Images and CNN applications in smart agriculture

Mohammad El Sakka[a*] and Josiane Mothe[a] and Mihai Ivanovici[b]

[a]*IRIT UMR5505 CNRS, Université de Toulouse, France*
{mohammad.el-sakka / Josiane.Mothe}@irit.fr

[b]*Transilvania University of Brasov, Romania*
Mihai.Ivanovici@unitbv.ro

Abstract

In recent years, the agricultural sector has undergone a revolutionary shift toward 'smart farming', integrating advanced technologies to strengthen crop health and productivity significantly. This paradigm shift holds profound implications for food safety and the broader economy. At the forefront of this transformation is deep learning, a subset of artificial intelligence based on artificial neural networks, has emerged as a powerful tool in detection and classification tasks. Specifically, Convolutional Neural Networks (CNNs), a specialized type of deep learning and computer vision models, demonstrated remarkable proficiency in analyzing crop imagery, whether sourced from satellites, aircraft, or terrestrial cameras. These networks often leverage vegetation indices and multispectral imagery to enhance their analytical capabilities. Such model contribute to the development of systems that could enhance agricultural outcomes (Tian et al., 2020). This review encapsulates the current state of the art in using CNNs in agriculture. It details the image types utilized within this context, including, but not limited to, multispectral images and vegetation indices. Furthermore, it catalogs accessible online datasets pertinent to this field. Collectively, this paper underscores the pivotal role of CNNs in agriculture and highlights the transformative impact of multispectral images in this rapidly evolving domain.

Keywords: Deep learning; Multispectral images; Vegetation index; Agricultural datasets; Smart farming; Smart agriculture

Introduction

Crops are not only used to produce food but also to feed livestock, create paper and furniture, or produce fuel. Plant diseases, pests, and weed invasions directly affect these productions or their yield. The Food and Agriculture Organization of the United Nations estimates that $220 billion USD are lost worldwide due to crop diseases, which accounts for more than 20% of crops produced a year (Food and Agriculture Organization of the United Nations, 2019).

Plants are subject to two main types of threats, (1) biotic stress, such as pathogens, fungi, bacteria, and insects, and (2) abiotic stress, such as climate change, weather, salinity, soils, and chemicals (Gull et al., 2019). In both cases, diseases appear on plants and can be spotted by human observation and measurement. For instance, examining leaves can reveal deformed shapes and discoloration in infected tissues. These tissues often show black, yellow, powdery, or blight spots (Angelo Randaci, Earth's Ally, 2021). Thus, spotting and detecting diseased spots, then diagnosing diseases is a major process to prevent plant disease outbreaks and to protect both the population and the economy. Traditionally, farmers rely on visual and physical properties of crops to identify anomalies such as plant stress, diseases, or weeds. Such properties can be captured nowadays by cameras or satellites, allowing machine learning models to help farmers make decisions on the collected images. To avoid propagation, threats should be detected as soon as possible and remedial actions should be as precise as possible. Machine learning models can help accelerate the detection and decision processes. Indeed, the world is leaning toward 'smart agriculture' (Pratyush Reddy et al., 2020). Smart agriculture, or smart farming (Jayaraman et al., 2016), is the use of advanced technologies to increase agricultural productivity and crop health using minimum resources. Smart farming includes intelligent watering systems, smart greenhouses, automatic or semi-automatic robots, drones, or user notifications (Jha et al., 2019). Machine learning includes models and algorithms that learn to perform certain tasks such as decision-making and prediction based on knowledge acquired from previously seen data (Burkov, 2019; Kelleher, 2019; Segaran, 2007). Different machine learning techniques have been applied to smart agriculture, for instance: Random Forest for crop prediction (Geetha et al., 2020), or Decision Tree (Rajesh et al., 2020) and K-Nearest Neighbors (Suresha et al., 2017) for plant leaf disease detection. More recently deep learning and neural networks are often used because of their high performance. Neural networks are named and structured as an inspiration of how biological neurons signal to one another (Goodfellow et al., 2016). They are made of layers of neurons which are units that make computations to perceive features and help make decisions. Different architectures of deep neural networks have been developed for various tasks. In computer vision and imagery, convolutional neural networks (CNNs) (Yoo, 2015) are a commonly used architecture because of their higher effectiveness than other machine learning methods (Liakos et al., 2018). As many applications in agriculture are related to image analysis, it is relevant to examine how CNNs can be used in different applications in agriculture. This survey focuses on remote sensing and image processing using CNNs in agriculture. There

have been several survey papers focusing on one specific agriculture task (Hasan et al., 2021; K. Hu et al., 2021; Kamarudin et al., 2021; Kok et al., 2021; J. Liu & Wang, 2021; Saleem et al., 2019; X. Wu et al., 2019), but none has covered multiple tasks in agriculture. This survey aims to fill this gap by covering several different tasks that CNNs solve in agriculture. In fact, many applications in agriculture have common properties and techniques to share. Thus, a survey is beneficial in revealing similarities and differences across different applications. This paper covers the analysis of more than 100 research papers published in reputable and high-impact journals and conferences in this field related to multispectral remote sensing, datasets, vegetation indexes, and CNN applications in agriculture. We explore in detail the methods and processes employed by more than 15 research papers on CNN applications and provide a comprehensive summary about 20 others. This survey aims to achieve three goals: firstly, to provide an overview of the types of images used to train machine learning models in agriculture, including ground-based or aerial, multispectral or RGB or with vegetation indexes; secondly, to present the various CNN applications in agriculture, while analyzing their implementation and architecture; and thirdly, to provide a list of useful resources to researchers such as references to common practices in machine learning or datasets useful in agriculture. In this, we cover the different phases for new developments for smart agriculture that use images as a source of decision-making.

The rest of this survey is organized as follows: first, we introduce the methodology we followed to retrieve research articles, reviews, and other relevant data for our study. Then, we present related work, consisting in reviews and surveys on machine learning in agriculture in general and CNNs in agriculture in particular, in addition to reviews that address the topics of image datasets. Then, we report on different types of images used in agriculture and we explain how the use case of these different images relies on their acquisition method, because different use cases may require different acquisition methods since the way images are acquired has an impact on the resolution, the quality, the distance to the subject, and the angle of shot of the images. In addition, we present multispectral images along with satellite images because of the importance and the value that they add to remote sensing and computer vision in agriculture. Then, we present a set of free and public image datasets in agriculture, which can be used to train machine learning models. The datasets that we describe fall into 6 categories: weed management, disease detection, pest detection, crop classification, yield estimation, and damage detection. Then, we review research articles and studies that use CNNs in agriculture for weed management, plant disease detection, yield prediction, crop type classification, crop counting, and water management. Finally, we end with a discussion on our findings with future prospects and identified gaps and a conclusion.
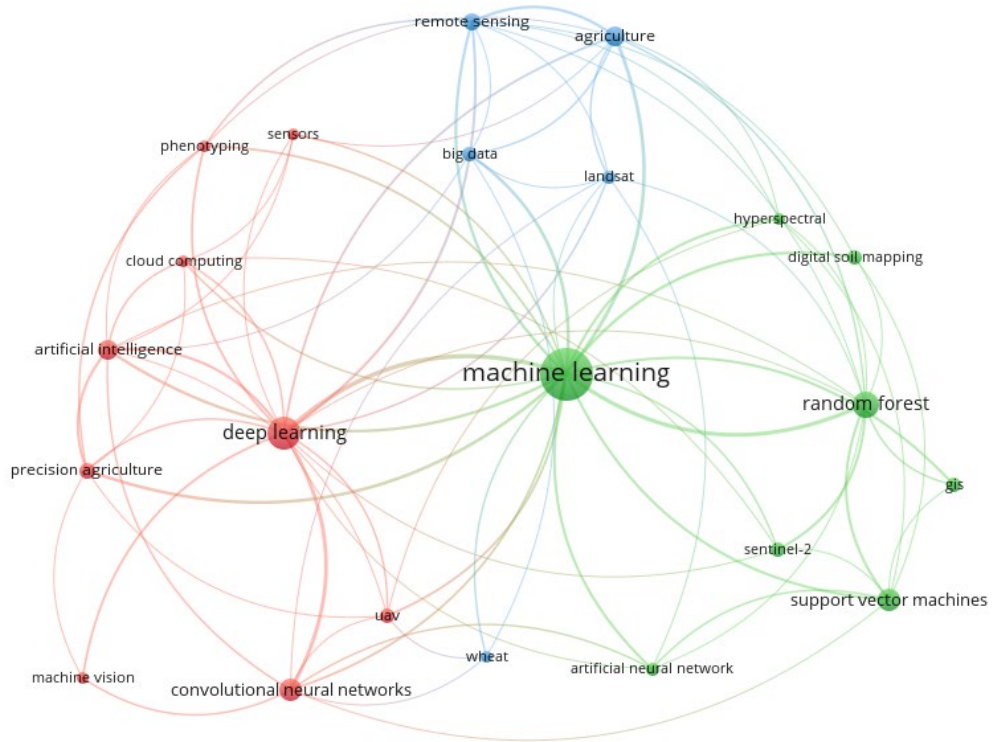
Figure 1. Keyword co-occurrence map based on the results of the query [machine learning] AND [agriculture] on the Web of Science
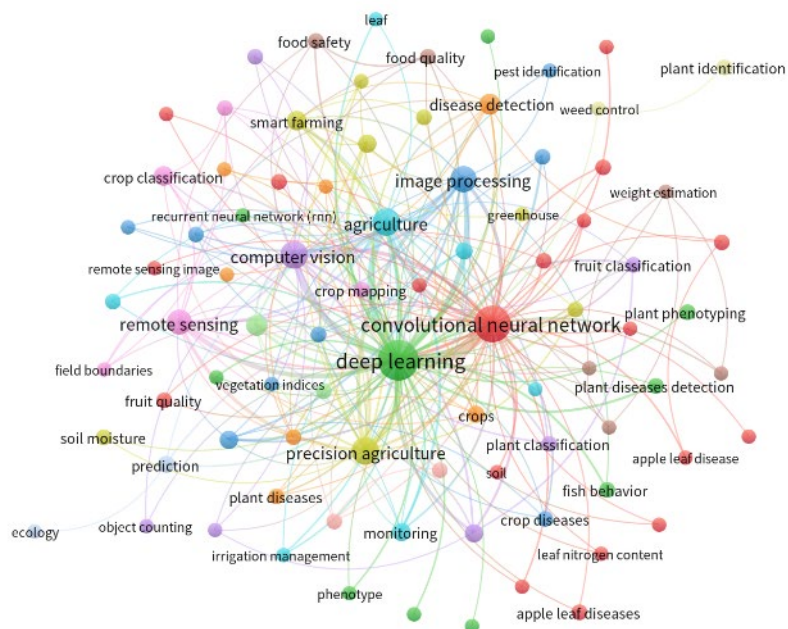

Figure 2. Keyword co-occurrence map based on the results of the query [convolutional neural networks] AND [agriculture] on the Web of Science

Methodology

This survey focuses on CNN applications in the field of agriculture. Our study revolves around two main axes: (1) analyzing the type and properties of the images used in the field, and (2) analyzing how CNNs contribute to agriculture-related tasks.

In the first axis, we search for publicly available agriculture image datasets and we analyze the properties of these images, such as their resolution, the acquisition method used, the number of channels they have, or the environment in which they were taken. Not all datasets are scholarly referenced, thus we manually search for agricultural datasets. This phase of our survey was performed by searching on the Web for datasets relevant to agriculture, on websites such as Kaggle (up to April 2023). On the other hand, we also searched the literature using the Web of Science and Google Scholar for datasets that are referenced by articles. We did this either by looking for scholar papers that present a survey or review of agricultural datasets, or studies relevant to agriculture that provide their experimental datasets. More specifically we search in journals dedicated to data, such as Data in Brief (up to April 2023). As a result of this strategy, we retrieved datasets that are useful in different subfields of agriculture.

We categorized our findings of freely available image datasets in agriculture into six categories: weed management, disease detection, pest detection, crop classification, yield estimation, and damage detection. From this large pool, we picked 25 datasets that are representative enough of the diverse identified applications. The selection process prioritized datasets based on their reliability and originality, as well as the creativity and uniqueness they offer. Given that datasets and studies contain multispectral images, vegetation indexes, and satellite images, we dedicate a special part of our paper to this type of images by searching on Google Scholar and the web for books and articles addressing the topic of multispectral remote sensing. With regard to the second axis, on how CNNs contribute to agriculture-related tasks, we used Google Scholar and the Web of Science. Our motivation to look for CNNs in agriculture comes from the hypothesis that CNNs are one of the most used deep learning methods in agriculture. We employed a bibliometric research method, which validated our hypothesis. For that, we used a keyword-based search on the Web of Science: [machine learning] AND [agriculture], then, by selecting papers with high impact, we collected featured author keywords and we kept the ones that were repeated the most and removed less relevant keywords.

Figure 1 shows a graphical representation of these keywords and their links. In that figure, the nodes represent keywords; the bigger the node, the more frequent the keyword. Each color corresponds to a cluster of nodes, that is, a set of keywords that appear often together. The edges represent the co-occurrences of keywords; the thicker an edge is, the more the keywords co-occur. What we notice in Figure 1 is that among machine learning methods, deep learning is commonly used, and, in deep learning itself, CNNs are the most distinguishable. This interpretation of the bibliometrics validates choice of the focus of this survey. We also noticed that Sentinel-2 and Landsat satellite images, hyperspectral images, and UAV (Unmanned Aerial Vehicles) images appear very frequently. This validates the importance of these topics in this field as well. We are also

interested in knowing which agriculture-related problems could be solved with CNNs. To this end, we used the same method again to acquire bibliometrics about CNN applications in agriculture.

We did another keyword-based search on the Web of Science and selected the papers with highest impact as well: [convolutional neural networks] AND [agriculture], and filtered the keywords of the resulting articles the same way we did for the first query. The corresponding bibliometric map is given in Figure 2. This map clearly shows 6 main categories for CNN applications: weed detection, disease detection, yield prediction, crop type classification, crop counting, and water management. Our survey will focus on these 6 categories of CNN applications. We conducted a search for research articles that have a high impact or that are highly cited falling in these categories using Google Scholar and the Web of Science. Our search resulted in 35 papers that we describe in Table 4 and Table 5, and other studies that we analyzed deeply to describe the implementation and the evaluation techniques of their CNN models.

