UNIVERSITÀ DEGLI STUDI "MEDITERRANEA" DI REGGIO CALABRIA
DIPARTIMENTO DI AGRARIA
Dottorato di Ricerca in
Scienze Agrarie, Alimentari e Forestali
Curriculum Produzioni Agrarie
Ciclo XXXVI, 2021/24 - SSD: AGR/09

Dr. Matteo Anello

Development of a mobile web application to assess the essential oil content of bergamot in the field from smartphone images using convolutional neural networks.

PH.D. THESIS

Tutor
Prof. Giuseppe ZIMBALATTI

Co-Tutor
Prof. Bruno BERNARDI

Ph.D. Coordinator
Prof. Leonardo SCHENA

Reggio Calabria 15/01/2024

CONTENTS

ABSTRACT

ABSTRACT (ITALIAN VERSION)

INTRODUCTION 1

CHAPTER 1 4

BERGAMOT, MOBILE APPS AND ARTIFICIAL INTELLIGENCE: A SINERGIC VISION FOR AGRICULTURE 4

1.1 An overview of bergamot citrus fruit 4

1.2 The use of mobile apps in agriculture 8

1.3 Artificial intelligence systems and machine/deep learning algorithms 11

1.4 Deep learning techniques and convolutional neural networks (CNN) 16

CHAPTER 2 20

MATERIAL AND METHODS 20

2.1 ***Field sampling and camera calibration*** 20

2.2 Essential oil extraction 27

2.3 Convolutional neural network architecture 28

2.3.1 CNN models built for 'Fantastico' during the year 2021/2022 28

2.3.1.1 Training and validation datasets 30

2.3.1.2 Models 31

2.3.1.3 Transfer learning models 34

2.3.1.4 Training parameters 36

2.3.2 CNN models built for 'Fantastico' and 'Femminello' during the year 2022/2023 37

2.3.2.1 Training and validation datasets 41

2.3.2.2 Models 42

2.3.2.3 Transfer learning models 44

2.3.2.4 Regression models 46

2.3.3 CNN models built to join data from 2021 to 2023 48

2.3.3.1 Training and validation datasets 50

2.3.3.2 Models 51

2.3.3.3 Transfer learning models 53

2.3.3.4 Regression models 55

CHAPTER 3 56

RESULTS AND DISCUSSION 56

3.1 CNN Models performance on 'Fantastico' during year 2021/2022 56
3.2 CNN models performance on 'Fantastico' and 'Femminello' during the year 2022/2023 61
3.3 CNN models performance on data from 2021 to 2023 67
3.4 Software implementation strategy and convenience 74
CHAPTER 4 75
MOBILE APPLICATION PROGRAMMING AND DEVELOPMENT 75
4.1 Interface implementation, programming, and development 75
CHAPTER 5 80
CONCLUSIONS AND FUTURE PERSPECTIVES 80
6 ACKNOWLEDGEMENTS 82
7 REFERENCES 83

ABSTRACT

The Essential oil (EO) extracted from bergamot skin (Citrus bergamia, Risso et Poiteau) is highly appreciated for its use in perfumery and gastronomy. Notably, 90% of the EO bergamot production is concentrated in the Province of Reggio Calabria (Southern Italy) under a protected designation of origin (PDO). Early yield estimation of the essential oil content of fruits is fundamental to helping growers make farming decisions, especially regarding harvesting operations. The application of advanced modelling techniques based on artificial intelligence and digital device technology can contribute to this goal. This thesis represents a three-year study dedicated to the development of an Android application aimed at determining the essential oil content in Bergamot fruits. The entire process is divided into three distinctive phases, which cover fruit sampling and oil extraction, the use of deep learning techniques, in particular convolutional neural networks (CNN), and the practical implementation of the Android application. The first phase involved the systematic sampling of Bergamot fruits and the extraction of the essential oil, establishing the basis for the subsequent analysis. In the second phase, advanced deep learning techniques, mainly CNNs, were used to process the collected data. This phase involved the development of a robust algorithm capable of accurately predicting essential oil content based solely on fruit color. The third and final year was dedicated to the practical implementation of this algorithm within an Android application. Using Android Studio, an accessible and easy-to-use application was created, allowing users to quickly and conveniently exploit the results obtained from the data analysis phase. The methodological approach adopted in this work not only offers an effective means to determine the essential oil content in Bergamot fruits through color recognition, but also represents a tangible demonstration of the practical application of deep learning techniques in the agricultural sector, together with the development of advanced applications for Android devices. This study not only contributes to the growing body of knowledge in the field of artificial intelligence and agriculture, but also constitutes a significant step towards the convergence of these disciplines, paving the way for new

perspectives in agricultural management and simplified access to solutions innovative technologies.

ABSTRACT (ITALIAN VERSION)

L'Olio Essenziale (EO) estratto dalla buccia di bergamotto (Citrus bergamia, Risso et Poiteau) è molto apprezzato per il suo utilizzo in profumeria e gastronomia. In particolare, il 90% della produzione di bergamotto OE è concentrato nella provincia di Reggio Calabria (Italia meridionale) sotto una denominazione di origine protetta (DOP). La stima anticipata della resa del contenuto di olio essenziale dei frutti è fondamentale per aiutare i coltivatori a prendere decisioni agricole, in particolare per quanto riguarda le operazioni di raccolta. L'applicazione di tecniche di modellazione avanzate basate sull'intelligenza artificiale e sulla tecnologia dei dispositivi digitali può contribuire a questo obiettivo. Questa tesi rappresenta uno studio triennale dedicato allo sviluppo di un'applicazione Android volta a determinare il contenuto di olio essenziale nei frutti di Bergamotto. L'intero processo è suddiviso in tre fasi distinte, che riguardano il campionamento dei frutti e l'estrazione dell'olio, l'utilizzo di tecniche di deep learning, in particolare le reti neurali convoluzionali (CNN), e l'implementazione pratica dell'applicazione Android. La prima fase ha previsto il campionamento sistematico dei frutti di Bergamotto e l'estrazione dell'olio essenziale, ponendo le basi per le successive analisi. Nella seconda fase sono state utilizzate tecniche avanzate di deep learning, principalmente CNN, per elaborare i dati raccolti. Questa fase ha comportato lo sviluppo di un robusto algoritmo in grado di prevedere con precisione il contenuto di olio essenziale basandosi esclusivamente sul colore del frutto. Il terzo ed ultimo anno è stato dedicato all'implementazione pratica di questo algoritmo all'interno di un'applicazione Android. Utilizzando Android Studio è stata creata un'applicazione accessibile e facile da usare, che consente agli utenti di sfruttare in modo rapido e comodo i risultati ottenuti dalla fase di analisi dei dati. L'approccio metodologico adottato in questo lavoro non solo offre un mezzo efficace per determinare il contenuto di olio essenziale nei frutti di Bergamotto attraverso il riconoscimento del colore, ma rappresenta anche una dimostrazione tangibile dell'applicazione pratica di tecniche di deep learning nel settore agricolo, insieme allo sviluppo di applicazioni avanzate per dispositivi Android. Questo studio non solo contribuisce al crescente corpus di conoscenze nel campo dell'intelligenza artificiale e

dell’agricoltura, ma costituisce anche un passo significativo verso la convergenza di queste discipline, aprendo la strada a nuove prospettive nella gestione agricola e all’accesso semplificato a soluzioni tecnologiche innovative.

Keywords: *cv. Fantastico, cv. Femminello, smartphone camera, distillation, deep learning, Convolutional neural networks (CNN), Android studio.*

INTRODUCTION

Bergamot (*Citrus bergamia*, Risso et Poiteau) is an evergreen plant of great interest for the essential oil (EO) contained in the flavedo and, with a growing trend, for the extraction of the juice, which is rich in antioxidants (Di Donna et al., 2020);(Giuffrè et al., 2019). Today, the principal cultivation area is a narrow coastal strip of about 150 km in the Reggio Calabria Province (Southern Italy), comprising 1000-400 ha. 90% of EO production is performed in this growing area (Amato et al., 2013), protected under the designation of origin (PDO) Bergamot of Reggio Calabria. The most important cultivars are 'Femminello', 'Castagnaro', and 'Fantastico', the latter being the most widespread (Dugo & Bonaccorsi, 2013). Depending on the geographical area, the hesperidium produced between November and March is slightly flattened and has a subglobose to pyriform shape. The peel is up to 6 mm thick and characterized by a shiny green at the beginning of the productive season and turning to pale yellow when the fruit is ripe. The peel contains numerous glands enclosing the EO (Quirino et al., 2022). The oil presents a greenish color at the beginning of the productive season and turns brownish-yellow at the end of the season (Navarra et al., 2015). EO is considered a natural product obtained from the raw material of plant origin using several methods, including steam distillation, mechanical processes from the epicarp, or dry distillation, after separation of the aqueous phase from any physical processes (ISO 9235:2013). It is highly appreciated in the perfumery, cosmetics, and food industries for its intense fragrance and freshness (Gioffrè et al., 2020). The fundamental procedures for EO isolation are distillation, selective solvent extraction, and mechanical expression (Yang et al., 2013). Bergamot interest and production are increasing, highlighted by the fact that many abandoned areas dedicated to bergamot cultivation are being recovered (Sicari et al., 2016). In this context, EO yield estimation using fast and modern technologies becomes fundamental for helping growers with farming decisions, especially for anticipating harvesting operations (Anderson et al., 2021), as identifying the optimal ripening degree according to essential oil content becomes fundamental. To this view, Hunter (2010) reported that essential oil spurts

out with a force that decreases with the length of fruit maturation, so it is easier to express fresh and immature fruits. Defining the finest harvesting period has been traditionally done based on the visual appreciation of the fruit color, which requires extensive grower experience but is very subjective and depends on personal criteria. An alternative is the use of technology to obtain and analyze fruit images to make objective decisions about the best harvesting moment. There are different sensors capable of acquiring images in the field, such as thermal (Cohen et al., 2022), multi or hyperspectral (Munera et al., 2021), stereo (Lohar et al., 2021), or color (Gonzalez-Gonzalez et al., 2021) cameras. Color cameras are the most used for price, simplicity, and similarity with human perception. Today, almost all smartphones are equipped with several cameras (Blahnik and Schindelbeck, 2021). Mobile devices incorporating powerful color cameras have become a valuable and affordable technology. The main advantage is accessibility, as almost any farmer can have a device in their pocket. But, above all, smartphones are powerful computing systems equipped with high-quality and resolution cameras, so we can understand how it has become a natural trend to push new analysis and elaboration methods toward mobile applications (Wang, 2018). However, in recent years, artificial intelligence (AI) has emerged as a powerful technique to emulate the human visual system to interpret and classify image content (Elgendy, 2020). A broad area of AI is closely related to image classification also in the agriculture field, where enormous progress has been made in many domains, including pest and disease classification, fruit recognition, precision agriculture, and decision support systems (Kujawa & Niedbała, 2021. Fazari et al., 2021. Jia et al., 2020; Kamilaris & Prenafeta-Boldú, 2018). Image classification is strictly related to the science of Convolutional Neural Networks (CNNs), a class of Artificial Neural Networks (ANNs) that is generally intended to perform visual analysis on computers (Carniage et al., 2020). According to Fajardo Muñoz et al., (2023) and El-Attar et al., (2020), CNNs can also contribute to solving several issues related to biological compounds extracted from different parts of aromatic plants such as flowers and leaves, such as predicting essential oil extraction yield. As reported by Carniage et al., (2020) there are ten main types of plants capable of producing useful essential oils that will be classified by the CNN, and one of the most important is

bergamot (*Citrus bergamia,* Risso et Poiteau). However, there are no previous studies for estimating EO content in the citrus peel directly in the field using automated and non-destructive means. Hence, this work aims to quantify for the first time bergamot EO in the field by i) image acquisition of the fruits using a smartphone camera coupled with a customized low-cost portable inspection chamber; ii) essential oil extraction in the laboratory by hydro-steam-distillation; iii) set fruit color categories related to the EO content, and iv) classify the fruit attending the color and, therefore, the EO content.

CHAPTER 1

BERGAMOT, MOBILE APPS AND ARTIFICIAL INTELLIGENCE: A SINERGIC VISION FOR AGRICULTURE

1.1 An overview of bergamot citrus fruit

Bergamot (*Citrus bergamia* Risso) is a plant that grows wild in southern Italy with a production of about 25,000 tonnes per year (Scerra et al., 2021), it is a citrus fruit that varies in color from green to yellow and is often used to produce perfumes and cosmetics. The herbal formulations derived from *Citrus bergamia* include bergamot essential oil (BEO) and bergamot juice (BJ) (Mannucci et al., 2017). According to Farmacopea Ufficiale Italiana (1991), bergamot essential oil (BEO) is obtained by cold pressing the rind and part of the middle rind of fresh fruit. This fruit is a citrus fruit classified as *Citrus bergamia*, Risso of the genus Rutacee Citrus, growing exclusively in Calabria (Italy), along the southeast coast, in the region of Reggio Calabria. This citrus fruit is grown for its essential oil, a product that is in high demand in the perfumery and cosmetic industries but is also used in the pharmaceutical, food, and confectionery industries. BEO consists of a volatile fraction (93-96% of the total) and a non-volatile fraction (4-7% of the total) containing coumarins and psoralens [such as bergapten (5-methoxypsoralen) and bergamottin (5-geranyloxypsoralen)](Bagetta et al., 2010). The definition of the International Organization for Standardization (ISO) (ISO 2020 document ISO 9235: 2013-2.11) states that EO is "a product obtained from natural plant raw materials by steam distillation, from citrus epicarps by mechanical methods or by dry distillation after separation of the water phase - if present - by physical with methods. It is about three production methods, one of which is reserved only for citrus EO, which is therefore clearly separated from other EO in terms of production method and therefore chemical composition. Distilled EO contains only volatile molecules, while cold-pressed EO can contain much heavier molecules and different chemical compositions than usually distilled EO. Another characteristic of citrus EO is that, from an economic point of view, the EO is in many cases a secondary product, while the fruit juice is the main product and the fruit can be used for other products (sugar peels, "cells", albedo, etc.) This applies to lemons, sweet

oranges, grapefruits, and tangerines, but not to bergamot and bitter orange, which usually produce little or no fruit juice. This fact partly explains the higher costs of both EOs. Due to its economic advantages, bergamot is particularly important for the cultivation area. Three cultivars (cv) (genotypes) are known: Castagnaro, Fantastico, and Femminello (Giuffrè, 2019). It belongs to the Rutaceae family, subfamily Esperidea, and has been common in the Mediterranean region for centuries. Its botanical and geographical origin is still uncertain; bergamot probably originated in Calabria through mutations of other citrus species or arrived there from Berga (hence the name bergamot) or according to other theories arrived from the Antilles, Greece, and the Canary Islands. The history of bergamot cultivation in Calabria is interesting from an ethnobotanical point of view, since various cultures (Greeks, Byzantines, Arabs) and the awareness of the ‘‘subsistence’’, due to the poorly developed road network and the isolation of villages (Di et al., 2013). BEO has been used in the pharmaceutical industry for its antibacterial and antioxidant properties, and in food as a flavoring agent (Xing et al., 2019). Over the decades, both the extraction of bergamot essential oil and analytical methods have changed, allowing better identification of the compounds present in bergamot essential oil (Gioffrè et al., 2021). Bergamot differs from other citrus fruits not only in its nutritional value but also in its particularly high content of flavonoids and the chemical composition of BEO that has been extensively studied. It is well known that BEO contains several bioactive molecules with potential health benefits. (Lamiquiz-moneo et al., 2020; Navarra et al., 2015). The literature suggests that bergamot plays an important role in various areas of interest such as the nervous system, cardiovascular health, inflammation, diabetes, bone metabolism, and skin. Preliminary results indicate that BEO extract can reduce cardiovascular disease, anxiety, and stress, improve cognitive function, and improve sleep (Perna et al., 2019). The characterization and confirmation of the biological effects of bergamot led to the use and research of bergamot phytocomplexes in various fields, which increased the demand for bergamot production. Calabrian growers therefore reassessed its potential as an important economic product in the province of Reggio Calabria, home to almost all the world's production of bergamot. Growing scientific interest led to many experimental and clinical studies that improved the therapeutic

profile of bergamot, revealing its wider range of uses and more significant roles in human health (Adorisio et al., 2023). Therefore, the cultivation of bergamot is an important pillar of the positive economic reality of the Calabria region. It needs protection because research on its biological functions continues and a systematic and rigorous approach to the study of potential phytotherapeutic agents is an achievement of the last decades (Maruca et al., 2017). An important by-product of the bergamot chain is pulp, which can be used as an energy supplement in ruminant diets, especially in bergamot production areas. As a result, the disposal of bergamot juice and pulp is a major economic and environmental challenge for the essential oil processing industry (By-products et al., 2023). However, the quality of bergamot must be preserved, as climate change also threatens this culture, in fact as shown by (Cautela et al., 2021), to mitigate the effects of abiotic stress on the quality of bergamot essential oil, it will be necessary to intervene by rationally increasing irrigation in bergamot groves, especially during the summer. Therefore, according to (Ukhurebor & Aidonojie, 2021) strategies like climate-smart agriculture (CSA) practices, sustainable water management, efficient waste management and recycling, safeguarding food resources, climate protection measures, and innovative machinery for analyzing climate conditions in agricultural processes, including the Farm Input Tracker (FIT), are crucial elements that require revitalization, evolution, and development. Clinical studies have also shown that bergamot oil is effective in addressing and alleviating psychological issues that arise post-COVID-19, as shown in a study conducted by (Wakui et al., 2023), its use before bedtime and upon awakening has been found to enhance feelings of alertness upon rising, improve sleep duration, and alleviate depressed mood. In comparison to a placebo, the study indicates that the aroma of bergamot essential oil effectively mitigates psychological stress. The findings suggest that incorporating bergamot essential oil into one's routine is a safe and cost-effective method to enhance both sleep quality and morning wakefulness. The cultivation and production of bergamot have become a developmental opportunity for the rural areas of the province of Reggio Calabria, renowned for the cultivation of this citrus fruit. The promotion of cultivation practices, coupled with the sustainable use of all its by-products, including waste materials, could serve as a catalyzing element for

economic and sustainable development in the region. From both a territorial and societal perspective, bergamot cultivation holds significant importance. It plays a pivotal role in creating economic opportunities that enable farmers to maintain a sustainable income. Presently, the substantial surge in the demand for fresh bergamot consumption has prompted an effective resolution for commercial distinctiveness, resulting in an expansion of the cultivated area (Campolo, 2014; Strano et al., 2017).

