to identify fruit. According to [Bulanon and Kataoka, 2010] those techniques can be divided into:

- Shaped-base techniques
- Spectral-base techniques

Several authors [Bulanon et al., 2002; Leemans, Magein, and Destain, 1999] used spectral analysis to identify apples while [Linker and Kelman, 2015; Linker, Cohen, and Naor, 2012; Stajnko, Lakota, and Hočevar, 2004] used shape, color and texture based analysis. [Swain et al., 2008] used the pixel index representing the ripe fruit in the images and achieved very high estimation coefficient. In another study [Adamsen et al., 2000] used colour digital images to count the number of flowers and ultimately to estimate yield.

Using new and innovative techniques of AI, [Cheng et al., 2017] presented an approach of using image analysis and tree canopy features to predict early yield with artificial neural networks. The novelty of this approach is the combination of fruit features, tree canopy features and foliage leaf area to develop models for early yield prediction.

3.3 Materials and Methodology

The study was carried out during 2017 in a 4 year apple orchard of the experimental farm of the University of Bologna, Cadriano, Italy. Data/images were collected from a 0.5 ha orchard of 'Gala' apples. The training system of the orchard is an adapted Spindle system with planting dimensions $3.3m \times 1m$.

100 images of apple tree canopies were taken during the whole season, twice per month. Different capture sensors (cameras), but this experiment we will consider the 100 images taken with a smartphone camera. The camera sensor was 16 mega pixels. Each photo that was taken was from a distance of 1.5 meter form the tree.

Each image was analysed, using fruit detection model presented in the previous chapter of this dissertation. Accuracy score and the number of apples on that tree were given by the model, then was proceeded with cropload calculations. To calculate cropload two methods were used: algorithmic method - AlgoCrop (a simple relation algorithm) and depp-learning method - DeepCrop (a CNN and DNN chained together to estimate the cropload).

The determination of cropload per tree, allowed to perform an estimate of the thinning intensity required to reach the desired cropload.

3.3.1 Fruit detection model

Despite being one of the most popular state-of-the art model, YOLO (You-Only-Look-Once model) has problems detecting small objects [Redmon et al., 2015]. The main problem with YOLO is that, the model can detect only one object class per cell, making it very difficult to detect small apples. And since in apple tree, we are dealing with small fruit in relation to the canopy, we needed t use a modified version based on YOLO.

The standard YOLO takes input image and divides into 13×13. In our case we scaled up to 26×26. This improved the detection, as at this division of input image, we approximately have the grid cell size approximately like the apple size. Same time, we removed more layers, making the model shallower but faster. We saw that, removing pooling layers and some other convolutional layers, we increase in speed while not loosing in accuracy. Two more blocks were added (one in entrance and one in the end, Figure 3.1). The entrance block, noted as "splitter" and the end one noted as "joiner". Splitter, takes the image and separates it into 4 individual images. The resolution before split is 1216×1216 and then it gets splits into exactly four 608×608 images. The joiner at the end, is responsible to put the four pieces together and output the results in 1216×1216 single image with detection. Ths technique of adding blocks to split and join the image introduced by [Chen, Zhang, and Ouyang, 2018], is very effective and accurate, however it decreases the speed of network, as this in essence in feeding four images instead of one. More details of this model were describes in previous chapter .

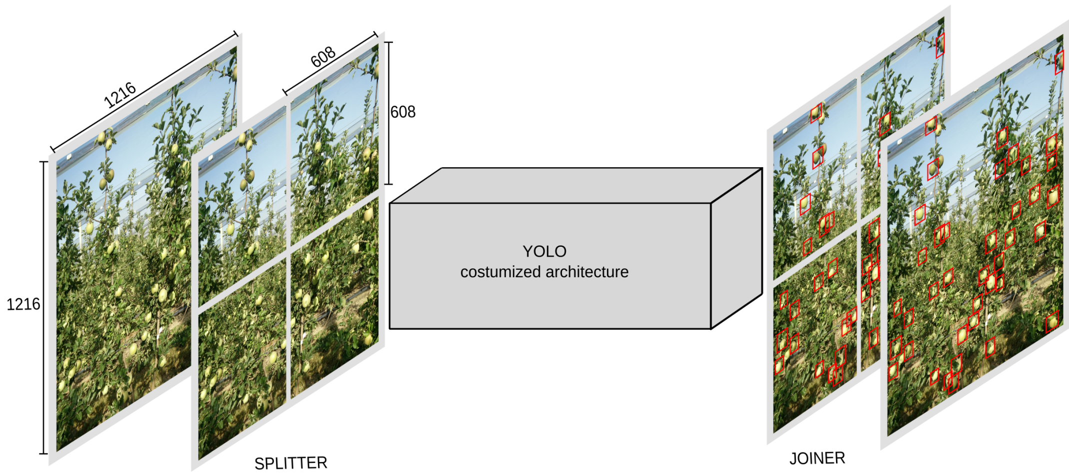


FIGURE 3.1: Fruit detection model

3.3.2 Estimating cropload

Estimating fruit number simply based on the fruit detected by the model is not very accurate, despite the model being more than 90% accurate. This not because of the model detection pipeline or architecture, but because of the nature of the tree itself. Within an image of the tree, there will be still apples that are not visible even by human looking at the image. This is a phenomenon well known in computer vision counting and tracking systems . Even looking at the tree from 1.5 meters (the same distance the images were taken) is difficult to see fruit inside canopy. To solve this, two techniques were developed and compared them to see the performace.

AlgoCrop

AlgoCrop is a simpler solution, which correlates the number of visible to that of hidden apples. Every fruit was counted in a tree, and fruit were counted again in the image, manually. Depending on the training system this number varies from 70% visible to 85% visible. Thus tree cropload calculates as:

$$cropload = d_i + (d_i \times (1 - F_1)) + ((t_i + m_i/2) \times (d_i + (d_i \times (1 - F_1))))$$

where d_i is detection number by the model, F_1 accuracy score, t_i is percentage of hidden in that training system, and m_i approximate canopy size.

DeepCrop

The second method uses a Convolutional Neural Network (GoogLeNet with 24 Layers) where the output of this CNN serves as input of another Deep-Neural Network (a simple classifier with 6 Layers). It is important to use CNN to capture the shape and the foliage of the tree, as fruit occlusion is highly dependant on the amount of tree foliage.

GoogLeNet is used as a CNN classifier. Because of the use of Inception modules (Figure 3.2), GoogLeNet is computationally very cheap but very accurate. However a simple 3x3 kernel with 256 input channels and 256 output, would have an amount of 9x256x256 calculations. Such a network where every output is connected to every input, is referred as dense connection. And while in most of CNNs activation layer for those connection is often either zero or not valuable, proving that not all those input channels are connected to output ones. Despite many techniques developed to cut off those unnecessary connections, the computation needed is huge. The inception module of the model we chose to work with approximates a sparse CNN with a normal dense construction, and since the effective number is low (because of zeros and unnecessary activations) the number of convolution filters is kept small. In addition it uses convolutions of different sizes to capture details at different scales (5x5, 3x3) and it uses the so called bottleneck layer 1x1, for reduction of computation requirements. GoogLeNet is a 27 layer deep CNN, with 22 convolution and inception layers and 5 pooling layers. However the overall number of the independent blocks is over 100. The images were downresolutioned to 512x512x3 inputs for GoogLeNet.

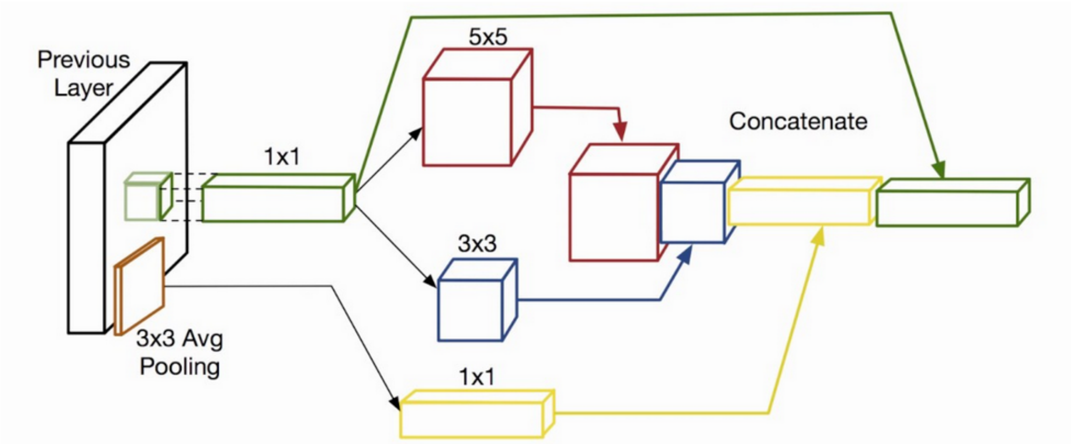


FIGURE 3.2: GoogLeNet's Inception module

The last layer of the network serves as input of the next neural network, a simple 6 layer DNN.. However only after normalization with other inputs: date of image taken, F1 score of detection model and number of apples detected by detection model.

3.3.3 Estimating thinning amount

This task is far more complex. In our experimental trial, chemical thinning was applied at the beginning of May (when fruit diameter was around 8mm) and manual thinning at the end of June (when fruit diameter was around 25mm). The aim of the thinning amount estimation is to define the effectiveness of chemical thinning and estimate if and how much manual (or mechanical) thinning should be applied. Again two methods were used: algorithmic method (AlgoThinning) and deep learning model (DeepThinning).

Images of fruit detected after chemical thinning (end of May) were used as input, and images detected after manual thinning were used as target (beginning of July). DeepCrop was used to estimate tree cropload after chemical thinning was performed. AlgoThinning and DeepThinning were evluated and compared to estimate how much manual thinning should be applied. To measure the precision of thinning DeepCrop was used again on images taken after manual thinning was performed. This provided data before and after thinning to train the models.

AlgoThinning

The first methodology used to estimate the effectiveness of chemical thinning is a simple relation between actual cropload after chemical thinning and the target. This is a very simple relation, where the target amount is subtracted from the mean of each tree cropload:

$$thinning = \sum_{i=1}^{100} c_i - T$$

where c_i is cropload and T target amount for manual thinning.

DeepThinning

The second method implements another CNN. The goal is to exploit tree shape in conjunction with cropload to define the thinning amount to be applied. This is important as in many orchards variables light and soil conditions render cropload and canopy quite different from tree to tree. Here a CNN estimated the vigour of the tree and related it to the amount of thinning. The same CNN we used for DeepCrop, the GoogLeNet model was used. The output of this CNN serves as one of the inputs (together with CropLoad) of a simpler DNN with 6 layers.

3.4 Results

3.4.1 Cropload estimation

The objective of this part is to evaluate and compare algorithms and deep learning architectures for the estimation apple fruit tree cropload before manual thinning is done and subsequently until harvest.

As shown in Figures 3.3, 3.6, 3.9, and 3.12 the apple detection accuracy varied for individual tree. However there it is clear that the latter the apple detection is performed the higher the accuracy is. Accuracy of detection in late June (Figure 3.3), the detection accuracy score (F1) is 86% while in early August (3.12), the F1 score reaches 92%.