Related work

Several reviews and surveys studied machine learning in agriculture, especially the applications of CNNs in computer vision and remote sensing. For instance, Liakos et al., 2018 presented a collection of machine learning models and algorithms that were categorized according to the problem they solved. They conducted their review on at least 40 existing studies that were released in the period 2004 - 2018 and they categorized their findings into four main categories: livestock management, crop management, water management, and soil management. They found that methods such as clustering, decision trees, regression, neural networks, support vector machines, and Bayesian models, are used to perform crop monitoring tasks such as yield prediction (Amatya et al., 2016), disease detection (Moshou et al., 2014), weed detection (Pantazi et al., 2016), crop quality (Maione et al., 2016), and species recognition (Grinblat et al., 2016), in addition to other types of tasks related to water (Mehdizadeh et al., 2017), soil (Johann et al., 2016), livestock management (Pegorini et al., 2015). According to their findings, problems related to crops are more commonly solved using machine learning than problems related to water, soil, or livestock management. They also found that artificial and deep neural networks are used more than other machine learning methods. While their review covers approaches in solving different problems in agriculture using machine learning techniques, it does not focus on a single type of algorithms, such as CNNs in our case, nor does it specifically discuss multispectral remote sensing or image datasets.

Kok et al., 2021 researched how Support Vector Machines (SVM) performed in precision agriculture. To do so, they first collected 60 research articles that employ SVM and other machine learning and deep learning models, then they proceeded to identify which model performed the best. The studies that they collected focused on six major fields of agriculture: nutrient estimation, disease detection, crop classification, yield estimation, quality classification, and weed detection. The authors concluded that SVM

performed lower than Random Forest in some fields and lower than deep learning methods in all fields.

The reviews presented above, show that deep learning and CNNs outperform classical machine learning methods. This encourages the scientific community to keep improving and researching deep learning and CNN models in order to come up with innovative solutions to be used in real life scenarios.

Kamilaris & Prenafeta-Boldú, 2018a reviewed CNN applications for agriculture. They searched on the Web of Science and Google Scholar based on keywords i.e., deep learning, agriculture, farmingand found and analyzed 23 scientific articles related to these topics. Next, they analyzed the solutions provided by each study, along with the methodologies and data employed to achieve these outcomes. They also presented how the approaches were evaluated and how they performed in comparison with other approaches. The addressed problems are the same as what other machine learning methods address: land cover, fruit counting, and other previously mentioned tasks. In addition to that, a small experiment on detecting missing vegetables in a sugar cane field using CNN was proposed. In their experiment, they used a VGG network (Simonyan & Zisserman, 2014), pretrained on ImageNet (Deng et al., 2009), and achieved 79.2% accuracy, which the authors considered low accuracy because of mislabeled images in their dataset.

Kamilaris & Prenafeta-Boldú, 2018b published a review addressing the topic of deep learning in agriculture. The method they adopted to find their information is similar to Kamilaris & Prenafeta-Boldú, 2018a, but it also contained details on image preprocessing, image augmentation, and how deep learning models are evaluated. They also presented a list of studies that use deep learning to accomplish multiple agricultural tasks.They also provided a short list of 14 publicly available datasets related to images in agriculture without categorizing them.

Other reviews focus more on one specific problem in agriculture. For example, J. Liu & Wang, 2021 wrote a literature review of the period 2014-2018 on deep learning methods to detect diseases and pests in plants. Their review aims to assist researchers in quickly understanding the methods used in this field. The authors detailed classification methods along with their advantages and their disadvantages. They also provided a list of 13 datasets of plant diseases and pests different from those cited in Kamilaris & Prenafeta-Boldú, 2018a. They found that deep learning (i.e., artificial neural networks and CNNs) performs better than classical image processing methods (i.e., Kmeans clustering, decision trees, support vector machines, K-Nearest Neighbors) in plant disease and pest detection. They concluded that there is still room to improve the models and datasets that are tested in laboratories before employing them on real world fields.

Saleem et al., 2019 reviewed plant disease detection and classification with CNNs. They presented the CNN architectures that are used the most in the literature (not only in agriculture). Then, they provided a list of research articles that employ these CNN architectures, alongside with the datasets used. What distinguishes this review is that the authors also discussed visualization methods such as segmentation map, saliency map, heatmap, etc. According to them, more datasets containing images of plants in real

conditions in different scenarios are needed. The authors also recommended employing hyperspectral and multispectral imaging for training deep learning models in the early detection of plant diseases, as these technologies can identify the diseases before symptoms become visually apparent.Aside from disease detection in plants, Kamarudin et al., 2021 wrote a review on deep learning in plant water stress tasks such as evapotranspiration forecasting, plant water status estimation, water stress identification, soil moisture estimation, and soil water modeling. To acquire their data, they applied a method which consists in finding answers to a custom series of problems and questions. In that way, they found the data that were used to train deep learning methods, the architectures of deep learning models that were applied, and how these methods compared to classical ones in accomplishing the same tasks. Their findings indicated that even though deep learning is still in its early stages in water stress assessment, it outperforms other models and it is still subject to a lot of improvement.

Other surveys focused on weed detection in crops using CNNs (Hasan et al., 2021; K. Hu et al., 2021; Z. Wu et al., 2021). Their reviews included CNN architectures and examples of publicly available datasets of weed management. Lu & Young, 2020 dedicated a review on publicly available datasets on agriculture and categorized them into two main categories: weed management and fruit detection. They also defined a third category that gathers other uses such as flower detection, yield estimation, canopy species and biomass prediction, as well as damage and disease detection.

According to the different reviews and surveys above, CNNs offer great solutions to problems in agriculture. While many earlier surveys have examined the applications of CNNs with an emphasis on specific uses, this paper offers a thorough review encompassing all relevant stages of employing CNNs in agriculture. This is particularly beneficial for newcomers to the field. We explain various methods of image acquisition, encompassing different types of image data with their respective advantages and drawbacks. Additionally, it provides a list of publicly available datasets used in different applications related to agriculture, serving as a resource for researchers seeking data for their studies. Then, we explain general and basic concepts about machine learning and specific ones about how CNNs are built. By covering all the study stages from image acquisition to model building and evaluation instead of focusing on one particular stage, we provide valuable resources that help in developing studies on the applications of CNNs in agriculture.

Images in Agriculture

CNNs are designed to process images; in this section, we present images that are acquired using different methods depending on the context, such as UAVs or satellites. We also present different types of images; RGB, multispectral images, and vegetation indexes that can derive from them.

Basic Image Acquisition

Depending on the needs, the crop types and the fields that the studies focus on, image acquisition methods can differ. Images directly taken from fields have been used to train CNNs, for instance, Crop/Weed Field Image Dataset (Haug & Ostermann, 2015)(CWFID), Early crop weed (Espejo-Garcia et al., 2020), and weedNet (Sa et al., 2018), Sugar beets 2016 (Chebrolu et al., 2017) are some publicly available datasets of field images.
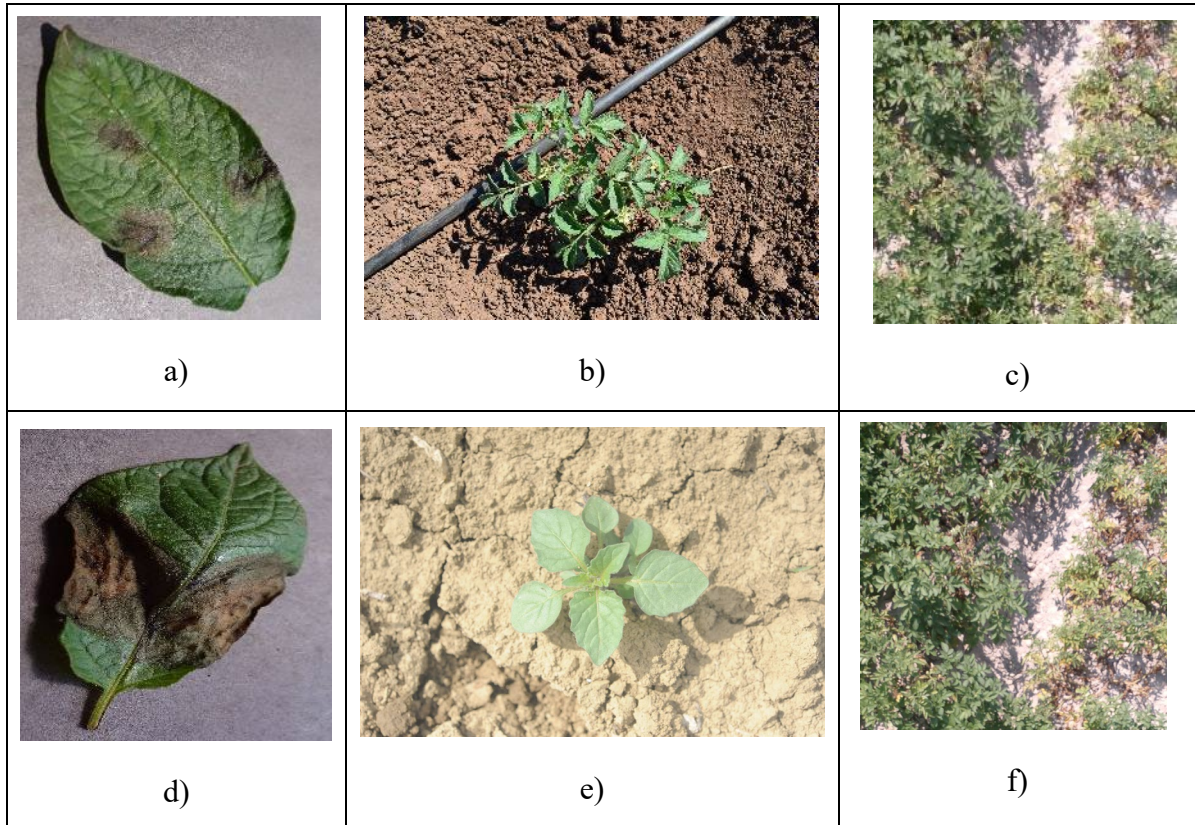

Figure 3. Examples of agricultural images that can be used to train convolutional neural networks. Source: (a)&(d) (Hughes et al., 2015), (b)&(e) (Espejo-Garcia et al., 2020), (c)&(f) (Butte et al., 2021)

Other images can consist of plants or leaves with plain background such as in PlantVillage (Hughes et al., 2015) and Rice Leaf Diseases Dataset (Prajapati et al., 2017). An advantage of having plain background in images is the reduction of noises in images and the increasing focus on plants (Figure 3).

This type of images can be acquired from ground robots such as BoniRob (Ruckelshausen et al., 2009) as in CWFID (Haug & Ostermann, 2015) images and Sugar beets 2016 (Chebrolu et al., 2017). Other images such as the ones of PlantVillage (Hughes et al., 2015) are classical images taken with a digital camera under different weather conditions, sunny and cloudy days.

Aerial images taken few meters above the ground may be required in agriculture. For this, unmanned aerial vehicles (UAV) can be equipped with cameras or sensors to monitor agricultural fields. For example, the images of a multispectral potato plant image dataset (Butte et al., 2021) were taken with a drone flying at an altitude of 3 meters. UAV are helpful for short and low altitude image acquisition missions over agricultural lands, but they get quickly limited by their short battery life and their field of view. Satellites are a complimentary solution to UAVs by offering a broader coverage area and longer imaging periods that span to multiple years. Their capability to cover large areas of land in a single pass and their more advanced instruments makes them suitable for different and wider range of applications in agriculture that are not possible on UAVs. However, the quality of satellite images can be affected due to heavy cloud coverage that can occur for example.

Multispectral Remote Sensing

In addition to differences in acquisition methods, the electromagnetic spectrum range in which the images are taken can differ depending on the needs. The most common type of images is color images; theyare a combination of three different wavelengths of light: red, green and blue (RGB). A main advantage of RGB images is the simplicity in their acquisition. They do not require special equipment to be taken, normal phone or digital cameras are good enough to build agricultural datasets. For instance, PlantVillage (Hughes et al., 2015) is a famous dataset that contains RGB images captured with smartphones.

RGB images visualize a wide range of colors and shapes, which is useful to analyze crop features. The colors of a plant can reveal a lot of information about its health. For example, a discoloration or distortion of leaves or fruits which is caused by diseases is directly seen in RGB images. RGB images contain spatial and visual information that can be exploited to deduce the state of a plant.

RGB images only capture the spectrum of light that is visible to the human eyes. However, RGB images do not provide information about crop stress and diseases that are not observable. These mainly include abiotic diseases especially in their early stages, such as water stress, nutrient stress, bad soil quality, and high salinity. Searching for information beyond visible light is essential for a more effective analysis of crop images. Images that contain such information are known as multispectral images.

Multispectral images are images that contain information of light across the electromagnetic spectrum that typically includes wavelength outside the visible range, such as near infrared (NIR) or shortwave infrared (SWIR), that commonly refer to the regions of light between 750 nm and 1400 nm, and 1400 nm and 3800 nm, respectively. Using multiple wavelengths provides valuable information about the composition, structure, and properties of crops or other objects. Healthy crops exhibit high reflectance of NIR light, distinguishing them from diseased plants or other objects (Vincent & Dardenne, 2021). Furthermore, SWIR light provides information about soil moisture

content, organic matter content, and water stress (Yue et al., 2019). Using the reflectance of vegetation and crop lands in NIR and SWIR is essential in agriculture due to the distinct responses exhibited by healthy and stressed vegetation or non-vegetation objects in these bands.

In addition to providing valuable information about the composition and the structure of crops, multispectral images can also be used to compute various vegetation indexes. These indexes are mathematical formulas applied to various spectral information captured in different wavelengths. Vegetation indexes can bring more information to help assessing crop stress and health. One of the most widely used vegetation indexes is the Normalized Difference Vegetation Index (Rouse Jr et al., 1973) which compares the reflectance of NIR to red light in order to quantify biomass density and crop health.

Multispectral images are captured using special remote sensing instruments that are sensitive to particular wavelengths of light. Advanced remote sensing instruments are more commonly found on Earth Observation satellites rather than UAVs or ground vehicles.

Satellite Imagery

In agriculture, satellite images have shown success in accomplishing multiple functions. For instance, Phiri et al.(2020) gave an exhaustive list of agricultural uses in satellite imagery using Sentinel-2 satellites (Gatti & Bertolini, 2013). Their review contains multispectral satellite image uses in management of water, crops, soils, forests, grasslands, and fields. Other similar use cases include rice field mapping using satellite images and deep learning by Nguyen et al., 2020. Sishodia et al. (2020) provided more details in their review on water stress, evapotranspiration, soil moisture, nutrient management, disease management, weed management, and crop monitoring and yield.