1.2 The use of mobile apps in agriculture

Information and communication technology (ICT) is a key component of corporate operations and offers a solid foundation for solving many common issues. Users of today desire immediate access to information that can be used. The use of mobile technologies has so significantly increased (Sopegno et al., 2016). A wide range of smartphone apps have been created in recent years and are available on the market to help in agriculture. Some mobile applications were created expressly to offer farmers information services (Mane & Kulkarni, 2019). A new intelligent intermediary layer between people and systems is created by smartphones driven by cutting-edge sensor technologies, artificial intelligence (AI), and machine learning (ML) algorithms to tackle complicated problems or even numerous common problems effectively. More and more users desire immediate access to relevant data and information. Because of this, the use of mobile communications has skyrocketed during the last years (Mendes et al., 2020). A smartphone is a device that is used to make phone conversations and even can send and receive email, connect to Wi-Fi and a modem, access the internet, open Office documents, operate simply with a touch screen, and most importantly, run customized software. One further crucial feature of smartphones is the user interface. A smartphone's touchscreen capabilities, ability to zoom in and out utilizing straightforward interface buttons, menus, and forms, and compatibility with a qwerty keyboard make them simple and easy to use for persons who are unfamiliar with ICT technology or who are even not sufficiently educated (Patel & Patel, 2016). New generation (G) mobile technology, like 2G and 3G smartphones, was first introduced to the world. A mobile device could only make calls, leave voicemails, and send text messages before smartphones. Smartphones can perform more than just the bare minimum thanks to improved technologies. With the inclusion of a hardware sensor, smartphones can take high-resolution photos and videos. With mobile data subscriptions using 2G, 3G, and 4G networks, wireless data access is the standout feature of every smartphone. Our lives have changed and become easier as a result of the smartphone being a part of our way of life. Due to the increased demand for smartphones, mobile applications such as step counters, are frequently utilized in daily life (Science, 2020). However, one of the burgeoning

areas in enhancing agricultural quality and rural development is the use of mobile apps for agriculture and rural development when they are employed to inform farmers and assist them in raising productivity (Series & Science, 2021). Mobile apps were originally designed to perform basic computer program tasks such as email, web browsing, calendars, contacts, and weather forecasts. Due to the current high demand for new mobile products and services, both businesses and organizations are developing mobile apps for commerce, banking, healthcare, and tourism to meet the specific needs of different business sectors. are compelled to do so. Above all, the agricultural sector is an important pillar of the economy and as a business, it covers the food needs of the world population. However, the development of mobile applications in agriculture is limited compared to other industries (Costopoulou et al., 2016). Agriculture plays an important role in economic and social development in most underdeveloped countries. This is due to food safety and human health problems, the demand to increase yields and improve food. The challenges of agricultural development in any country are great, not only to meet the growing demand for food but also to reduce poverty and malnutrition (Milovanović, 2014). Mobile phones have a multifaceted positive impact on sustainable poverty reduction, and accessibility is seen as a key challenge in realizing their full potential. Therefore, there is a great opportunity to use ICT to improve the dissemination of agricultural information received by farmers (Mandi & Patnaik, 2019). When applying these new technologies, the data is useless on its own, only numbers or images, so the challenge in getting data from crops is to deliver something of consistent value. Companies that choose to leverage technology in some way offer valuable benefits such as: saving money and effort, increasing production or reducing costs with minimal effort, or producing higher quality food in a more environmentally friendly way (Saiz-rubio, 2020). As a result, intelligent agriculture, characterized by intelligence, connectivity, and digitization, has emerged, becoming a new direction for modern agricultural development. Combining modern technology with agriculture to avoid problems such as pollution and land degradation. Developed countries, especially the United States, began researching smart agriculture in the 1980s, and over the past two decades, many scientific companies have been established to provide smart farming solutions (Xia

et al., 2023). However, new technologies and software designed to digitize agribusiness alone cannot solve all the challenges of digital transformation in the supply chain. Infrastructure, continuing education and qualification, a sufficient structural and legislative operating environment, and the desire to adopt new technologies are also important. For Farming 4.0 to work, modern telecommunications infrastructure is essential in rural areas. In addition, the ability to use structured and unstructured data throughout the agricultural supply chain will prove essential for the successful transformation of existing agricultural processes into agriculture in the era of Industrie 4.0. (Braun et al., 2023). In agriculture as regards citrus fruits, (Cubero et al., 2017) have developed an Android application for mobile devices to estimate the standard Citrus Color Index (CCI) using image processing in mobile devices. This work shows us the possibility of determining the color of fruits using the computer vision system, an area of AI. Artificial intelligence systems have vast automation capabilities thanks to a field of research that replaces human vision with current technologies involving a camera and computer. Machine vision technology-based expert and intelligent systems have been extensively used in agricultural operations in recent years (Dhanya et al., 2022). The large amount of data produced by digital technologies, often called "big data", requires a large storage capacity in addition to editing, analysis, and interpretation (Benos et al., 2021). Therefore, efficient feature extraction methods are needed to obtain accurate data. Recently, techniques based on deep learning (DL) and convolutional neural networks (CNN) have become effective tools for processing large amounts of data (Bernardi et al., 2021).

1.3 Artificial intelligence systems and machine/deep learning algorithms

Artificial intelligence (AI) is the concept of creating intelligent machines that can imitate human thinking and behavior. Machine learning is an application of artificial intelligence that allows machines to learn from data without special programming (Pimenov et al., 2023). The existence of artificial intelligence is increasing, and approximately 70% of companies worldwide plan to adopt artificial intelligence by 2030. Therefore, artificial intelligence is expected to change all aspects of society (Brasse et al., 2023). As proven by (Kumar, 2017) McCarthy invented artificial intelligence in 1956, researching a machine that could think, solve problems, and improve itself as a human. It has key features such as adaptive control, better handling of stored data, and reusability. According to the development, the AI system can perform human actions such as observation, interpretation, reasoning, learning, communication, and decision-making to find a solution to a given problem. From the beginning, the artificial intelligence system has been set various development goals that expand its applications such as pattern recognition, automation, computer vision, virtual reality, diagnostics, image processing, non-linear control, robotics, automated reasoning, data mining, process engineering, intelligent agent and control, manufacturing, etc. Many researchers focus on the field of artificial intelligence, which makes the study of artificial intelligence rich and diverse. Research areas in artificial intelligence include search algorithms, data graphs, natural language processing, expert systems, evolutionary algorithms, machine learning (ML), and deep learning (DL) (Ai, 2021). (Guliyev, 2023) examined the rise of data science, machine learning, and artificial intelligence in the Google Trend Index in 24 high-tech countries, showing the trend of data science, machine learning, and artificial intelligence GTI volumes in high-tech countries for the period 2005-2021. Until 2018, AI had an upward trend and a slight decrease after that, while machine learning and data science increased from 2005 to 2019 and slightly decreased after that. According to (Zhang & Lu, 2021), Applications have found widespread use in diverse domains, including automated programming, expert systems, knowledge systems, and intelligent robotics. Artificial intelligence (AI) necessitates not only logical reasoning and emulation but also encompasses the integration of emotional components as an essential facet.

The forthcoming advancement in the realm of AI is poised not only to enhance computers' logical reasoning capabilities but also to endow them with emotional capacities. The potential for machines to exhibit intelligence may soon transcend that of humans. As in several sectors, agriculture represents a dynamic domain characterized by situations that defy generalization, precluding the imposition of universally applicable solutions. The implementation of artificial intelligence (AI) techniques has empowered us to intricately capture the nuanced details inherent in each situation, thereby facilitating the derivation of solutions tailored to the specific parameters of the problem at hand. The progressive evolution of diverse AI techniques is contributing to the resolution of increasingly complex agricultural challenges (Bannerjee et al., 2018). According to (Mahesh, 2020), Machine learning (ML) is used to teach machines to process data more efficiently. After reviewing the data, you may not be able to interpret the information extracted from the data. In this case, we use machine learning. With so much information, the need for machine learning is growing. Many industries use machine learning to extract relevant information. The purpose of machine learning is to learn from data. A lot of research has been done on how to make machines learn on their own without having to program them individually. Many mathematicians and programmers use multiple approaches to solve this problem using large datasets. During this time, many machine learning techniques were developed, with names such as logistic regression, neural networks, decision trees, support vector machines, Kalman filters, and many others. This multidisciplinary effort involved statistics, artificial intelligence, optimization, signal processing, speech, vision, and control theory, as well as the machine learning community itself. In the traditional approach to solving a new machine learning problem, the practitioner must choose an appropriate algorithm or technique from among those known to him and then either use existing software or write his implementation (Binkhonain & Zhao, 2023). Our generation is experiencing an unprecedented combination of vast and increasing amounts of information. advances in computing hardware and reductions in computing, data storage, and transmission costs; advanced algorithms; a lot of open-source software and benchmark issues; and the industry's significant and continued investment in data-driven problem-solving. These advances have in turn generated interest and

progress in the field of machine learning (ML) to extract insights from this data. These learning algorithms can be classified as supervised, semi-supervised, and unsupervised learning depending on the information available to the learning machine (LM) (Brunton et al., 2020). According to (Nasteski, 2018), Machine learning techniques are used to operate multidimensional data occurring in several different application domains. Hence, machine learning algorithms based on the desired result of the algorithm are divided into the following groups:

- Supervised learning - Different algorithms create a function that maps inputs to desired outputs. One standard form of a supervised learning task is a classification problem: the student must learn (to approximate behavior) a function that maps a vector to something in several classes by looking at several examples of input-output functions.
- Unsupervised learning - model inputs: no labeled examples
- Semi-supervised learning - combines both labeled and unlabelled examples to create a suitable function or classifier.
- Reinforcement learning - the algorithm learns a policy on how to act in each situation perception of the world. Every activity has some effect on the environment, and the environment provides feedback that guides the learning algorithm.
- Transduction - Similar to guided learning but does not build direct function instead try to predict new outputs based on training inputs, educational achievements, and new contributions.
- Learning to learn – where the algorithm learns based on its inductive bias based on previous experience. Apart from these groups of machine learning algorithms, they are divided into two general groups, supervised and unsupervised learning.

The development of artificial intelligence and machine learning over the last two decades has led to a significant increase in the number of agricultural projects. Since the beginning of the 21st century, artificial intelligence has been used more and more in this field, and one of the growth indicators is the number of scientific articles on the corresponding topic. However, smart agriculture is essential to meet

the challenges of agricultural productivity, environmental impact, food security, and sustainability as the world's population continues to grow, making a strong increase in food production necessary to keep up with it. time availability and the high quality of restaurants around the world, protect natural ecosystems through sustainable agricultural practices (K. Dokic et al., 2020) (Kamilaris & Prenafeta-boldú, 2018). DL has received much attention in agriculture. One of its applications in agriculture is image recognition, which has overcome many obstacles that limit the rapid development of robotics and mechanized agriculture and farming. These improvements can be seen in many aspects of agriculture, such as plant disease detection, weed control, and plant counting (Zhu et al., 2018). (Bochtis, 2021) in his review of the recent literature identified four general categories for the period 2004–2018. These categories relate to crops, water, soil, and livestock. Regarding plant care, it represented the most articles of all categories (61% of all articles) and was further divided into:

- Yield prediction;
- Disease detection;
- Weed detection;
- Crop recognition;
- Crop quality.

The general categories of water and soil management were found to be less researched, cumulatively accounting for 20% of the total number of publications (10% in each category). Finally, two main subcategories were identified for livestock-related submissions, corresponding to a total of 19% of journal entries:

- Breeding;
- Animal welfare.

(Coulibaly et al., 2022) shows us through a bibliometric analysis how the benefits of digital agriculture are extraordinary. Artificial intelligence and deep learning applications provide many opportunities to strengthen different phases of agriculture to meet different challenges and achieve goals:

1. Agricultural Data Processing: Condition Monitoring of plants and animals is essential for agricultural production.

2. The models make disease detection a feasible process and increase the production potential of healthy plants.

3. Optimal control of agricultural production systems: Control strategies for agricultural production systems are often interdependent the experience or knowledge of the farmer is not considered the physiological state of plants.

4. Plant care practices have reached new heights and it has been achieved it has become quite practical for farmers to manage as few crops as possible a problem.

5. Intelligent Agricultural Machinery: Agricultural production involves many tasks.

6. Deep learning models to effectively make accurate predictions and analyze agricultural data.

7. Management of the agricultural economic system: Agriculture just the result is not enough. There are many other factors to consider, such as prices and quality of agricultural products. Forecasting the prices of agricultural products is very important.

8. Artificial intelligence has significantly changed the weather forecasting system, which plays a key role in agriculture.

Some non-linear transformations are used to model higher-level abstractions of data and are the basis of DL. One of the main advantages of DL is the automatic extraction of features from raw data or feature learning. Generating properties from lower-level components creates properties for higher-level components. Recurrent neural networks (RNNs) and CNNs are two types of DL networks commonly used in agriculture (Albahar, 2023). This type of neural network is widely used in agricultural research, especially CNNs due to their powerful image processing capabilities. DL technology is extensively used in plant and crop classification, including yield prediction, pest control, and disaster monitoring (Craik et al., 2019).

1.4 Deep learning techniques and convolutional neural networks (CNN)

Recently, machine learning (ML) has become very widespread in scientific research and has been used in several applications, including text mining, spam detection, video recommendation, image classification, and multimedia concept retrieval. Among the various ML algorithms, deep learning (DL) is very often used in these applications. Another name for DL is representational learning (RL). The constant emergence of new research in the fields of deep and distributed learning is due to the unpredictable growth of the ability to acquire both information and hardware technologies, e.g. High-Performance Computing (HPC) (Alzubaidi et al., 2021). Deep learning is now considered a core technology of the Fourth Industrial Revolution (4IR or Industry 4.0). Due to its ability to learn from data, DL technology originates from the artificial neural network (ANN) and has become a hot topic in the context of computing and is widely used in various application fields such as healthcare, visual recognition, text analytics, cyber security, agriculture, and many others (Sarker, 2021). Inspired by biological neural networks, ANNs are massively parallel computing systems consisting of a very large number of simple processors with many connections. ANN models try to use some of the "organizing" principles that are supposed to be used by humans. One type of network sees nodes as "artificial neurons". These are called Artificial Neural Networks (ANN). An artificial neuron is a computer model that has been influenced by natural neurons. Since the task of ANNs is information processing, they are mainly used in related fields. A wide variety of ANNs are used to model real-world neural networks and study the behavior and control of animals and machines, but there are also ANNs used in design applications such as pattern recognition, prediction, and data compression. The brain is a very complex organ that controls the entire body. Even the most primitive animal brain has more capabilities than the most advanced computer. Its role is not only to control the physical parts of the body but also more complex functions such as thinking, visualizing, dreaming, imagining, learning, etc., activities that are not physically described. An artificial thinking machine still exceeds the capabilities of the most advanced supercomputers. (Journal et al., 2017) (Gupta, 2013). The basic idea of such networks is (somewhat) inspired by the functioning of the biological nervous system, processing data and information to

learn and generate knowledge. The central element of this idea is the creation of new structures for the data processing system. The system consists of a large number of closely interconnected processing elements called neurons, which together solve a problem and transmit information through synapses (electromagnetic connections). Neurons are closely connected and are organized in layers. The input layer receives the data, while the output layer provides the final result. There is usually one or more secret layers between the two. This arrangement makes it difficult to predict or know the exact data flow (Dastres et al., 2021). Neurons are divided into groups called layers and are precisely connected to form a network. As mentioned, when the number of layers is large, the neural network is defined as deep (Franco & Santurro, 2021). As mentioned, the basic building block of any artificial neural network is an artificial neuron, or a simple mathematical model (function). Such a model has three simple rules: multiplication, addition, and activation. In an artificial neuron, the inputs are weighted, meaning that each input value is multiplied by the weight of an individual. At the heart of an artificial neuron is a sum function that sums all the weighted inputs and biases. At the output of the artificial neuron, the preweighted sum of the inputs and the bias passes through an activation function, also called a transfer function (Suzuki, 2017). In the world of computing, neural networks have many uses. Their ability to learn by example makes them very flexible and efficient. Artificial Neural Networks (ANN) have come a long way, as have other deep models such as Feedforward Neural Networks (FNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) (Minar, 2006) (Zakaria et al., 2014). Convolutional neural networks (CNNs) are similar to traditional ANNs in that they are composed of neurons that optimize themselves through learning. Each neuron receives input and performs the underlying operations of a myriad of ANNs. Specifically, CNN is a deep learning algorithm that consists of multiple convolutional, pooling, and fully connected layers (Altalak et al., 2022) (Shea & Nash, n.d.). However, CNN or ConvNet is a popular discriminative deep learning architecture that can learn directly from input objects without the need for human feature extraction (Directions, 2023). In traditional neural networks (NNs), the complete connectivity between layers often leads to computationally slow and parameter overfitting. Unlike NN, CNN uses

specific layers to perform convolutions and avoids common multiplications, resulting in faster computation. However, CNN passes the input image through many deep layers such as convolution layers, pooling layers, and activation layers for feature extraction and performs classification using fully connected layers with nonlinear classifiers (Sony et al., 2021). Hence, the CNN also consists of the input layer and the output layer. As shown by (Phung & Rhee, 2019) these layers are divided into two parts, feature extraction and classification:

1. Feature extraction consists of an input layer, convolution layer, and pooling layer.
2. Classification consists of a fully connected layer and an output layer.

However, the decline in agricultural output, resulting from the impact of global warming, the drying up of wetlands, unsustainable irrigation practices, and heedless agricultural methods, presents a significant threat to the global population. These factors have led to a parallel increase in food scarcity alongside the world's rapid population growth. Given these developments, it appears that the adoption of smart farming practices has become imperative to address these challenges (Bal, 2021). Smart farms encompass farm management systems that leverage emerging technologies to enhance food safety, quality, and quantity while simultaneously diminishing environmental repercussions. Enhanced production control translates into improved cost management and decreased wastage (Alibabaei et al., 2022). Hence, it is crucial to implement alternative production systems as a means of mitigating these challenges and establishing a sustainable food supply chain (Ojo & Zahid, 2022). Smart agriculture leverages deep learning to manage functions like categorization, identification, partitioning, and guidance. Deep learning methods find application in multiple facets of agriculture, including the identification of pests, detection of diseases, assessment of soil nutrient levels, recognition of leaf stress, identification of weeds, recommendations for pesticides, disease management, suggestions for herbicides, assessment of crop yields, and recommendations for optimizing yield (Saranya et al., 2023). (Lu & Tan, 2021), (Zhao et al., 2022), (Rakhmatulin & Kamilaris, 2021), (Afonso & Barros, 2021), (Wakchaure et al., 2023) have particularly utilized these methods in the realm of agriculture by employing Convolutional Neural Networks (CNNs). However,

digital images in various formats are becoming the most prevalent type of data. The digitization of agriculture can assist farmers in comprehensively monitoring agricultural lands for tailored management. With the proliferation of satellite remote sensing, aerial remote sensing, and proximal sensing using field robots in agriculture, the volume of collected images has surged. A critical concern is a method for swiftly extracting the most valuable details from these images. Traditionally, decoding this information necessitates specialized knowledge and extensive expertise. According to (Saleem et al., 2022) many researchers are currently dedicated to evaluating the architecture of CNNs. Through this work, using neural network architectures and deep learning techniques, a mobile application designed for Android devices was developed. The primary objective is to quantify the essential oil content in bergamot fruits based on their color.

CHAPTER 2

MATERIAL AND METHODS

2.1 *Field sampling and camera calibration*

Two bergamot plantations with trees of cv 'Fantastico' and 'Femminello', aged 15–20 years, were studied. The two research areas were located at 37° 58'49" N, E 16° 05'25"E and 38° 01′ 56.4″N 15°39′47.3″E, respectively. Tests were conducted on flat terrain under similar weather conditions, with trees in good physiological and health. The rootstocks used were bitter orange (Citrus aurantium L.) for Fantastico and Citrange Troyer (Poncirus trifoliata (L.) Rat. × Citrus sinensis (L.) Osbecke). The soil contained 23% (by volume) of rock fragments larger than 2 mm in diameter. The texture comprised 26.4% sand, 42.1% silt, and 31.5% clay. The concentration of organic matter in the soil was 5.1%, pH 8, and the total calcium (CaCO3) was 4%. The average monthly temperature is 19 °C, and the monthly rainfall is 49.5 mm. Mineral fertilization (1 kg/plant) was carried out using a ternary fertilizer (11 N–22 P2O5 - 16 K2O), fertigation by potassium nitrate (2.500 kg/ha twice a year), and foliar fertilization with algae extracts and boron. Between May and October, 80 L of water was provided to each plant twice a week. Phytosanitary management includes uprooting in the event of Mal Secco disease. The insecticide "Movento" was used twice a year in early and late spring to control sucking pests (aphids, leaf miners, and mealybugs) at a dose of 200 g/100 L of water. Pest control included mowing three times a year. On average, the production yield of the fresh product was approximately 35 tons/ha, whereas that of the EO was 16 kg/ha. Sampling was performed weekly with image acquisition of bergamot fruit on nine dates during the harvesting period, i.e., from 03/11/2022 to 04/01/2023. On each date, sampling was conducted in the morning (8:00–12:00 a.m.), selecting 10 healthy fruits per plant from 10 bergamot trees, totaling 100 fruits. Fruits were selected based on the ripening evolution evaluated by the Citrus Colour Index (CCI) defined by Jiménez-Cuesta et al. (1981) using the standard Citrus Colour Chart (Fig. 2.1) developed by the Valencian Institute of Agricultural Research (IVIA). This index is used in the citrus industry to assess the optimal harvesting period or determine the suitable period to apply post-harvest treatments such as degreening

(Vidal et al., 2013; Benalia 2015). It is calculated considering the three coordinates in the HunterLab color space, which can be read directly from a photoelectric colorimeter (tristimulus method), according to the equation when parameters “a” (green to red), and “b” (blue to yellow) are the two chromatic components, varying from –120 to +120 while “L” indicates the brightness level between 0 and 100 (black and white).

$$CCI = \frac{1000 \times a}{L \times b}$$

Figure 2.1 Fruit Classification using Citrus Color Chart.

A Citrus Color Index (CCI) with a negative value indicates that the fruits exhibit shades of dark green to green. Values close to zero suggest that the fruits display a range of colors from green to yellow and orange, small positive values mean a yellow color, and high positive values mean red or orange. Figure 2.2 shows an

example of some fruits selected based on the CCI from -9 to +6 using the Citrus Color Chart. For each sample, the fruits had the same CCI values.

Figure 2.2 Fruits classification according to Citrus Color Index.

Five CCI values (-9, -5, -3, +1, and +3) were established according to the PDO specifications, which reported that the fruit must be harvested when the color ranges from green to yellow. The previous year, sampling was also carried out on the only 'Fantastico' on four dates: 02/12/2021, 16/12/2021, 04/01/2021, and 20/01/2022. Three CCI values were recorded (-5, -3, and +3). Images were captured using a smartphone (Huawei P9 Lite) with 3 GB RAM (Kirin 650); the main features of the smartphone are listed in Table 2.1.