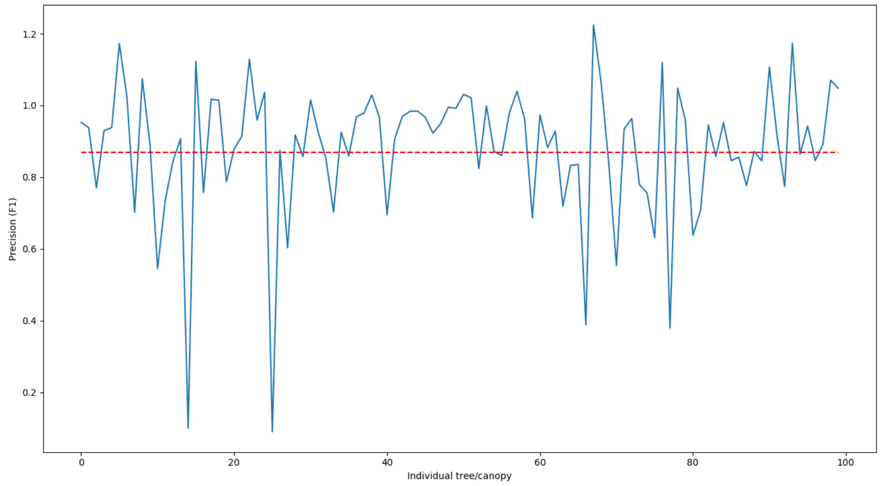


FIGURE 3.3: Accuracy of apple detection (Late June)

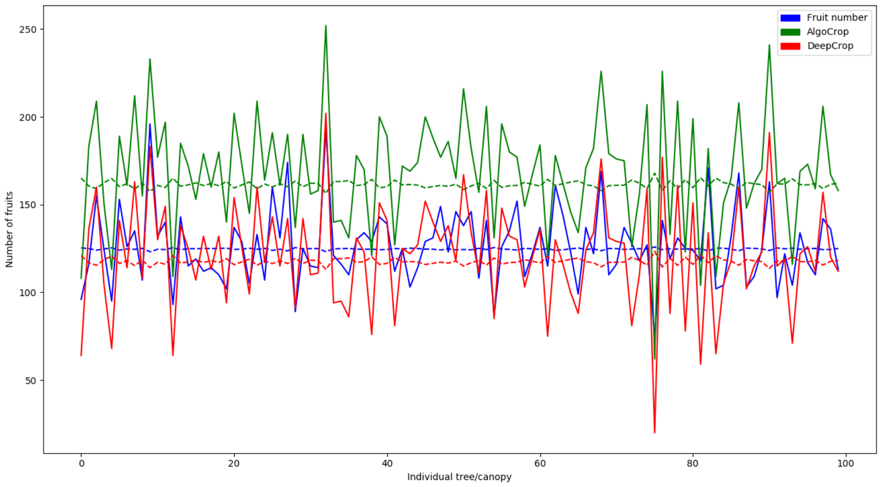


FIGURE 3.4: Cropload estimation per individual tree (Late June)

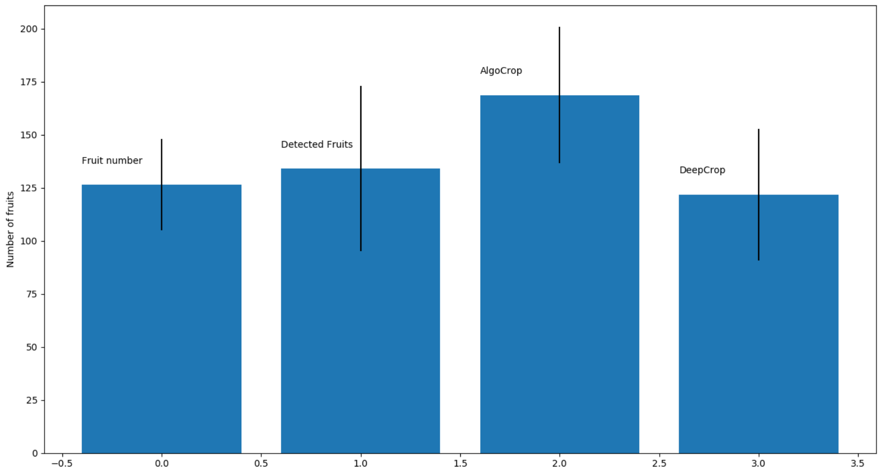


FIGURE 3.5: Comparison of cropload estimation methods (Late June)

However the detection rate here does not account for the fruit occlusions due to canopy and foliage, that can be up to 30% of hidden fruits. And that varies highly inside the orchard, due t the canopy structure. AlgoCrop and DeepCrop models are compared and evaluated in 100 individual trees as shown in Figures 3.4, 3.7, 3.10, and 3.13. The blue line indicates the real number of fruit in the tree (visible and occluded). When we simply used the AlgoCrop algorithm (green line), with 0.87 F1 score, 0.8 occlusion per Spindle training system and 0.7 score of average canopy size for late June (Figure 3.4) we have quiet disappointing results. The whole algorithmic approach always assumes that there those variables are constant, thus, we have a trend line that follows the patterns of detected fruit, which in our case is 30% higher than the real number 3.5.

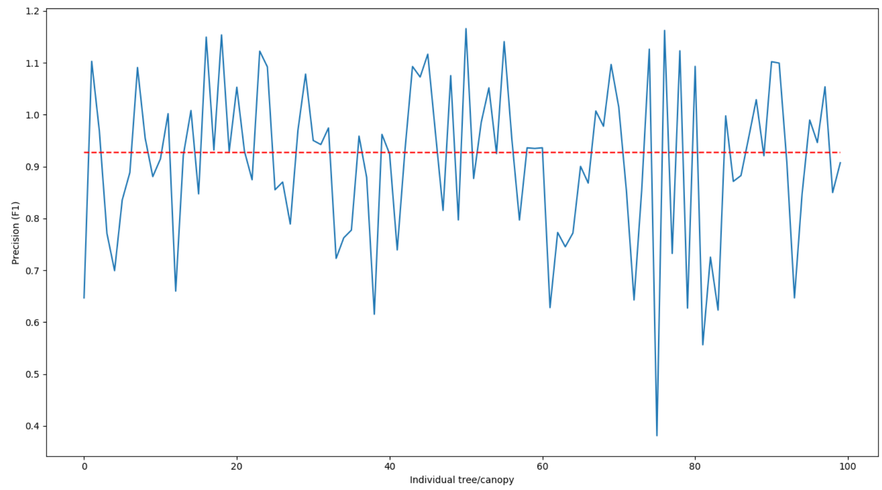


FIGURE 3.6: Accuracy of apple detection (Early July)

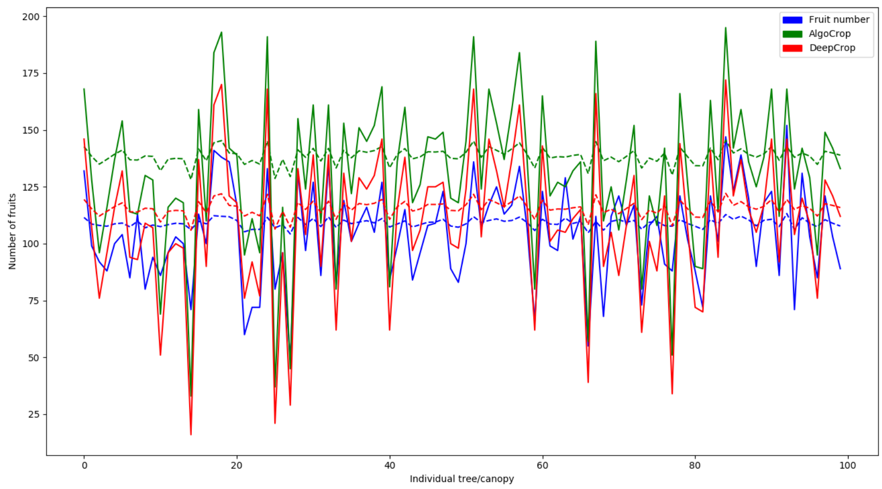


FIGURE 3.7: Cropload estimation per individual tree (Early July)

Using DeepCrop (red line), the model was trained in multiple images, which was able to better correlate the real number of fruit, with occluded number of fruit and detected number of fruit, we got better results. As shown in Figures 3.4, 3.7, 3.10, and 3.13 the model is very accurate in predicting the overall cropload more than 90% few weeks before harvest (early August, Figure 3.14).

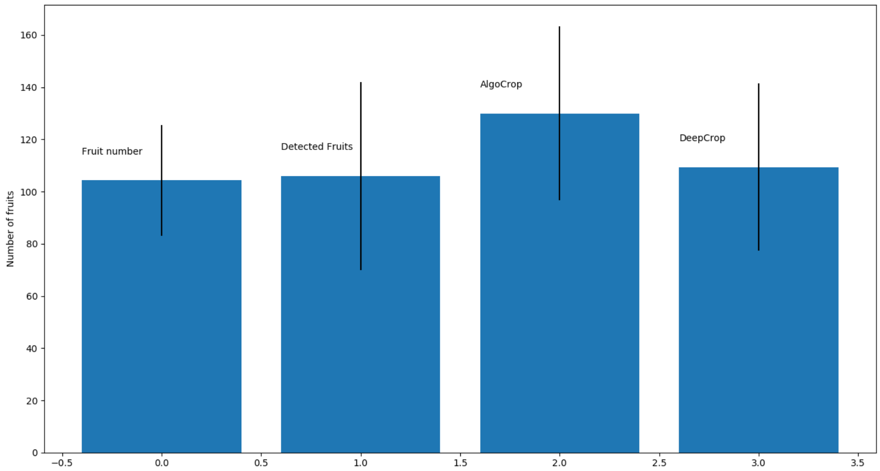


FIGURE 3.8: Comparison of cropload estimation methods (Early July)

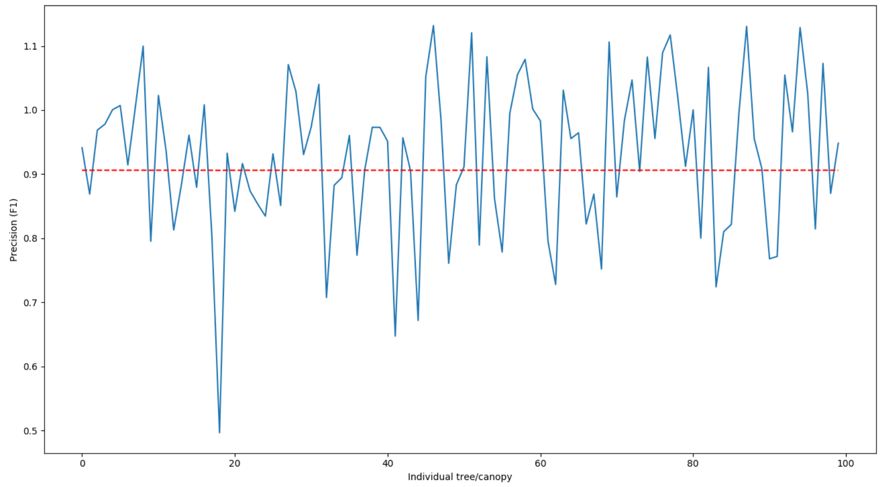


FIGURE 3.9: Accuracy of apple detection (Late July)