Band	Description	Resolution(m)	Wavelength(nm)
B1	Coastal aerosol	30	430-450
B2	Blue	30	450-510
B3	Green	30	530-590
B4	Red	30	640-670
B5	Near infrared	30	850-880
B6	Shortwave infrared	30	1570-1650
B7	Shortwave infrared	30	2110-2290
B8	Panchromatic	15	500-680
B9	Cirrus	30	1360-1380
B10	Thermal infrared	100	10600-11190
B11	Thermal infrared	100	11500-12510

Table 1. Landsat-8 multispectral bands

Band	Description	Resolution(m)	Wavelength(nm)
B1	Ultra Blue	60	443
B2	Blue	10	490
B3	Green	10	560
B4	Red	10	665
B5	Visible and near infrared	20	705
B6	Visible and near infrared	20	740
B7	Visible and near infrared	20	783
B8	Visible and near infrared	10	842
B8A	Visible and near infrared	20	865
B9	Shortwave infrared	60	940
B10	Shortwave infrared	60	1375
B11	Shortwave infrared	20	1610
B12	Shortwave infrared	20	2190

Table 2. Sentinel-2 multispectral bands

Landsat-8 (Roy et al., 2014) and Sentinel-2 (Gatti & Bertolini, 2013) are two well known satellites in remote sensing. They are both Earth Observation satellites deployed in 2013 by NASA and in 2015 by ESA respectively. Data they captured is freely available online. Contrary to RGB images, images captured by Landsat-8 or Sentinel-2 satellites have more than 3 channels. Landsat-8 has 11 channels whereas Sentinel-2 has 13 channels. Each channel represents a band, which is a part of the electromagnetic spectrum the satellites capture. In remote sensing, satellite bands are used to capture information of the Earth's atmosphere and surface. Both Landsat-8 and Sentinel-2 capture light that ranges from the aerosol part of the electromagnetic spectrum to shortwave infrared while passing by the visible spectrum of light and near infrared.

Both satellites have a high-resolution panchromatic band which are B8-Panchromatic and B8 for Landsat-8 and Sentinel-2 respectively. Panchromatic bands are used to create images with higher resolution. Sentinel-2's bands generally have higher resolution than Landsat-8. For instance, most Sentinel-2's bands have a resolution of 10 m or 20 m while most of Landsat-8's bands have a resolution of 30 m. A major difference is that Landsat-8 has thermal infrared bands that capture thermal radiation of the Earth's surface whereas Sentinel-2 does not have such bands. Table 1 and Table 2 resume the properties of Landsat-8's and Sentinel-2's bands.

Vegetation indexes

Jackson & Huete (1991) defined vegetation indexes as spectral data combinations that characterize vegetation and its changes over time. Basic vegetation indexes include ratios, differences, sums and linear transformations of at least two spectral data. Bannari get al. (1995) presented a large selection of indexes and how they reveal more information than individual spectral bands, especially for green plants. Sun et al., 2021 presented a selection of indexes related to soil moisture. There are hundreds of vegetation indexes that are obtained from linear operations between spectral data and others that are computed with more complex operations such as logarithms (Serrano et al., 2002). This paper provides the description and the formulas of representative vegetation indexes in Appendix A.

Vegetation indexes can be used when doing image processing using machine learning generally and CNNs specifically. For example, Yaloveha et al., 2022 used CNNs to do land cover mapping on EuroSat dataset (Helber et al., 2019) augmented with NDVI (Rouse Jr et al., 1973), NDWI (Gao, 1996), and GNDVI (Gitelson et al., 1996) and the accuracy that they obtained was higher than RGB images alone. Zhang et al. (2020) used high resolution NDVI for land cover mapping. Their method shows better performance when compared to other state-of-the-art methods. Weng et al., 2022 used a multispectral crop dataset from which they constructed a number of vegetation indexes that include the mentioned indexes: NDVI, SAVI (Huete, 1988), SR (Birth & McVey, 1968), NDWI; in addition to Red-Edge Chlorophyll Index (Thompson et al., 2019), Normalized Difference Red Edge Index (Thompson et al., 2019), Modified SAVI (Rouse Jr et al., 1973), and Green Chlorophyll Index (Thompson et al., 2019). Their method showed better performance than other methods such as Random Forest in crop classification tasks.

The above studies show that adding vegetation indexes leads to better results than raw data. They also show that the suitability of vegetation indexes depends on the type of vegetation, soil moisture content, and atmospheric conditions. Different vegetation indexes have different purposes. For example, NDWI is more suitable in studies on water content than NDVI. Choosing or combining vegetation indexes depends on the study environment and the available data.

Dataset	Size	Type	URL
Weed management			
Food crops and weeds dataset	1118	RGB	https://data.mendeley.com/datasets/nj4vtk4tt6/1
CWF-788	788	RGB	https://github.com/ZhangXG001/Real-Time-Crop-Recognition
Crop vs weed discrimination	40	Multispectral	https://lcas.lincoln.ac.uk/wp/research/data-sets-software/crop-vs-weed-discrimination-dataset/
Early crop weed	504	RGB	https://github.com/AUAgroup/early-crop-weed
The GlassCover Image Dataset	39600	RGB	https://vision.eng.au.dk/grass-clover-dataset/
Rice seedlings and weeds	224	RGB	https://figshare.com/articles/dataset/rice_seedlings_and_weeds/7488830
WeedNet	465	Mutlispectral	https://github.com/inkyusa/weedNet
Plant seedlings dataset	960	RGB	https://vision.eng.au.dk/plant-seedlings-dataset/
Soybean and weed dataset	15336	RGB	https://data.mendeley.com/datasets/3fmjm7ncc6/2
Sugar beets 2016	300	Mutlispectral	http://www.ipb.uni-bonn.de/data/sugarbeets2016/
CWFID	60	Mutlispectral	https://github.com/cwfid/dataset
Plant disease detection			
PlantDoc	2598	RGB	https://github.com/pratikkayal/PlantDoc-Dataset
New Plant Diseases Dataset	87000	RGB	https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset
Maize disease	18222	RGB	https://osf.io/p67rz/
PDDB	2326	RGB	https://www.digipathos-rep.cnptia.embrapa.br/
XDB	46513	RGB	https://www.digipathos-rep.cnptia.embrapa.br/
Rice Leaf Disease Dataset	120	RGB	https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases
PlantVillage	50000	RGB	https://data.mendeley.com/datasets/tywbtsjrjv/1
Multispectal Potato Plants	3720	Mutlispectral	https://www.webpages.uidaho.edu/vakanski/Multispectral_Images_Dataset.html
Crop and quality classification			
Fresh and Rotten Fruits	12335	RGB	https://data.mendeley.com/datasets/bdd69gyhv8/1
FruitNet	14700	RGB	https://data.mendeley.com/datasets/b6fftwbr2v/1
Pest detection			

Soybean leaf dataset	6140	RGB	https://data.mendeley.com/datasets/bycbh73438/1
8 common tomato pests dataset	609	RGB	https://data.mendeley.com/datasets/s62zm6djd2/1
IP102	75000	RGB	https://github.com/xpwu95/IP102
Damage detection			
Sugarcane billets	152	RGB	https://github.com/The77Lab/SugarcaneBilletsDataset
Yield prediction			
Date fruit dataset	8000	RGB	https://ieee-dataport.org/open-access/date-fruit-dataset-automated-harvesting-and-visual-yield-estimation

Table 3. Public image datasets in weed management, plant disease detection, crop classification, pest detection, damage detection, and yield prediction. The columns show the dataset name, its size, its data type, and its download URL.

Publicly available image datasets for agriculture related tasks

For this section, we collected a large number of datasets related to agriculture. However, many datasets have been constructed without rigorous control and lack testing and reliability. Other datasets can be very similar to each other. Therefore, we selected 25 image datasets that offer the most variety and are freely available online. The characteristics of datasets, such as the type of imaging device used, the resolution of the images, the number of images included in the dataset, types of annotations or labels, and other details may vary depending on the specific purpose or intended use of the dataset. Given that datasets for weed management and plant diseases are more present in the state of the art, we will describe them separately from other datasets. The datasets are presented in chronological order and their summary is in Table 3.

Weed management datasets

Food crops and weed (2020)

Food crops and weeds dataset (Sudars et al., 2020) includes 1118 RGB images of six food crop species and eight weed species under different resolutions, i.e., 720x1280, 1000x750, 640x480, 640x360, and 480x384 pixels. It also contains 7853 XML annotations used to generate bounding boxes.

Crop in weedy field (CWF-788) (2019)

CWF-788 (Li et al., 2019) is a dataset of field images containing 788 images of cauliflowers in high weed presence. The dataset provides images of two resolutions: 400x300 pixels and 512x384 pixels, in addition to pixel level annotation that represents a mask of crops without weeds.

Crop vs weed discrimination dataset (2019)

The images in this dataset (Bosilj et al., 2019) are acquired using an RGB camera and a NIR camera mounted 5 cm apart on a ground cart. The cameras produced high resolution (2464x2056 px) images of carrots (20 images) and onions (20 images) and to align the results of the two cameras, an adjustment step was needed, thus resulting in 2428x1985 px images for carrots and 2419x1986 px images for onions. The dataset provides RGB and NIR as well as NDVI images and pixelwise annotations (crop, weed, soil/other).

Early crop weed (2019)

This dataset (Espejo-Garcia et al., 2020) dataset consists of RGB images of two crops (tomato and cotton) and two weed types (velvet leaf and black nightshade) taken at early crop growth stage. The dataset contains a total of 504 images taken at 1m above the ground in varying resolutions. Classification can be applied to these images to identify each type of plants.

The GrassClover Image Dataset (2019)

The GrassClover Image Dataset (Skovsen et al., 2019) has 8000 synthetic images annotated at pixel level and 31,600 unlabeled images, all in RGB. Among the annotated images, we distinguish six classes that are red clover, white clover, grass, weeds, soil and unknown.

Rice seedlings and weeds (2019)

This dataset (Ma et al., 2019) has 224 912x1024 pixels RGB images of rice seedlings in paddy fields combined with weeds at their early growth stages, captured by a camera that was 80 cm to 120 cm above the water surface. In the paddy fields, the rice seedling rows were separated by 30 cm and the plants themselves had 14 cm to 16 cm between them. The dataset is designed to be segmented and ground truth pixel annotation exists in a .mat file. Each pixel is classified as rice seedlings, weeds or background.

WeedNet (2018)

WeedNet dataset (Sa et al., 2018) contains images taken from a sugar beet field that were captured by a multispectral camera mounted on a UAV at 2 m above the plants. It has 465 images divided into 132, 243 and 90 images of crops, weeds and a mixture of crops

and weeds respectively. Each image has a copy in Red, NIR and NDVI, and is annotated at pixel level for crops, weed and background (soil).

Plant seedlings dataset (2017)

Plant seedlings dataset (Giselsson et al., 2017) is a dataset of 960 high resolution (5184x3456 px) annotated RGB images of seedlings planted in styrofoam boxes. The images show 12 species at several growth stages with a physical resolution of 10 px/mm. Each image shows a unique species among: Maize, Common Wheat, Sugar Beet, Scentless Mayweed, Common Chickweed, Shepherds' Purse, Cleavers, Redshank, Charlock, Fat Hen, Small-flowered Cranesbill, Field Pansy, Black-grass, and Loose Silky-bent.

Soybean and weed dataset (2017)

Images of soybean and weed dataset (dos Santos Ferreira et al., 2017) were captured with UAVs on a surface area of one hectare at an average altitude of 4 m above the ground. There are 3249 soil images, 7376 soybean images, 3520 grass images and 1191 broadleaf images, resulting in 15,336 images in total in RGB annotated at pixel level.

Sugar beets 2016 (2016)

This dataset (Chebrolu et al., 2017) was acquired by a field robot from early sugar beet plant emergence until their growth. The images are in PNG format with a resolution of 1296x966 pixels in RGB and NIR. They are provided alongside their annotations to pixel level of sugar beets, nine types of weeds and background. The dataset provides around 300 images and their annotations.

Crop/Weed Field Image Dataset (2014)

Crop/Weed Field Image Dataset (CWFID) (Haug & Ostermann, 2015) is a dataset for weed detection among carrot plants. It is composed of 60 images of 1296 x 966 pixels resolution that were captured using a multispectral camera mounted on a ground robot (Ruckelshausen et al., 2009) that captures visible and near infrared light (NIR). The images have 3 channels (Red-NIR-Red) and are annotated on pixel level on 3 channels. The first two channels have binary annotation and represent weed and crop presence respectively, and the third channel is always equal to 0, representing soil.

Plant disease detection datasets

PlantDoc (2020)

PlantDoc (Singh et al., 2020) is a dataset of 2598 annotated field RGB images of 13 crop species and 17 types of diseases. The images are distributed across 27 classes of healthy crops and diseases specific to each crop type. This dataset was created to do leaf detection

and disease classification tasks in real-life conditions unlike PlantVillage (Hughes et al., 2015) specifically.

New Plant Diseases Dataset (2019)

This dataset (Kaggle, 2019) is an augmented version of PlantVillage (Hughes et al., 2015) that contains 87,000 RGB images of healthy and unhealthy crops.

Maize disease (2018)

This dataset (Wiesner-Hanks & Brahimi, 2018) has 18,222 images of maize leaves that are labeled to detect damage caused by Northern Leaf Blight, which is a common disease that is devastating for maize. Among the images, 1787 have been taken with a handheld camera, 8766 with a boom camera, and 7669 with a drone.

Image Database of Plant Disease Symptoms (PDDB) (2018)

PDDB (Barbedo et al., 2018) has 2326 images of 171 diseases across 21 plant species. 715 images are field images while the remaining 1611 are images captured in a controlled environment. An extended dataset called XDB has been created from subdivided PDDB images, resulting in 46,513 images.

Rice Leaf Disease Data Set (2017)

This dataset (Prajapati et al., 2017) contains 120 JPG images of rice leaf showing 3 types of diseases: leaf smut, brown spot, and bacterial leaf blight. The images are in RGB on a white background under direct sunlight and each class counts 40 instances.