Parameter	Description
CPU	HiSilicon Kirin 650
GPU	Mali-T830MP2
RAM	3 GB Kirin 650
Cameras	13 MP
Operating System	Android 6.0

Table 2.1 Detailed parameters of the smartphone.

A low-cost portable inspection chamber was constructed to prevent interference from natural light during image acquisition. The inspection chamber consisted of an internal black plastic tube containing a diffuse light-emitting diode (LED) ring light (6500 K), allowing the maintenance of image acquisition under controlled conditions (Fig. 2.3). For each fruit, four images were acquired to capture the color variations, for a total of 400 images per sampling date.

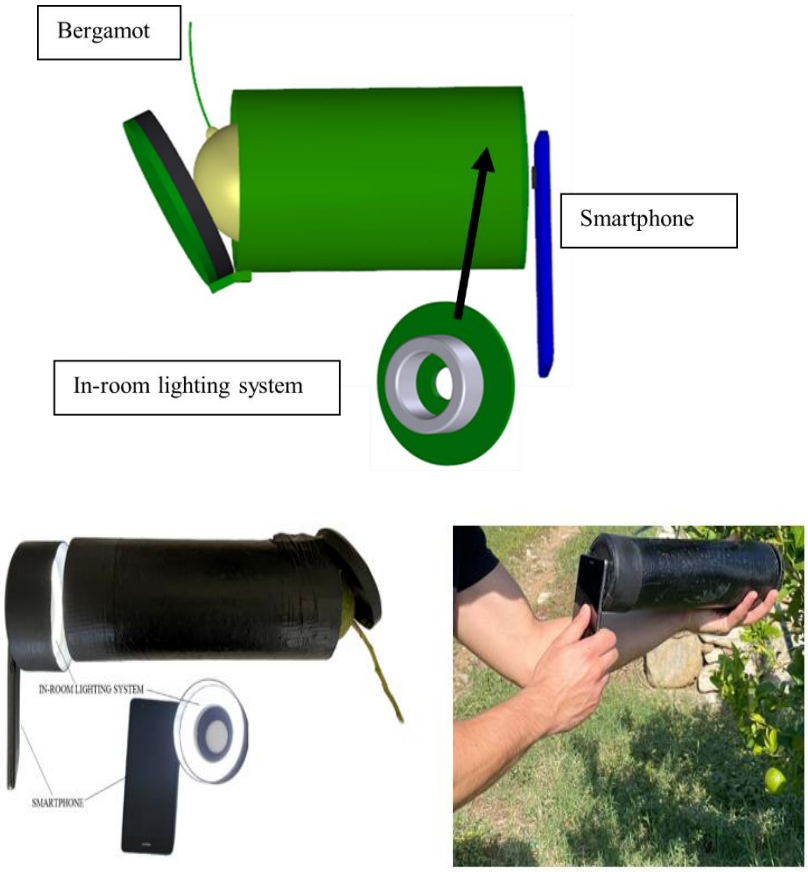


Figure 2.3 Portable inspection chamber.

First, the CCI was calculated using a portable spectrophotometer (CM-700d, Konica Minolta, Tokyo, Japan) to optimize the accuracy of color acquisition by the mobile camera, using the Citrus Color Chart developed by the Valencian Institute of Agricultural Research (IVIA). The smartphone camera was set up manually using different settings about ISO, f/stop, and exposure time. ISO, f-stop (aperture),

and exposure time are all key parameters in photography that influence exposure and the appearance of the captured image. Settings compared were:

- ISO 100, f/stop 2, exposure time 1/200
- ISO 100, f/stop 2, exposure time 1/250
- ISO 100, f/stop 2, exposure time 1/320
- ISO 100, f/stop 2, exposure time 1/400

ISO represents the sensitivity of the camera's sensor to light. A higher ISO value makes the sensor more light-sensitive, allowing photography in low-light conditions. However, a higher ISO can introduce noise or grain into the image. A lower ISO is ideal for optimal lighting conditions. The aperture, measured in f-stop values, controls the amount of light entering the camera. A lower f-stop value indicates a wider aperture, allowing a lighter and shallower depth of field. Conversely, a higher f-stop value indicates a narrower aperture, reducing the amount of light and increasing the depth of field. Exposure time, measured in seconds or fractions of a second, determines how long the camera's sensor captures light. Short exposure times capture quick snapshots and reduce the risk of motion blur. Longer exposure times capture more light and can create intentional long-exposure or motion blur effects. The choice of appropriate values depends on the desired effect and shooting conditions. In combination, these three parameters allow control over exposure, depth of field, and motion effects in a photo. The images were stored in a JPEG format. In addition, the color of each fruit was further analyzed in the laboratory using the portable spectrophotometer Konica Minolta CM-700d considering the CIE Illuminant D65 and the 10° observer standard for comparison purposes. The instrument was calibrated according to a white tile reference (L^*=99.32, a^*=-0.12, b^*=-0.13). The Konica Minolta CM-700d colorimeter is a precise instrument used to measure the color of objects or surfaces. It works by illuminating the target with a known light source and then detecting and processing the reflected light using a sensitive sensor. Every fruit with a mean of three measurements in different zones represented a replicate. To evaluate the accuracy of the color acquisition by the mobile camera, the CCI index was

calculated considering data from both conventional color analysis and bergamot image analysis. This was performed later using Food Color Inspector software (Fig. 2.4), a tool specially developed for this purpose at the Laboratory of Artificial Vision of IVIA (Spain). This application converts RGB values of the pixels of the entire image or a selected region of interest (ROI) to HunterLab space. The first step consists of the conversion of RGB values to XYZ values, and then from XYZ to HunterLab (Cubero, 2012). The obtained CCI values (portable spectrophotometer vs mobile camera) were compared using the coefficient of determination (R^2).

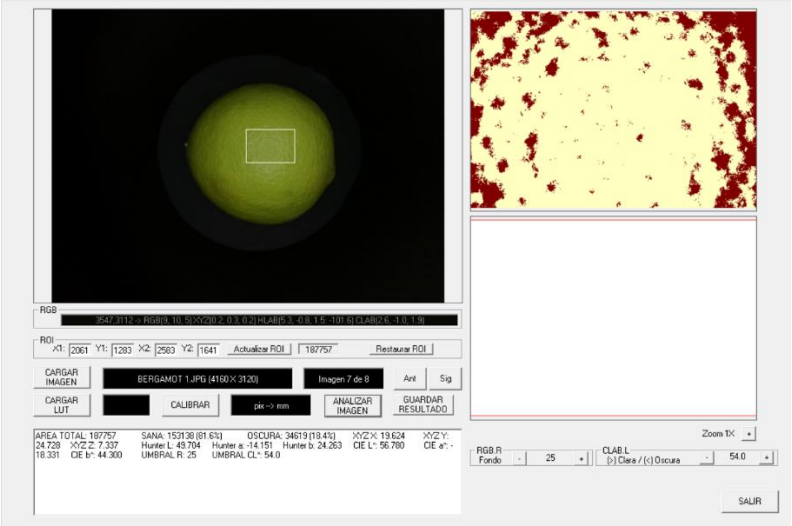


Figure 1.4 Inspector program used for image analysis.

Concerning the accuracy of the CCI, the best linear regression value (R^2 = 0.93), when comparing the image analysis of the mobile camera and the spectrophotometric results, was found with the mobile camera set to ISO 100, f-stop 2 and an exposure time of 1/200 (Figure 2.5). This result is to be expected as the lighting conditions in the field are more stable thanks to the use of the low-cost portable inspection chamber. This tool is specifically designed to eliminate the effects of natural light on image capture, optimize data collection from smartphones, and create an environment that limits the influence of external factors and standardizes the data collected (Anello et al., 2023).

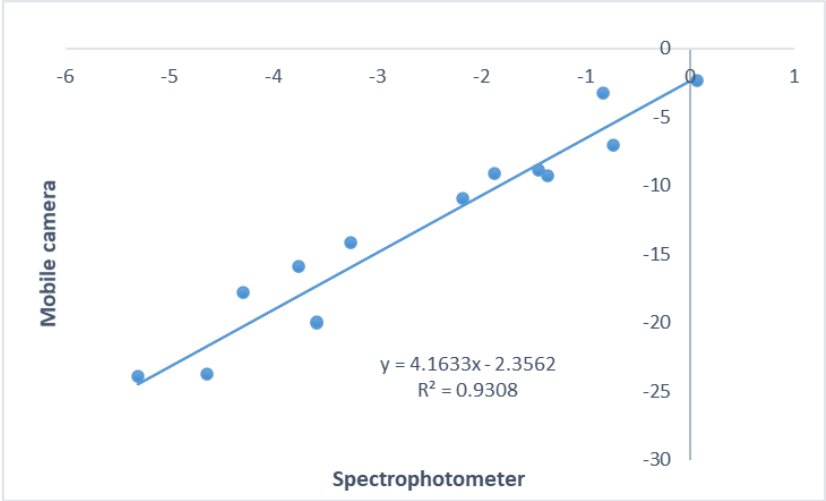


Figure 2.5 Linear regression between CCI obtained using the mobile camera image analysis vs the spectrophotometer.

2.2 Essential oil extraction

After image acquisition, 50 previously sampled fruits were randomly collected by hand and transported to specific plastic containers for EO extraction. The goal of laboratory EO extraction was to quantitatively determine the amount (in grams) of essential oil obtained. The flavedo was gently removed from each fruit using a knife and reduced to square pieces (1 cm2). Peel EO was extracted by hydro-distillation using a Clevenger apparatus (Fig. 2.6). The vegetable material was introduced into a 2-L glass flask at a ratio of 250 g of bergamot peel and 1 L of deionized water. The glass flask was heated to the boiling point and maintained for 2 h. The hydro-steam-distillation is based on the principle that two immiscible liquids boil at a temperature lower than the boiling points of either pure component. Subsequently, the vapor mixture was condensed and distilled into a separator, and the volatile concentrate was collected. EO extraction was conducted in duplicates. The EO was recovered and quantified. The final result is expressed as the mean of two replicates.

Figure 2.6 Hydro-steam-distillation of bergamot fruits.

The dataset was created considering the sampling date. All bergamot images were matched with the CCI values and extracted oil content was recorded on that date, as shown in Table 2.2. The R, G, and B coordinates of the pixels belonging to the fruit were used as inputs to the models.

2.3 Convolutional neural network architecture

2.3.1 CNN models built for 'Fantastico' during the year 2021/2022

The dataset referring to the first year of sampling is shown in Table 2.2.

Extraction date	CCI value	Oil content (g)	Class label
02/12/2021	-5	2.8788	0
16/12/2021	-3	4.5158	1
04/01/2022	+3	3.4497	2
20/01/2022	+3	3.4450	2

Table 2.2 Sampling dates, CCI value, oil content, and the class labels for 'Fantastico' (2021/2022).

Preprocessing steps were performed on the images to prepare them for the deep learning algorithm. The transfer learning models were pre-trained using the ImageNet dataset (Deng et al., 2009; Paymode & Malode, 2022). This dataset requires input images of size 224 x 224 pixels, and the color bands are normalized between 0 and 1. Then, all images were resized and normalized by dividing the RGB values of each pixel by 255. Figure 2.7 shows some examples of the images after resizing and normalization.

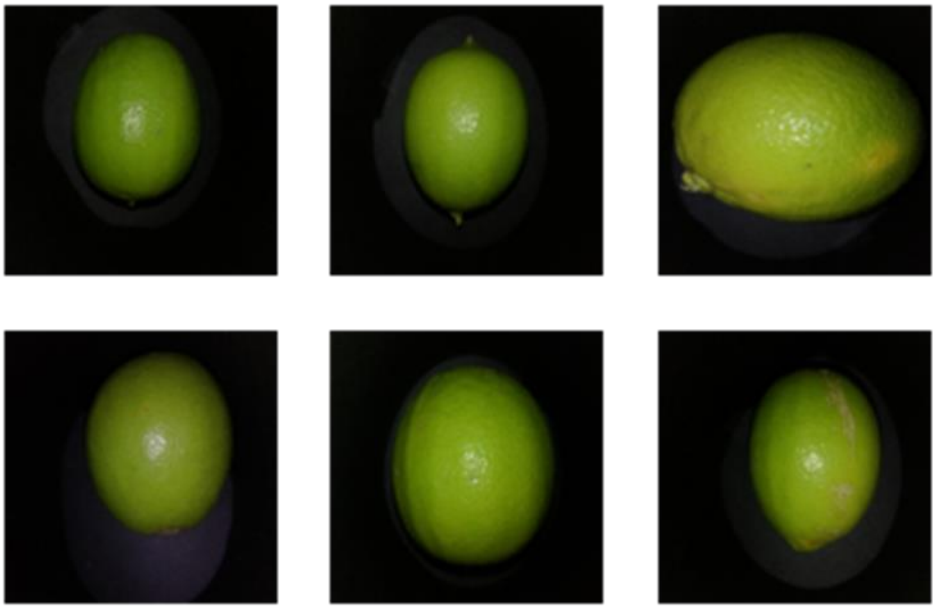


Figure2.7 Example of six bergamot fruit images after the pre-processing stages.

One-hot encoding of outputs

The labels' values were also processed to remove the effect of cardinality (i.e. relative distances between class representations which could encode some artificial ordering) and also because, in our models, we had one output node per label, each one with a softmax activation function that produced probability levels between 0 and 1. Therefore, the output for each image in the sample was a vector composed of four elements, following a "one-hot" representation. In this kind of representation, only one of the four elements of the vector is 1 (the index corresponding to the class value), while the rest are zeros. The four possible encodings in our data are listed in Table 2.3. Every time a prediction for a sample was obtained from the CNN, the arg max function was applied to the four outputs to determine the predicted class index.

Class	Encoding
0	(1, 0, 0)
1	(0, 1, 0)
2	(0, 0, 1)

Table 2.3 One-hot encoding of the outputs.

2.3.1.1 Training and validation datasets

Tensorflow/Keras libraries from the Anaconda3 package were used to develop the model (see last item from Section 2: Hardware and software). In these libraries, the training data is used to make the model learn by adjusting its weights iteratively, while the validation data is used to assess the generalization on an external dataset after each weight adjustment. The process can continue for a predefined number of iterations (epochs) or be stopped when a stop criterion is reached (for example, the validation error does not improve for ten consecutive epochs). The approach we followed was to train for enough epochs (100) until we verify that the validation error reaches a plateau and does not further improve, and then keep the best model weights as the ones that produce the highest accuracy in the validation set history by using the ModelCheckpoint function of Keras. Finally, the model is evaluated on a separate test set for unbiased accuracy estimation. To correctly assess the model generalization capabilities, the data were split randomly into a training set containing 60% of the data (1008 images), a validation set of 20% of the data (336 images), and a test set of 20% of the data (336 images). The final tensor sizes used to create the models are listed in Table 2.4.

DATA SUBSET	SHAPE
TRAINING INPUT	(1008, 224, 224, 3)
TRAINING OUTPUT	(1008, 1)
VALIDATION INPUT	(336, 224, 224, 3)
VALIDATION OUTPUT	(336, 1)
TEST INPUT	(336, 224, 224, 3)
TEST OUTPUT	(336, 1)

Table 2.4 Tensor sizes for training, validation, and test subsets.

2.3.1.2 Models

Custom CNN model

The first approach was to create a custom CNN model from scratch. A CNN for image classification typically applies several 2D convolution operations followed by pooling operations and stacks these blocks several times. Finally, the obtained feature maps are flattened as a 1D array and fed into a traditional fully connected network that produces the desired outputs. A deep learning model in a classification problem has one output node per class, typically using a softmax activation function that returns a probability value for each class. Then, the class with the maximum probability was selected as the predicted one. Convolutional layers in a convolutional neural network systematically apply learned filters to input images to create feature maps that summarise the presence of those features in the input. Convolutional layers prove very effective, and stacking convolutional layers in deep models allows layers close to the input to learn low-level features (e.g. lines) and layers deeper in the model to learn high-order or more abstract features, like shapes or specific objects. A limitation of the feature map output of convolutional layers is that they record the precise position of features in the input. This means that small movements in the position of the feature in the input image may result in a different feature map. The pooling layers perform downsampling of the feature maps. They keep only a representative value for a block of data (maximum, mean, sum, etc.), thus creating a lower-resolution version of an input image that still contains the large or important structural elements without the fine detail that may not be as useful to the task. The custom CNN model architecture proposed for our particular task is described in Figure 5. The architecture follows the same philosophy of VGG architecture but with lower complexity. It contains three functional blocks that contain: i) two consecutive 2D convolution layers that add depth to the feature maps, both using 3 x 3-pixel filters and "same" padding to avoid reducing the feature map width and height; and ii) a max pooling layer with 2 x 2 filters with stride = 2, that keeps the maximum value from each 2 x 2-pixel square in the feature map, thus halving the width and height of the feature map at each step. The top of the CNN is composed of a flattened layer and a fully connected structure with 128 hidden nodes with ReLU activation, a dropout layer with

probability p = 0.2, and 4 output nodes with softmax activation that provide the class probabilities.

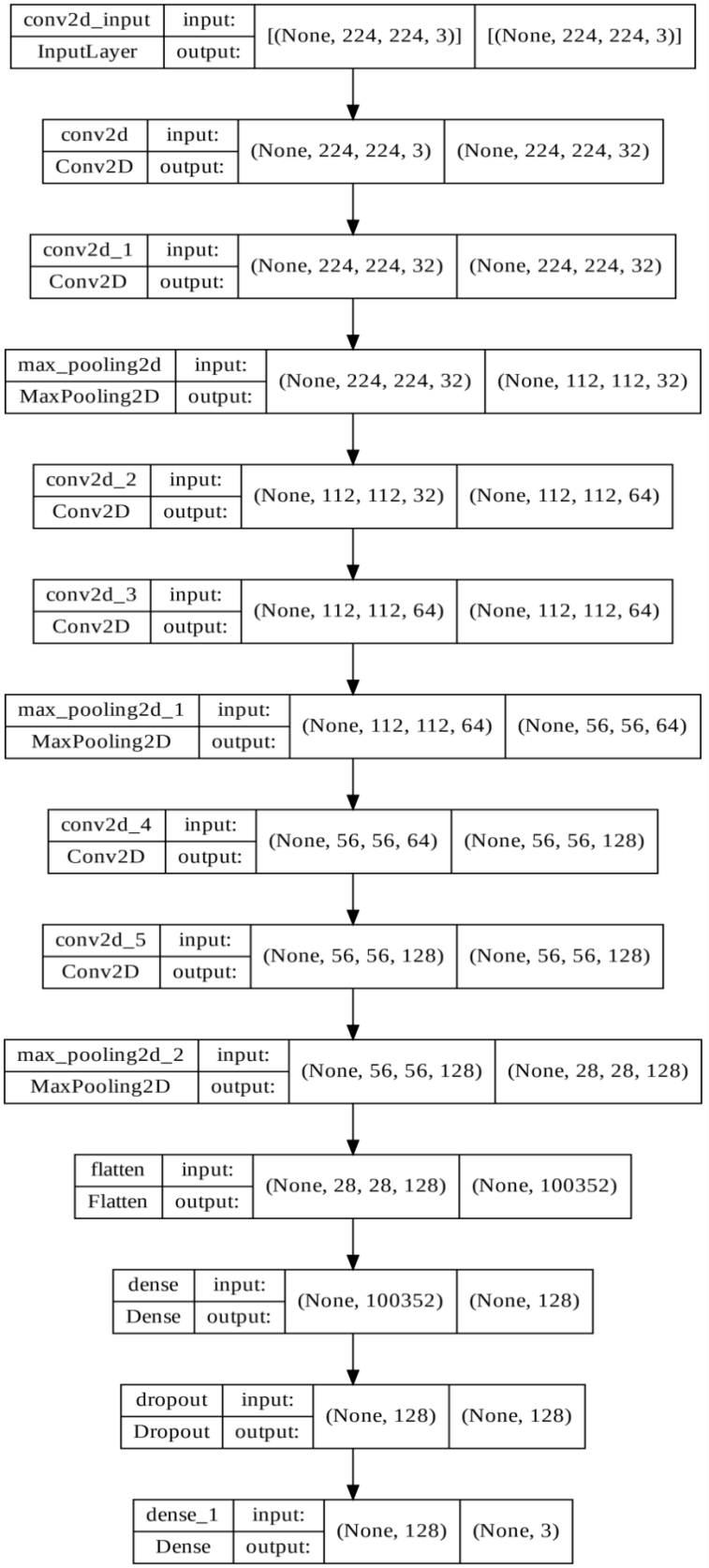
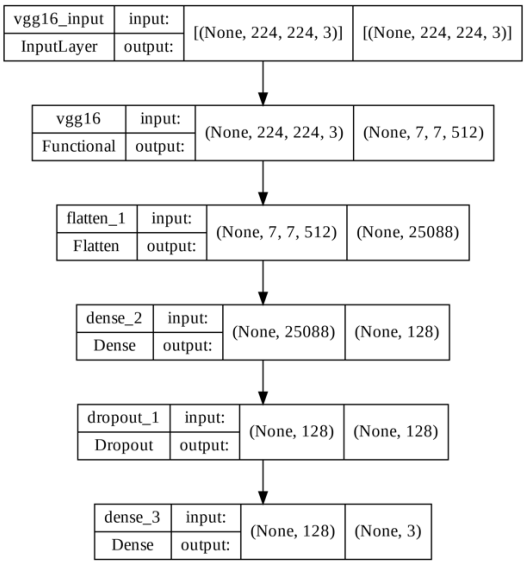


Figure 2.8 Proposed custom CNN model architecture.