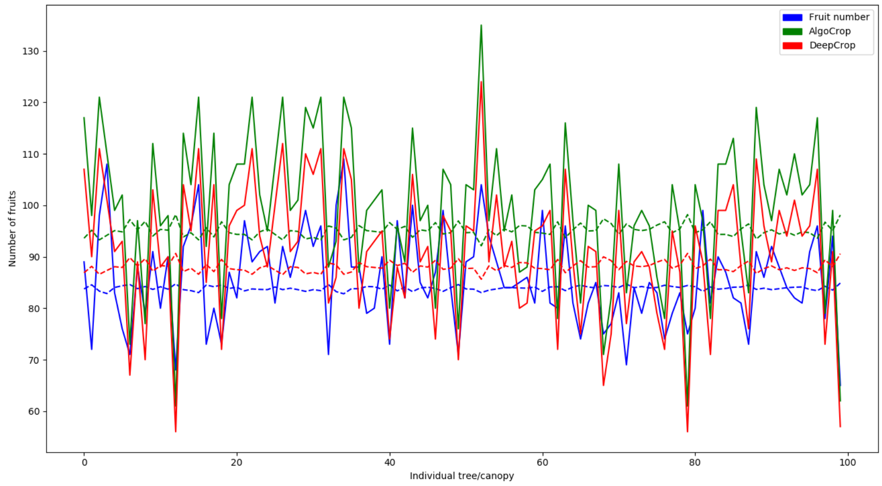


FIGURE 3.10: Cropload estimation per individual tree (Late July)

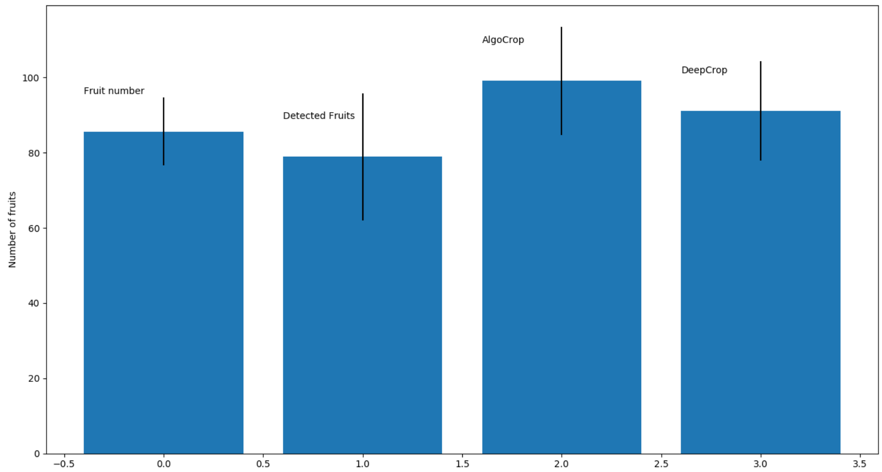


FIGURE 3.11: Comparison of cropload estimation methods (Late July)

Figures 3.5, 3.8, 3.11, and 3.14, show the real number of fruit in the first column, and the detected number of fruit in the second. As seen here, the number of detected fruit varies quite a lot from tree to tree the earlier in season (Figure 3.5) the detection is performed, and it appears that the number of detected apple is higher than the number of real apples, which shows that despite many apples being occluded, the model still detects more apples. This is due to size of apples and miss-detection. When fruit are around 20mm (and at that size they are green colored) is very difficult to separate them from leaves. This factor was not considered initially in AlgoCrop, thus making it very poor at predicting overall cropload. On the other hand, DeepCrop, was able to learn during model training to relate the canopy size to the real number, thus allowing the model to perform much better than AlgoCrop (Figure 3.5).

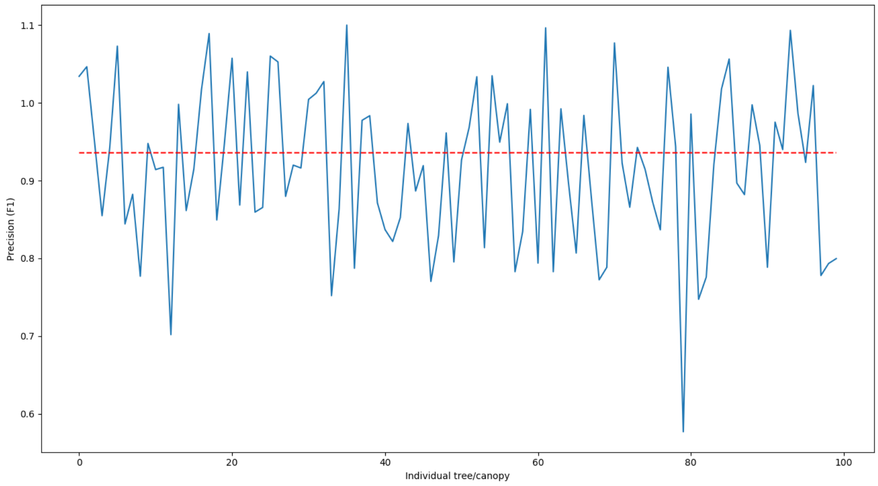


FIGURE 3.12: Accuracy of apple detection (Early August)

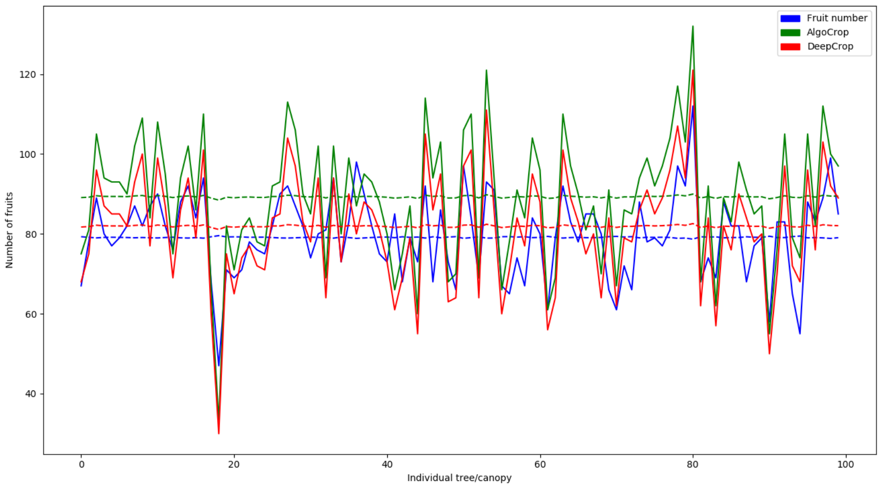


FIGURE 3.13: Cropload estimation per individual tree (Early August)

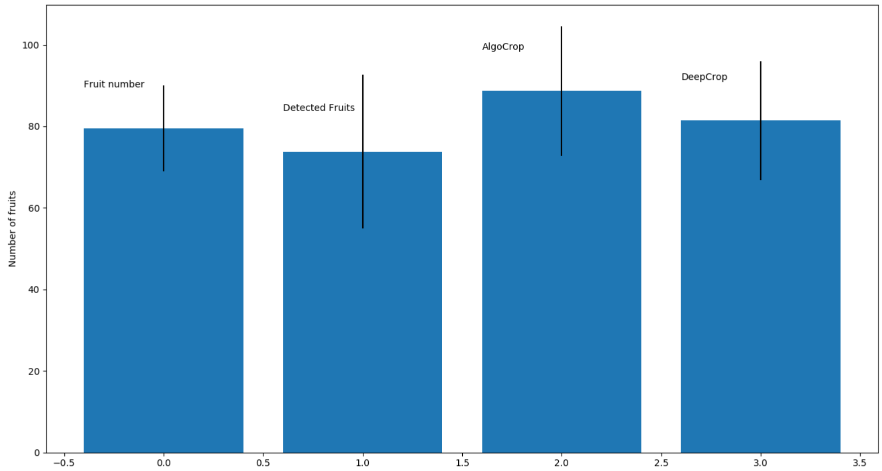


FIGURE 3.14: Comparison of cropload estimation methods (Early August)

3.4.2 Thinning estimation

Using DeepCrop estimation for cropload in mid/late June, two methods, AlgoThinning and DeepThinning, were compared to estimate amount of thinning (Figure 3.15). As shown, when using AlgoThinning, we subtracted from each tree the unnecessary amount of fruit from the average of DeepCrop. Again here, as it was with AlgoCrop, the AlgoThinning simply follows the DeepCrop trend line. When we used DeepThinning, which in turn takes in consideration the shape of the canopy of the tree, then the trend line seems very close to the desired amount of each individual tree.

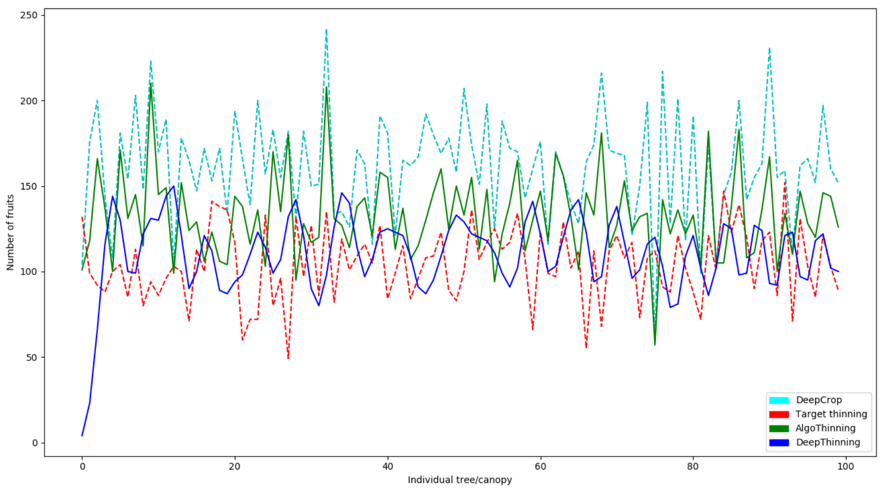


FIGURE 3.15: Thinning estimation per individual tree

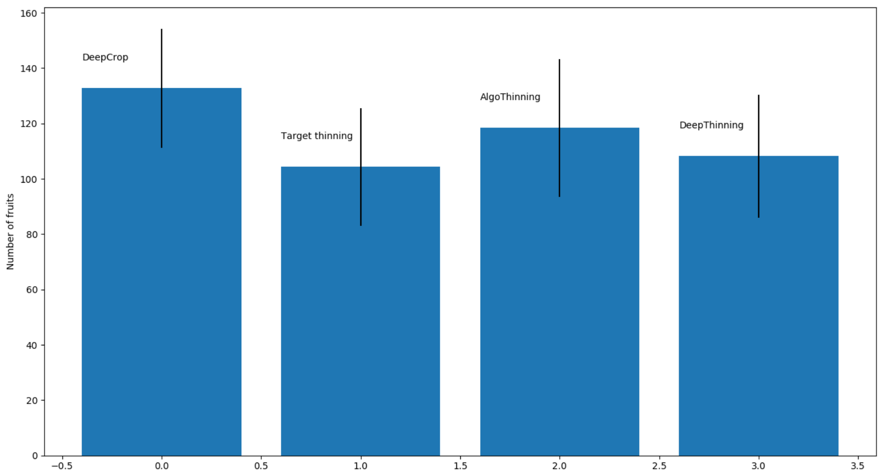


FIGURE 3.16: Comparison of thinning estimation methods

At first glance, it might not look effective to create models for fruit thinning. When cropload is known at more than 90% accuracy, then reaching the desired cropload is a simple matter of subtraction of unnecessary fruit. However cropload varies considerably from tree to tree making a per-tree model very desirable. Another point to be made is related to the cost of thinning. Manual thinning of each tree to a specific number of fruit would be expensive. However, a mechanical thinner, would load a GIS map containing each individual tree thinning instructions, and could apply the amount required much more economically.

3.5 Conclusions

This study presented some models to estimate the cropload on apple trees and estimate the effectiveness of chemical thinning. The novelty of those models is the use of convolutional neural networks to capture canopy size and shape, and relate it to the actual number of fruit on the tree and the number of fruit detected by another CNN model. These models show more than 95% accuracy in estimating tree cropload and more than 90% accuracy in relating cropload to the effectiveness of chemical thinning. The CNNs were able to accurately classify the canopy shape of smaller or larger trees, and estimating the possible amount of fruit occlusion and mis-detection.