PlantVillage (2015)

PlantVillage (Hughes et al., 2015) is a widely used dataset of over 50,000 RGB images of healthy and unhealthy leaf images across 38 crop species and types of diseases. It is designed to be used in classification tasks. The number of classes and their labels vary depending on the crop type, for instance, potato leaves are classified either as "healthy", "early blight", or "late blight", while apple leaves are classified as "healthy", "Cedar Rust", "Black Rot", or "Scab". The images have a resolution of 256x256 pixels and the leaves are placed on plain background under different light conditions. The dataset is still evolving and is famous among classification studies. Many versions of PlantVillage exist with synthetic images as in (Pandian & Gopal, 2019) that has 61,486 images, or the New Plant Diseases Dataset (Kaggle, 2019).

Multispectral Potato Plants Images

This dataset (Butte et al., 2021) contains potato crop images in field capture by a drone. It is divided into 4 categories, all annotated with bounding boxes to detect healthy and stressed crops. The categories include 360 RGB images of 750x750 pixels resolution, an

augmented version containing 1500 images, 360 multispectral images of red, green, red-edge, and near infrared of a 416x416 pixels resolution (360 images of each spectral band), and an augmented version of 1500 multispectral images.

Others

Crop classification: Fresh and Rotten Fruits Dataset (2022)

Fresh and Rotten Fruits Dataset (Sultana et al., 2022) contains 12,335 augmented images of 16 types of fruit classes in two quality state: fresh or rotten. Originally there were 3200 images on which they applied augmentation techniques that include: rotation, flipping, zooming, and shearing. This dataset is suitable to do crop type classification tasks by machine learning models.

Crop classification: FruitNet (2021)

FruitNet (Meshram & Patil, 2022) is a dataset of 6 different fruits of 3 qualities. Each image contains at least one fruit of the same type, which may be of good or bad quality, or there may be a mix of good and bad quality fruits in the image. The dataset contains more than 14,700 high quality 256x256 images of one of the following fruit classes: banana, apple, guava, lime, orange, and pomegranate. This dataset can be used to train, test, and validate models in quality or crop type classification tasks as well.

Pest detection: Soybean leaf dataset (2022)

This dataset (Mignoni et al., 2022) contains 6140 images that show healthy soybean leaves, or leaves that were damaged by caterpillars, or Diabrotica speciosa. The image are available in jpeg format of 500x500 size. Applying classification tasks on this dataset can help the detection of pests in a field, even if the pests are not in the images themselves.

Pest detection: A database of eight common tomato pest images (2020)

This dataset (M. Huang & Chuang, 2020) has 609 images of 8 types of common tomato pests: Tetranychus urticae, Bemisia argentifolii, Zeugodacus cucurbitae, Thrips palmi, Myzus persicae, Spodoptera litura, Spodoptera exigua, and Helicoverpa armigera. The images were augmented using 90, 180 and 270 degrees rotation, horizontal and vertical flip, and crop. The image augmentation result is a dataset of 299x299 pixel jpg RGB images.

Pest Detection: IP102 - A Large-Scale Benchmark Dataset for Insect Pest Recognition (2019)

This dataset (X. Wu et al., 2019) has more than 75,000 images of 102 insect type where 19,000 images have bounding box annotations for object detection tasks.

Damage detection: Sugarcane billets (2018)

Harvesting sugarcane using mechanized methods can induce damage to sugarcane billets, therefore the damage shown in this dataset is not caused by diseases, hence this category. This dataset (Alencastre-Miranda et al., 2018) provides images of 5 types of damage that sugarcane billets can sustain and images of undamaged ones, so in total 6 labels. This dataset contains 152 RGB images of high resolution (2448x2048 pixels) of sugarcane billets.

Yield prediction: Date fruit dataset (2019)

This dataset of (Altaheri et al., 2019) serves two purposes: crop type classification, and yield prediction. The dataset is divided into two sub-datasets. The first one has more than 8000 labeled images of 5 date fruit varieties that can be used to train models on date fruit type classification. Whereas the second dataset has 152 images of date belonging to the same type, mainly destined to do yield prediction.

Discussions

In this section, we have presented datasets that are available online for different tasks of agriculture. However, it is important to note that there exist other datasets that are less accessible, known, or used in studies than the presented ones. We excluded datasets that are not public from this survey. For example, some of on-request datasets are AgricultureVision (Chiu et al., 2020) for land use analysis or Perennial ryegrass dataset (Yu, Schumann, et al., 2019) for weed detection.

In addition to that, it is noticeable that datasets can have multiple purposes, e.g., crop type classification and yield prediction or crop type and quality classification. Datasets also have different properties. For example, some datasets also make use of vegetation indexes such as NDVI, while others include multispectral channels but without explicitly computing vegetation indexes. The environment conditions vary as well. For instance, some images are taken in laboratories, some others are taken in crop fields, and some datasets have synthetic images.

There exist more datasets related to weed management than for yield prediction,

damage detection, or soil moisture. The descriptions of the existing datasets offer an

insight on dataset properties that can inspire dataset creators.

CNN for agriculture

In this section, we present 35 studies where CNNs are used in agriculture. The studies we present employ different methods to solve different problems. We categorized the addressed issues into 6 categories: weed detection, disease detection, yield prediction, crop type classification, crop counting, and water management. Some of the studies use

datasets that were previously presented in this review (Hughes et al., 2015) while others have their own private datasets (Hu et al., 2019a). This section offers a diversity of studies that show the extent of what can be done using CNNs. Weed detection and plant disease detection are very common in the literature (Kamilaris & Prenafeta-Boldú, 2018b, 2018a; Liakos et al., 2018) and there are a larger number of studies. Table 4 and Table 5 include additional studies.

A brief on convolutional neural networks

CNNs are a type of neural networks in deep learning that are powerful in image processing. CNNs are designed to learn and extract features within an image (O'Shea & Nash, 2015).

Machine learning algorithms with images include three types of problem: classification (binary or multiclass), segmentation, and regression. Classificationaims to categorize images into predefined classes, either two (binary) or multiple (multiclass) (Chen et al., 2021). Segmentation involves the classification of each pixel within an image into a particular class (Minaee et al., 2020). In regression tasks, the goal is to predict a numerical value based on the image (Lathuilière et al., 2020).

In agriculture applications, binary classification is used, for example, in detecting whether a plant is stressed or not. Identifying the type of disease a plant suffers from can be cast into a multiclass classification problem (Atila et al., 2021; Jiang et al., 2019). On the other hand, segmentation is used, for instance, to separate weeds from crops in an image (Sa et al., 2018). Finally, predicting yield or soil moisture content can be considered as a regression problem (Hegazi et al., 2023). Moreover, a specific task can be solved in multiple ways depending on the context. For example, detecting weed might consist in segmenting the image into regions that contain weeds and others that do not, or classifying the whole image as either weed containing or not. CNNs can be used to solve any of these three types of problems: classification, segmentation and regression.
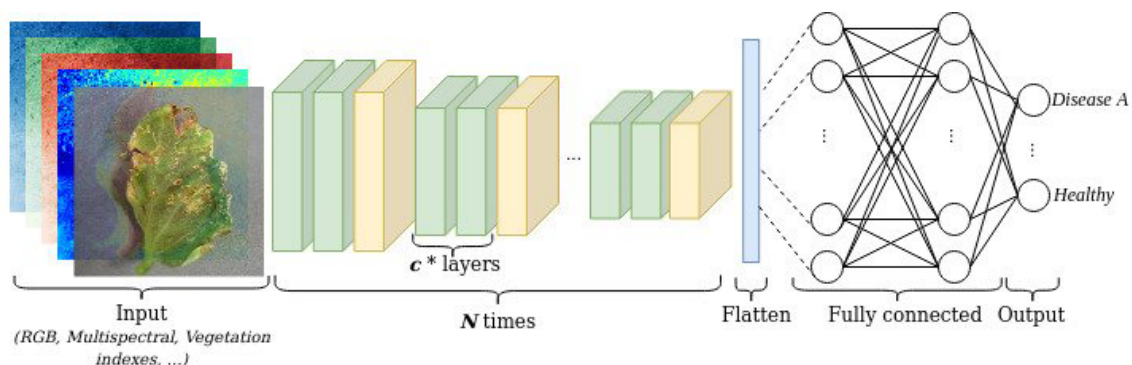

Figure 4a. General architecture of a CNN for multi-class classification. Here is the case of disease identification in a leaf image. The output is the type of diseases the leaf is showing or if it is healthy.
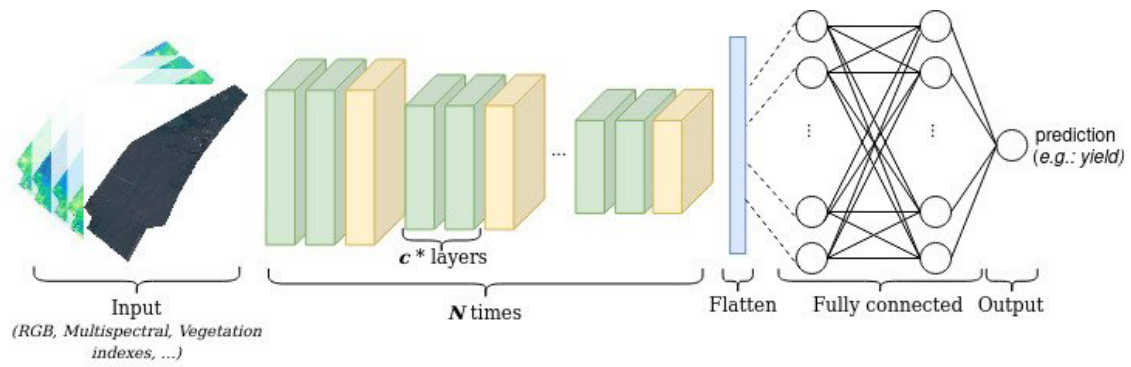
Figure 4b. General architecture of a CNN in the case of regression. Here is the example of yield prediction from an aerial or satellite image. The output is a numerical and continuous value that indicates the yield of the field in input.
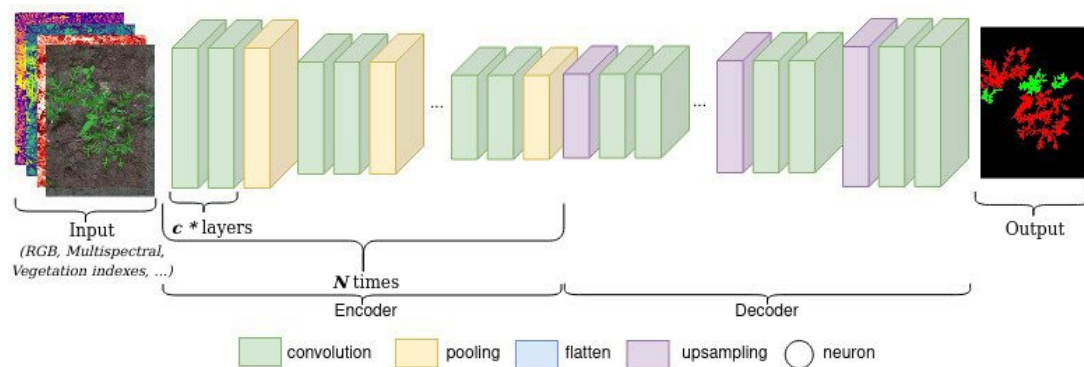

Figure 4c. General architecture of a CNN for image segmentation. In this case, the task is weed detection. Each pixel in the output image is classified as weeds (red), crops (green), or soil (black). Segmentation relies on an encoder-decoder architecture where all the layers are fully convolutional.

Figure 4. Three types of architectures of CNNs, with an example of each. The input images can be RGB, multispectral, vegetation indexes, or a combination of these. The output and overall architecture varies depending on the problem.

CNNs that solve these problems usually have similarities in their input and architecture but differ in their output. These architectures are presented in a simplified way in Figure 4.

As seen in previous sections, CNN inputs are images that belong to different color spaces such as RGB, multispectral, vegetation indexes, or combinations of these.

To process these images, CNNs typically have a combination of convolutional layers that that apply learnable filters (or kernels) to the input image. These kernels extract features from the images and produce a feature map during the process. Each kernel is a matrix of learnable values that are optimized during the training phase of the CNN. A

summed dot product is performed between the kernels and the input image in order to product an output value (O'Shea & Nash, 2015).

Another type of layers that is usually found in CNNs is the pooling layer. Pooling layers reduce the dimensions of the image while preserving features. The pooling operation reduces the number of parameters of the model. Max-pooling and average-pooling are two typical pooling methods used in CNNs (O'Shea & Nash, 2015).

Contrary to traditional Artificial Neural Networks (ANNs) where each neuron is connected to all other neurons, each neuron of convolutional and pooling layers is connected to a small number of neurons around it, which effectively reduces the number of parameters and speeds up the convergence (O'Shea & Nash, 2015). Moreover, in traditional ANNs, more neurons are required to accommodate higher resolution images with multiple channels, and consequently the number of trainable parameters increases, which is not the case of CNNs which benefit from sliding kernels, weight sharing, and pooling layers (Z. Li et al., 2020; O'Shea & Nash, 2015).

When solving classification or regression problems, it is necessary to convert the multidimensional image data into a single dimension array by using flattening layer. Typically, the result of the flattened layer is passed to one or several fully connected neural layers (Chen et al., 2021).

On the other hand, segmentation applies classification on each pixel of the image rather than classifying the image as a whole. To accomplish this task, an encoder-decoder architecture is commonly used. A basic encoder includes convolutional and pooling layers and the decoder typically has upsampling and convolutional layers. Upsampling layers increase the resolution of the input image and they are seen as the opposite pooling layers. The decoder is optimized to create a segmented image from features extracted by the decoder. An encoder-decoder architecture is fully convolutional, all layers in the model are convolutional layers, i.e., there are no fully connected layers at the output of the network, which is the case of classification or regression (Sa et al., 2018).

Evaluation metrics

Evaluating CNNs is important to measure the model's performance. Various metrics are used to evaluate the performance of a CNN (or any other machine learning model) depending on the task that they achieve (Kuhn & Johnson, 2013; Taha & Hanbury, 2015). These metrics vary depending on which task the CNN is performing. Below is a list of commonly used evaluation metrics in the literature (Agarwal, Singh, et al., 2020; Atila et al., 2021; Bosilj et al., 2020; Fawakherji et al., 2019; M & M.N, 2015; Srivastava et al., 2022; Subeesh et al., 2022; Veeranampalayam Sivakumar et al., 2020; Yu, Schumann, et al., 2019).