2.3.1.3 Transfer learning models

The second approach was to profit from pre-trained models that already achieve very high accuracy on extensive specific datasets, such as ImageNet, and reuse them with slight modifications for our particular task. In the literature, this is called transfer learning. In the literature, examples where transfer learning models achieve comparable accuracies to custom-built models for various image classification tasks while generally keeping lower training times can be found (Kim et al., 2020; Hossain et al., 2018, Shao et al., 2018). These models were initially trained for a different purpose. Still, we can benefit from the learned features encoded into their weights to build a powerful model to tackle a different problem by just fine-tuning the last layer or layers of the model. These layers that need to be built from scratch and trained are usually the top layers, i.e., the fully connected part of the model. This means that the number of parameters that need to be tuned is reduced drastically concerning the total amount of parameters in the model, saving valuable time at the model training stage. In the present work, the VGG-16 and VGG-19 (Simonyan & Zisserman, 2014) and the Xception (Chollet, 2017) model architectures were used as starting points to build transfer learning models. In each one, the top layers were removed and replaced by a block that produces the mapping from the original feature map depth of 512 (for VGG-16/19) or 2048 (for Xception) to only four outputs corresponding to our model classes. The new top layers are made of fully connected structures composed of a flattening layer, a dense layer of 128 neurons with ReLU activation function, followed by a dropout layer with a dropout probability of 0.3, to help reduce overfitting, and finally, the output layer of 4 nodes with softmax activations. The VGG-16 and Xception-based architectures are summarised in Figure 2.9. The VGG-19 architecture is very similar to a) but replacing the pre-trained InputLayer and Functional layer with those of VGG-19.

a)



b)

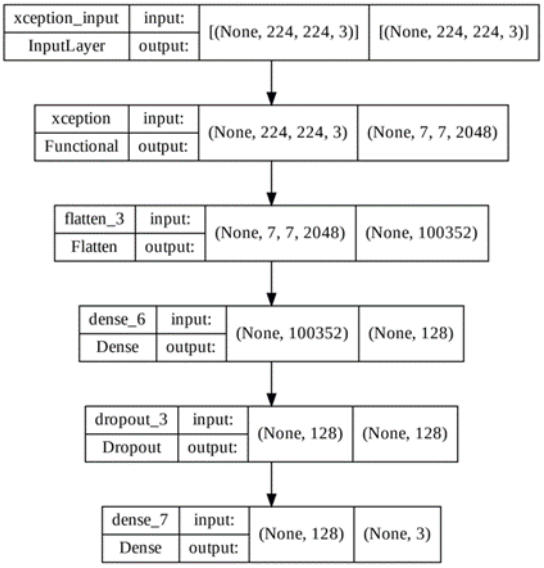


Figure 2.9 Proposed transfer learning models based on: a) VGG-16; and b) Xception.

2.3.1.4 Training parameters

All training was carried out for 100 epochs, using the Adam optimizer, the categorical cross-entropy loss function, and a batch size of 64. The best weights, corresponding to the maximum validation accuracy, were saved in each case.

Error metrics

As we considered the problem a multiclass classification one, the typical error metric that can be used to compare model performances is accuracy, defined as:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

Additionally, to provide insight into the distribution of the errors among classes in the test data, a confusion matrix may be calculated for each model architecture.

Hardware and software

For the data preprocessing as well as for the construction, training, and testing of the models, Python 3.10.4 and Anaconda3 (version 2022.05) were used. The following Anaconda3 packages were used too: pandas, TensorFlow, Keras, sklearn, cv2, and matplotlib. The code was run on a Quadro RTX 6000 GPU (24 GB GDDR6X memory).

2.3.2 CNN models built for 'Fantastico' and 'Femminello' during the year 2022/2023

The dataset resulting from the second year of sampling was created taking into account the Fantastico variety together with the Femminello variety, as shown in the tables below.

Date	CCI value	Oil content (g)	Class label
03/11/2022	-9	4.18	0
10/11/2022	-9	3.56	1
17/11/2022	-5	5.00	2
24/11/2022	-5	4.79	3
01/12/2022	-3	4.72	4
06/12/2022	-3	4.13	5
15/12/2022	1	4.02	6
20/12/2022	3	4.52	7
04/01/2023	3	3.88	8

Table 2.5 Sampling dates, CCI value, oil content, and the class labels for Fantastico.

Date	CCI value	Oil content (g)	Class label
03/11/2022	-9	4.16	0
10/11/2022	-5	3.65	1
17/11/2022	-5	3.91	2
24/11/2022	-3	3.67	3
01/12/2022	1	3.57	4
06/12/2022	3	3.59	5
15/12/2022	3	3.35	6
20/12/2022	3	3.97	7
04/01/2023	3	3.41	8

Table 2.6 Sampling dates, CCI value, oil content, and the class labels for Femminello.

CNNs were built for a) estimating the extraction date as a class, which is a classification problem, and b) estimating the EO content as an actual number, which is a regression or function approximation problem. Independent models were built and trained for the two cultivars, using the acquired images as inputs.

Preprocessing

Before the creation of the predictive models, a preprocessing stage was performed on the input images to correct possible acquisition errors, noise, or deviations in the pixel values, and to adapt them to the learning models so that they can learn optimally.

Image rescaling

The first preprocessing was applied to rescale the three channels of all the images between 0 and 1. This helps uniformize the possible uneven brightness level of the different images so that this difference in pixel range is not used as a feature to minimize the loss of the model during the training stage. As the RGB images are coded with 8 bits per pixel in each channel (representing a pixel value range in [0,255]), the normalization of each channel (cR, cG, cB) is:

$$c_i^{'}=c_i/255, \quad i=R,G,B$$

Image resizing

Another preprocessing step was to resize the images from the original 4160 x 3120 or 3120 x 4160 pixels to a unique size. This is necessary to uniformize all input image sizes that were used as input features to the developed models. The original high-resolution images contain too much fine detail that is irrelevant to the prediction goal, and therefore, the image sizes can be reduced with the benefit of lower training times and without a reduction of performance. Firstly, as 2D convolutions with square filters were used, the images must also be square. Secondly, as one of the goals of this work is to try transfer learning techniques to predict the oil content and the transfer learning models used were pre-trained with the ImageNet dataset (Deng et al., 2009), the image sizes were adapted to the standard ImageNet size: 224 x 224 pixel. This is often adopted in the literature for optimal compatibility (Paymode & Malode, 2022).

One-hot encoding of outputs

When using classification models, the labels' values must also be processed to remove the effect of cardinality (i.e. relative distances between class representations which could encode some artificial ordering) and also because, in our models, we had one output node per label, each one with a softmax activation function that produced probability levels between 0 and 1. Therefore, the output for each image in the sample was a vector composed of nine elements, following a "one-hot" representation. In this kind of representation, only one of the nine elements of the vector is 1 (the index corresponding to the class value), while the rest are zeros. The four possible encodings in the data are listed in Table 2.7. Every time a prediction for a sample was obtained from the CNN, the arg max function was applied to the four outputs to determine the predicted class index.

Class	Encoding
0	(1, 0, 0, 0, 0, 0, 0, 0, 0)
1	(0, 1, 0, 0, 0, 0, 0, 0, 0)
2	(0, 0, 1, 0, 0, 0, 0, 0, 0)
3	(0, 0, 0, 1, 0, 0, 0, 0, 0)
4	(0, 0, 0, 0, 1, 0, 0, 0, 0)
5	(0, 0, 0, 0, 0, 1, 0, 0, 0)
6	(0, 0, 0, 0, 0, 0, 1, 0, 0)
7	(0, 0, 0, 0, 0, 0, 0, 1, 0)
8	(0, 0, 0, 0, 0, 0, 0, 0, 1)

Table 2.7 One-hot encoding of the class outputs.

2.3.2.1 Training and validation datasets

For the model development, the Tensorflow/Keras Python libraries were used. In these libraries, the training data makes the model learn by adjusting its weights iteratively, while the validation data is used to assess the generalization on an external dataset after each weight adjustment. The approach consisted of training for enough epochs (100) until the validation error reached a plateau and did not further improve, and then keeping the best model weights as the ones that produced the highest accuracy in the validation set history by using the ModelCheckpoint function of Keras. Finally, the model was evaluated on the test set for unbiased accuracy estimation. To correctly assess the model generalization capabilities, the 3600 images captured for each cultivar were randomly split into a training set containing 60% of the data (2160 images), a validation set of 20% of the data (720 images), and a test set of 20% of the data (720 images). The final tensor sizes used to create the models are listed in Table 2.8.

Data subset	Shape
Training input	(2160, 224, 224, 3)
Training Output	(2160, 9)
Validation input	(720, 224, 224, 3)
Validation output	(720, 9)
Test input	(720, 224, 224, 3)
Test output	(720, 9)

Table 2.8 Tensor sizes for training, validation, and test subsets.

2.3.2.2 Models

Custom classifier

The first approach was to create a classifier to estimate the extraction date among the 9 possible dates encoded as the target classes. A custom CNN model was created from scratch and compared its performance with other pre-trained classifiers, using transfer learning to fine-tune them. The architecture follows the same philosophy of VGG CNN architectures but with lower complexity. It contains three functional blocks that contain: i) two consecutive 2D convolution layers that add depth to the feature maps, both using 3 x 3-pixel filters and "same" padding to avoid reducing the feature map width and height; and ii) a max pooling layer with 2 x 2 filters with stride = 2, that keeps the maximum value from each 2 x 2-pixel square in the feature map, thus halving the width and height of the feature map at each step. The top of the CNN is composed of a flattened layer and a fully connected structure with 128 hidden nodes with ReLU activation, a dropout layer with probability p = 0.2, and 9 output nodes with softmax activation that provide the class probabilities. The custom CNN model architecture proposed for the classifier is shown in Figure 7, including input/output tensor sizes for each block. The same architecture was used for the two cultivars, but the training input data was different for each one.

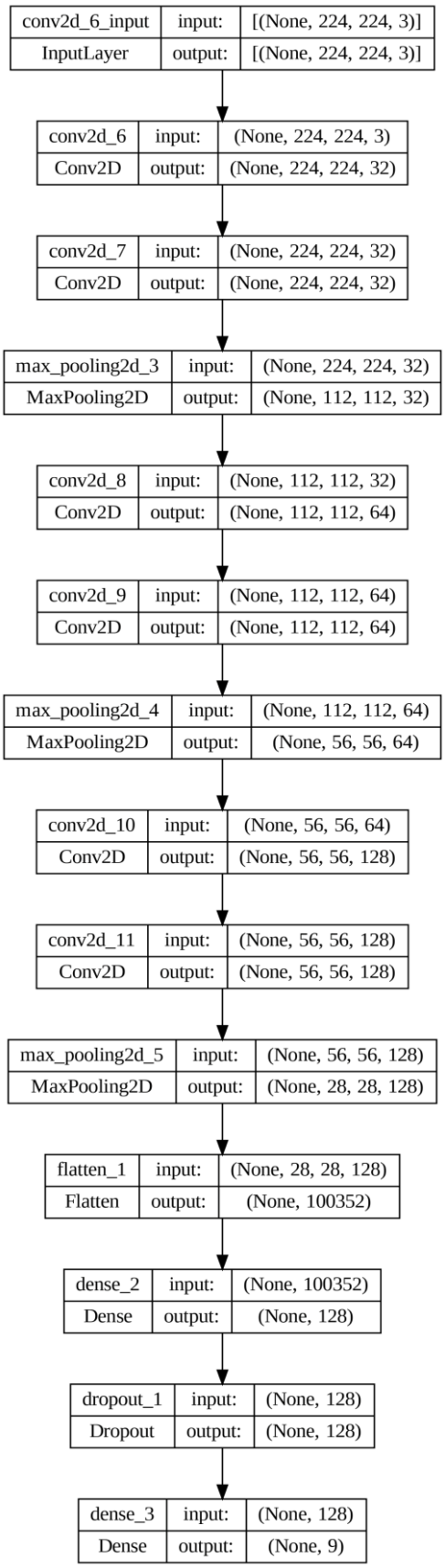


Figure 2.10 Architecture of the proposed custom CNN classifier.

2.3.2.3 Transfer learning models

We compared the performance of the proposed architectures against popular pre-trained CNN architectures using transfer learning. These pre-trained architectures achieve very high accuracy on extensive specific datasets, such as ImageNet, and we can reuse them with slight modifications for our particular goal. In the literature, one can find examples where transfer learning models achieve comparable accuracies to custom-built models for various image classification tasks while keeping generally lower training times (Kim et al., 2020; Hossain et al., 2018, Shao et al., 2018). These models were initially trained for a different purpose, but one can benefit from the learned features encoded into their weights to build a powerful model to tackle a different problem by just fine-tuning the last layer or layers of the model. These layers that need to be built from scratch and trained are usually the top layers, i.e., the fully connected part of the model. This means that the number of parameters that need to be tuned is reduced drastically concerning the total amount of parameters in the model, saving valuable time at the model training stage. In the present work, the VGG-16 and VGG-19 (Simonyan & Zisserman, 2014) and the Xception (Chollet, 2017) model architectures, were used as starting points to build transfer learning models. In each one, the top layers were removed and replaced by a block that produces the mapping from the original feature map depth 512 (for VGG-16/19) or 2048 (for Xception) to only 9 outputs representing our model classes. The new top layers are made of fully connected structures composed of a flattening layer, a dense layer of 128 neurons with ReLU activation function, followed by a dropout layer with a dropout probability of 0.3 to help reduce overfitting, and finally, the output layer of 9 nodes with softmax activations. The resulting architectures, including the input/output tensor sizes for each block, are shown in Figure 2.11.

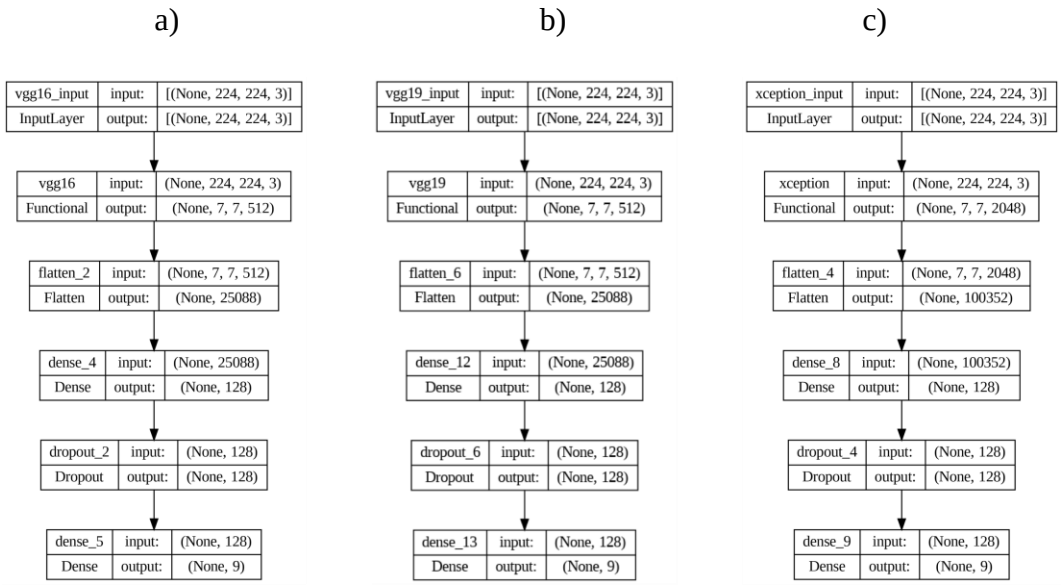


Figure 2.11. Proposed transfer learning architectures for classification: a) VGG-16 based; b) VGG-19 based; c) Xception based.

2.3.2.4 Regression models

For the regression approach, we compared two custom CNNs. The first one (REG1) derives from our custom classifier, replacing the last layer with a single node with linear activation to yield a single real value. The second architecture (REG2) also reuses its architecture but inserts a batch normalization (Ioffe & Szegedy, 2015) and a dropout (p=0.2) operation in each convolutional and fully-connected block, aiming for a possible improvement in speed and generalization. Both architectures are shown in Figure 2.12.

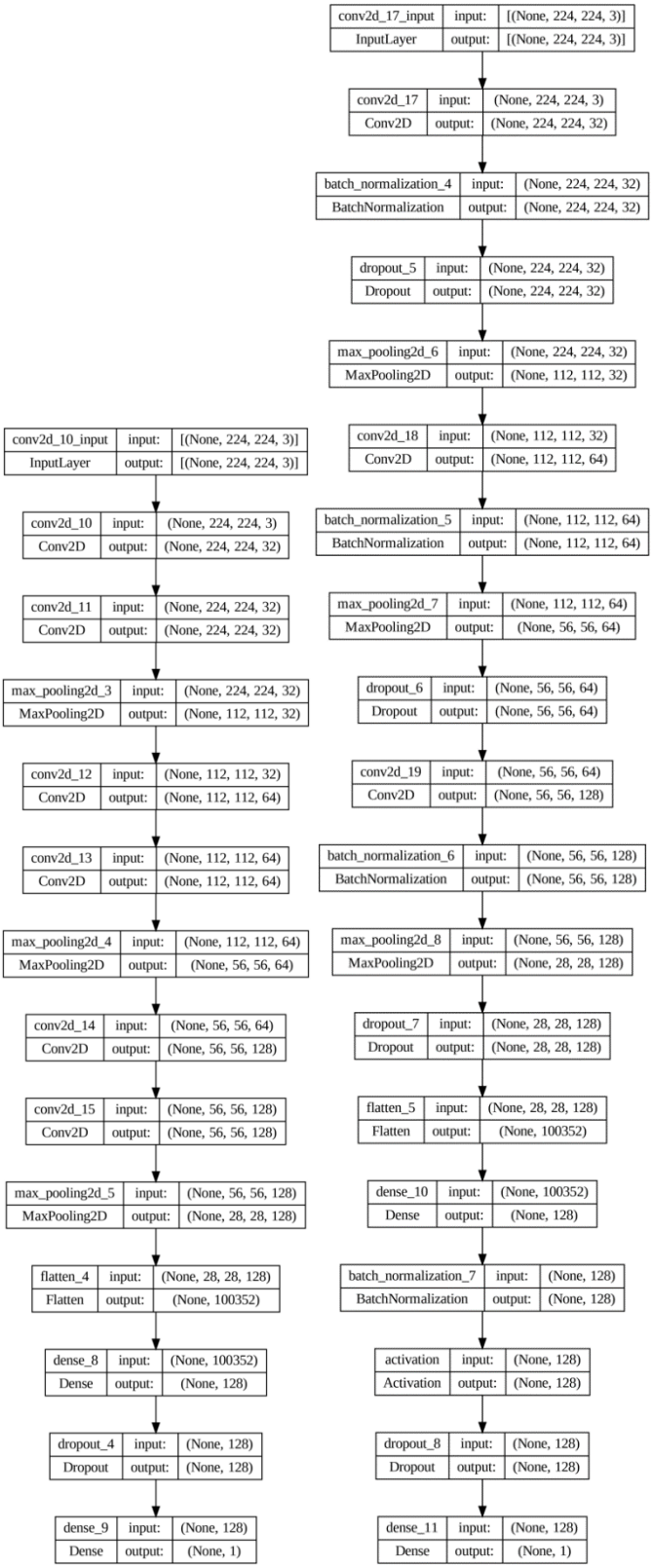


Figure 2.12 Architecture of the proposed regression models: a) REG1, b) REG2.

2.3.3 CNN models built to join data from 2021 to 2023

CNNs were built to a) estimate the extraction data as a class, which is a classification problem, and b) estimate the EO content as an actual number, which is a regression or functional approximation problem.

In this work, two approaches were considered to build the models:

1) *One general model to classify the extractions of both cultivars*. This classifier attempts to determine one of the combinations of cultivar/extraction dates (22 classes).

2) *Two separate models to classify the extractions of each cultivar*. These classifiers attempted to classify the extracts of each cultivar independently (13 classes for Fantastico and nine for Femminello).

These two approaches were also considered for the regression task but aimed to predict a real output value instead of a class.

Pre-processing

Before the creation of the predictive models, a pre-processing stage was performed on the input images to correct possible acquisition errors, noise, or deviations in the pixel values and to adapt them to the learning models so that they could learn optimally.