However, more accurate and definitely earlier forecast of yield will increase the confidence of growers to perform early chemical thinning, which will minimize the needs for hand thinning, and in turn will increase the fruit size and potential flowering in subsequent seasons.

References

Adamsen, FJ et al. (2000). "Method for using images from a color digital camera to estimate flower number". In: *Crop Science* 40.3, pp. 704–709.

Addicott, F T and R S Lynch (1955). "Physiology of Abscission". In: *Annual Review of Plant Physiology* 6.1, pp. 211–238. DOI: 10.1146/annurev.pp.06.060155.001235. URL: https://doi.org/10.1146%2Fannurev.pp.06.060155.001235.

Aggelopoulou, A. D. et al. (2010). "Yield prediction in apple orchards based on image processing". In: *Precision Agriculture* 12.3, pp. 448–456. DOI: 10.1007/s11119-010-9187-0. URL: https://doi.org/10.1007%2Fs11119-010-9187-0.

Ampatzidis, Y.G. and S.G. Vougioukas (2009). "Field experiments for evaluating the incorporation of RFID and barcode registration and digital weighing technologies in manual fruit harvesting". In: *Computers and Electronics in Agriculture* 66.2, pp. 166–172. DOI: 10.1016/j.compag.2009.01.008. URL: https://doi.org/10.1016%2Fj.compag.2009.01.008.

Aquino, Arturo et al. (2018). "Automated early yield prediction in vineyards from on-the-go image acquisition". In: *Computers and Electronics in Agriculture* 144, pp. 26–36. DOI: 10.1016/j.compag.2017.11.026. URL: https://doi.org/10.1016%2Fj.compag.2017.11.026.

Bulanon, DM and T Kataoka (2010). "Fruit detection system and an end effector for robotic harvesting of Fuji apples". In: *Agricultural Engineering International: CIGR Journal* 12.1.

Bulanon, D.M. et al. (2002). "AE—Automation and Emerging Technologies". In: *Biosystems Engineering* 83.4, pp. 405–412. DOI: 10.1006/bioe.2002.0132. URL: https://doi.org/10.1006%2Fbioe.2002.0132.

Cheein, Fernando A. Auat et al. (2015). "Real-time approaches for characterization of fully and partially scanned canopies in groves". In: *Computers and Electronics in Agriculture* 118, pp. 361–371. DOI: 10.1016/j.compag.2015.09.017. URL: https://doi.org/10.1016%2Fj.compag.2015.09.017.

Chen, Zhong, Ting Zhang, and Chao Ouyang (2018). "End-to-End Airplane Detection Using Transfer Learning in Remote Sensing Images". In: *Remote Sensing* 10.1, p. 139. DOI: 10.3390/rs10010139. URL: https://doi.org/10.3390%2Frs10010139.

Cheng, Hong et al. (2017). "Early Yield Prediction Using Image Analysis of Apple Fruit and Tree Canopy Features with Neural Networks". In: *Journal of Imaging* 3.1, p. 6. DOI: 10.3390/jimaging3010006. URL: https://doi.org/10.3390%2Fjimaging3010006.

Cohen, Oded, Raphael Linker, and Amos Naor (2011). "Estimation of the Number of Apples in Color Images Recorded in Orchards". In: *Computer and Computing Technologies in Agriculture IV*. Springer Berlin Heidelberg, pp. 630–642. DOI: 10.1007/978-3-642-18333-1_77. URL: https://doi.org/10.1007%2F978-3-642-18333-1_77.

Dorj, Ulzii-Orshikh, Malrey Lee, and Sang seok Yun (2017). "An yield estimation in citrus orchards via fruit detection and counting using image processing". In: *Computers and Electronics in Agriculture* 140, pp. 103–112. DOI: 10.1016/j.compag.2017.05.019. URL: https://doi.org/10.1016%2Fj.compag.2017.05.019.

Escolà, Alexandre et al. (2011). "Performance of an Ultrasonic Ranging Sensor in Apple Tree Canopies". In: *Sensors* 11.3, pp. 2459–2477. DOI: 10.3390/s110302459. URL: https://doi.org/10.3390%2Fs110302459.

Forshey, Chester (1986). "Chemical fruit thinning of apples". In:

Gautam, Rahul et al. (2017). "Traceability using RFID and its formulation for a kiwifruit supply chain". In: *Computers & Industrial Engineering* 103, pp. 46–58. DOI: 10.1016/j.cie.2016.09.007. URL: https://doi.org/10.1016%2Fj.cie.2016.09.007.

Gebbers, Robin et al. (2013). "Opti thin - implementation of precision horticulture by tree-specific mechanical thinning". English. In: vol. P-211. Potsdam, Germany, pp. 95 –98.

Gongal, A. et al. (2016). "Apple crop-load estimation with over-the-row machine vision system". In: *Computers and Electronics in Agriculture* 120, pp. 26–35. DOI: 10.1016/j.compag.2015.10.022. URL: https://doi.org/10.1016%2Fj.compag.2015.10.022.

Ji, Wei et al. (2012). "Automatic recognition vision system guided for apple harvesting robot". In: *Computers & Electrical Engineering* 38.5, pp. 1186–1195. DOI: 10.

1016/j.compeleceng.2011.11.005. URL: https://doi.org/10.1016%2Fj.compeleceng.2011.11.005.

Jiménez, A.R. et al. (1999). "Automatic fruit recognition: a survey and new results using Range/Attenuation images". In: *Pattern Recognition* 32.10, pp. 1719–1736. DOI: 10.1016/s0031-3203(98)00170-8. URL: https://doi.org/10.1016/s0031-3203(98)00170-8.

Kim, Yunseop, John F. Reid, and Qin Zhang (2008). "Fuzzy logic control of a multispectral imaging sensor for in-field plant sensing". In: *Computers and Electronics in Agriculture* 60.2, pp. 279–288. DOI: 10.1016/j.compag.2007.09.008. URL: https://doi.org/10.1016%2Fj.compag.2007.09.008.

Koike, Hiroo et al. (2003). "Influence of time of thinning on yield, fruit quality and return flowering of'Fuji'apple". In: *Journal of the American Pomological Society* 57.4, p. 169.

Layne, D. and D. Bassi, eds. (2008). *The peach: botany, production and uses*. CABI. DOI: 10.1079/9781845933869.0000. URL: https://doi.org/10.1079%2F9781845933869.0000.

Leemans, V, H Magein, and M.-F Destain (1999). "Defect segmentation on 'Jonagold' apples using colour vision and a Bayesian classification method". In: *Computers and Electronics in Agriculture* 23.1, pp. 43–53. DOI: 10.1016/s0168-1699(99)00006-x. URL: https://doi.org/10.1016%2Fs0168-1699%2899%2900006-x.

Linker, Raphael, Oded Cohen, and Amos Naor (2012). "Determination of the number of green apples in RGB images recorded in orchards". In: *Computers and Electronics in Agriculture* 81, pp. 45–57. DOI: 10.1016/j.compag.2011.11.007. URL: https://doi.org/10.1016%2Fj.compag.2011.11.007.

Linker, Raphael and Eliyahu Kelman (2015). "Apple detection in nighttime tree images using the geometry of light patches around highlights". In: *Computers and Electronics in Agriculture* 114, pp. 154–162. DOI: 10.1016/j.compag.2015.04.005. URL: https://doi.org/10.1016%2Fj.compag.2015.04.005.

Llorens, Jordi et al. (2011). "Ultrasonic and LIDAR Sensors for Electronic Canopy Characterization in Vineyards: Advances to Improve Pesticide Application Methods". In: *Sensors* 11.2, pp. 2177–2194. DOI: 10.3390/s110202177. URL: https://doi.org/10.3390%2Fs110202177.

Manfrini, L., J.A. Taylor, and L.C. Grappadelli (2009). "Spatial analysis of the effect of fruit thinning on apple crop load". In: *European Journal of Horticultural Science* 74.2. cited By 5, pp. 54–60. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-67649510901&partnerID=40&md5=97cfa3c2408b54db6924c817659a2aae.

Mehl, Patrick M et al. (2004). "Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations". In: *Journal of Food Engineering* 61.1, pp. 67–81. DOI: 10.1016/s0260-8774(03)00188-2. URL: https://doi.org/10.1016%2Fs0260-8774%2803%2900188-2.

Mizushima, Akira and Renfu Lu (2011). "Cost benefits analysis of in-field presorting for the apple industry". In: *Applied engineering in agriculture* 27.1, pp. 33–40.

Okamoto, Hiroshi and Won Suk Lee (2009). "Green citrus detection using hyperspectral imaging". In: *Computers and Electronics in Agriculture* 66.2, pp. 201–208. DOI: 10.1016/j.compag.2009.02.004. URL: https://doi.org/10.1016%2Fj.compag.2009.02.004.

Pan, Gang, Feng min Li, and Guo jun Sun (2007). "Digital camera based measurement of crop cover for wheat yield prediction". In: *2007 IEEE International Geoscience and Remote Sensing Symposium*. IEEE. DOI: 10.1109/igarss.2007.4422917. URL: https://doi.org/10.1109%2Figarss.2007.4422917.

Pfeiffer, Sebastián Arriagada et al. (2018). "Mechatronic terrestrial LiDAR for canopy porosity and crown surface estimation". In: *Computers and Electronics in Agriculture* 146, pp. 104–113. DOI: 10.1016/j.compag.2018.01.022. URL: https://doi.org/10.1016%2Fj.compag.2018.01.022.

Pla, F. and F. Juste (1995). "A thinning-based algorithm to characterize fruit stems from profile images". In: *Computers and Electronics in Agriculture* 13.4, pp. 301–314. DOI: 10.1016/0168-1699(95)00029-1. URL: https://doi.org/10.1016%2F0168-1699%2895%2900029-1.

Rahnemoonfar, Maryam and Clay Sheppard (2017). "Deep Count: Fruit Counting Based on Deep Simulated Learning". In: *Sensors* 17.4, p. 905. DOI: 10.3390/s17040905. URL: https://doi.org/10.3390%2Fs17040905.

Redmon, Joseph et al. (2015). "You Only Look Once: Unified, Real-Time Object Detection". In: *CoRR* abs/1506.02640. arXiv: 1506.02640. URL: http://arxiv.org/abs/1506.02640.

Rodriguez, J. C. et al. (2004). "Wheat yield estimation using remote sensing and the STICS model in the semiarid Yaqui valley, Mexico". In: *Agronomie* 24.6-7, pp. 295–304. DOI: 10.1051/agro:2004037. URL: https://doi.org/10.1051%2Fagro%3A2004037.

Sanz, Ricardo et al. (2018). "LIDAR and non-LIDAR-based canopy parameters to estimate the leaf area in fruit trees and vineyard". In: *Agricultural and Forest Meteorology* 260-261, pp. 229–239. DOI: 10.1016/j.agrformet.2018.06.017. URL: https://doi.org/10.1016%2Fj.agrformet.2018.06.017.

Stajnko, D., M. Lakota, and M. Hočevar (2004). "Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging". In: *Computers and Electronics in Agriculture* 42.1, pp. 31–42. DOI: 10.1016/s0168-1699(03)00086-3. URL: https://doi.org/10.1016/s0168-1699(03)00086-3.

Swain, Kishore C. et al. (2008). "Estimation of rice yield and protein content using remote sensing images acquired by radio controlled unmanned helicopter". English. In: *American Society of Agricultural and Biological Engineers Annual International Meeting 2008, ASABE 2008*. Vol. 1. American Society of Agricultural and Biological Engineers, pp. 134–144. ISBN: 9781605605364.