In classification (binary or multiclass):

- **FP, TP, FN, TN:** these represent abbreviations of four possible outcomes of a classification prediction. False Positive (FP) or Negative (FN) is when the classifier incorrectly predicts the positive or negative class for an instance,

respectively. True Positive (TP) or Negative (TN) is when the classifier correctly predicts the positive or negative class for an instance, respectively.

- **Accuracy:** rate of correct predictions. $Accuracy = \frac{TN+TP}{TN+TP+FN+FP}$

- **Precision:** rate of true positive among all positive predictions. $Precision = \frac{TP}{TP+FP}$

- **Recall:** (or sensitivity) rate of true positive predictions among all predictions that should be positives. $Recall = \frac{TP}{TP+FN}$

- **F1 score:** harmonic mean of precision and recall. It provides a score that quantifies the trade-off between these two metrics. $F1 = \frac{2*Precision*Recall}{Precision+Recall}$

- **Specificity:** rate of true negative among all predictions that should be negative. $Specificity = \frac{TN}{TN+FP}$

- **Cohen's Kappa coefficient (Cohen, 1960) :** this score serves to measure the agreement between two evaluators when dealing with categorical items. Its advantage lies in its consideration of the possibility of agreement occuring by chance. $Kappa = \frac{P_{agree}-P_{chance}}{1-P_{chance}}$

- **Matthews correlation coefficient (Phi):** MCC is a statistical rate particularly useful in the context of imbalanced datasets. It gives a better assessment of classification than metrics such as accuracy and F1 score (Chicco & Jurman, 2020). $MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

In multiclass classification, these metrics can be micro-averaged, which consists in summing up TP, FP, TN, and FN across all classes, and then computing metrics such as recall or precision based on these values. On the other hand, macro-average involves computing metrics for each class and then averaging them.

In segmentation:

- **Intersection over Union (IoU):** it is used to measure the overlap between the predicted and the ground truth segmentation. It is calculated by the following formula: $IoU = \frac{Area_{overlap}}{Area_{union}}$

Metrics such as accuracy, precision, recall, F1 score, specificity can also be used in segmentation because segmentation can be reduced to a pixel-wise classification task.

In regression:

- **Mean Squared Error:** MSE measures the average squared difference between predicted and ground truth values and it commonly used to evaluate regression tasks.

- **Mean Absolute Error:** MAE is another metrics used in evaluating regression tasks. It measures the average absolute difference between the predicted and true values.
- **Root Mean Squared Error:** RMSE is the square root of MSE. Interpreting the performance of the model in regression is easier with RMSE because its value is in the same units as the predicted and true values.

These metrics are quality oriented which measure the quality of predictions that deep learning models make, e.g. CNNs. Other metrics measure the footprint of the models, such as their size, computational requirements, and energy consumption. Some of these metrics include:

- **Model Size:** the memory required to store the model.
- **Inference Time:** the time it takes to make predictions on new data.
- **FLOPS (FLoating-point Operations Per Second):** the quantity of computational workload required during inference
- **Energy Consumption:** an estimation of the power or energy consumption of the model during inference.

Balancing quality and footprint metrics is crucial for efficient evaluation if CNN models. These metrics ensure an accurate model that is resource-efficient.

Reference	Description	Data	Architectures	Metrics
Subeesh et al., 2022	Weed identification in polyhouse grown bell peppers	polyhouse RGB images	AlexNet, GoogLeNet, **InceptionV3**, Xception	Precision, Recall, F1 score, Accuracy
Razfar et al., 2022	Weed detection in soybean plantations	Soybean and weed dataset	MobileNet, ResNet, **custom CNNs**	Accuracy, Runtime, Memory usage
Espejo-Garcia et al., 2020	Weed detection in early growth stage in field	Early crop weed dataset	Xception, InceptionResNet, MobileNet, **DenseNet**, VGG	F1 score
J. Gao et al., 2020	Weed detection in sugar beet fields	RGB images of sugar beet fields	YOLO-v3, **tiny YOLOv3**	Mean average precision (mAP)
Veeranampalayam Sivakumar et al., 2020	UAV imagery for weed detection in soybean fields	UAV images of soybean at different growth stages	**Faster RCNN**, Single Shot Detector	Precision, Recall, F1 score, Intersection over Union (IoU)
Bosilj et al., 2019	Applying transfer learning to detect weed across different plant speciess	Crop vs weed discrimination dataset	SegNet	Cohen's kappa coefficient, Precision, Recall
Petrich et al., 2020	Detecting toxic flowering weeds in high quality drone images	RGB drone images of sites containing weeds	U-Net	Precision, Recall, F2 score
Kounalakis et al., 2019	Transfer learning in weed detection in grasslands	field images acquired by a robot	AlexNet, VGG, **InceptionV1**, ResNet	Accuracy, False Positive Rate, False Negative Rate, Precision, Recall
Ma et al., 2019	Detecting weeds in paddy field images	Rice seedlings and weeds dataset	U-Net, FCN, **SegNet**	Mean pixel accuracy, Mean Intersection over Union
Yu, Schumann, et al., 2019	Detecting several weed types in perennial ryegrass	Perennial ryegrass and weed dataset	**AlexNet**, GoogLeNet, VGG, DetectNet	Precision, Recall, F1 score, Phi coefficient (MCC)
Yu, Sharpe, et al., 2019	Detecting several weed types in bermudagrass	Field images	GoogLeNet, VGG, **DetectNet**	Precision, Recall, F1 score
Fawakherj	plant/soil	RGB images with	VGG-SegNet,	Mean

| i et al., 2019 | segmentation and weed/crop classification on sunflower field images using transfer learning | vegetation indexes | U-Net, **VGG-UNet**, BonNet, FCN8, **VGG** | Intersection over Union, Accuracy, Sensitivity, Specificity, Precision |

Table 4. Research articles (reference and short description of the research objectives) that use CNNs in weed management along with the datasets, CNN architectures and the metrics used to train and evaluate their models. In studies that compare multiple CNNs, the model in **bold** has the best performance..

Weed detection

Appropriate weed management results in higher yield (Soltani et al., 2016). Moreover, mapping weed in an area allows to control weed at early growth stages, hence, reducing the use of herbicides. However, weed detection is challenging because of their similar physical properties with food crops.

Detecting weed using deep learning among food crops has gained research interest over the years. This is due to the increase in agricultural production to meet the needs of the rapidly expanding population faced with a heavy loss of crops due to weeds that can be up to 50% on some crops (Sothearith et al., 2021). The concept behind both segmentation and classification of crops is quite similar: identify and locate each plant within an image or pixel and classify it as either a plant or a weed.

Suh et al., 2018 used a SegNet based (Badrinarayanan et al., 2017) CNN to segment a sugar beet field. They used images of 3 channels (NIR - Red - NDVI) from the weedNet dataset (Sa et al., 2018) as inputs to train the network, and the output was a map with pixel-wise annotation as crops, weeds, or background.

Another segmentation task (Milioto et al., 2018) was accomplished differently by using 14 channel images as input, that includes: RGB, Excess Geen Index (ExG), Excess Red Index, Color Index of Vegetation Extraction, Normalized Difference Index, HSV, Sobel X, Sobel Y, Laplacian, and Canny Edge Detector, on ExG (W. Gao et al., 2010; Mlsna & Rodríguez, 2009; Song et al., 2017; Woebbecke et al., 1995).

The authors showed that these channels improve the model's performance in separating the vegetation from soil as well as speeding up the convergence process during training. The proposed CNN is a lightweight auto-encoder with less than 30,000 parameters that can be used to do real-time segmentation when running in hardware that is attached on a ground or aerial robot for instance.

In other studies (Czymmek et al., 2019; J. Gao et al., 2020; R. Zhang et al., 2020), YOLO-v3 (Redmon & Farhadi, 2018) or its variant tiny YOLO-v3 (Adarsh et al., 2020) or faster YOLO-v3 (Yin et al., 2020) were used. YOLO-v3 (You Only Look Once) is real-time object detection neural network, a variant of CNN, that predicts the bounding boxes and class probabilities in an image. It works by dividing the image into a grid of

cells where each one is responsible in detecting objects in its region of the image. Combining YOLO-v3 with UAV and a user interface, R. Zhang et al., 2020 showed the best performance related to real-time weed detection in videos at 45 frames per second (fps).

In Table 4, we summarize the CNN architectures used in different studies. The best architectures are in bold. As we can see, it is difficult to determine whether there is a best architecture. Different studies suggest different best architectures. We found that accuracy and precision values are almost equal in most comparisons. We cannot conclude which models are the best, because they are tested in different environments. The results of an experiment depend on the hyperparameters, and other setup conditions that are employed (Koutsoukas et al., 2017). There is a lack of comprehensive study comparing all the available CNN architectures for tasks in agriculture. Also, the use of pretrained and fine-tuned state-of-the-art networks leads to better results than training from scratch (see Table 4 for examples of studies).

Reference	Description	Data	Architectures	Metrics
Atila et al., 2021	Disease detection and classification in plant leaves using transfer learning	PlantVillage	**EfficientNet**, AlexNet, ResNet, VGG, InceptionV3	Sensitivity, Specificity, Accuracy, Precision
Agarwal et al., 2020	Tomato crop disease identification	PlantVillage	**light-VGG**, VGG, InceptionV3, MobileNet	Accuracy, F1 score, Precision, Recall, True Positive Rate, False Positive Rate
Barbedo, 2019	Plant disease classification	PDDB and XDB	GoogLeNet	Accuracy
G. Hu et al., 2019	Identification of diseases in tea leaves	RGB tea leaf images	VGG	Accuracy
Jiang et al., 2019	Real time detection and identification of apple leaf diseases	Laboratory and field images of apple leaves	AlexNet, GoogLeNet, InceptionV3, ResNet, VGG, **VGG-INCEPTION**	Accuracy
Türkoğlu & Hanbay, 2019	Plant disease classification	Plant disease and pest images	AlexNet, VGG, GoogLeNet, **ResNet**, Inception, InceptionResNetv2, SqueezeNet	Specificity, Sensitivity, F1 score, Accuracy
Zhong & Zhao, 2020	Apple leaf disease detection	Images of apple leaves	DenseNet	Precision, Sensitivity, F1 score
K. Zhang et al., 2018	Identiying tomato leaf diseases	PlantVillage	AlexNet, GoogLeNet, **ResNet**	Accuracy, Time
B. Liu et al., 2017	Apple leaf diseases identification	Apple leaf images	**AlexNet**, GoogLeNet, ResNet, VGG	Accuracy
Amara et al., 2017	Classification of banana leaf diseases	RGB images of banana leaves in the field	LeNet	Accuracy, Precision, Recall, F1 score

Table 5. Research articles that use CNNs in plant disease detection. Table structure and conventions are similar to Table 4.

Disease detection

Detecting plant diseases at early stages is important for several reasons. In fact, it can help farmers to control the spread of the disease, thus increasing agricultural productivity by avoiding reduced crop yields. Some plant diseases might also threaten food security

and the environment, leading ecosystem disrupt or soil erosion.

Plants affected with diseases might show discoloration, wilting, blight, rust, leaf

spots or other symptoms related to a change in plant phenotype.

Deep learning approaches have been implemented to detect and classify diseased plants showing symptoms. Similarly to weed detection, some researchers have proposed real-time disease detection methods using CNNs (Jiang et al., 2019). Their approach involved collecting images of both diseased and healthy apple leaves to train a CNN that is based on SSD (Single shot multibox detector) (W. Liu et al., 2016) that they named INAR-SSD (SSD with Inception module and Rainbow concatenation (Jeong et al., 2017)). Their network is designed to detect apple leaf diseases quickly and accurately, which could help apple growers to reduce crop loss and improve yield. INAR-SSD provided a better performance when it came to detecting small objects (small spots of apple leaf disease in their case) as compared to the original SSD that has difficulty in correctly detecting them. This is due to rainbow concatenation which concatenates multiple layer outputs before moving to classification.

Türkoğlu & Hanbay, 2019 suggested two methods for plant disease and pest detection that they applied to their own dataset of 4000x6000 RGB images. First, they applied deep feature extraction from various fully connected layers. They extracted certain layers and feature vectors from the pretrained deep learning models (see Table 5). Then, they used three classifiers: Support Vector Machine (SVM) (Noble, 2006), K-Nearest Neighbor (KNN) (Peterson, 2009), and Extreme learning machine (ELM) (G.-B. Huang et al., 2006) in the classification phase. They obtained the best performance with the architecture consisting of ResNet50+SVM at an accuracy of 97.86% ± 1.56. On the second hand, they used transfer learning of the same neural networks. They used pre-trained models as a starting point for their problem then only changed the last three layers in order to adapt the models to their classification task. VGG16 was the best performing network with an accuracy score of 96.92% ± 1.26 .

CNNs show promising results in detecting plant diseases. Most studies rely on crop leaves to diagnose a plant because in slight anomaly in the leaf color can be detected at early stages by CNNs. The choice of the architecture depends on the conditions of the study, real-time detection or high-resolution images for example.

Yield prediction

The detection of weeds and diseases in crops constitutes a major part of CNN applications in agriculture. However, there are various other uses of CNNs in this field as well.

These applications include crop yield production for instance. Crop yield is the amount of crops that are harvested per unit area of land. In the following studies, it is shown that crop yield can be predicted from crop images.

Nevavuori et al., 2019 used multispectral data acquired with a UAV equipped with a NIR sensor. In 2017, they selected nine different crop varieties of wheat and barley and took multispectral images of fields covering 90 hectares during the growth season

which spanned from June to August. They also acquired harvest yield data in September of the same year. The yield production can be directly measured by harvesting a sample of the crop and determining its weight per hectare (kg/ha). In order to obtain their dataset, they sub-sampled their data and shuffled it, resulting in approximately 15,200 images from which they took 15% for testing their model respectively. They further divide the other 85% into 3 subsets: training, validation, and testing, to apply k-fold cross-validation, which is a technique that partitions the data into k equal sized folds, and trains the model iteratively on k-1 folds while using the remaining fold for validation and testing, then, they test the resulting model on the initial testing set. Among the architecture tests they did, their best performing network had 5 convolutional layers of 64 kernels followed by batch normalization and ReLU then 1 convolutional layer of 128 kernels also followed by batch normalization, ReLU and max pooling. The output of the final convolutional layer is passed through two linear fully connected layers that each have a single output at the end which is the yield estimation in kg/ha. By training and testing their network on NDVI and RGB images separately, their results show that their model performed better on RGB images.