Image rescaling

The first pre-processing step was applied to rescale the three channels of all images between zero and one. This helps to uniformize the possible uneven brightness levels of the different images so that this difference in pixel range is not used as a feature to minimize the loss of the model during the training stage. Because the RGB images are coded with 8 bits/pixel in each channel (representing a pixel value range in [0,255]), the normalization of each channel (c_R, c_G, c_B) is

$$c_i^{'} = \frac{c_i}{255}, \quad i = R, G, B$$

Image resizing

Another pre-processing step was to resize the images from the original 4160 × 3120 or 3120 × 4160 pixels to a unique size. This was necessary to ensure that all input image sizes were used as input features for the developed models. The original high-resolution images contain too much fine detail, which is irrelevant to the prediction goal. Therefore, the image sizes can be reduced with the benefit of shorter training times and without a reduction in performance. First, because 2D convolutions with square filters are used, the images must also be squared. Second, as one of the goals of this work is to try transfer learning techniques to predict oil content, the transfer learning models used were pre-trained with the ImageNet dataset (Deng et al., 2009), which was adapted to a standard ImageNet size of 224 × 224 pixels. This approach has often been adopted in the literature for optimal compatibility (Paymode & Malode, 2022). An example of an image obtained after the pre-processing stage is shown in Figure 4. When using classification models, the label values must also be processed to remove the cardinality effect (i.e. relative distances between class representations which could encode some artificial ordering); this is because, in our models, we had one output node per label, each with a softmax activation function that produced probability levels between 0 and 1. Therefore, the output for each image in the sample is a vector composed of as many elements as classes, following a "one-hot" representation. In this type of representation, only one of the elements of the vector is 1 (the index corresponding to the class value), whereas the rest are zeros. Every time a prediction for a sample was obtained from the CNN, the Argmax function was applied to the outputs to determine the predicted class index.

2.3.3.1 Training and validation datasets

TensorFlow/Keras Python libraries were used for model development. In these libraries, the training data enable the model to learn by adjusting its weights iteratively, whereas the validation data are used to assess the generalization on an external dataset after each weight adjustment. The approach consisted of training for a maximum of 100 epochs until the validation error reached a plateau and did not improve further. An early stopping procedure was implemented to avoid overfitting, and the training process was stopped if the validation loss did not improve for 10 consecutive epochs. After training, the best model weights (those that produced the lowest cross-entropy loss in the validation set history) were saved using the Keras model checkpoint function. Finally, the model is evaluated using a test set for an unbiased accuracy estimation. To correctly assess the model generalization capabilities, the captured images were randomly split into training (60%), validation (20%), and test (20%) sets. All the classes were equally distributed among the three groups. Depending on the approach, the final tensor sizes used to create the models are listed in Table 2.9.

	General model	**Separate models**	
Data subset	Shape	Shape for model 1	Shape for model 2
Training input	(5328, 224, 224, 3)	(3168, 224, 224, 3)	(2160, 224, 224, 3)
Training Output	(5328, 22)	(3168, 13)	(2160, 9)
Validation input	(1776, 224, 224, 3)	(1056, 224, 224, 3)	(720, 224, 224, 3)
Validation output	(1776, 22)	(1056, 13)	(720, 9)
Test input	(1776, 224, 224, 3)	(1056, 224, 224, 3)	(720, 224, 224, 3)
Test output	(1776, 22)	(1056, 13)	(720, 9)

Table 2.9 Tensor sizes for training, validation, and test subsets for both classification approaches.

2.3.3.2 Models

Custom classifier

The first approach is to create a classifier (or two classifiers for the second classification approach) to estimate the extraction date among the possible dates encoded as the target classes. A custom CNN model was created from scratch, and its performance was compared with that of the other pre-trained classifiers using *transfer learning* for fine-tuning. The architecture followed the same philosophy as the VGG CNN architectures, but with lower complexity. It comprises three functional blocks with (i) two consecutive 2D convolution layers that added depth to the feature maps, both using 3 × 3-pixel filters and "same" padding to avoid reducing the feature map width and height and (ii) a max pooling layer with 2 × 2 filters with stride = 2, that kept the maximum value from each 2 × 2-pixel window in the feature map, thus halving the width and height of the feature map at each step. The top of the CNN comprised a flattened layer and a fully connected structure with 128 hidden nodes with ReLU activation, a dropout layer with probability p = 0.2, and nine output nodes with softmax activation that provided the class probabilities. The custom CNN model architecture proposed for the general classifier is shown in Figure 2.13, including input/output tensor sizes for each block. The same architecture was used for separate classifiers with a difference in the number of outputs in the last dense layer (13 or 9 instead of 22).

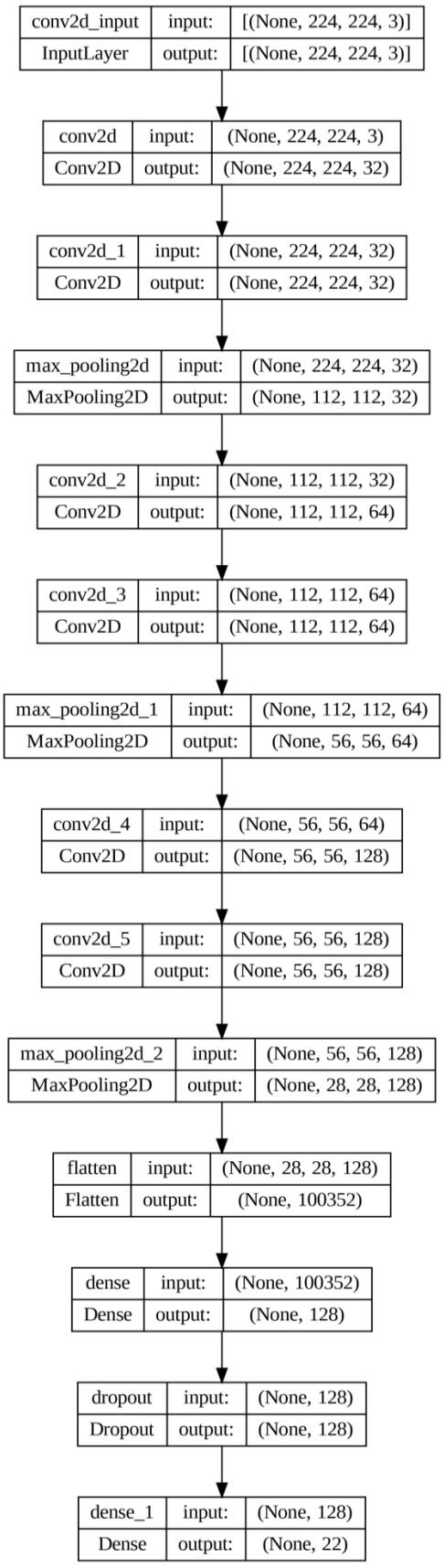


Figure 2.13 Architecture of the proposed custom CNN classifier.

2.3.3.3 Transfer learning models

We compared the performances of the proposed architectures with those of popular pre-trained CNN architectures using *transfer learning*. These pre-trained architectures achieve high accuracy on extensive specific datasets, such as ImageNet, and we can reuse them with minor modifications for our goal. In the literature, one can find examples where transfer learning models achieve accuracies comparable to custom-built models for various image classification tasks, while generally maintaining lower training times (Kim et al., 2020; Hossain et al., 2018; Shao et al., 2018). These models were initially trained for different purposes; however, one can benefit from the learned features encoded into their weights to build a powerful model for tackling different problems by fine-tuning the last layer or layers of the model. These layers, built and trained from scratch, are usually the *top* layers, that is, the fully connected parts of the model. This implies that the number of parameters that are required to be tuned is drastically reduced relative to the total number of parameters in the model, thereby saving valuable time during the model training stage. In the present work, VGG-16, VGG-19 (Simonyan & Zisserman, 2014), and Xception (Chollet, 2017) model architectures were used as starting points to build the transfer learning models. In each layer, the top layers were removed and replaced by a block that produced the mapping from the original feature map depth of 512 (for VGG-16/19) or 2048 (for Xception) to only 22 outputs (in the general model), representing our model classes. The new top layers are made of fully connected structures composed of a flattening layer, two dense layers with 1024 and 512 neurons, respectively, with ReLU activation functions, and dropout layers with a dropout probability of 0.3 to help reduce overfitting and an output layer of as many nodes as classes with softmax activations. The resulting architectures for the general model, including the input and output tensor sizes for each block, are shown in Figure 2.14. For the separate models, the only difference was in the number of outputs in the last dense layer.

a) b) c)

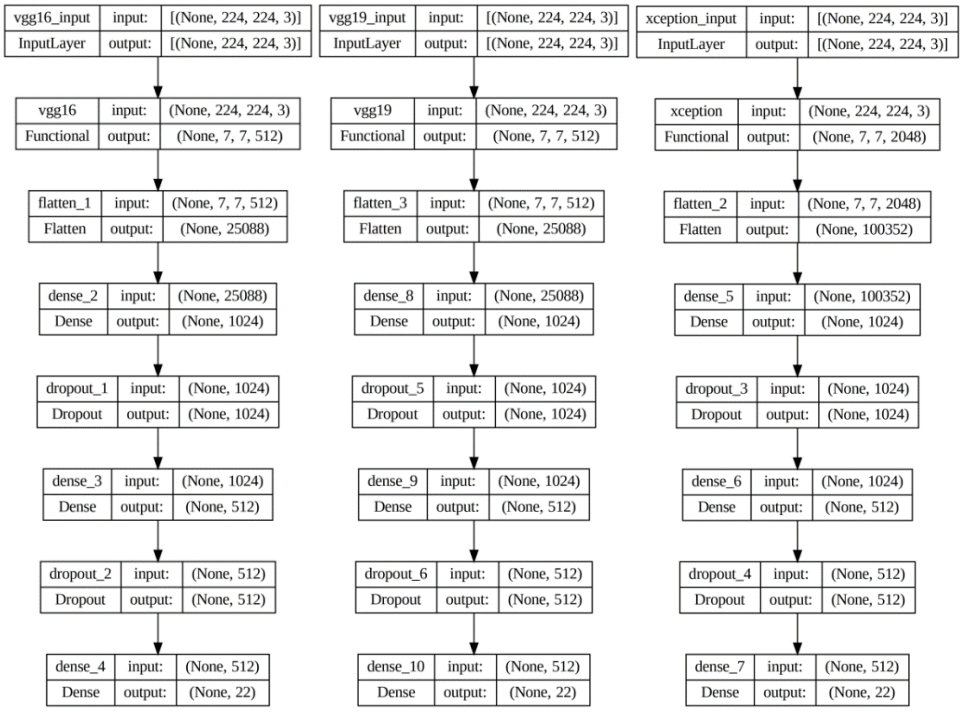


Figure 2.14 Proposed transfer learning architectures for classification: a) VGG-16 based, b) VGG-19 based, c) Xception based.

2.3.3.4 Regression models

For the regression approach, we implemented a model derived from a custom classifier, replacing the last fully connected part with a single layer of 128 nodes and ReLU activation, followed by a single node with linear activation to yield a real value. The architecture is shown in Figure 2.15.

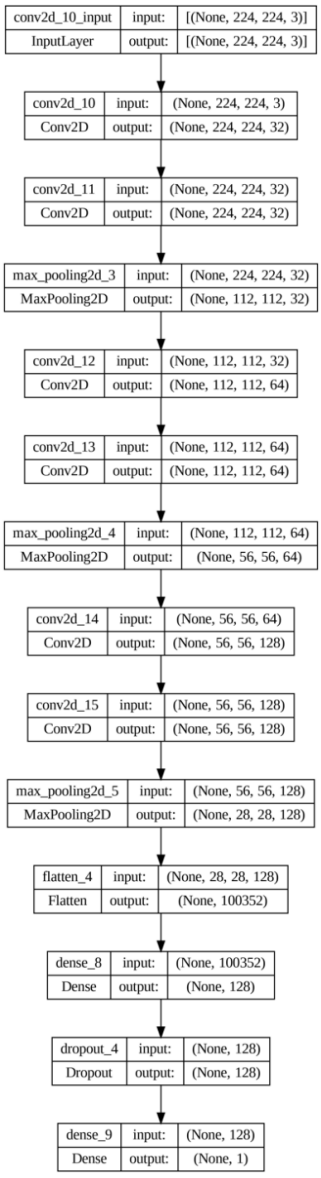


Figure 2.15 Architecture of the proposed regression model.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 CNN Models performance on 'Fantastico' during year 2021/2022

Table 3.1 summarises the obtained accuracies (training, validation, and test) for all CNN architectures. The four tested CNN architectures perform reasonably well, correctly classifying more than 98% of the test samples and achieving perfect classification for the training set. Three of the models produced test accuracies above 99%. The best-performing architectures were the VGG-based CNNs, with a test accuracy of 0.994 and a perfect training accuracy of 1 in the case of the VGG-16-based architecture.

MODEL	TRAINING ACCURACY	VALIDATION ACCURACY	TEST ACCURACY
CUSTOM	1	0.955	0.991
VGG-16 BASED	1	0.935	0.994
VGG-19 BASED	0.999	0.929	0.994
XCEPTION BASED	1	0.979	0.982

Table 3.1 Classification accuracies for the four tested CNN architectures.

The accuracy measures for the four models along the different epochs are shown in Figure 3.1. As observed, 100 epochs were enough in all cases to reach the accuracy plateau where validation accuracy did not further improve, and cross-entropy loss did not further decrease.

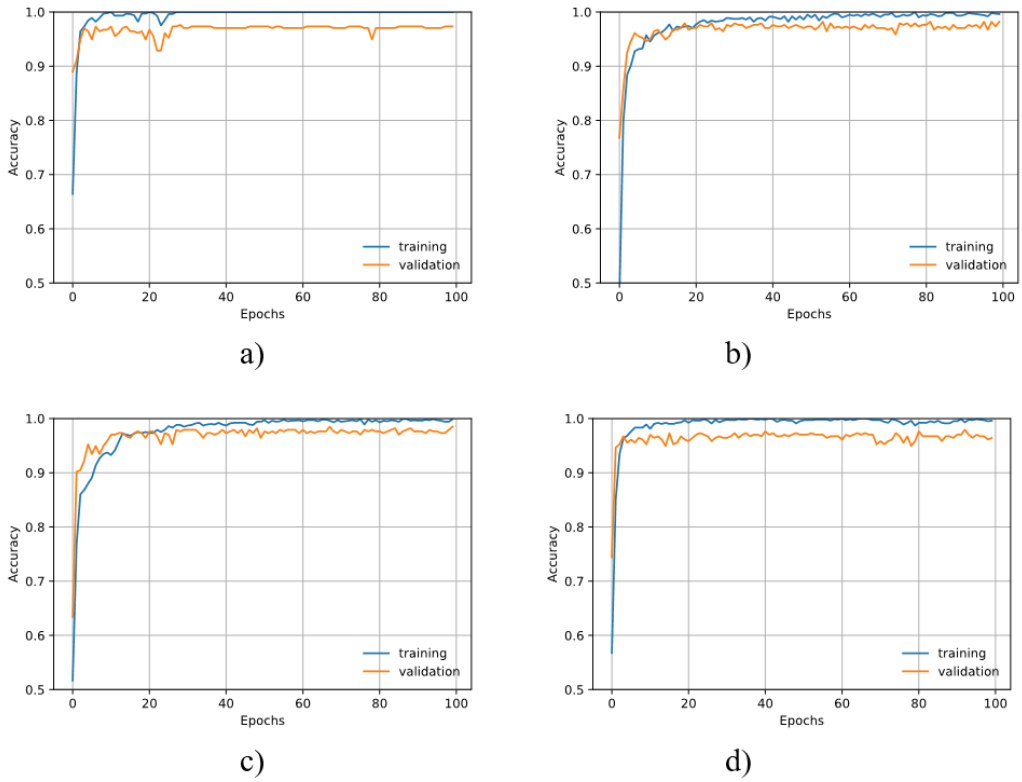


Figure 3.1 Accuracy graphs for a) custom model; b) VGG-16-based model; c) VGG-19-based model and d) Xception-based model.

Besides accuracy, it was interesting to analyze to which class were assigned those misclassified samples. This could be analyzed using confusion matrices, which aggregated all pairwise combinations of true and predicted classes. Each trained model used their best weights to generate a confusion matrix. The result is shown in Figure 3.2.

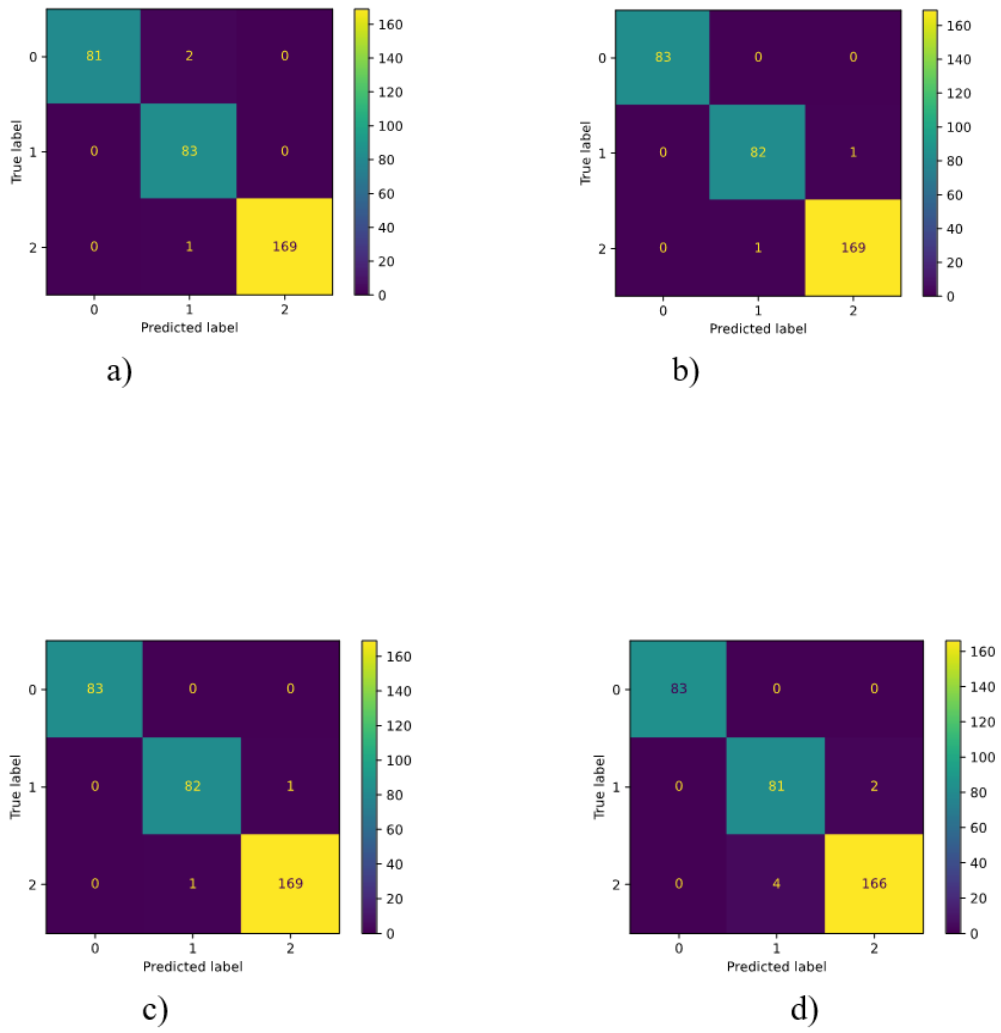


Figure 3.2 Confusion matrices in test data set for a) custom CNN model; b) VGG-16 based model; c) VGG-19 based model; d) Xception based model.