Volz, RK (1988). “Regulation and estimation of crop load on ‘Gala’apple trees”. In: *New Zealand Journal of Experimental Agriculture* 16.1, pp. 47–53.

Wei, J. and M. Salyani (2004). “DEVELOPMENT OF A LASER SCANNER FOR MEASURING TREE CANOPY CHARACTERISTICS: PHASE 1. PROTOTYPE DEVELOPMENT”. In: *Transactions of the ASAE* 47.6, pp. 2101–2107. DOI: 10.13031/2013.17795. URL: https://doi.org/10.13031%2F2013.17795.

Westwood, MN, AN Roberts, et al. (1970). “The relationship between trunk cross-sectional area and weight of apple trees.” In: *Journal of the American Society of Horticultural Science* 95, pp. 28–30.

Wright, AH et al. (2006). “Fruit mass, colour and yield of ‘Honeycrisp’™ apples are influenced by manually-adjusted fruit population and tree form”. In: *The Journal of Horticultural Science and Biotechnology* 81.3, pp. 397–401.

Wright, H, D Nichols, and C Embree (2006). “Evaluating the accountability of trunk size and canopy volume models for determining apple tree production potential across diverse management regimes”. In: *Acta horticulturae*.

Wulfsohn, Dvoralai et al. (2011). “Multilevel systematic sampling to estimate total fruit number for yield forecasts”. In: *Precision Agriculture* 13.2, pp. 256–275. DOI: 10.1007/s11119-011-9245-2. URL: https://doi.org/10.1007%2Fs11119-011-9245-2.

Yang, Chenghai et al. (2004). “Airborne Hyperspectral Imagery and Yield Monitor Data for Mapping Cotton Yield Variability”. In: *Precision Agriculture* 5.5, pp. 445–461. DOI: 10.1007/s11119-004-5319-8. URL: https://doi.org/10.1007%2Fs11119-004-5319-8.

Ye, Xujun et al. (2007). “Prediction of citrus yield from airborne hyperspectral imagery”. In: *Precision Agriculture* 8.3, pp. 111–125. DOI: 10.1007/s11119-007-9032-2. URL: https://doi.org/10.1007%2Fs11119-007-9032-2.

Zaman, Q. U. et al. (2008). “Estimation of Wild Blueberry Fruit Yield Using Digital Color Photography”. In: *Transactions of the ASABE* 51.5, pp. 1539–1544. DOI: 10.13031/2013.25302. URL: https://doi.org/10.13031%2F2013.25302.

Chapter 4

Sensor-fusion and deep neural networks for autonomous UAV navigation within orchards

Abstract

With the increase of population in the world, the demand for quality food is increasing too. In recent years, increasing demand and environmental factors have heavily influenced the agricultural production. Automation and robotics for fruit and vegetable production/monitoring have become the new standard. This paper discusses an autonomous Unmanned Aerial Vehicle (UAV) able to navigate through rows orchard rows. The UAV is comprised of a flight controller (AP stack), a microcontroller for analog reading of different sensors, and an On-Board Computer (OBC). Pictures are taken through a camera and streamed through WiFi to a Ground Control Computer (GCC) running a convolution neural network model. Based on prior training, the model outputs three directions: RIGHT, LEFT and STRAIGHT. A moving average of multiple frames per second is extracted and sent to a build-in Proportional-Integral-Derivative (PID) controller on the UAV. After error correction from this feedback, controller sends the direction to the flight controller using MAVLink protocol's radio channel overrides, thus performing autonomous navigation.

4.1 Introduction

Agriculture as one of the oldest occupation in the world has seen many changes. Before the Industrial Revolution more than 80% of the population were estimated to be working as farmers, while now that number is estimated to be less than 2%. One of the dominant changes that characterizes a growing economy is the proportionate decline in agriculture sector [Fröder, 2014]. This phenomenon is commonly attributed to two facts: food is not as demanding as other goods and services, or/and the rapid development of new farming technologies leads to expanding food supplies per hectare and per worker [Lima and Souza, 2017].

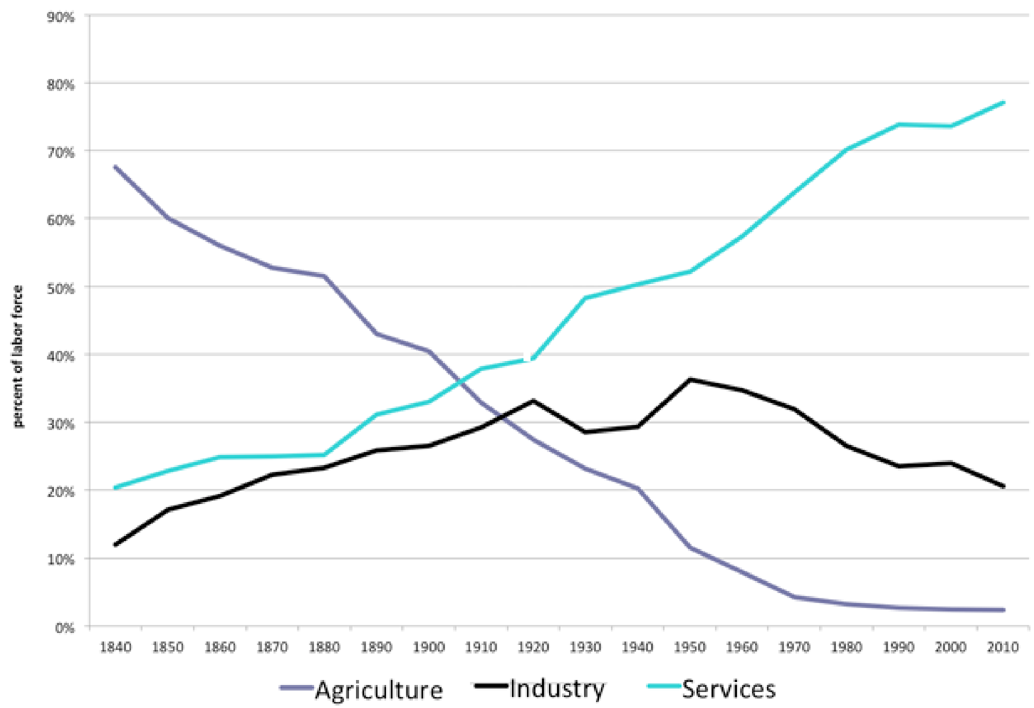


FIGURE 4.1: Distribution of the labour force by sector, 1850-2000

There is a significant change at every step of the industrial revolution. Industry 1.0 brought mechanization, water management etc.; Industry 2.0 brought engines, electricity, protected environments etc.; Industry 3.0 brought computers and smart appliances, GPS tractors and so on; while Industry 4.0 is foreseen to be the most significant one, bringing AI, Automation and Robotics [Harrison and Wolyniak, 2015].

We face the challenge of how to feed the growing world population in the future: it should certainly be sustainable, cost-effective and most importantly environmentally friendly. In order to feed 9.5 billion people that the Food and Agriculture Organization (FAO) predicts to inhabit the planet by 2050 while climate change is making more difficult to grow crops - is going to be done by smart farming, a high-tech and AI driven agricultural management system [Colbert, Yee, and George, 2016]. Agriculture is highly repetitive, and as such, many tasks can be automated. Each agricultural activities on the farm takes a lot of resource: for example planting, maintaining, and harvesting crops need money, energy, labour and resources [Aubert, Schroeder, and Grimaudo, 2012; Mulla, 2013]. What if we can use technology to replace some of the human activities and guarantee efficiency? That's where artificial intelligence and modern robotics is looking for changes.

Fully autonomous vehicles have been studied for many years, with a number of innovations explored as early as the 1920s. The concept of fully autonomous agricultural vehicles is far from new; examples of early driverless tractor prototypes using leader cable guidance systems date back to the 1950s and 1960s [[Basu et al., 2018]. More recently the potential for combining computers with image sensors provided huge opportunities for machine vision based guidance systems.

In Agriculture, autonomous systems can be grouped in three main categories [Emmi et al., 2014]:

1. Big autonomous tractors
2. Small specialized robots
3. Swarm or fleet robotics

Autonomous tractors have been studied and have been in use in agriculture for many years. Precision agriculture helped propel vehicle guidance in terms of providing position information that is required for vehicle guidance [Reid et al., 2000]. Agricultural tractors usually operate on all terrains, and there are a lot of unpredictable disturbances and noise sources to the signals from navigation sensors. Therefore, it is necessary to have an effective means for signal conditioning and system state estimation in the sensor fusion modules [Noguchi et al., 2001]. The topography, vegetation landscape, soil composition, texture and structure, air visibility, illumination, light quality and atmospheric conditions change at rates varying from seconds to months and on scales from millimeters to kilometers [Bechar and Vigneault, 2016]. In order to perform well, the next-gen agriculture autonomous robots (Agbots) must be able to recognize and understand the physical properties of each specific object encountered, and also be able to work under both varying field and controlled environment conditions [Eizicovits and Berman, 2014]. Therefore, sensing systems, robotics arms, specialized manipulators, effectors should be able to work under different and unstable environmental conditions [Bechar and Vigneault, 2017].

Robots have wide applications in PA, ranging from soil analysis, seedling, weed control, environmental monitoring, harvesting and so on, but in a broader perspective they can be grouped as in [Tab 4.1]:

	Purpose
1	Harvest Management
2	Autonomous navigation
3	Pest management and spraying
4	Weed management and mowing
5	Soil Management
6	Irrigation Management
7	Remote camera sensing UAVs
8	Pruning and Thinning
9	Sorting and packing
10	Seedling and nursery
11	Transporting and cleaning
12	Other general purposes

TABLE 4.1: Robots in Agriculture

According to [Bechar, Meyer, and Edan, 2009; Oren, Bechar, and Edan, 2011] a robot, to perform a fully autonomous agricultural action needs to go through four continuous steps: 1)the robot senses and acquires raw data from and about the environment, task and/or its state using various sensors; 2) the robot processes and analyses the data received from its sensors to generate reasoning and a perception of the environment, the task or its state to some level of situation awareness; 3) the robot generates an operational plan based on its perception of the environment and state, or the task objectives; and 4) the robot executes the required actions included in the operational plan [Tab 4.2].

Stage	Task
Sensing	Sensory inputs like: RGB cameras, LIDAR, Sonar, rotary encoders, potentiometers, resistors...
Analyzing	Landmark detection, Point-cloud analyzing, Kinematics and Inverse kinematics of manipulator...
Planning	Trajectory estimation, object voidance...
Action	Performing the planned action, triggering other actions based on location, time, another action ...

TABLE 4.2: Continuous stages of an Agbot [Bechar and Vigneault, 2016]

A farmer robot, to be fully autonomous, needs to navigate through very diverse and harsh environments, where the changes are instant and imminent, and those changes need to be observed and logged without human supervision. Then, it should perform a set of actions at specific location like: picking fruits, evaluates a site, spray pesticide, cut branches, plant a seed, image scan a whole plant or takes specific measurement. Controlled environments like greenhouses are more manageable because of "controllable environment" and better engineered infrastructure, and as result, sensor measurements produce less noise. Whereas outdoor environment are much harsher and generally not controllable, thus making far more difficult for mobile robots. Most of outdoor robot are equipped with GPS for sensing the location, but due to the need of higher accuracy, they are often together with other sensors like Inertial Measurement Unit (IMUs), 3DCameras, Rotary Encoders to create a sensor fusion for a much more precise action taking process. Robots nowadays are wirelessly connected to a central operator to both receive updated instructions regarding the mission, and report status and data. However, a truly autonomous farm robot requires better controllers, localization, communication and action taking systems. The technology is similar to that of autonomous cars applied to agitech. Where it differs is that farming robots often need to manipulate their environment, picking vegetables or fruits, applying pesticides in a localized manner, or planting seeds.