Srivastava et al., 2022 proposed a custom CNN architecture that aggregates daily weather and meteorological data from 1999 to 2019 in Germany to predict winter wheat yield. Their proposed CNN uses 1D convolutions (Kiranyaz et al., 2021) on periodic numerical data. Their weather data included features such as minimum and maximum temperature, relative humidity, precipitation, solar radiation, and wind speed. They also used soil data such as soil categories, water availability, saturation point, and bulk density. For ground truth they relied on crop yield data that is recorded between 1999 and 2019. In addition to that, they used crop phenology data that describes the winter wheat (sowing, flowering, and harvested). Given that daily data over 20 years require a large number of parameters, the data were down-sampled by averaging and aggregating the values of weather features to 45 weekly samples. This reduced significantly the number of inputs to the model. Thus, on one hand, the CNN receives 45 samples of each weather feature, then it applies alternate convolutions and average pooling until the data becomes one-dimensional. Then, the feature vectors are concatenated together to form one vector of weather features. On the other hand, soil and phenology data are concatenated and processed in an artificial neural network, then the output is concatenated to the weather feature vector. The newly created vector is then processed in several fully connected layers that predict the yield. The proposed model was compared to other machine learning algorithms that include Random Forest, K-Nearest Neighbors, LASSO and Ridge Regression, Regression Tree, Support Vector Regression, XGBoost, and Deep Neural Networks. It performed better than all the others when evaluated using MSE, RMSE, and correlation coefficient metrics.

Although in-site measurements help predict crop yield, Khaki et al., 2020 had a different approach. They propose YieldNet, a model that has the ability to predict soybean and corn yield. They used temporal data of 30 multispectral satellite images taken during growing season provided by MODIS products. For yield data, they used data from 13 American states that produce the most corn. Then, they focused on corn and soybean

lands by excluding non-croplands from satellites images with the help of the USDA-NASS cropland data layers which offers land cover data based on satellite imagery. To avoid making a model with a large number of parameters because of the large yield data size, the data was discretized into bins, where a bin is a range of values. Thus, each image was reduced to a histogram of values that shows the frequency of pixel that fall into different bins. By concatenating the different histograms of the same field, an image that has Time x Bin as dimensions is constructed. By applying this to each band of the multispectral image, a cube of data is obtained, which is the input of the CNN. YieldNet has a backbone of 5 convolutional layers that produce an output that goes in two smaller CNNsswhich produce an output for each corn and soybean yield separately. YieldNet performed better than other machine learning methods such as Random Forest, Deep feed forward neural network, Regression Tree, LASSO, and Ridge. The metrics used in evaluation were mean absolute error, the root-mean-square error, and the correlation coefficient. It is more complicated to do yield prediction than weed or disease detection using CNNs. This is mainly due to the fact that yield cannot be directly predicted based on visual input alone, but it requires additional data such as in-situ measurements, multispectral data, temporal data, or meteorological data. The three studies presented above used different CNN methods and additional data to achieve yield prediction and succeeded to perform better than other traditional machine learning methods.

Crop type classification

In addition to yield prediction, crop type classification is an important task for precision agriculture, food security and smart farming. Crop type classification refers to detecting the presence of crops in an image and their species. Crop type classification is often applied to aerial or satellite images of whole fields in the context of land cover. Knowing the different types of plantations in fields helps crops management and monitoring, e.g. in food security (Kordi & Yousefi, 2022) . Kussul et al., 2017 conducted a study on a large area of 28,000km2 over Kyiv region of Ukraine as an attempt to classify 11 classes (water, forest, grassland, bare land, winter wheat, winter rapeseed, spring cereals, soybeans, maize, sunflowers and sugar beet). Images were acquired using Landsat-8 and Sentinel-1 images during the 2015 vegetation season that spans from October 2014 till September 2015. As for ground truth data, they conducted a ground survey and collected polygons of different classes that were divided equally between training and validation sets. Their approach started with preprocessing images to deal with missing data in cloudy area for instance. Then, they applied supervised classification using two different CNNs that explore spectral and spatial features, respectively. On average, the use of CNN reached an accuracy of 85% on major crops and as a final result, the CNN that explored spatial features outperformed the other one even though small objects were misclassified.

In another similar study, Z. Sun et al., 2020 collected Landsat-8 images over North Dakota (USA) for model inputs and field surveys as well as roadside for ground truth labels. They calculated NDVI from band 4 and band 5 (Table 1) and counted pixels which NDVI values are greater that a threshold of 0.4. Then, they considered images from 2013,

2014 and 2015 as training images while images from 2016 and 2017 are testing images. The identified crops are: winter wheat, spring wheat, rice, corn, cotton, soybeans, barley, oata, peanuts, sorghum and range and pasture. The proposed network is based on the SegNet model (Badrinarayanan et al., 2017) with 7 input channels that are fed the 7 Landsat bands and 132 output channels that represent the 132 crop classes that the USDA NASS crop classification system has. To increase the efficiency of the training, they took extra steps in input preprocessing by applying a sieve process by removing pixels that result from noises, thus, reducing uncertainties. Overall, they achieved more than 82% in accuracy.

Crop type classification is vital for land cover analysis that employs remote sensing techniques and temporal satellites images to identify different crop types. As seen, CNNs are efficient in land cover over large areas in satellite images with a large number of classes.

Crop counting

Another task that is important in crop management is crop counting. It helps farmers to know the quantity of crops they have in their fields and thus it allows them to do better stock management and yield prediction. Crop counting tasks include counting entire crops or the fruits of plants.

W. Li et al., 2016 used a satellite image of palm tree field in Malaysia that dates back to 2009. The image was captured by the QuickBird satellite which is a high-resolution satellite with a 0.6m/pixel panchromatic black and white band and 4 2.4m/pixel multispectral bands (red, green, blue, and near infrared). For the study, they used a fusion of the 5 bands by applying the Gram-Schmidt fusion process (Laben & Brower, 2000). To create the dataset, they divided the initial image into 9000 RGB patches of 17x17 pixels by using a sliding window of 17x17 in size and a step of three pixels. The size of the window was chosen experimentally in order to fit exactly one palm tree in the center of the patch. Then, they manually annotated the patches in palm tree samples and background samples and randomly selected 80% of the patches for the training set and 20% for the test set. A patch is labeled as "palm" only if it contains a palm tree in its center. Then, they used a LeNet (LeCun et al., 1998) architecture to achieve an average accuracy of 97%. When their network classified each sample, a postprocessing stage took place in order to merge the samples and add boundary circles around predicted palm trees.

The above approach requires dividing an image into samples and then classifying the samples. However, Ribera et al., 2017 took another approach of the crop counting problem by considering it as a regression problem instead of a classification problem. In fact, they input images of sorghum plants into a CNN then a fully connected layer that has one output, that is the predicted number of crops. Experimentally they used 546x146 pixels field images as inputs to a slightly modified version of InceptionV3 (Szegedy et al., 2015) where the last pooling layer was removed, resulting in the last convolution layer outputting a 2048x15x1 feature map that is then flattened and input to a fully connected

network. This method scored 6.7% in the Mean Absolute Percentage Error, making it a well-performing method.

Rahnemoonfar & Sheppard, 2017 proposed a different method by training an Inception-ResNet (Szegedy et al., 2017) based neural network on synthetic images only. They artificially generated 24,000 images by creating a green and brown blurred background with different size red circles to simulate tomato plants in fields. They then tested their neural network on 100 real images of tomato plants and achieved an accuracy of 91%. This method can be applied for other sorts of plants when an image dataset is not easily available.

Different approaches were taken in crop counting using CNNs. Some of them divided a satellite images into patches, while others used regression backed with convolution on whole images, or created fully synthetic images. Overall, CNNs have shown success in counting crops.

Water management

Lack of water for plants can lead to reduced yield or crop death, which is a concerning problem in agriculture, especially in regions facing drought. CNNs are used in water management for tasks such as estimating soil moisture and identifying water stress in crops. An et al., 2019 compared the performance of two CNN models: ResNet50 and ResNet152 (He et al., 2016) at identifying and classifying maize water stress. Their model identified whether maize plants had a light drought, a moderate drought, optimum moisture (normal growth). They trained their models on 640x480 RGB images of maize in fields under different light conditions at different times during the day. The pictures contained images of maize crops at two growth stages: seedling and jointing. Overall, they carried out their study on RGB images and their gray versions, at seedling and jointing stages, and with ResNet50 and ResNet152, with transfer learning and training from scratch. They found that RGB images performed better than gray ones and that transfer learning increased performance. Both CNN models performed well with little difference in accuracy but ResNet50 trained faster than ResNet152. Their models achieved 98.14% accuracy at identifying drought and 95.95% at classifying it.

Chandel et al., 2021 used CNN architectures to detect water stress in three types of crops: maize, okra, and soybean. They compared GoogLeNet (Szegedy et al., 2015), AlexNet (Krizhevsky et al., 2017), and Inception V3 (Szegedy et al., 2016) at detecting whether field images of the three plants showed water stress or not. In their experiments, GoogLeNet performed better than Inception V3 and AlexNet for all three crops.

Hegazi et al., 2023 proposed using Sentinel-2 images to predict soil moisture content. The authors solved this as a CNN-based regression problem. They compared multiple types of inputs to understand which one describes the best the soil moisture content. Their predictions were compared with ground truth values taken from soil moisture monitoring stations. The tested inputs are: Sentinel-2 bands individually, Sentinel-2 bands combined, NDVI, GVMI, NDWI, indexes combination (NDVI, GVMI, and NDWI), Sentinel-2 bands (Table 2) and indexes combination, Green-RedB8, B7-

B8a-B11 S2 bands, B8-B11-B12 S2 bands, Green-Red-B8-B11-B12 S2 bands, B7-B8-B8a-B11 S2 bands, and Red-B5-B6-B7-B8-B8a S2 bands. They found out that the "Red-B5-B6-B7-B8-B8a" does the best predictions of soil moisture content and that NDWI is the most sensitive to soil moisture when compared to NDVI and GVMI.

In this category, we distinguish two types of methods in detecting water stress. The first one relies on the physical properties that appear on plants when they are stressed because of lack of water. In this case, the problem is similar to crop disease detection. The second one is studying the composition of the soil by looking at multispectral data. Water related vegetation indexes are very useful in this task.

Discussion

In this review, we presented remote sensing and CNNs for applications in agriculture.

Images

We discussed the different types of imaging techniques and image types used in agriculture. We showed the importance of multispectral imaging, especially in the NIR part of the electromagnetic spectrum.

We also showed the importance of multispectral imaging especially in the NIR part of the electromagnetic spectrum. We explained how different types of images are exploited using CNNs in agriculture. Multispectral images and vegetation indexes have proven to improve CNN performance in most cases. Additionally, satellite images are very useful for temporal studies over large areas of land. We have focused on optical sensors but there are other types of sensors that were not discussed, such as radar or thermal sensors.

We found that vegetation indexes are not exclusive to a single agricultural task, instead, they are used throughout all of them. However, this does not exclude the fact that some indexes are more useful than others in specific fields. For instance, water indexes are more useful to solve problems related to water issues, since they are specifically designed for this purpose. However, we did not come across any studies that evaluate how individual vegetation indexes perform across different agricultural tasks. Such study would give more insight into which agriculture task each index performs the best, which could lead to the development of more precise solutions.

Datasets

In order to illustrate examples of agriculture image datasets we listed and described 25 datasets related to agriculture that are free and available online. We also mentioned other datasets that are private or inaccessible that were used in several studies. We gave examples of datasets that are intended to be used in different types of tasks (weed management, plant disease detection, crop and quality classification, pest detection,

damage detection, and yield prediction).

We found out that PlantVillage (Hughes et al., 2015) is a very frequently used dataset. It is composed of more than 50,000 RGB labeled images of 256x256 pixels resolution, which makes it one of the largest datasets. It encompasses a variety of plants and threat types, which explains its popularity, especially when a big dataset is needed when using deep learning approaches. Most of the neural networks trained with PlantVillage had classification accuracy scores above 98%. However, Noyan, 2022 demonstrated that his model classified PlantVillage image backgrounds with 49% accuracy, which is very high compared to random guessing and therefore the images contained bias. This means that we should be careful and aware of this bias when using this dataset.

In agriculture, collecting a sufficiently diverse and balanced dataset is often challenging because of the time consuming and heavy nature of that task. This problem makes models struggle to generalize well enough to newly seen data. Data scarcity results in overfitting (Roelofs et al., 2019), where the models learn to only perform very well on the training images instead of learning patterns that are useful in real scenarios. To overcome this issue, image augmentation is applied (Roelofs et al., 2019), which is a way to artificially expand datasets by generating new images by transforming original ones. Simple image augmentation techniques such as image scaling, cropping, flipping, padding, rotation or translation and change of brightness, contrast, saturation or hue, add more samples in the dataset. Furthermore, data augmentation is also a way to introduce diversity that simulate different environmental conditions such as sunny or cloudy days. More advanced image augmentation techniques include Conditional Generative Adversarial Network (C-GAN) (Mirza & Osindero, 2014) which purpose is to generate synthetic images close to reality. Increasing the size of datasets by generating new images and including diversity exposes CNN models to more images, which makes their training more robust and their predictions more robus and closer to real scenarios.

Architecture	# parameters	Description	Example
InceptionV3 (Szegedy et al., 2016)	24 million	42 layers that are built on the inception module	• https://github.com/josemenber/image-based-crop-anomaly-detection • https://github.com/angshumanroy77/cropdiseaseclassification
AlexNet (Krizhevsky et al., 2017)	62.3 million	5 convoluional and 3 pooling layers with ReLU activation followed by fully connected layers	https://github.com/Dharmendra444/Crop-Desease-Classification-Using-Alexnet https://github.com/Prajwal10031999/Plant-Diseases-Classification-using-AlexNet
VGG16 (Simonyan & Zisserman, 2014)	138 million	A deep 16-layer model with stacked convolutional layers and pooling layers that feed 3 dense layers	• https://github.com/anirudhjak06/Crop-Disease-Detection • https://github.com/LeadingIndiaAI/Weed-Detection-in-Dense-Culture-using-Deep-Learning-
ResNet50 (He et al., 2016)	25 million	Uses residual blocks to solve the vanishing gradient problem	• https://github.com/josemenber/image-based-crop-anomaly-detection • https://github.com/shreyas-muralidhara/Leaf-Wilting-Classification-Transfer-Learning
DenseNet121 (G. Huang et al., 2017)	7.6 million	121 layers where all layers are connected to each other	• https://www.kaggle.com/code/marquis03/densenet-121-plants-classification
Unet (Ronneberger et al., 2015)	30.1 million	Encoder-decoder model that use skip connection to prevent resolution loss	• https://www.kaggle.com/code/atharvakadeth/farm-plot-detection-unet

Table 6. State-of-the-art CNN description with usage examples in agriculture (March 19th 2024).