In summary, the general performance of the created models was good. Test accuracies were higher than 0.98 for all models, which is a satisfactory rate, given the number of captured images. All methods reached convergence (validation accuracy plateau) in less than 100 epochs. Training accuracies reached a maximum of 1, but this does not imply any overfitting because the validation performance did not drop during training. The test errors on unseen data were also consistent with the validation ones, although slightly lower. This was because the validation error was maximized to select the best weights of the network, but the test set was completely independent of the training procedure and therefore provided a more

realistic performance measure on unseen data. The best-performing architectures were the ones based on VGG-16 architecture (test accuracy = 0.994), as they failed to classify only 2 test examples out of the total 336 (see Figure 8b and 8c). Despite being simpler in terms of several parameters than transfer learning alternatives, the custom CNN model exhibits comparable results in test accuracy (0.991), failing to classify only one sample more than the transfer learning variants. This is probably due to the similar CNN architecture to VGG but with fewer convolutional and pooling blocks. Some authors who compare custom-built models with fine-tuned transfer learning variants also report higher accuracies in favor of the latter (Thenmozhi & Reddy, 2019). However, this depends on the data and the model design. The Xception-based model performs worse than the custom one despite producing near state-of-the-art performance in major image classification challenges such as ImageNet or JFT (Chollet, 2017). Thus, the present application does not seem to benefit from depthwise separable convolutions. The VGG-19-based architecture, however, achieves a test accuracy of only 0.6 % lower than the custom model. Although being different tasks, in general, the obtained accuracies are in line with other recent studies regarding fruit classification and grading using CNN models (Anupriya & Sri, 2022; Ibrahim et al., 2018; Parashar et al., 2022; Sustika et al., 2018). By looking at the confusion matrices, helpful information could be extracted too. On the one hand, the best-performing classes are 0 and 1 (extractions made in December 2021), with perfect class separations in some cases. This was probably due to the different levels of EO content (2.8788 g for the first extraction and 4.5158 g for the second). On the other hand, some models (especially the Xception-based one) become more confused between the second and the third/fourth extractions (corresponding to class labels 1 and 2, respectively). This is probably caused by the similarity between the features of bergamot fruits' images with higher EO contents. However, the accuracy rates are still highly satisfactory in this case. As reported by Marzocchi et al. (2019), during the fruit ripening stage, size, shape, color, taste, flavor, and aroma change, affecting significantly the composition of bergamot essential oil. It is mainly composed of volatile compounds (93-96%) and differs from those other citrus fruits in the high content of oxygenated compounds and lower hydrocarbon content (Giuffrè et al., 2020). Non-volatile

compounds, representing the remaining 4-7%, behave as a natural fixative of the perfumes, providing further indications and valuable information on the quality of the essential oil (Statti et al., 2004). Several studies have assessed the volatile and non-volatile compound composition of bergamot oil by evaluating several requirements (Dugo et al., 2013; Verzera et al., 2000; Giuffrè et al., 2020; Bourgou et al., 2012; Sawamura et al., 2006; Rowshan & Najafian 2015, Cautela et al., 2021). All these studies put evidence that obtaining an annual product with constant values of its component is very complex, depending on many factors such as cultivar, ripening stage, climate, the year of production, the age of the plant, agronomic techniques, etc. Preliminary analysis should be carried out to program harvesting when EO composition meets the best features. Analysis which are complex as well as expensive. However, quickly determining the period when the EO is highest in quantity makes it possible to have more essence available to make, if necessary, a mix of the amounts of EO in different batches in the right proportions to prepare a final product with the characteristics desired by customers (Giuffrè et al., 2020). Moreover, the encouraging results of this first study bode well for the possibility of extending this methodology to investigate qualitative characteristics of EO as well.

3.2 CNN models performance on 'Fantastico' and 'Femminello' during the year 2022/2023

The models were trained for 100 epochs in batches of 64 samples, using the Adam optimizer and minimizing the cross-entropy loss. In all cases, 100 epochs are enough to reach the accuracy plateau where validation accuracy does not further improve, and cross-entropy loss does not further decrease. The classification metrics on the test set (20 % of the samples) for Fantastico and Femminello cultivars are listed in Tables 3.2 and 3.3, respectively.

Table 3.2 Classification summary for Fantastico. The best results are highlighted in bold.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Custom	**0.768**	**0.782**	**0.768**	**0.769**
VGG-16 based	0.697	0.704	0.697	0.698
VGG-19 based	0.585	0.593	0.585	0.567
Xception based	0.600	0.625	0.600	0.600

Table 3.2 Classification summary for Femminello. The best results are highlighted in bold.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Custom	0.754	**0.762**	0,754	**0.755**
VGG-16 based	**0.756**	**0.762**	**0.756**	0.752
VGG-19 based	0.683	0.705	0.683	0.682
Xception based	0.468	0.460	0.468	0.440

Despite being simpler in architecture than the transfer learning alternatives, the custom model offered the best results in most cases. In particular, it outperformed all other models in all metrics for Fantastic and obtained the best precision and F1 score for Femminello. In this last case, it performs almost identically to the VGG-16 architecture on which it is based. One possible cause is that larger models tend to overfit the data while simpler ones do not. The fact that VGG-16 performs better than VGG-19 despite being simpler supports this hypothesis too. It is also convenient to analyze to which class assigned those misclassified samples. This can be visualized by composing confusion matrices, which aggregate all pairwise combinations of true and predicted classes. A confusion matrix was generated for each trained model, using their best weights. The results for Fantastico and Femminello are shown in Figures 3.3 and 3.4, respectively. These results show that

the main source of misclassification is the confusion between adjacent classes, i.e., samples obtained on close dates. This is a reasonable finding, as fruit images acquired in close time frames tend to be more similar among them and are, therefore, more likely to confuse the classifiers. Notably, for the Femminello cultivar, the Xception model confuses almost all images from the second extraction date (10/11/2022) with those of the first date (3/11/2022). In general, the Xception model performs poorly. This is because this particular learning problem does not benefit from the Depthwise Separable Convolutions of this method. Regarding regression, the same preprocessing and dataset splitting was done as for classification, except that a target variable is a real number (EO content in g), and the loss function was the minimization of mean squared error (MSE) between the target and the predicted value. We trained the models for 100 epochs, using a batch size of 64 as in the classification approach. Tables 3.4 and 3.5 summarise the performance results of the two regression models on the test set. Besides MSE, other metrics have been included: mean absolute error (MAE) and mean absolute percentage error (MAPE).

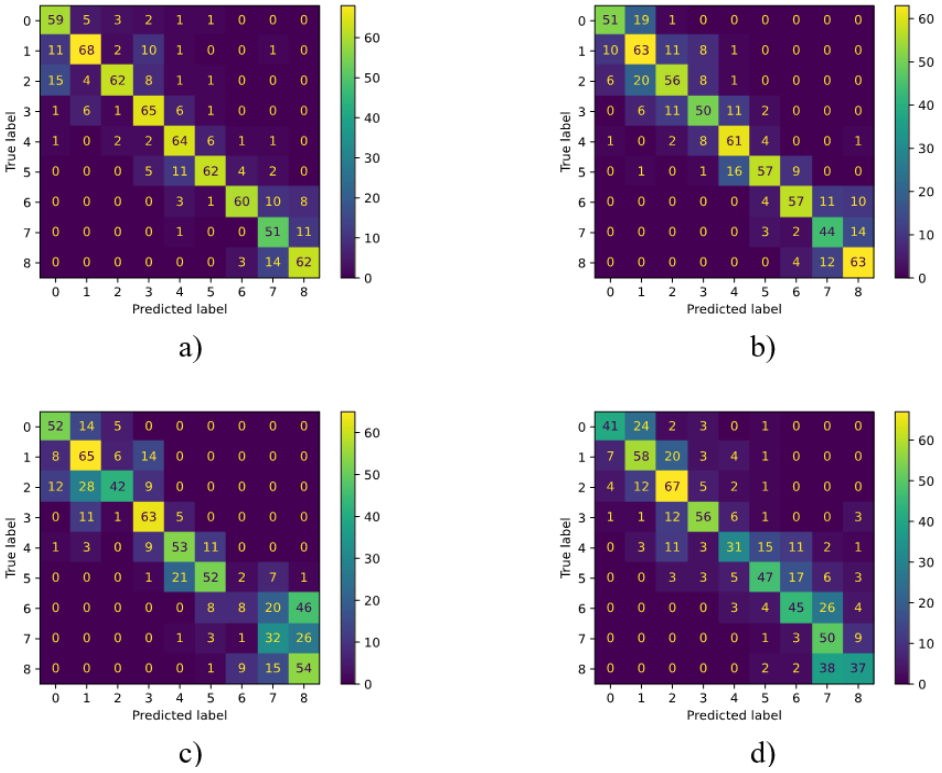


Figure 3.3 Confusion matrices for Fantastico using the different methods (test set): a) Custom model; b) VGG-16 based; c) VGG-19 based; d) Xception based.

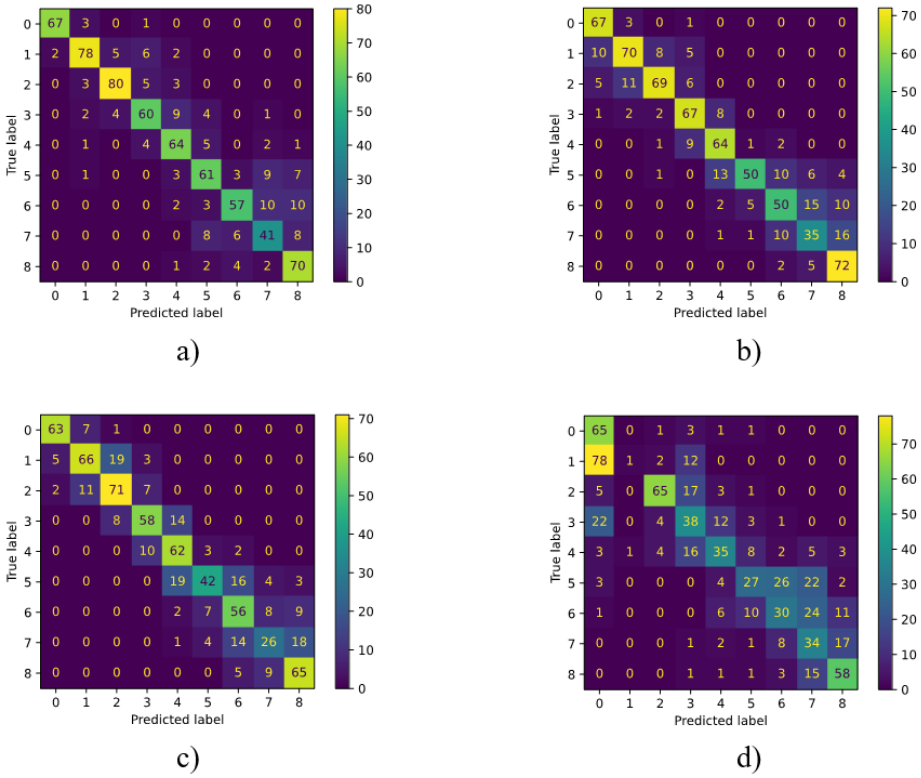


Figure 2.4 Confusion matrices for Femminello using the different methods (test set): a) Custom model; b) VGG-16 based; c) VGG-19 based; d) Xception based.

Table 3.4 Regression performance summary for Fantastico. The best results are highlighted in bold.

Model	MSE	MAE	MAPE
REG1	0.151	0.309	7.043
REG2	**0.138**	**0.289**	**6.668**

Table 3.5 Regression performance summary for Femminello. The best results are highlighted in bold.

Model	MSE	MAE	MAPE
REG1	**0.035**	**0.142**	**3.770**
REG2	0.105	0.267	7.264

From the results, it can be concluded that none of the models always performed better than the other. On the one hand, for the Fantastico cultivar, the REG1 model performed slightly better, but, on the other hand, for the Femminello data, the REG2 model performed substantially better (nearly 50 % reduction in absolute error metrics). This shows how, depending on the problem, applying dropout operations may help increase the generalization by avoiding local error minima. Concerning BEO, if the chemical composition of the volatile and non-volatile fractions has been extensively studied and is well known, as reported by Valussi et al. (2021) and Navarra et. al (2015), the same in-depth information is not available for the oil yield present in the fruit peel, which, in addition, also depend by the extraction method. In this perspective, Uddin et al., (2023) compared conventional extraction methods (hydro-distillation, steam distillation, etc.) and advanced methods (Supercritical CO_2, Subcritical water, Microwave-assisted, etc.) on Ginger essential oil. Mohanty et al., (2023) performed a comparative evaluation of hydro-distillation and solvent-free microwave extraction methods on three Curcuma species finding the last one to be more efficient in terms of oil yield. da Silva et al., (2022), show the advantages and disadvantages of the main methods proposed for the EO extraction of basil (*Ocimum basilicum* L.) and the extraction of essential oils from other sources.

However, there are no previous studies to establish a better method to estimate EO yield content in the bergamot peel. In the present research, Clevenger essential oil extraction was preferred to other extraction techniques because it allows very high precision as regards the measurement of the essential oil quantity and is cheap and easy to conduct in a laboratory. Other systems, such as Supercritical Extraction, are expensive in terms of the cost of the apparatus and extractive procedure, besides it is more difficult to manage in a laboratory. Industrial essential oil extraction involves hundreds of kilos of fruits for each harvest date and cannot be applied in experiments like the one we have designed here. Furthermore, it is not possible to extract the essence from a single fruit unless it is rubbed and pressed, but the error rate can be very high. Several studies have shown that the ripening stage, size, shape, and color of the fruit have a significant influence on the chemical composition of the essential oil (Marzocchi et al., 2019; Torbati et al., 2023), but little is known about the essential oil yield. The CCI values tested allowed a precise classification of the variability of essence content in the two cultivars studied, covering the whole harvesting period. The research also shows that the amount of essential oil is higher in the CV. Fantastico and the highest level occur in mid to late November. Whereas for cv Femminello this period is between the beginning and the middle of November. There may be many reasons for this condition, which is not the purpose of this research. However, it should be noted that this increase is associated with more intense rainfall days. From a general point of view, it is possible to observe a gradual and never abrupt transition in the quantity of the EO content. Finally, according to Ming et al. (2020), using smartphone cameras can allow an easy acquisition of high-quality images directly in the field. The most recent studies in similar fields are by Dos Santos et al. (2023), who determined eugenol in clove essential oil using a smartphone, an endoscopic camera, and a 3D printed device, and by Lebanov & Paull (2021), who used a smartphone-based miniaturized Raman spectrometer and machine learning for rapid identification and discrimination of adulterated essential oils (EOs). Hakonen & Beves (2018) developed a simple method for the authentication of edible oils using a 405 nm LED flashlight and a smartphone. This can contribute to introducing digital technology solutions to improve the agricultural sector's competitiveness and help stakeholders

make smarter decisions for their production processes. The challenge we faced was to calibrate the learning algorithms so that they would perform well when they were applied to unknown input data.

3.3 CNN models performance on data from 2021 to 2023

The models were trained for a maximum of 100 epochs in batches of 64 samples using the Adam optimizer and minimizing cross-entropy loss. In all cases, fewer than 100 epochs were sufficient to stop the training using early stopping, with a patience of 10 epochs. The classification metrics for the test set (20% of the samples) for the general, Fantastico, and Femminello models are presented in Tables 3.6, 3.7, and 3.8, respectively.

Table 3.6 Classification summary for the general model. The best results are highlighted in bold.

Model	Accuracy	Precision	Recall	F1-Score
Custom	0.694	**0.719**	0.694	0.694
VGG-16 based	**0.697**	0.716	**0.697**	**0.697**
VGG-19 based	0.648	0.675	0.648	0.643
Xception based	0.654	0.672	0.654	0.656

Table 3.7 Classification summary for Fantastico model. The best results are highlighted in bold.

Model	Accuracy	Precision	Recall	F1-Score
Custom	**0.811**	**0.814**	**0.811**	**0.810**
VGG-16 based	0.787	0.794	0.787	0.787
VGG-19 based	0.768	0.779	0.768	0.769
Xception based	0.708	0.732	0.708	0.707

Table 3.8 Classification summary for Femminello model. The best results are highlighted in bold.

Model	Accuracy	Precision	Recall	F1-Score
Custom	0.719	0.758	0.719	0.725
VGG-16 based	**0.778**	**0.802**	**0.778**	**0.780**
VGG-19 based	0.672	0.705	0.672	0.660
Xception based	0.679	0.707	0.679	0.680

We observed a clear advantage of the custom and VGG-16-based models. Overall, the best performer was the VGG-16-based model, which yielded the best results for the general and femminello models. Despite being simpler in architecture than the transfer learning alternatives, the custom model offers the best results for the Fantastico model. In particular, it outperformed all other models for Fantastico in all metrics and obtained the best precision for the general model. The general model performed almost identically to the VGG-16 architecture, using fewer parameters. The most complex models, such as VGG-19 and Xception, are less accurate. One possible reason is that larger models tend to overfit the data, whereas simpler models do not. The fact that VGG-16 performed better than VGG-19 despite being simpler supports this hypothesis. It is also convenient to analyze the class assigned to misclassified samples. This can be visualized by composing confusion matrices that aggregate all the pairwise combinations of the true and predicted classes. A confusion matrix is generated for each trained model using the best weights. The results of the Fantastico and Femminello models are shown in Figures 3.5, 3.6, and 3.7, respectively. These results indicate that the main source of misclassification is confusion between adjacent classes, that is, samples obtained on close dates. This is a reasonable finding, as fruit images acquired in close time frames tend to be more similar and are, therefore, more likely to confuse the classifiers. In general, the Xception model performed poorly. This is because this particular learning problem does not benefit from the depth-wise separable convolutions of this method. Regarding regression, the same pre-processing and dataset splitting was performed for classification (Table 2.5), except that the target variable was a real number (EO content in grams), and the loss function was the minimization of mean squared error (MSE) between the target and the predicted value. We trained the models for a maximum of 100 epochs, implementing the same early stopping procedure and using a batch size of 64, similar to that in the classification approach. Table 3.9 summarises the performance results of the general, Fantastico, and Femminello models for the test set. Besides MSE, other metrics have been included: mean absolute error (MAE) and mean absolute percentage error (MAPE).

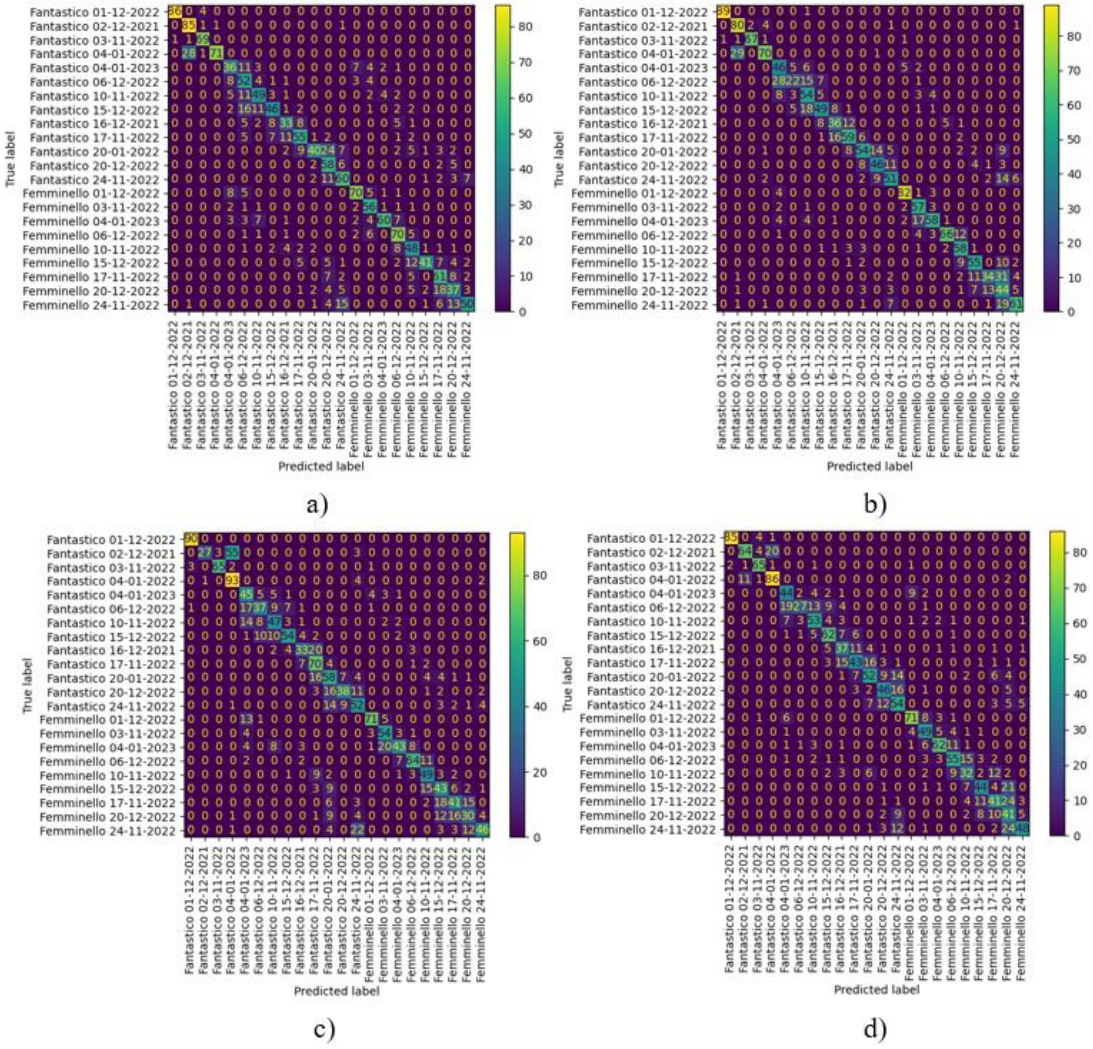


Figure 3.5 Confusion matrices for the general model using the different methods (test set): a) Custom model; b) VGG-16 based; c) VGG-19 based; d) Xception based.