Autonomous navigation robots for agriculture based on camera systems have been in development since late the 80s [Reid and Searcy, 1987; Reid et al., 2000]. They were popular due to the availability of low cost cameras and the plethora of computer vision techniques that could be readily applied [Hiremath et al., 2014]. The most adopted approach in auto-navigation is GPS-based navigation [Pérez-Ruiz et al., 2012]. However fruit production, the trees are often protected with anti-hail netting that makes GPS lock unpredictable, ruling out to relay solely on GPS.

In fruit production, crop monitoring is extremely important. This paper discusses an autonomous Unmanned Areal Vehicle (UAV) flying at tree canopy level, between adjacent rows and under anti-hail nets. To successfully follow the row, the robot it has firstly to know the orchard and where is the starting row, then has to determine the path between two rows while maintaining the altitude and avoid any collision with lateral branches from the trees and void any other obstacles. Machine learning techniques have been used for a long time, to map visual inputs to actions [Giusti et al., 2016]. This paper considers the whole navigation as a classification task: analyzing the front face camera images, by using a convolution neural network to classify the video frames stream into direction with respected weight on a single shot.

4.2 Formulation

Current orchards are very complex with components and management procedures tht change during the growing season. There are many management decisions that often change the structure and visuals of the orchard, thus making it an ever changing organism. In this perspective, hard-coding algorithms for specific tasks where randomness is infinite, is not a effective approach. The path between two rows of trees, is maintained by farmers in different ways, differently during the season, the same goes with the canopy, where plants starts without leaves but then later they are covered with them. The UAV itself has to fly under anti-hail nets, 1.5m above the ground, making headway along the path, resembling a long corridor. Using deep learning approach the model is able to accommodate for changes and progressively learn how to navigate even when new scenarios are being dealt with.

4.3 Materials

The UAV uses a RaspberryPi gen.3 as a On-Board Computer (OBC) which is connected to the flight controller (PixHawk board with ArduPlot software stack) through a serial link. The OBC due to performance limitations can not run deep-learning algorithms on its own, however its an excellent, power efficient computer running full LINUX inside. This allows it to act as intermediary layer between flight controller and other devices: Ground Control Computer (GCC), microcontroller (Arduino) boards with different sensors, camera (RasPi Cam gen.2), and other radio-communication devices. OBC and Flight controller communicate through MAVLink protocol. MAVProxy is used to broadcast device info and status to other devices through wireless hotspot. Since it runs full LINUX and Robotics Operating System (ROS), with the help of MAVROS package, the whole UAV is controlled as any other robot inside ROS. OBC's camera is a RasPi Cam gen.2 which processes and encodes in its own chips the image, making it very easy to directly stream in network.

The input images taken from front-facing camera, are sent to Ground Control Computer GCC (CUDA capable) though WiFi. The computer runs the picture stream through a trained model that has three outputs: right, left, and straight. The moving average of three outputs is sent as MAVLink RC_CHANEL_OVERRIDE through OBC to PixHawk.

4.3.1 Data Acquisition

Data collection is done through the camera of the OBC in the UAV moving inside the orchard, controlled manually by radio controller. At UAV start up, it automatically starts some scripts on OBC (the process is managed with crontab):

1. Connect to known WiFi if any exist, else create hotspot
2. Use MAVproxy protocols to sent and receive flight plans and commands
3. Automatically output camera stream to network

In the GCC, the camera feed is piped from the network with netcat to a python program. The program takes the camera input, divides it in frames, labels dnd saves it in a proper dataset.

The UAV is flown very carefully inside the rows while streaming the camera feed to GCC. The flight is repeated many times, in different rows and directions. However there have been three modes/categories, and for each mode a thousand pictures/frames were taken and labelled accordingly:

(A) LEFT

(B) STRAIGHT

(C) RIGHT

FIGURE 4.2: Images taken from UAV - winter (trees without leaves)

1. LEFT: The drone would fly closer to the left row, and/or yawed (facing) the left side. Data collected were labelled as LEFT, model would return LEFT and RIGHT CHANNEL OVERRIDE would be sent. Fig 4.2a and 4.3a.
2. STRAIGHT: The drone would fly the best position as much as it can, in the middle of the row, facing straight and having both rows symmetrical to each other. Data collected were labelled as STRAIGHT, model would return STRAIGHT and FORWARD CHANNEL OVERRIDE would be sent. Fig 4.2b and 4.3b.
3. RIGHT: The drone would fly closer to the right row, and/or yawed (facing) the right side. Data collected were labelled as RIGHT, model would return RIGHT and LEFT CHANNEL OVERRIDE would be sent. Fig 4.2c and 4.3c.

Pictures were captured during daytime in late winter of 2018. Daytime is important as the RasPi Cam is very sensitive to light quality and light exposure.

In addition to this dataset, another set of images is used. The later set is taken

(A) LEFT

(B) STRAIGHT

(C) RIGHT

FIGURE 4.3: Images taken manually - autumn (trees with leaves)

manually (using a smartphone camera), but during autumn of 2017, while the tree had leaves and chlorophyll was still green. The set has 100 images per mode/category but this number proved to be very small. The same model is run through both sets separately, and then together too.

4.3.2 The Model

To better manage different datasets and models, Nvidia's Deep Learning GPU Training System (DIGITS) is used. DIGITS is not in itself a machine-learning framework, rather is a wrapper for most used frameworks available. It simplifies the common machine-learning tasks such as managing dataset including train/validation/test splitting, designing and training different neural networks (on CUDA capable GPUs), real-time monitoring of the training process and visualisation of the process.

GoogLeNet is used as a Convolutinal Neural Network (CNN) classifier. Because of the use of Inception modules, GoogLeNet is more versatile and computationally less expensive. A simple 3x3 kernel with 256 input channels and 256 output, would have an amount of 9x256x256 calculations. Such a network where every output is connected to every input, is referred as dense connection. And while in most of CNNs activation layer for those connection is often either zero or not valuable, proving that not all those input channels are connected to output ones. Despite many techniques developed to cut off those unnecessary connections, the computation needed is huge. The inception module of the model we chose to work with approximates a sparse CNN with a normal dense construction, and since the effective number is low (because of zeros and unnecessary activations) the number of convolutional filters is kept small. In addition it uses convolutions of different sizes to capture details at different scales (5x5, 3x3) And it uses the so called bottleneck layer 1x1, for reduction of computation requirements. GoogLeNet is a 27 layer deep CNN (Fig 4.4), with 22 convolution and inception layers and 5 pooling layers. However the overall number of the independent blocks is well over 100.

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

FIGURE 4.4: GoogLeNet model structure

4.3.3 Training

Nvidia's DIGITS dealt with splitting training, validation and testing sets. 15% of images was kept for training and 5% for testing. Each images was of size 256x256.

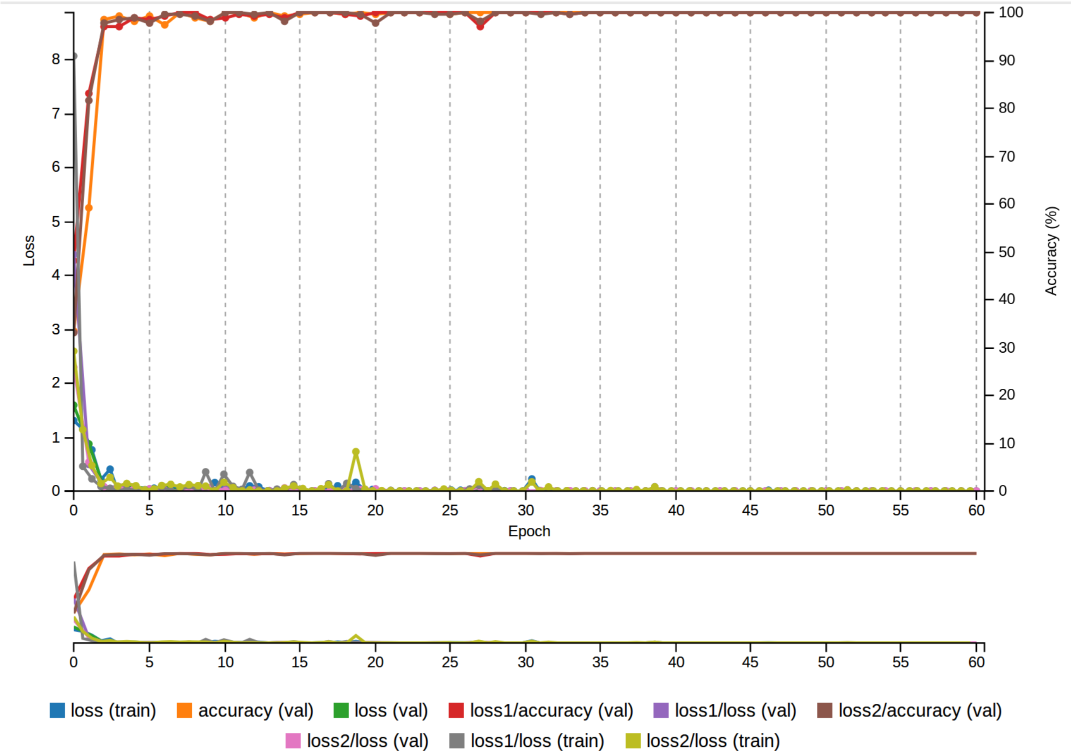


FIGURE 4.5: Training on 60 epochs

Through image augmentation (manually with a script), all LEFT images could be mirrored and produce RIGHT images and vice versa, while STRAIGHT images when mirrored created another STRAIGHT image. The model was trained for 60 epochs (Fig 4.5) in Amazon's AWS S3 instance with NVIDIA's Tesla K80 with 12GB Memory. An Adaptive Movement Solver (ADAM Optimiser) with 0.002 base learning rate was chosen. A sigmoid decay of gamma 0.08 learning rate (Fig 4.6) with 60 steps is used. After training for few hours, the model is downloaded as a pre-trained model, and transferred to the ground control computer (with CUDA capabilities) that runs the transferred model through Nvidia's GTX 960M GPU.

4.3.4 Navigation policy

As briefly mentioned above, the OBC, after capturing the data,it sends them through WiFi to the GCC to run through the model. Since the OBC is not a powerful enough computer, this step is necessary. Pictures received from the network are piped to the Coffe2 model through Netcat (a GNU/Linux utility for reading from and writing to network connections using TCP and UDP protocol). OpenCV captures the Netcat pipe and serves to the model as frames (a single image where the model performs the recognition).