Convolutional neural networks

We listed at least 20 studies that use remote sensing and CNN in agriculture, and discussed the methods used at least in 15 others. By deeply analyzing these studies we pointed out that there is not a single general method to accomplish all tasks but it rather depends on the study's goal.

We also found out that transfer learning is a common practice that increases model performance. Transfer learning is an approach that leverages the knowledge of a pretrained model, which have already trained on extracting features and patterns in images of other datasets (Zhuang et al., 2021). By exploiting what the pretrained model has learnt, models effectively captured relevant features when trained on agricultural datasets, especially limited ones that contain few images. Transfer learning happens in two stages. The first one consists in training a model from scratch on a vast and general

dataset or importing one that is already pretrained. In the second stage, the same model is trained and specialized on the final dataset.

We presented a set of studies that propose CNNs real-time detection where they are implemented in remote software on systems fit on drones or ground robots or cameras. R. Zhang et al., 2020 achieved a speed of 2fps with a mobile phone device, while Jeong et al., 2017; Jiang et al., 2019; Milioto et al., 2018 reached best records of 32.2fps, 35.0fps and 23.13fps, respectively, with more sophisticated hardware acceleration. CNN can achieve agricultural tasks in real-time on fields, which is expected to increase monitoring speed and productivity for farmers, food technologists, and agricultural engineers.

We also notice that state-of-the-art CNNs from computer vision such as AlexNet (Krizhevsky et al., 2017), VGG (Simonyan & Zisserman, 2014), Inception (Szegedy et al., 2017), GoogLeNet (Szegedy et al., 2015), and ResNet (He et al., 2016) often appear in studies in agriculture. In this case, state-of-the-art CNNs refer to model architectures that incorporate an innovative design that often represents a cutting edge of research and development. Continuous research results in the evolution of CNNs over time as new technologies are proposed and validated through experimentation. In Table 6 we describe some state-of-the-art architectures and we provide examples in agriculture that use and implement them.

The success of CNNs in computer vision inspired researchers to include the convolution operation in different deep learning families such as Graph Neural Networks (GNNs) (Zhou et al., 2020) which became Graph Convolutional Networks (GCNs) (S. Zhang et al., 2019). In the latter, the convolution operation is known as message passing, during which each node receives vectors of values from its neighbors in order to update its vector by aggregating them. Some studies used GCNs in agriculture to accomplish computer vision tasks, by considering the patches of an image as the nodes of a graph (K. Hu et al., 2020) or by representing the low-level features of an image as a graph (H. Jiang et al., 2020). Recently, more complex technologies gained a lot of popularity in computer vision, especially Transformers, which are based on the multi-head attention mechanism (Khan et al., 2022; Vaswani et al., 2023).

Transformer models were initially designed to process text (Vaswani et al., 2023) and their success inspired their adaptation to computer vision, especially since 2020 (Moutik et al., 2023). Despite their large success, it is still unclear if transformers are more performant and could replace CNNs (Pinto et al., 2022). In fact, studies have shown that recent CNN architectures can be as robust and reliable as transformers (Pinto et al., 2022) and in some situations outperform them (Bai et al., 2021; Matsoukas et al., 2021; Pinto et al., 2022). It is difficult to determine which family of models is the best because of their similar performances and vulnerabilities (Deininger et al., 2022; Matsoukas et al., 2021; Pinto et al., 2022). Some studies have suggested that using hybrid CNN-transformer models is the best solution because these models could fill each others gaps (Moutik et al., 2023). Overall, there is no absolute best architecture, as their performance varies depending on the dataset, the hyperparameters, and the study context. Additionally, researchers sometimes develop CNNs that perform just as well as well-known state-of-the-art CNN architectures.

CNNs, similarly to other deep learning models, often present a black box nature which is due to the large number of parameters that a model has and to the various hyperparameter combinations (X. Li et al., 2022). This makes explaining and interpreting the predictions and the decisions of a model more difficult. Interpretable deep learning aims to solve this problem by providing a set of tools that help interpreting deep learning models, for example, by visualizing the regions of importance of an image in computer vision (Linardatos et al., 2021). Various methods such as Gradients (Simonyan et al., 2014), DeepLIFT (Shrikumar et al., 2017), Class Activation Maps (CAMs) (B. Zhou et al., 2016), Grad-CAM (Selvaraju et al., 2017), or Grad-CAM++ (Chattopadhay et al., 2018), make it possible to interpret deep learning models (Linardatos et al., 2021). Having robust models with good performance is not enough, especially in critical applications of agriculture such as food safety, but trustworthiness and interpretation are also needed.

Future perspectives

Although CNNs showed good capabilities at accomplishing different applications in agriculture using different types of data, some challenges and questions remain prevalent. First, the interpretability of CNNs in agriculture remains a significant challenge that is not often addressed in research papers. Understanding how CNNs make their predictions by interpreting their decision-making process is important for improving their performance. Additionally, CNNs often have a large number of parameters and require a lot of computational power to do training. The computational cost in terms of energy of CNNs in agriculture has still to be considered in order to conclude on the energy efficiency of these models in this field. Additionally, CNNs were designed to process images but they can be used with numerical and non-image data as well (Srivastava et al., 2022). In agriculture, a lot of data come from in-situ measurements, such as moisture content or pH, therefore, CNNs could be beneficial with multimodality in agriculture. Lastly, with the rapid emergence of advanced technologies, especially Transformers and GCNs, it is relevant to review the literature for use cases that employ CNNs in comparison with other technologies.

Conclusion

In this literature review, we discussed the importance of multispectral remote sensing and vegetation indexes, and we provided a list of different image datasets. We also presented how convolutional neural networks (CNNs) are useful in accomplishing tasks in agriculture. These tasks consist mainly of: weed detection, disease detection, crop type and quality classification, yield prediction, and water management. There is not an absolute best CNN architecture within each of these tasks nor is there a best one among all the tasks. Instead, current studies rely on a set of state-of-the-art CNN architectures that achieve good performance; data preprocessing is used to improve their results and

transfer learning or fine tuning is also considered.

   This review article summarizes the current state of utilization of CNN for agriculture applications. Moving forward, there is a need to run experiments on multiple tasks to compare different CNN architectures with additional features such as vegetation indexes in a more systematic manner.

Disclosure statement

The authors declare that they have no financial or non-financial competing interest. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Data availability

Not applicable

References

Adarsh, P., Rathi, P., & Kumar, M. (2020). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. *2020 6th International*

Conference on Advanced Computing and Communication Systems (ICACCS), 687–694.

Agarwal, M., Gupta, S. K., & Biswas, K. (2020). Development of Efficient CNN model for Tomato crop disease identification. *Sustainable Computing: Informatics and Systems*, *28*, 100407.

Agarwal, M., Singh, A., Arjaria, S., Sinha, A., & Gupta, S. (2020). ToLeD: Tomato leaf disease detection using convolution neural network. *Procedia Computer Science*, *167*, 293–301.

Alencastre-Miranda, M., Davidson, J. R., Johnson, R. M., Waguespack, H., & Krebs, H. I. (2018). Robotics for Sugarcane Cultivation: Analysis of Billet Quality using Computer Vision. *IEEE Robotics and Automation Letters*, *3*(4), 3828–3835. https://doi.org/10.1109/LRA.2018.2856999

Altaheri, H., Alsulaiman, M., Muhammad, G., Amin, S. U., Bencherif, M., & Mekhtiche, M. (2019). Date fruit dataset for intelligent harvesting. *Data in Brief*, *26*, 104514. https://doi.org/10.1016/j.dib.2019.104514

Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. *Datenbanksysteme Für Business, Technologie Und Web (BTW 2017)-Workshopband*.

Amatya, S., Karkee, M., Gongal, A., Zhang, Q., & Whiting, M. D. (2016). Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting. *Biosystems Engineering*, *146*, 3–15. https://doi.org/10.1016/j.biosystemseng.2015.10.003

An, J., Li, W., Li, M., Cui, S., & Yue, H. (2019). Identification and classification of maize drought stress using deep convolutional neural network. *Symmetry*, *11*(2), 256.

Angelo Randaci, Earth's Ally. (2021, March 10). *Common Plant Diseases & Disease Control for Organic Gardens*. https://earthsally.com/disease-control/common-plant-diseases.html

Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2021). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, *61*, 101182.

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495.

Bai, Y., Mei, J., Yuille, A. L., & Xie, C. (2021). Are Transformers more robust than CNNs? *Advances in Neural Information Processing Systems*, *34*, 26831–26843. https://proceedings.neurips.cc/paper/2021/hash/e19347e1c3ca0c0b97de5fb3b69 0855a-Abstract.html

Bannari, A., Morin, D., Bonn, F., & Huete, A. (1995). A review of vegetation indices. *Remote Sensing Reviews*, *13*(1–2), 95–120.

Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, *180*, 96–107.

Barbedo, J. G. A., Koenigkan, L. V., Halfeld-Vieira, B. A., Costa, R. V., Nechet, K. L., Godoy, C. V., Junior, M. L., Patricio, F. R. A., Talamini, V., Chitarra, L. G., &

others. (2018). Annotated plant pathology databases for image-based detection and recognition of diseases. *IEEE Latin America Transactions*, *16*(6), 1749–1757.

Birth, G. S., & McVey, G. R. (1968). Measuring the Color of Growing Turf with a Reflectance Spectrophotometer1. *Agronomy Journal*, *60*(6), 640–643. https://doi.org/10.2134/agronj1968.00021962006000060016x

Boegh, E., Soegaard, H., Broge, N., Hasager, C. B., Jensen, N. O., Schelde, K., & Thomsen, A. (2002). Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sensing of Environment*, *81*(2), 179–193. https://doi.org/10.1016/S0034-4257(01)00342-X

Bosilj, P., Aptoula, E., Duckett, T., & Cielniak, G. (2019). Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. *Journal of Field Robotics*, *to be determined (published online)*.

Bosilj, P., Aptoula, E., Duckett, T., & Cielniak, G. (2020). Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. *Journal of Field Robotics*, *37*(1), 7–19.

Burkov, A. (2019). *The Hundred-Page Machine Learning Book*. Andriy Burkov. https://books.google.fr/books?id=0jbxwQEACAAJ

Butte, S., Vakanski, A., Duellman, K., Wang, H., & Mirkouei, A. (2021). Potato crop stress identification in aerial images using deep learning-based object detection. *Agronomy Journal*, *113*(5), 3991–4002. https://doi.org/10.1002/agj2.20841

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., & Grégoire, J.-M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sensing of Environment*, *77*(1), 22–33. https://doi.org/10.1016/S0034-4257(01)00191-2

Chandel, N. S., Chakraborty, S. K., Rajwade, Y. A., Dubey, K., Tiwari, M. K., & Jat, D. (2021). Identifying crop water stress using deep learning models. *Neural Computing and Applications*, *33*, 5353–5367.

Chattopadhay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 839–847. https://doi.org/10.1109/WACV.2018.00097

Chebrolu, N., Lottes, P., Schaefer, A., Winterhalter, W., Burgard, W., & Stachniss, C. (2017). Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields. *The International Journal of Robotics Research*, *36*(10), 1045–1052. https://doi.org/10.1177/0278364917720510

Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. (2021). Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sensing*, *13*(22), Article 22. https://doi.org/10.3390/rs13224712

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification

evaluation. *BMC Genomics*, *21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Chiu, M. T., Xu, X., Wei, Y., Huang, Z., Schwing, A., Brunner, R., Khachatrian, H., Karapetyan, H., Dozier, I., Rose, G., & others. (2020). Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. *arXiv Preprint arXiv:2001.01306.*

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Czymmek, V., Harders, L. O., Knoll, F. J., & Hussmann, S. (2019). Vision-based deep learning approach for real-time detection of weeds in organic farming. *2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 1–5.

Deininger, L., Stimpel, B., Yuce, A., Abbasi-Sureshjani, S., Schönenberger, S., Ocampo, P., Korski, K., & Gaire, F. (2022). *A comparative study between vision transformers and CNNs in digital pathology* (arXiv:2206.00389). arXiv. http://arxiv.org/abs/2206.00389

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255.

dos Santos Ferreira, A., Freitas, D. M., da Silva, G. G., Pistori, H., & Folhes, M. T. (2017). Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture*, *143*, 314–324.

Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vasilakoglou, I. (2020). Towards weeds identification assistance through transfer learning. *Computers and Electronics in Agriculture*, *171*, 105306.

Fawakherji, M., Youssef, A., Bloisi, D., Pretto, A., & Nardi, D. (2019). Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation. *2019 Third IEEE International Conference on Robotic Computing (IRC)*, 146–152.

Food and Agriculture Organization of the United Nations. (2019). *FAO - News Article: New standards to curb the global spread of plant pests and diseases.* https://www.fao.org/news/story/en/item/1187738/icode/

Gao, B.-C. (1996). NDWI?A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sensing of Environment*, *58*, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gao, J., French, A. P., Pound, M. P., He, Y., Pridmore, T. P., & Pieters, J. G. (2020). Deep convolutional neural networks for image-based Convolvulus sepium detection in sugar beet fields. *Plant Methods*, *16*(1), 1–12.

Gao, W., Zhang, X., Yang, L., & Liu, H. (2010). An improved Sobel edge detection. *2010 3rd International Conference on Computer Science and Information Technology*, *5*, 67–71. https://doi.org/10.1109/ICCSIT.2010.5563693

Gatti, A., & Bertolini, A. (2013). Sentinel-2 products specification document. *Available Online (Accessed February 23, 2015) Https://Earth. Esa.*

Int/Documents/247904/685211/Sentinel-2+ Products+ Specification+ Document.

Geetha, V., Punitha, A., Abarna, M., Akshaya, M., Illakiya, S., & Janani, A. (2020). An effective crop prediction using random forest algorithm. *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 1–5.

Giselsson, T. M., Dyrmann, M., Jørgensen, R. N., Jensen, P. K., & Midtiby, H. S. (2017). A Public Image Database for Benchmark of Plant Seedling Classification Algorithms. *arXiv Preprint*.