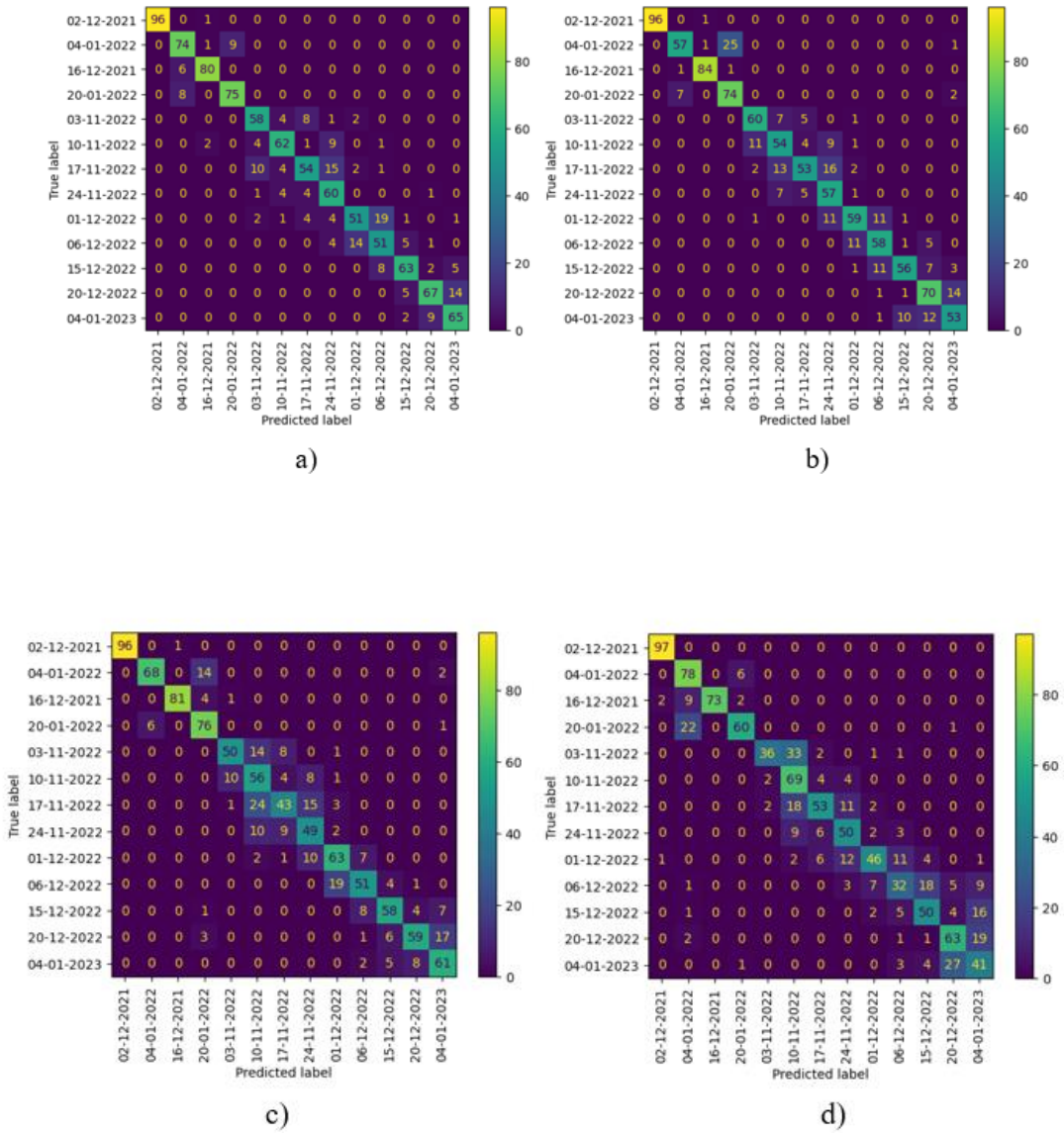


Figure 3.6 Confusion matrices for Fantastico using the different methods (test set): a) Custom model; b) VGG-16 based; c) VGG-19 based; d) Xception based.

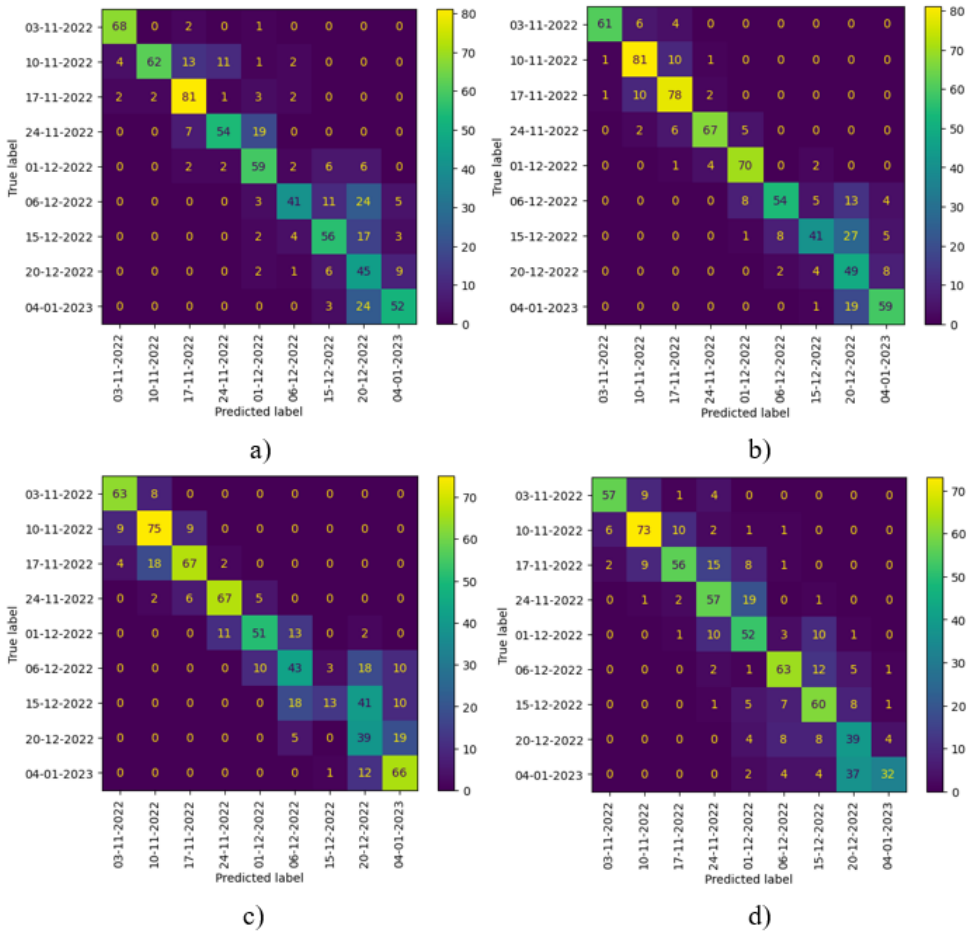


Figure 3.7 Confusion matrices for Femminello using the different methods (test set): a) Custom model; b) VGG-16 based; c) VGG-19 based; d) Xception based.

Metric	General model	Fantastico model	Femminello model
MSE	0.112	0.117	0.035
MAE	0.246	0.257	0.149
MAPE	6.326	6.386	4.024

Table 3.9 Regression performance summary on the test set.

The results showed that using separate regression models for the two cultivars was more convenient than using a single model alone. In particular, MSE on Femminello was reduced by a factor of 3.2 (> 2% MAPE improvement), whereas the Fantastico cultivar obtained roughly the same performance regardless of the approach. Concerning BEO, several studies (Dugo & Bonaccorsi, 2013; Verzera et al., 2000; Giuffrè et al., 2020; Bourgou et al., 2012; Sawamura et al., 2006; Rowshan & Najafian 2015; Cautela et al., 2021) show that obtaining an annual product with constant values is complex, depending on many factors such as cultivar, ripening stage, climate, the year of production, the age of the plant, agronomic techniques, etc. Furthermore, the chemical compositions of the volatile and non-volatile fractions have been extensively studied and are well known, as reported by Valussi et al. (2021) and Navarra et al. (2015); however, the same in-depth information is not available for the oil yield present in fruit peels, which also depends on the extraction method. From this perspective, Uddin et al. (2023) compared conventional extraction methods (such as hydro distillation and steam distillation) with advanced methods (such as Supercritical CO_2, Subcritical water, and microwave-assisted) for ginger essential oil. Mohanty et al. (2023) performed a comparative evaluation of hydro-distillation and solvent-free microwave extraction methods on three Curcuma species and found that the latter was more efficient in terms of oil yield. However, no previous studies have established a better method for estimating the EO yield in bergamot peel. In the present study, Clevenger essential oil extraction was preferred over other extraction techniques because it allows very high precision in the measurement of essential oil quantity, is cheap, and is easy to conduct in a laboratory. Other extraction methods, such as supercritical extraction, are expensive in terms of the cost of the apparatus and extractive procedure and are more challenging to manage in a laboratory. Industrial essential oil extraction involves hundreds of kilos of fruit for each harvest date and cannot be applied in experiments such as the one designed here. Furthermore, extracting the essence from a single fruit is impossible unless the fruit is rubbed and pressed. However, the error rate can be very high. The CCI values tested allowed for the precise classification of the variability of essence content in the two cultivars studied, covering the entire harvesting period. The study also showed that the

amount of essential oil was higher in Fantastico, and the highest level occurred in mid-to-late November. Meanwhile, for Femminello, this period was between the beginning and middle of November. There may be many reasons for this, from agronomic to weather-related, but this is beyond the scope of this study. However, it should be noted that this increase was associated with more intense rainfall. From a general perspective, it is possible to observe a gradual and never abrupt transition in the EO content. Finally, according to Ming et al. (2020), smartphone cameras can easily acquire high-quality images directly in the field. A low-cost portable inspection chamber was designed to eliminate the effects of natural light on image capture, optimize data collection from smartphones, create an environment that limits the influence of external factors, and standardize the data collected. This can be a helpful tool for the development of new technologies and postharvest operations. The most recent studies in similar fields were by Dos Santos et al. (2023), who determined eugenol in clove essential oil using a smartphone, an endoscopic camera, and a 3D printed device, and by Lebanov and Paull (2021), who used a smartphone-based miniaturized Raman spectrometer and machine learning for the rapid identification and discrimination of adulterated EO. Hakonen and Beves (2018) developed a simple method to authenticate edible oils using a 405 nm LED flashlight and a smartphone.

3.4 Software implementation strategy and convenience

Several studies on the relationship between essential oils or oil content and fruit ripening have recently been conducted to determine the optimal harvest date. Balasundram et al. (2006) found a significant correlation between total palm fruit oil and all color components, with black having the strongest correlation (r = -0.85), followed by red (r = 0.81), orange (r = 0.62), and yellow (r = 0.48). Wannes et al. (2009) reported that the essential oil and fatty acid compositions of myrtle fruits at different stages of ripening show large differences, suggesting that these differences could be due to the effects of harvest time and environmental conditions. Using a digital RGB camera, Tan et al. (2010) found that digital imaging techniques for determining the oil content of fresh fruit bunches could be successfully applied to a homogeneous population of palms that exhibit similar color changes during the ripening process. Bourgou et al. (2012) and Bhuyan et al. (2015) studied the variation in essential oils during citrus peeling and found that it depended on the stage and species. The stage of maximum EO yield was immature for C. limon, semi-mature for *C. sinensis* and *C. reticulata*, and mature for C. aurantium. Wu et al. (2013) observed a significant increase in EO during fruit development and ripening in C. medica. Rowshan and Najafian (2015) investigated the effect of the ripening stage on the chemical composition of essential oils extracted from the peel of bitter orange (Citrus aurantium). The results showed that the ripening stage affected the essential oil composition. Di Rauso et al., 2020 showed that the variation of peel EOs during fruit ripening depends on the origin and cultivar of lemons (maximum yield in November for 'Ovale di Sorrento' and 'Sfusato Amalfitano' lemons, whereas the peak was reached in December for 'Femminello Cerza' and 'Femminello Adamo'). Ghani et al. (2021) studied the essential oil content of grapefruit peels (Citrus paradisi var. red blush) during the color change stages and found significant changes during the fruit ripening stages. This can contribute to the introduction of digital technology solutions to improve the agricultural sector's competitiveness and help stakeholders make smart decisions regarding their production processes. The challenge is to calibrate learning algorithms to perform well when applied to unknown input data.

CHAPTER 4

MOBILE APPLICATION PROGRAMMING AND DEVELOPMENT

4.1 Interface implementation, programming, and development

The final step of the present research was to develop a mobile application for an Android device capable of providing a reliable estimate of essential oil content in bergamot fruits, thereby contributing significantly to research and production in the agricultural sector. The ultimate goal is to make neural network codes (example in Figure 4.1) directly usable via mobile applications. Android is an operating system employed in smartphones that are based on the Android platform. The development of Android applications utilizes the Android Studio tool, which has been created by Google (Ukhurebor & Aidonojie, 2021). The application was developed using the Android Studio development environment and taking advantage of the features of the Android framework. During the development process, advanced image analysis techniques were integrated to precisely evaluate the color and degree of ripeness of bergamot fruits. Artificial intelligence algorithms have been implemented to enhance the precision of the estimates obtained.

```
model = Sequential()
model.add(Conv2D(32, (3, 3), activation='relu', kernel_initializer='he_uniform', padding='same', input_shape=(224, 224, 3)))
model.add(Conv2D(32, (3, 3), activation='relu', kernel_initializer='he_uniform', padding='same'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(64, (3, 3), activation='relu', kernel_initializer='he_uniform', padding='same'))
model.add(Conv2D(64, (3, 3), activation='relu', kernel_initializer='he_uniform', padding='same'))
model.add(MaxPooling2D((2, 2)))
model.add(Conv2D(128, (3, 3), activation='relu', kernel_initializer='he_uniform', padding='same'))
model.add(Conv2D(128, (3, 3), activation='relu', kernel_initializer='he_uniform', padding='same'))
model.add(MaxPooling2D((2, 2)))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.2))
model.add(Dense(4, activation='softmax'))
print(model.summary())
tf.keras.utils.plot_model(model, to_file='Figure5.png', dpi=300)
# compile model

model.compile(optimizer="adam", loss='categorical_crossentropy', metrics=['accuracy'])
mcp_save = ModelCheckpoint('.model_custom_wts.hdf5', save_best_only=True, monitor='val_accuracy', mode='max')

history = model.fit(X_train, y_train, batch_size=64, epochs=100, callbacks=mcp_save,
                    validation_data=(X_val, y_val))
```

```
vgg16_model = Sequential()
vgg16_model.add(pretrained_model_vgg16)
vgg16_model.add(Flatten())
vgg16_model.add(Dense(128, activation='relu'))
vgg16_model.add(Dropout(0.3))
vgg16_model.add(Dense(4, activation='softmax'))

vgg16_model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

mcp_save = ModelCheckpoint('.model_vgg16_wts.hdf5', save_best_only=True, monitor='val_accuracy', mode='max')

history_vgg16 = vgg16_model.fit(X_train, y_train, batch_size=64, epochs=100, callbacks=mcp_save,
                validation_data=(X_val, y_val))
print("history_vgg16" , history_res)
plt.plot(history_vgg16.history['accuracy'], label="training")
plt.plot(history_vgg16.history['val_accuracy'], label="validation")
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.ylim([0.5, 1])
plt.grid()
plt.legend(loc='lower right')
plt.show()
vgg_accuracy = vgg16_model.evaluate(X_test, y_test)
vgg_accuracy
vgg16_model.save('modelo_vgg16')
from keras.utils  import Xception

pretrained_model_xc = Xception(weights="imagenet", include_top=False, input_shape=(224,224,3))
# congelamos pesos
for layer in pretrained_model_xc.layers:
        layer.trainable=False
```

Figure 4.1 Part of the code sequence for programming CNN

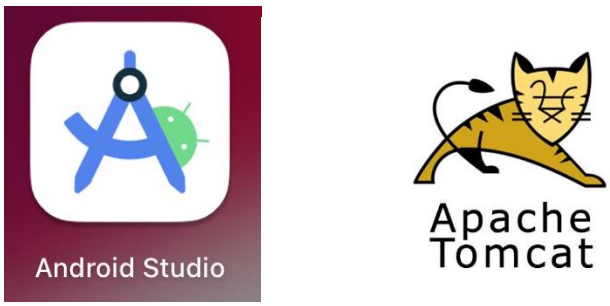


Figure 4.2 Android studio logo and Tomcat used for implementation.

During development, the Python programming language and the Tomcat web server were also used. Tomcat played the role of the webserver to host certain functionalities while Python was used to exploit its potential with particular emphasis on the use of the Keras, Matplotlib, TensorFlow, and Panda's libraries. Keras was used for neural network development, TensorFlow as a machine learning framework, Matplotlib for data visualization, and Pandas for efficient data manipulation. This combination of libraries enabled advanced management of machine learning operations and detailed analysis of the results obtained during application development. The application was designed and implemented using

RESTful API techniques, a software architecture approach that promotes the creation of scalable and interoperable web services. RESTful APIs allow communication and data exchange between the application client and the server through HTTP requests, following clear and standardized principles.

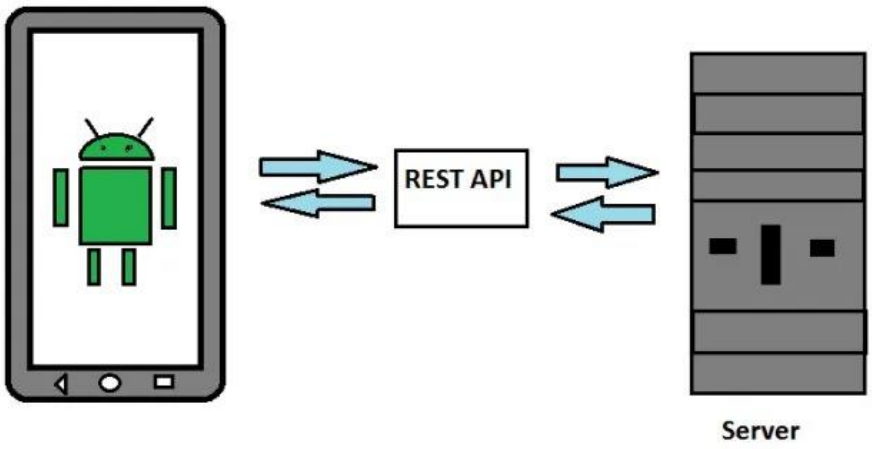


Figure 4.3 Example of communication between client (App) and server via API rest

In the RESTful API architecture, the client (application) initiates communication by sending requests to the server to access or manipulate resources through standard HTTP operations. The server holds and manages resources, responds to requests by performing specific operations on the resources, and provides agreed-upon data representations to the client. Communication occurs via URIs (Uniform Resource Identifiers), and the server maintains an information-free state between requests. This standardized model simplifies client-server interaction in RESTful APIs. During the development of the application, the Java programming language played a crucial role in implementing key functionalities. The `MainActivity` in Java emerged as the central component, managing the initial user interaction and initialization operations. Activity is a core component in the Android framework and refers to an application screen that users can interact with. Through the MainActivity file, the initialization operations, the arrangement of the user interface elements, and the management of events are defined and managed. The `build. Gradle' file was used to configure dependencies and compilation settings, ensuring

consistency in the development environment. The `AndroidManifest.xml` document provided essential information about the app, and the integration of services in Java-supported background operations, ensuring the continuity of functionalities even when the app was not actively in the foreground. The synergistic use of these elements contributed to an organized and well-functioning structure of the application. In conclusion, the application was successfully developed, testimony to the effectiveness of the methodologies adopted and the techniques employed by the development team. Strategic decisions, from programming language usage to user interface design choices, have proven to be carefully considered and beneficial to the development of the App. The approach and use of advanced tools contributed to an efficient development cycle and the creation of an application that meets the needs of farmers in this sector. Fig. 4.4 shows the final interface of the App and its use. Once the photo of the bergamot fruit has been taken, the color class that comes closest to that fruit is returned, and the respective oil content indicates a useful estimate for harvesting. As can be seen, the image below also shows the name given to the Mobile App, called "*Essential*".