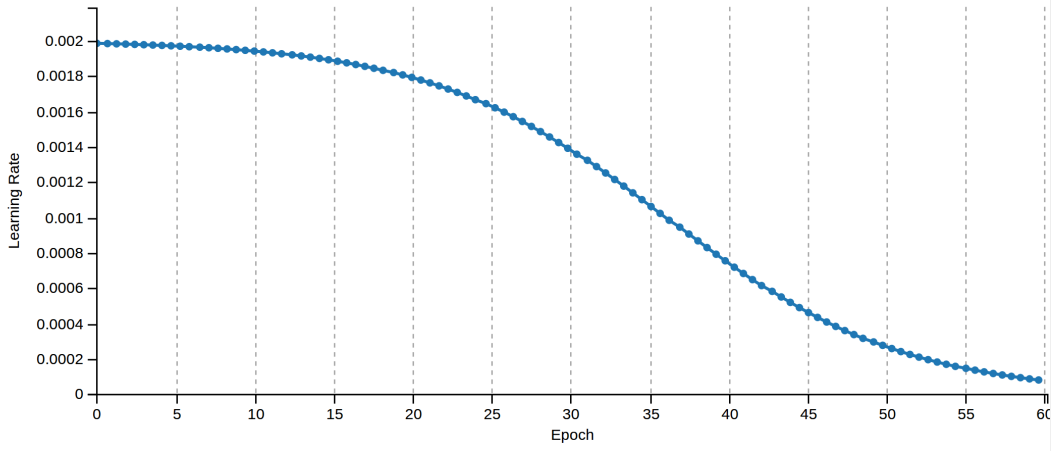


FIGURE 4.6: Learning rate

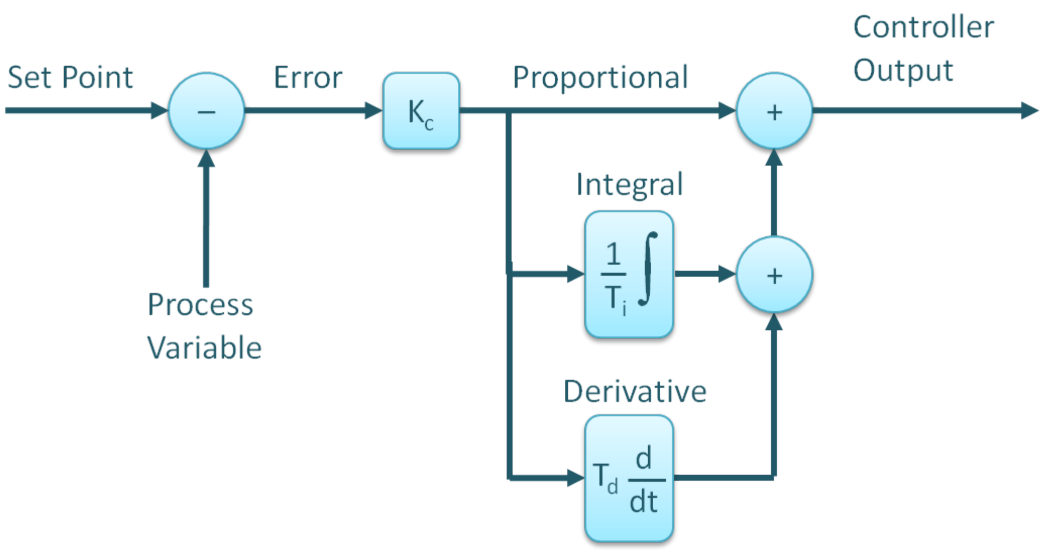


FIGURE 4.7: PID controller for finer control

The model has three outputs (LEFT, RIGHT, STRAIGHT), and based on recognition/classification confidence assigns to each frame the values from 0 (not confident) to 1 (very confident) per class. The outputs are changed into RC_CHANEL_OVERRIDE MAVLink protocol (with values from 1000 to 2000 per channel). Those values go through a moving average script, that takes the values for several frames (in our case all frames per second) and extract the moving average or running mean values for each channel. Running mean of each channel is sent through ZeroMQ (an highly efficient, high-performance and reliable asynchronous messaging library) to the PID (Proportional Integral Derivative) controller that was built in the OBC.

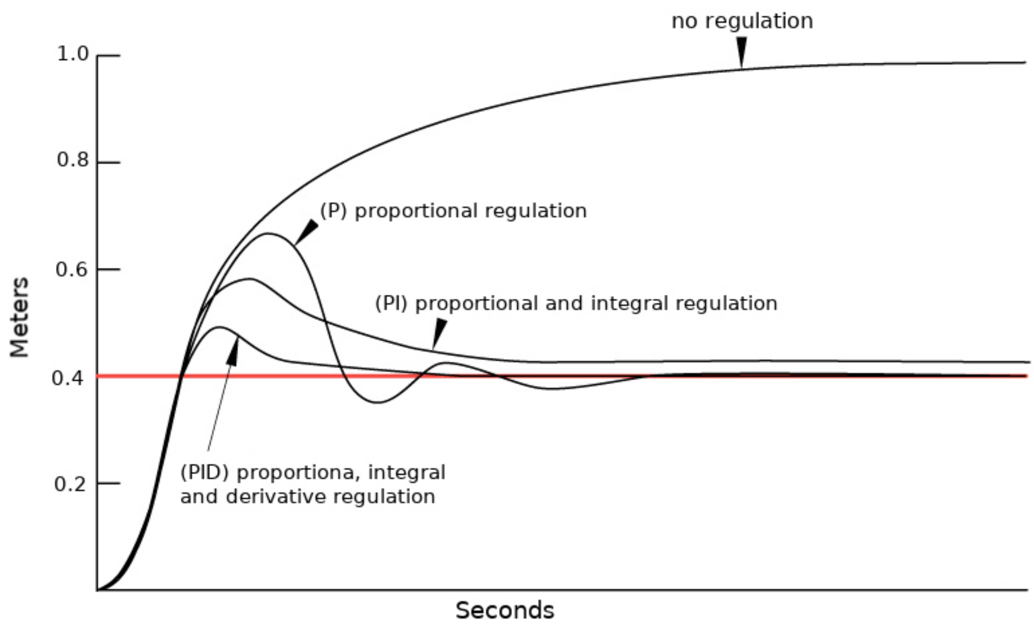


FIGURE 4.8: PID stabilizing mode

4.3.5 PID and Sensor Fusion

PID stands for Proportional, Integral, Derivative, it's an algorithm that reads the data from sensors and calculates how fast the motors should spin in order to attain the desired rotation, angle and speed of the UAV. The goal of the PID controller is to correct the "error", the difference between a measured value (gyro sensor measurement), and a desired set-point. This error can be minimized by adjusting the control inputs in every loop, which is the speed of the motors.

There are 3 values in a PID controller, they are the P term, I term, and D term (Figure 4.7):

- "P" looks at present error – the further it is from the set-point, the harder it pushes
- "D" is a prediction of future errors – it looks at how fast a set-point is approached and counteracts P when it is getting close to minimize overshoot
- "I" is the accumulation of past errors, it looks at forces that happen over time; for example if the UAV constantly drifts away from a set-point, it will spool up motors to counteract it

PID is essential in stabilizing the UAV when the signal for channel override comes from the CNN. As shown in Figure 4.8, using the channel overrides without PID regulation results in the UAV moving into an undesired position or correct position but with a very unstable shift.

Other sensors (UltraSonic, LIDARv3 and Flow camera - facing down) have been used for a finer control. UltraSonic sensors facing all four sides (forward, backwards and both sideways) act as safe triggers when the UAV shifts very far and tends to hit the row of trees. They are controlled by an Atmega microcontroller found in Arduino and is connected to OBC through SPI.

For UAV in-air stabilizion, a specialized high resolution downward pointing camera module is used and a 3-axis gyro that uses the ground texture and visible features to determine aircraft ground velocity. In addition a downward facing Lidar is used in flight modes which to have height control (altitude).

4.4 Restrictions

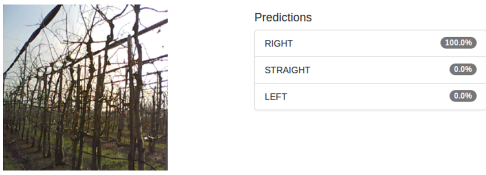
It does not go without mentioning that one of the major restrictions in this research is the uncontrolled environment. Since the navigation relies the most on camera input, light and sun exposure, time of the year, and the vegetative growth of the trees change, make the performance very challenging. In this context, more images under different weather conditions, different seasons and when tree leaves have different colours would make the model fit more scenarios and be more accurate.

4.5 Results

We ran the model through "test" dataset and reached accuracy of more than 85%. In the first models, we trained the winter batch (set 1 of images) and autumn batch (set 2 of images) separately, and tested images from the second dataset in the first model the very low accuracy, under 20%. And the same vice versa. The third model, that runs on all images together, reaches an accuracy greater than 76%. The testing set compromised of well over 200 images that were manually separated from DIGITS data frame. In cases where the front path is not visible in the image, the model gives equal (or almost equal) confidence score to both RIGHT and LEFT, as, when path is not in the frame, they are almost identical. To obviate this, PID on the OBC would then make a decision as to where to go.

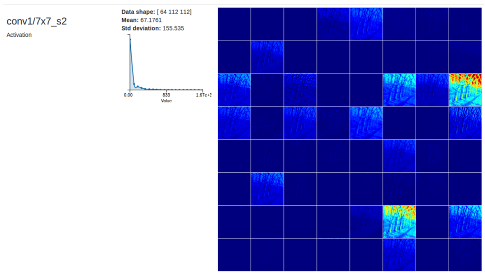
(A) LEFT

(B) STRAIGHT

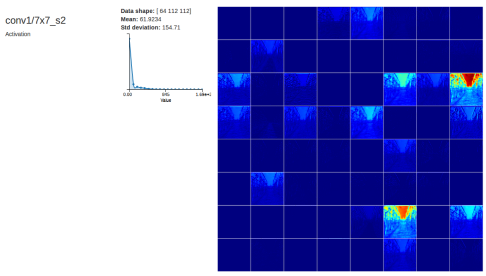


(C) RIGHT

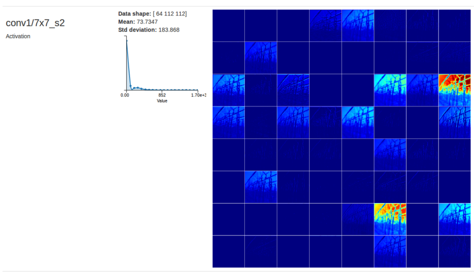
FIGURE 4.9: Predictions per class



(A) LEFT



(B) STRAIGHT



(C) RIGHT

FIGURE 4.10: Activations of last convolutional layer

4.6 Conclusions

Using deep learning and convolutional neural networks makes it possible for UAVs and other mobile robots to navigate paths and spaces not explicitly programmed. In orchard environments, where the top of the orchard is usually covered with hail nets, it is impossible to assess trees and production from above. Our work shows the possibility of using a UAV inside the orchard. Thiw will facilitate handeling specific sensors for precision measurement of fruit quality, fruit size estimation, tracking maturity of the fruits, controlling for diseases. Qualitative and quantitative results computed on large dataset used here show that this approach already yields an accuracy comparable to the accuracy of humans that are tested on the same image classification task.

Since the navigation direction are a moving average of many frames per second, this makes the model quite robust to even small errors or when misclassification occurs. In our case, one can argue that inference time is more important than accuracy.

A better model, taking into consideration all the restrictions mentioned, will contribute to a robustness of classifications (in turn better navigation). In this research we relied strongly on transferring the data from OBC to GCC for classification. By using a better on-board computer better results would be gained as this transfer would be eliminated. A NVIDIA's Jetson platform or Intel's Movidious Compute Stick would run the model locally (in the edge, no need to relay on GCC), and would make the platform more reliable and more responsive.

References

Aubert, Benoit A., Andreas Schroeder, and Jonathan Grimaudo (2012). "IT as enabler of sustainable farming: An empirical analysis of farmers adoption decision of precision agriculture technology". In: *Decision Support Systems* 54.1, pp. 510–520. DOI: 10.1016/j.dss.2012.07.002. URL: https://doi.org/10.1016%2Fj.dss.2012.07.002.

Basu, Subhajit et al. (2018). "Legal framework for small autonomous agricultural robots". In: *AI & SOCIETY*. DOI: 10.1007/s00146-018-0846-4. URL: https://doi.org/10.1007%2Fs00146-018-0846-4.