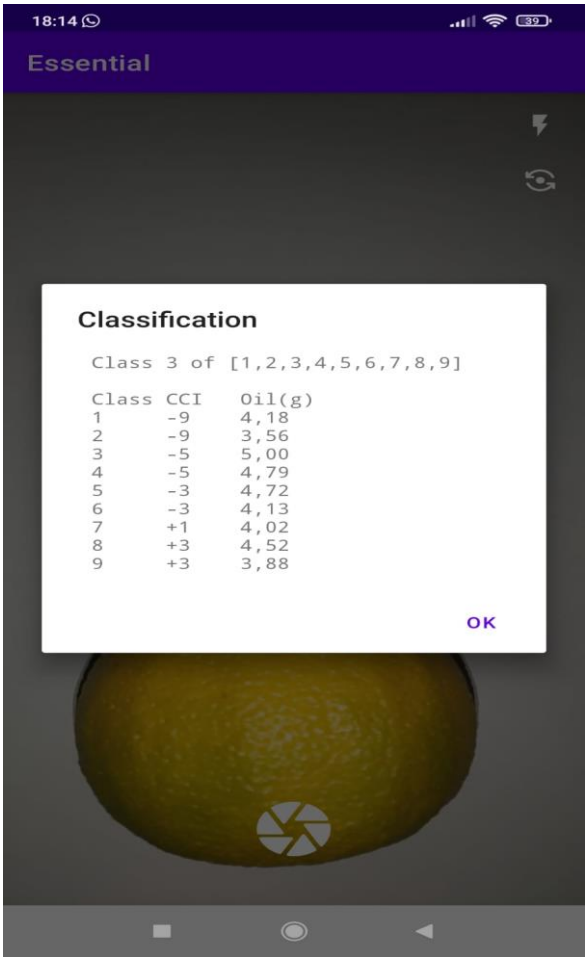


Figure 4.4 Opening the camera and returning the classes and their quantities in essential oil.

CHAPTER 5

CONCLUSIONS AND FUTURE PERSPECTIVES

In this study, we have undertaken an exploration into the cutting-edge domain of employing deep learning techniques, specifically concentrating on convolutional neural networks (CNNs), in the development of an application tailored for Bergamot. Through a meticulous analysis of specific datasets and the implementation of advanced models, we have achieved promising outcomes in the classification and processing of images associated with this invaluable agricultural resource. The resultant application underscores the efficacy of deep learning within the agricultural context, thereby laying the foundation for novel prospects in crop management. The precision and expeditiousness achieved in classifying Bergamot images suggest that deep learning techniques can be effectively harnessed to optimize agricultural processes, consequently enhancing productivity, and mitigating human errors. Looking towards the future, the utilization of deep learning techniques in agriculture emerges as a promising field for further development. The perspectives revolve around the broader application and widespread adoption of such methodologies to elevate agricultural management. The application's capacity to predict oil content based on fruit color represents a notable advancement in the application of deep learning techniques within the agricultural context. This particular feature introduces an innovative approach to estimating the oil content of Bergamot fruits solely based on their color characteristics. The incorporation of deep learning models, specifically convolutional neural networks (CNN), empowers the application to autonomously learn and recognize intricate patterns associated with the oil content in fruits. The capability to make accurate predictions solely from fruit color is a significant boon for farmers, enabling swift and non-invasive assessments of potential yield and crop quality. Finally, this study has some limitations, mainly related to the impossibility of quantifying the essence yield of every fruit and accounting for all agronomic variables linked to yield variation. In addition, mobile phone cameras only operate in the visible spectrum; therefore, developing systems that operate in non-visible absorption bands, such as NIR or SWIR, would certainly improve the accuracy of

the results. Moreover, the output of this prediction model can be implemented through further research. Beyond simplifying the evaluation process, this functionality has the potential to optimize harvesting and post-harvest management phases by providing timely and detailed insights into fruit properties. Moreover, the use of visual data for predicting oil content can serve as a pivotal decision-making tool for farmers, supporting targeted choices in the marketing and processing of agricultural products. Ultimately, the application's proficiency in predicting oil content based on fruit color epitomizes the innovative application of deep learning in the agricultural sector, offering practical and sophisticated support to optimize the management of Bergamot crops.

6 ACKNOWLEDGEMENTS

This research study was funded by the Italian Ministry of University and Research (MUR) in the framework of a doctorate grant funded within the National Operational Program "Research and Innovation 2014-2020, Innovative Doctorates with Industrial Characterisation". PhD course in Agricultural, Food, and Forestry Sciences (SAAF) at the University Mediterranea of Reggio Calabria (XXXVI cycle). Thankful to "Azienda Agricola Fratelli Foti" and "Azienda Agricola Patea" for providing bergamot fruit free of charge during the whole period of sampling.

I wish to extend my sincere appreciation to ETSE - Escola Tècnica Superior d'Enginyeria (Universitat de València) and Nanosoft S.r.l. (Cosenza) for their invaluable contributions characterized by high standards of excellence and unwavering support at every juncture of our collaboration. Their steadfast commitment has been instrumental in enhancing the quality and success of this project.

7 REFERENCES

Adorisio, S., Muscari, I., Fierabracci, A., Thuy, T. T., Marchetti, M. C., Ayroldi, E., Delfino, D. V., Adorisio, S., Muscari, I., Fierabracci, A., Thuy, T., Marchetti, M. C., Ayroldi, E., & Delfino, D. V. (2023). Biological effects of bergamot and its potential therapeutic use as an anti-inflammatory, antioxidant, and anticancer agent. *Pharmaceutical Biology*, *61*(1), 639–646. https://doi.org/10.1080/13880209.2023.2197010

Afonso, P., & Barros, F. De. (2021). *Plant diseases recognition on images using convolutional neural networks : A systematic review*. *185*(April). https://doi.org/10.1016/j.compag.2021.106125

Ai, C. (2021). *Artificial intelligence : A powerful paradigm for scientific research Arti-fi social intelligence : A powerful paradigm for scientific research*. https://doi.org/10.1016/j.xinn.2021.100179

Albahar, M. (2023). *A Survey on Deep Learning and Its Impact on Agriculture : Challenges and Opportunities*.

Alibabaei, K., Gaspar, P. D., Monteiro, J., & Lopes, C. M. (2022). *A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities*. 1–43.

Altalak, M., Ammad, M., Alajmi, A., & Rizg, A. (2022). *Applied Sciences Smart Agriculture Applications Using Deep Learning Technologies : A Survey*.

Alzubaidi, L., Zhang, J., Humaidi, A. J., Dujaili, A. Al, Duan, Y., Shamma, O. Al, Santamaría, J., Fadhel, M. A., Amidie, M. Al, & Farhan, L. (2021). Review of deep learning : concepts, CNN architectures, challenges, applications, future directions. In *Journal of Big Data*. Springer International Publishing. https://doi.org/10.1186/s40537-021-00444-8

Amato, A., Castellotti, T., Gaudio, F., Gaudio, G., Lovecchio, R., Pupo D'Andrea, M.R., Peluso, R., 2013.L'agricoltura nella Calabria in cifre 2012; INEA– Istituto Nazionale di Economia Agraria: Roma, Italy.

http://dspace.crea.gov.it/handle/inea/742 .Accessed May 25, 2022

Anderson, N.T., Walsh, K.B.,Wulfsohn, D., 2021. Technologies for Forecasting Tree Fruit Load and Harvest Timing—from Ground. Agronomy (Basel). 11:1409.

Anello, M., Mateo, F., Bernardi, B., Benalia, S., Zimbalatti, G., Blasco, J., Gómez-Sanchis, J. Is it possible to do a reliable assessment of bergamot color in the field with a smartphone camera? VII International Conference on Safety, Health, and Welfare in Agriculture and Agro-food Systems (Ragusa SHWA). 6 - 9 September 2023 Ragusa Ibla, Italy. "Unpublished results".

Anupriya, K. and Sri, G.M., 2022. Fruit Freshness Detection Using Machine Learning. In Cognitive Informatics and Soft Computing (pp. 633-642). Springer, Singapore.

Bagetta, G., Antonio, L., Rombolà, L., Amantea, D., Russo, R., Berliocchi, L., Sakurada, S., Sakurada, T., Rotiroti, D., & Tiziana, M. (2010). Fitoterapia Neuropharmacology of the essential oil of bergamot. *Fitoterapia*, *81*(6), 453–461. https://doi.org/10.1016/j.fitote.2010.01.013

Bal, F. (2021). *Review of machine learning and deep learning models in agriculture*. *05*(02), 309–323. https://doi.org/10.35860/iarej.848458

Balasundram, S.K., Robert P.C. and Mulla, D.J., 2006. Relationship Between Oil Content and Fruit Surface Color in Oil Palm (Elaeis guineensis Jacq.). Journal of Plant Sciences, 1: 217-227.

Bannerjee, G., Sarkar, U., Das, S., & Ghosh, I. (2018). *Artificial Intelligence in Agriculture : A Literature Survey*. *June*.

Benalia, S., Bernardi, B., Cubero, S., Leuzzi, A., Larizza, M., Blasco, J., 2015. Preliminary trials on Hyperspectral imaging implementation to detect Mycotoxins in dried figs. Chemical Engineering Transactions, 44, pp. 157 – 162. DOI: 10.3303/CET1544027

Benos, L., Tagarakis, A., Dolias, G., & Remigio, B. (2021). *Machine Learning in*

Agriculture : A Comprehensive. *May*. https://doi.org/10.3390/s21113758

Bernardi, B., Cubero, S., Benalia, S., Zimbalatti, G., & Blasco, J. (2021). *Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS / NIR hyperspectral images*. *187*. https://doi.org/10.1016/j.compag.2021.106252

Bhuyan, N., Barua, P.C., Kalita, P., Saikia, A., 2015. Physico-chemical variation in peel oils of Khasi mandarin (Citrus reticulata Blanco) during ripening. Indian J. Plant. Physiol. 20, 227–231.

Binkhonain, M., & Zhao, L. (2023). Machine Learning with Applications A machine learning approach for hierarchical classification of software requirements. *Machine Learning with Applications*, *12*(February), 100457. https://doi.org/10.1016/j.mlwa.2023.100457

Blahnik, V. and Schindelbeck, O. 2021. Smartphone imaging technology and its applications. Advanced Optical Technologies, vol. 10, no. 3, pp. 145-232. https://doi.org/10.1515/aot-2021-0023

Bochtis, D. (2021). *Machine Learning in Agriculture : A Comprehensive Updated Review*. 1–55.

Bourgou, S., Rahali, F.Z., Ourghemmi, I., Saïdani Tounsi, M., Ourghemmi, I., Tounsi, M.S. 2012. Changes of peel essential oil composition of four Tunisian citrus during fruit maturation. The Scientific World Journal.

Brasse, J., Rebecca, H., Maximilian, B., Mathias, F., & Irina, K. (2023). Explainable artificial intelligence in information systems : A review of the status quo and future research directions. *Electronic Markets*, 1–30. https://doi.org/10.1007/s12525-023-00644-5

Braun, A., Colangelo, E., & Steckel, T. (2023). ScienceDirect Farming in the Era of Industrie 4 . 0. *Procedia CIRP*, *72*(March), 979–984. https://doi.org/10.1016/j.procir.2018.03.176

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). *Machine Learning for Fluid Mechanics*.

By-products, B., Vedova, L. Della, Gado, F., Vieira, T. A., Grandini, N. A., Pal, T. L. N., Siqueira, J. S., Carini, M., Bombardelli, E., Correa, C. R., Aldini, G., & Baron, G. (2023). *Chemical, Nutritional, and Biological Evaluation*.

Campolo, D. (2014). *The Recovery of the Historic Centres in the Metropolitan Area of Reggio Calabria through the Use of Bergamot*. *11*, 514–517. https://doi.org/10.4028/www.scientific.net/AEF.11.514

Carnagie, J.O., Prabowo, A.R., Budiana, E.P., Singgih, I.K. (2022). Essential Oil Plants Image Classification Using Xception Model, Procedia Computer Science, 204, 395-402, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2022.08.048

Cautela, D., Pastore, A., Ferrari, G., Laratta, B., Onofrio, N. D., Luisa, M., Servillo, L., & Castaldo, D. (2021). Industrial Crops & Products Global warming threatens the world production of bergamot essential oil. *Industrial Crops & Products*, *172*(August), 113986.

https://doi.org/10.1016/j.indcrop.2021.113986

Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).

Cohen, B.; Edan, Y.; Levi, A.; Alchanatis, V., 2022. Early Detection of Grapevine (Vitis vinifera) Downy Mildew (Peronospora) and Diurnal Variations Using Thermal Imaging. Sensors, 22, 3585. https://doi.org/10.3390/s22093585

Costopoulou, C., Ntaliani, M., & Karetsos, S. (2016). *Studying Mobile Apps for Agriculture Studying Mobile Apps for Agriculture*. *December*. https://doi.org/10.9790/0050-03064449

Coulibaly, S., Kamsu-foguem, B., Kamissoko, D., & Traore, D. (2022). Intelligent Systems with Applications Deep learning for precision agriculture : A bibliometric analysis. *Intelligent Systems with Applications*, *16*(July), 200102. https://doi.org/10.1016/j.iswa.2022.200102

Craik, A., He, Y., Pal, H., Tripathi, S., Khatri-chhetri, A., & Sapkota, T. B.

(2019). *Deep learning and its role in smart agriculture Deep learning and its role in smart agriculture*. https://doi.org/10.1088/1742-6596/1399/4/044109

Cubero S., 2012. Diseño e implementación de nuevas tecnologías basadas en visión artificial para la inspección no destructiva de la calidad de fruta en campo y mínimamente procesada. Tesis Doctoral. Universidad Politecnica de Valencia.

Cubero, S., Albert, F., Prats-moltalb, M., & Fern, D. G. (2017). *ScienceDirect Application for the estimation of the standard citrus color index (CCI) using image processing in mobile devices*. *7*, 2–13. https://doi.org/10.1016/j.biosystemseng.2017.12.012

Dastres, R., Soori, M., Neural, A., Systems, N., & Journal, I. (2021). *Artificial Neural Network Systems To cite this version : HAL Id : hal-03349542*. *21*(2), 13–25.

Deng, J. et al., 2009. Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on computer vision and pattern recognition. pp. 248–255.

Dhanya, V. G., Subeesh, A., Kushwaha, N. L., Kumar, D., Kumar, T. N., Ritika, G., & Singh, A. N. (2022). Artificial Intelligence in Agriculture Deep learning-based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture*, *6*, 211–229. https://doi.org/10.1016/j.aiia.2022.09.007

Di, L., Iacopetta, D., Cappello, A. R., Gallucci, G., Martello, E., Fiorillo, M., Dolce, V., & Sindona, G. (2013). Hypocholesterolemic activity of 3-hydroxy-3- methyl-glutaryl flavanones enriched fraction from bergamot fruit (Citrus bergamia): ‘“ In vivo ”’ studies. *Journal of Functional Foods*, *7*, 558–568. https://doi.org/10.1016/j.jff.2013.12.029

Di, L., Lucia, D., Luciana, B., Maria, D. V., Angelo, G., Clotilde, M. G., Capocasale, M., Poiana, M., Urso, S. D., & Caridi, A. (2020). Vinegar production from Citrus bergamia - products and preservation of bioactive compounds. *European Food Research and Technology*, *0123456789*. https://doi.org/10.1007/s00217-020-03549-1

Di Donna, L., Bartella, L., De Vero, L., Gullo, M., Giuffrè, A.M., Zappia, C., Capocasale, M., Poiana, M., D'Urso, S., Caridi, A., 2020. Vinegar production from Citrus bergamia by-products and preservation of bioactive compounds. Eur. European Food Research and Technology 246, 1981-1990. https://doi.org /10.1007/s00217-020-03549-1

Di Rauso Simeone, G., Di Matteo, A., Rao, M.A., Di Vaio, C., 2020. Variations of peel essential oils during fruit ripening in four lemons (Citrus limon (L.) Burm. F.) cultivars. J. Sci. Food Agric. 100, 193–200.

Directions, F. (2023). *Theoretical Understanding of Convolutional Neural Network :*

Dugo, G.; Bonaccorsi, I, 2013. Citrus Bergamia: Bergamot and its Derivatives; CRC Press: Boca Raton, FL, USA, 2013; ISBN 9781439862278.

Dos Santos, I. C., Schlesner, S. K., de Moraes, D. P., Fereira, D. F., Voss, M., Giacomelli, S. R., & Barin, J. S. (2023). Smartphone-based rapid and low-cost method for the determination of eugenol content of clove essential oil. [Método rápido e de baixo custo empregando smartphone para a determinação de eugenol em óleo essencial de cravo] Ciencia Rural, 53(10). https://doi:10.1590/0103-8478cr20220498

El-Attar, N.E., Hassan, M.K., Alghamdi, O.A., Award, A.W. 2020. Deep learning model for classification and bioactivity prediction of essential oil-producing plants from Egypt. Sci Rep 10, 21349. https://doi.org/10.1038/s41598-020-78449-1

Elgendy, M. Deep Learning for Vision Systems; Simon and Schuster: New York, NY, USA, 2020

Fajardo Muñoz, S.E., Freire Castro, A.J., Mejía Garzón, M.I., Páez Fajardo, G.J. Páez Gracia, G.J. (2023), Artificial intelligence models for yield efficiency optimization, prediction, and production scalability of essential oil extraction processes from citrus fruit exocarps. Front. Chem. Eng. 4:1055744. doi: 10.3389/fceng.2022.1055744

Fazari, A., Pellicer-Valero, O. J., Gómez-Sanchıs, J., Bernardi, B., Cubero, S., Benalia, S., Zimbalatti, G., Blasco, J. (2021). Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images. Computers and Electronics in Agriculture, 187 doi:10.1016/j.compag.2021.106252

Franco, G. Di, & Santurro, M. (2021). research. *Quality & Quantity*, *55*(3), 1007–1025. https://doi.org/10.1007/s11135-020-01037-y

Ghani, A., Mohtashami, S., Jamalian, S., 2021. Peel essential oil content and constituent variations and antioxidant activity of grapefruit (Citrus × paradisi var. red blush) during color change stages. J. Food Meas. Charact., 15, 4917–4928.

Gioffrè, G., Ursino, D., Laura, M., Labate, C., & Giuffrè, A. M. (2021). *The peel essential oil composition of bergamot fruit (Citrus bergamia, Risso) of Reggio Calabria (Italy): a review*. *February*. https://doi.org/10.9755/ejfa.2020.v32.i11.2197

Giuffrè, A. M. (2019). *antioxidants Bergamot (Citrus bergamia, Risso): The effects of Cultivar and Harvest Date on Functional Properties of Juice and Cloudy Juice*. *July*. https://doi.org/10.3390/antiox8070221

Gonzalez-Gonzalez, M.G., Blasco, J., Cubero, S., Chueca, P., 2021. Automated detection of tetrarchs urticae Koch in citrus leaves based on color and vis/nir hyperspectral imaging. Agronomy, 11 (5), art. no. 1002. DOI: 10.3390/AGRONOMY1105100

Guliyev, H. (2023). Research in Globalization Artificial intelligence and unemployment in high-tech developed countries : New insights from the dynamic panel data model. *Research in Globalization*, *7*(May), 100140. https://doi.org/10.1016/j.resglo.2023.100140

Gupta, N. (2013). *Artificial Neural Network*. *3*(1), 24–29.

Hakonen, A., & Beves, J. E. (2018). Hue parameter fluorescence identification of edible oils with a smartphone. ACS Sensors, 3(10), 2061-2065.

https://doi:10.1021/acssensors.8b00409

Hossain, M.S., Al-Hammadi, M. and Muhammad, G., 2018. Automatic fruit classification using deep learning for industrial applications. IEEE Transactions on Industrial Informatics, 15(2), pp.1027-1034.

Hunter, M., 2010. Essential Oils: Art, Agriculture, Science, Industry, and Entrepreneurship; Nova Science Publishers, Inc.: Hauppauge, NY, USA.

Ibrahim, Z., Sabri, N. and Isa, D., 2018. Palm oil fresh fruit bunch ripeness grading recognition using convolutional neural network. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(3-2), pp.109-113.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning (pp. 448-456).

ISO 9235:2013 Aromatic natural raw materials - Vocabulary (ISO 9235:2013) International Organization for Standardization (ISO), Geneva.

Jia, W., Tian, Y., Luo, R., Zhang, Z., Lian, J., & Zheng, Y. (2020). Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot. Computers and Electronics in Agriculture, 172 doi:10.1016/j.compag.2020.105380

Jiménez-Cuesta, M.J., Cuquerella, J., Martínez-Jávega, J.M., 1981. Determination of a color index for citrus fruit degreening. In *Proceedings of the International Society of Citriculture*, Vol. 2, 750-753.

Journal, I., Advance, O., & Shiruru, K. (2017). *AN INTRODUCTION TO ARTIFICIAL*. *September 2016*.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). A review of the use of convolutional neural networks in agriculture. Journal of Agricultural Science, 156(3), 312-322. https://doi.org /10.1017/S0021859618000436

Kamilaris, A., & Prenafeta-boldú, F. X. (2018). *Deep learning in agriculture : A*

La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005),
Fondo Sociale Europeo, Azione I.1 “Dottorati Innovativi con caratterizzazione Industriale”

UNIONE EUROPEA
Fondo Sociale Europeo