Bechar, A., J. Meyer, and Y. Edan (2009). "An Objective Function to Evaluate Performance of Human–Robot Collaboration in Target Recognition Tasks". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 39.6, pp. 611–620. DOI: 10.1109/tsmcc.2009.2020174. URL: https://doi.org/10.1109%2Ftsmcc.2009.2020174.

Bechar, Avital and Clément Vigneault (2016). “Agricultural robots for field operations: Concepts and components”. In: *Biosystems Engineering* 149, pp. 94–111. DOI: 10.1016/j.biosystemseng.2016.06.014. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.06.014.

– (2017). “Agricultural robots for field operations. Part 2: Operations and systems”. In: *Biosystems Engineering* 153, pp. 110–128. DOI: 10.1016/j.biosystemseng.2016.11.004. URL: https://doi.org/10.1016%2Fj.biosystemseng.2016.11.004.

Colbert, Amy, Nick Yee, and Gerard George (2016). “The Digital Workforce and the Workplace of the Future”. In: *Academy of Management Journal* 59.3, pp. 731–739. DOI: 10.5465/amj.2016.4003. eprint: https://journals.aom.org/doi/pdf/10.5465/amj.2016.4003. URL: https://journals.aom.org/doi/abs/10.5465/amj.2016.4003.

Eizicovits, Danny and Sigal Berman (2014). “Efficient sensory-grounded grasp pose quality mapping for gripper design and online grasp planning”. In: *Robotics and Autonomous Systems* 62.8, pp. 1208–1219. DOI: 10.1016/j.robot.2014.03.011. URL: https://doi.org/10.1016%2Fj.robot.2014.03.011.

Emmi, Luis et al. (2014). “New Trends in Robotics for Agriculture: Integration and Assessment of a Real Fleet of Robots”. In: *The Scientific World Journal* 2014, pp. 1–21. DOI: 10.1155/2014/404059. URL: https://doi.org/10.1155%2F2014%2F404059.

Fröder, Christoph (2014). *Employment in European Agriculture: Labour Costs, Flexibility and Contractual Aspects*. Institut der deutschen Wirtschaft Köln, Geopa-Copa.

Giusti, Alessandro et al. (2016). “A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots”. In: *IEEE Robotics and Automation Letters* 1.2, pp. 661–667. DOI: 10.1109/lra.2015.2509024. URL: https://doi.org/10.1109%2Flra.2015.2509024.

Harrison, Peter and Joseph Wolyniak (2015). “The History of ‘Transhumanism’”. In: *Notes and Queries* 62.3, pp. 465–467. DOI: 10.1093/notesj/gjv080. URL: https://doi.org/10.1093%2Fnotesj%2Fgjv080.

Hiremath, Santosh A. et al. (2014). “Laser range finder model for autonomous navigation of a robot in a maize field using a particle filter”. In: *Computers and Electronics in Agriculture* 100, pp. 41–50. DOI: 10.1016/j.compag.2013.10.005. URL: https://doi.org/10.1016%2Fj.compag.2013.10.005.

Lima, Yuri Oliveira de and Jano Moreira de Souza (2017). “The future of work: Insights for CSCW”. In: *2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE. DOI: 10.1109/cscwd.2017.8066668. URL: https://doi.org/10.1109%2Fcscwd.2017.8066668.

Mulla, David J. (2013). “Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps”. In: *Biosystems Engineering* 114.4, pp. 358–371. DOI: 10.1016/j.biosystemseng.2012.08.009. URL: https://doi.org/10.1016%2Fj.biosystemseng.2012.08.009.

Noguchi, Noboru et al. (2001). "Autonomous Agricultural Tractor with an Intelligent Navigation System". In: *IFAC Proceedings Volumes* 34.19, pp. 197–202. DOI: 10.1016/s1474-6670(17)33136-1. URL: https://doi.org/10.1016%2Fs1474-6670%2817%2933136-1.

Oren, Y., A. Bechar, and Y. Edan (2011). "Performance analysis of a human–robot collaborative target recognition system". In: *Robotica* 30.05, pp. 813–826. DOI: 10.1017/s0263574711001020. URL: https://doi.org/10.1017%2Fs0263574711001020.

Pérez-Ruiz, M. et al. (2012). "Automatic GPS-based intra-row weed knife control system for transplanted row crops". In: *Computers and Electronics in Agriculture* 80, pp. 41–49. DOI: 10.1016/j.compag.2011.10.006. URL: https://doi.org/10.1016%2Fj.compag.2011.10.006.

Reid, J. and S. Searcy (1987). "Vision-based guidance of an agriculture tractor". In: *IEEE Control Systems Magazine* 7.2, pp. 39–43. DOI: 10.1109/mcs.1987.1105271. URL: https://doi.org/10.1109%2Fmcs.1987.1105271.

Reid, John F. et al. (2000). "Agricultural automatic guidance research in North America". In: *Computers and Electronics in Agriculture* 25.1-2, pp. 155–167. DOI: 10.1016/s0168-1699(99)00061-7. URL: https://doi.org/10.1016%2Fs0168-1699%2899%2900061-7.

Chapter 5

Conclusions

Technological innovation propels society forward in leaps and bounds, and one area that this is particularly true is in agriculture and farming. In the not-so distant future, highly-advanced drones and tractors, as well as other technologies, will be running farms and putting food in the mouths of billions. Spraying, mowing, watering, fertilizing, pruning, harvesting are all computer-controlled. Even the soil they grow in is monitored to within an inch of its life. Farms, then, are becoming like factories: tightly controlled operations for turning out reliable products, immune as far as possible from the vagaries of nature. The third industrial revolution was defined by the pioneering of automation. What separates this fourth iteration is that not only is the technology automated, but it is interconnected and has the capacity for machine learning and artificial intelligence.

Monitoring huge fields of crops is the main task of growers. New sensor and geo-mapping technologies are allowing growers to get a much higher level of data about their crops than they have in the past. Ground robots and drones provide a way to collect this data autonomously. Drones can measure stress on plants, irrigation water management and plant population counts, and as innovation drives improvements, the potential to pinpoint problem areas and make decisions only increases. Ground based robots, provide even more detailed monitoring as they are able to get closer to the crops, and can also be used for other tasks like weeding and fertilizing. The data collected can provide growers insight into what different areas of their farms need. Rather than using a one-size-fits-all approach to applying chemical thinning, for example, growers can put resources where they are needed most and save on areas that may need less management. The future of autonomous systems in precision agriculture comes down to growers being ready and willing to try out the technology for themselves. Regulation will continue to evolve and new advancements will keep changing the paradigm around what those new technologies can do. Getting involved now helps growers acquire an understanding around the potential autonomous systems posses and also allows them to guide and determine their own way forward.

In this study, we showed a potentially autonomous pipeline for an early cropload estimation for fruit trees. Cropload estimation is essential for efficient orchard management starting with optimizing chemical thinning, planning the labor force, harvest equipment, and vehicles for transportation of fruit from field to the packing plant. Management of cropload is a balancing act between reducing cropload sufficiently to achieve optimum fruit size and adequate return bloom and not reducing yield excessively. And since there are not any accurate cropload estimation systems, growers tend to perform insufficient thinning, preferring to be on the safe side rather than causing irreversibly low yield. This has significant repercussions later in the season when hand thinning is required to compensate for the insufficient chemical thinning. By using Artificial Intelligence models, we have created this pipeline where an autonomous UAV navigates the orchards, under anti-hail nets, takes images from the trees and estimates the cropload.

Today, more and more agricultural processes are becoming automated as precision farming technology becomes increasingly sophisticated. As precision agriculture advances, fleets/swarm of smaller, self-driving tractors will replace the colossal machinery on which many growers now rely. Steps are currently being made to develop technologies that will enable the automation of individual tasks and general purpose tasks that will empower growers to run operations in a fulfilling and efficient way.

Chapter 6

Future work

Understanding growth patterns of various types of fruit and different cultivars of a given crop is another area of research that affects the potential of machine vision system. For example, stem length and fruit shape may vary substantially between various apple cultivars. If a good understanding of such parameters is available, machine vision system can use such knowledge to improve its effectiveness in detection and/or reconstruction the whole tree canopy. In addition, future research and development in machine vision system should focus on detecting precise features from tree canopy training system and relate the fruit distribution for better understanding of occlusion.

One of the most important factors on any deep neural network is the amount of data together with the quality. More images on different stages of growth, images of different training systems and different cultivars as well as different fruit species would certainly make the platform more appealing and robust. Further studies and data collection, annotation and training are foreseen in the future, by creating a whole dataset of many fruit species.

Fruit detection for yield estimation presents a challenge that is different from fruit detection for robotic harvesting, and has some unique characteristics. First, in order to minimize problems related to image registration and to enhance performance, each image should contain a large portion of the tree, so that fruit appear as rather small objects in the image. Also, unlike for robotic harvesting, for UAV and in-line robots it is not possible to bring the camera closer to a specific suspected fruit, to illuminate it or to improve its image in any way. The envisioned system will consist of a geo-referenced systems equipped with one or several cameras mounted sideways that will capture images on-the-fly while traveling at low speed.

By using better, more powerfully and yet smaller on-board computer, with dedicated artificial intelligence processing units (APUs) will help one making the whole platforms (recognition, auto-navigation, cropload estimation) into a single unit. Or the other way around, the OBC would server only as intermediary to send large amounts of data to the cloud for further processing. In this case, the trans-receivers for data need to be very fast and very responsive so the sigal from the UAV to the cloud and back should be in matter of seconds. This is what computing-on-the-edge is all about, and can be a very content solution.

Acknowledgements

Generally speaking, most of acknowledgements I've seen are very robotic and kinda look like templates. Well, I will try not to be that robotic, and this is definitely no template (even chose different fonts for this). At the end, this part is where i thank all people dear to me that aided me not just for thesis, but were there whenever i needed something, whatever that might have been.

Not because i have to put first my supervisor, but actually i would like to thank him the most. My thesis is particular, its more technical rather physiological. My supervisor, Prof. Corelli, not only did he trust on what i am doing, but every time he had a chance he would come up to me and discuss about the progress in AI and Robotics is influencing Agriculture and see how and what are the things that we are contributing to this. That was extremely important for me, it was reassuring enough to know that i am doing something useful and contributing to something bigger than myself.

It seems that i am kinda following a template, because now i want to thank my cosupervisor. Apart from all the professional help i asked from Gigi, I always found him amazing. He does so much with very little time. He is able to manage a very successful career, finds time every early morning for his hobbies (for 10 years he has ran more than 20.000km and biked 100.000km)and enjoys at maximum his time with his amazing family. I always get inspired from him, for being always young and full of energy and passion for what he does. Same, i would like to thank Bru, for always being the go-to person if i ever needed scientific advice or approach for anything (and for asking me to give classes to her students so i can get more teaching/presenting experience).

I would like to thank my father for always pushing me to do more in life (and for updating me about politics sometimes, and yelling to shave my beard). My mother for always asking me if i slept well and pushing me eating better food (and nagging me there is a girl in my life already). My sisters Zana, Fitesa and Fjolla for being always awesome and inspiring with their successful life, family and careers.

Whenever i felt the need for beers or for a "bro", i would give a call to Alex (the one bro to rule them all) and go hang downtown for beers, and that is just awesome. Whenever i was in need for anything, i knew to relay on "super" Giulio (no cryptonite had to around though). Talking about space, i would like to thank Gianmarco for our common interests about: more or less everything.

And, lastly but not the least, i would like to thank Flavia for being so amazing with me. She made me feel comfortable and made me call home a country not home of mine.

‡‡‡