Gitelson, A. A., Merzlyak, M. N., & Lichtenthaler, H. K. (1996). Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements Near 700 nm. *Journal of Plant Physiology*, *148*(3), 501–508. https://doi.org/10.1016/S0176-1617(96)80285-9

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). Deep learning for plant identification using vein morphological patterns. *Computers and Electronics in Agriculture*, *127*, 418–424. https://doi.org/10.1016/j.compag.2016.07.003

Gull, A., Lone, A. A., & Wani, N. U. I. (2019). Biotic and abiotic stresses in plants. *Abiotic and Biotic Stress in Plants*, 1–19.

Hasan, A. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. (2021). A survey of deep learning techniques for weed detection from images. *Computers and Electronics in Agriculture*, *184*, 106067.

Haug, S., & Ostermann, J. (2015). A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. *European Conference on Computer Vision*, 105–116.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Hegazi, E. H., Samak, A. A., Yang, L., Huang, R., & Huang, J. (2023). Prediction of Soil Moisture Content from Sentinel-2 Images Using Convolutional Neural Network (CNN). *Agronomy*, *13*(3), 656.

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2217–2226.

Hu, G., Wu, H., Zhang, Y., & Wan, M. (2019). A low shot learning method for tea leaf's disease identification. *Computers and Electronics in Agriculture*, *163*, 104852.

Hu, K., Coleman, G., Zeng, S., Wang, Z., & Walsh, M. (2020). Graph weeds net: A graph-based deep learning method for weed recognition. *Computers and Electronics in Agriculture*, *174*, 105520. https://doi.org/10.1016/j.compag.2020.105520

Hu, K., Wang, Z., Coleman, G., Bender, A., Yao, T., Zeng, S., Song, D., Schumann, A., & Walsh, M. (2021). Deep learning techniques for in-crop weed identification: A Review. *arXiv Preprint arXiv:2103.14872*.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708.

Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, *70*(1–3), 489–501.

Huang, M., & Chuang, T. (2020). A database of eight common tomato pest images. *Mendeley Data*.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295–309.

Hughes, D., Salathé, M., & others. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv Preprint arXiv:1511.08060*.

Jackson, R. D., & Huete, A. R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*, *11*(3), 185–200. https://doi.org/10.1016/S0167-5877(05)80004-2

Jayaraman, P. P., Yavari, A., Georgakopoulos, D., Morshed, A., & Zaslavsky, A. (2016). Internet of things platform for smart farming: Experiences and lessons learnt. *Sensors*, *16*(11), 1884.

Jeong, J., Park, H., & Kwak, N. (2017). Enhancement of SSD by concatenating feature maps for object detection. *arXiv Preprint arXiv:1705.09587*.

Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture*, *2*, 1–12.

Jiang, H., Zhang, C., Qiao, Y., Zhang, Z., Zhang, W., & Song, C. (2020). CNN feature based graph convolutional network for weed and crop recognition in smart farming. *Computers and Electronics in Agriculture*, *174*, 105450. https://doi.org/10.1016/j.compag.2020.105450

Jiang, P., Chen, Y., Liu, B., He, D., & Liang, C. (2019). Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*, *7*, 59069–59080.

Johann, A. L., de Araújo, A. G., Delalibera, H. C., & Hirakawa, A. R. (2016). Soil moisture modeling based on stochastic behavior of forces on a no-till chisel opener. *Computers and Electronics in Agriculture*, *121*, 420–428. https://doi.org/10.1016/j.compag.2015.12.020

Kaggle. (2019). *New Plant Diseases Dataset*. https://www.kaggle.com/datasets/vipooool/new-plant-diseases-dataset

Kamarudin, M. H., Ismail, Z. H., & Saidi, N. B. (2021). Deep learning sensor fusion in plant water stress assessment: A comprehensive review. *Applied Sciences*, *11*(4), 1403.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018a). A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science*, *156*(3), 312–322.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018b). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70–90.

Kelleher, J. D. (2019). *Deep learning*. MIT press.

Khaki, S., Pham, H., & Wang, L. (2020). Yieldnet: A convolutional neural network for simultaneous corn and soybean yield prediction based on remote sensing data. *bioRxiv*, 2020–12.

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in Vision: A Survey. *ACM Computing Surveys*, *54*(10s), 1–41. https://doi.org/10.1145/3505244

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, *151*, 107398. https://doi.org/10.1016/j.ymssp.2020.107398

Kok, Z. H., Shariff, A. R. M., Alfatni, M. S. M., & Khairunniza-Bejo, S. (2021). Support vector machine in precision agriculture: A review. *Computers and Electronics in Agriculture*, *191*, 106546.

Kordi, F., & Yousefi, H. (2022). Crop classification based on phenology information by using time series of optical and synthetic-aperture radar images. *Remote Sensing Applications: Society and Environment*, *27*, 100812. https://doi.org/10.1016/j.rsase.2022.100812

Kounalakis, T., Triantafyllidis, G. A., & Nalpantidis, L. (2019). Deep learning-based visual recognition of rumex for robotic precision farming. *Computers and Electronics in Agriculture*, *165*, 104973.

Koutsoukas, A., Monaghan, K. J., Li, X., & Huan, J. (2017). Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data. *Journal of Cheminformatics*, *9*(1), 42. https://doi.org/10.1186/s13321-017-0226-y

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Kuhn, M., & Johnson, K. (2013). Measuring performance in regression models. *Applied Predictive Modeling*, 95–100.

Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, *14*(5), 778–782.

Laben, C. A., & Brower, B. V. (2000). *Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening*. Google Patents.

Lathuilière, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2020). A Comprehensive Analysis of Deep Regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(9), 2065–2081. https://doi.org/10.1109/TPAMI.2019.2910523

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Li, N., Zhang, X., Zhang, C., Guo, H., Sun, Z., & Wu, X. (2019). Real-time crop recognition in transplanted fields with prominent weed growth: A visual-attention-based approach. *IEEE Access*, *7*, 185310–185321.

Li, W., Fu, H., Yu, L., & Cracknell, A. (2016). Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sensing*, *9*(1), 22.

Li, X., Xiong, H., Li, X., Wu, X., Zhang, X., Liu, J., Bian, J., & Dou, D. (2022). Interpretable deep learning: Interpretation, interpretability, trustworthiness, and beyond. *Knowledge and Information Systems*, *64*(12), 3197–3234. https://doi.org/10.1007/s10115-022-01756-8

Li, Z., Yang, W., Peng, S., & Liu, F. (2020). *A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects* (arXiv:2004.02806). arXiv. http://arxiv.org/abs/2004.02806

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674.

Lin, T., Wang, Y., Liu, X., & Qiu, X. (2022). A survey of transformers. *AI Open*, *3*, 111–132. https://doi.org/10.1016/j.aiopen.2022.10.001

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy*, *23*(1), Article 1. https://doi.org/10.3390/e23010018

Liu, B., Zhang, Y., He, D., & Li, Y. (2017). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, *10*(1), 11.

Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: A review. *Plant Methods*, *17*, 1–18.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. *European Conference on Computer Vision*, 21–37.

Lu, Y., & Young, S. (2020). A survey of public datasets for computer vision tasks in precision agriculture. *Computers and Electronics in Agriculture*, *178*, 105760.

M, H., & M.N, S. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 01–11. https://doi.org/10.5121/ijdkp.2015.5201

Ma, X., Deng, X., Qi, L., Jiang, Y., Li, H., Wang, Y., & Xing, X. (2019). Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. *PloS One*, *14*(4), e0215676.

Maione, C., Batista, B. L., Campiglia, A. D., Barbosa, F., & Barbosa, R. M. (2016). Classification of geographic origin of rice by data mining and inductively

coupled plasma mass spectrometry. *Computers and Electronics in Agriculture*, *121*, 101–107. https://doi.org/10.1016/j.compag.2015.11.009

Matsoukas, C., Haslum, J. F., Söderberg, M., & Smith, K. (2021). *Is it Time to Replace CNNs with Transformers for Medical Images?* (arXiv:2108.09038). arXiv. http://arxiv.org/abs/2108.09038

Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Computers and Electronics in Agriculture*, *139*, 103–114. https://doi.org/10.1016/j.compag.2017.05.002

Meshram, V., & Patil, K. (2022). FruitNet: Indian fruits image dataset with quality for machine learning applications. *Data in Brief*, *40*, 107686. https://doi.org/10.1016/j.dib.2021.107686

Mignoni, M. E., Honorato, A., Kunst, R., Righi, R., & Massuquetti, A. (2022). Soybean images dataset for caterpillar and Diabrotica speciosa pest detection and classification. *Data in Brief*, *40*, 107756. https://doi.org/10.1016/j.dib.2021.107756

Milioto, A., Lottes, P., & Stachniss, C. (2018). Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs. *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2229–2235.

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2020). *Image Segmentation Using Deep Learning: A Survey* (arXiv:2001.05566). arXiv. http://arxiv.org/abs/2001.05566

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv Preprint arXiv:1411.1784*.

Mlsna, P. A., & Rodríguez, J. J. (2009). Chapter 19—Gradient and Laplacian Edge Detection. In A. Bovik (Ed.), *The Essential Guide to Image Processing* (pp. 495–524). Academic Press. https://doi.org/10.1016/B978-0-12-374457-9.00019-6

Moshou, D., Pantazi, X.-E., Kateris, D., & Gravalos, I. (2014). Water stress detection based on optical multisensor fusion with a least squares support vector machine classifier. *Biosystems Engineering*, *117*, 15–22. https://doi.org/10.1016/j.biosystemseng.2013.07.008

Moutik, O., Sekkat, H., Tigani, S., Chehri, A., Saadane, R., Tchakoucht, T. A., & Paul, A. (2023). Convolutional Neural Networks or Vision Transformers: Who Will Win the Race for Action Recognitions in Visual Data? *Sensors*, *23*(2), Article 2. https://doi.org/10.3390/s23020734

Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Computers and Electronics in Agriculture*, *163*, 104859.

Nguyen, T. T., Hoang, T. D., Pham, M. T., Vu, T. T., Nguyen, T. H., Huynh, Q.-T., & Jo, J. (2020). Monitoring agriculture areas with satellite images and deep learning. *Applied Soft Computing*, *95*, 106565.

Noyan, M. A. (2022). Uncovering bias in the PlantVillage dataset. *arXiv Preprint arXiv:2206.04374*.

O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv Preprint arXiv:1511.08458*.

Pandian, J. A., & Gopal, G. (2019). *Data for: Identification of Plant Leaf Diseases Using a 9-layer Deep Convolutional Neural Network. 1.* https://doi.org/10.17632/tywbtsjrjv.1

Pantazi, X.-E., Moshou, D., & Bravo, C. (2016). Active learning system for weed species recognition based on hyperspectral sensing. *Biosystems Engineering*, *146*, 193–202. https://doi.org/10.1016/j.biosystemseng.2016.01.014

Pegorini, V., Zen Karam, L., Pitta, C. S. R., Cardoso, R., Da Silva, J. C. C., Kalinowski, H. J., Ribeiro, R., Bertotti, F. L., & Assmann, T. S. (2015). In Vivo Pattern Classification of Ingestive Behavior in Ruminants Using FBG Sensors and Machine Learning. *Sensors*, *15*(11), Article 11. https://doi.org/10.3390/s151128456

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, *4*(2), 1883.

Petrich, L., Lohrmann, G., Neumann, M., Martin, F., Frey, A., Stoll, A., & Schmidt, V. (2020). Detection of Colchicum autumnale in drone images, using a machine-learning approach. *Precision Agriculture*, *21*(6), 1291–1303.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. *Remote Sensing*, *12*(14), 2291.

Pinto, F., Torr, P. H. S., & K. Dokania, P. (2022). An Impartial Take to the CNN vs Transformer Robustness Contest. In S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Eds.), *Computer Vision – ECCV 2022* (pp. 466–480). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-19778-9_27

Prajapati, H. B., Shah, J. P., & Dabhi, V. K. (2017). Detection and classification of rice plant diseases. *Intelligent Decision Technologies*, *11*(3), 357–373.

Pratyush Reddy, K. S., Roopa, Y. M., Rajeev L.N., K., & Nandan, N. S. (2020). IoT based Smart Agriculture using Machine Learning. *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 130–134. https://doi.org/10.1109/ICIRCA48905.2020.9183373

Rahnemoonfar, M., & Sheppard, C. (2017). Deep count: Fruit counting based on deep simulated learning. *Sensors*, *17*(4), 905.

Rajesh, B., Vardhan, M. V. S., & Sujihelen, L. (2020). Leaf disease detection and classification by decision tree. *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 705–708.

Razfar, N., True, J., Bassiouny, R., Venkatesh, V., & Kashef, R. (2022). Weed detection in soybean crops using custom lightweight deep learning models. *Journal of Agriculture and Food Research*, *8*, 100308.

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv Preprint arXiv:1804.02767*.

Ribera, J., Chen, Y., Boomsma, C., & Delp, E. J. (2017). Counting plants using deep learning. *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 1344–1348.

Roelofs, R., Shankar, V., Recht, B., Fridovich-Keil, S., Hardt, M., Miller, J., & Schmidt, L. (2019). A Meta-Analysis of Overfitting in Machine Learning. *Advances in Neural Information Processing Systems*, *32*. https://proceedings.neurips.cc/paper/2019/hash/ee39e503b6bedf0c98c388b7e85 89aca-Abstract.html

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Rouse Jr, J. W., Haas, R. H., Schell, J., & Deering, D. (1973). *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation*.

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., & others. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172.

Ruckelshausen, A., Biber, P., Dorna, M., Gremmes, H., Klose, R., Linz, A., Rahe, F., Resch, R., Thiel, M., Trautz, D., & others. (2009). BoniRob–an autonomous field robot platform for individual plant phenotyping. *Precision Agriculture*, *9*(841), 1.

Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., & Siegwart, R. (2018). weedNet: Dense Semantic Weed Classification Using Multispectral Images and MAV for Smart Farming. *IEEE Robotics and Automation Letters*, *3*(1), 588–595. https://doi.org/10.1109/LRA.2017.2774979

Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, *8*(11), 468.

Segaran, T. (2007). *Programming Collective Intelligence: Building Smart Web 2.0 Applications*. O'Reilly Media. https://books.google.fr/books?id=fEsZ3Ey-Hq4C

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). *Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization*. 618–626. https://openaccess.thecvf.com/content_iccv_2017/html/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.html

Serrano, L., Penuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2–3), 355–364.

Serrano, L., Peñuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2), 355–364. https://doi.org/10.1016/S0034-4257(02)00011-1

Shrikumar, A., Greenside, P., & Kundaje, A. (2017). Learning Important Features Through Propagating Activation Differences. *Proceedings of the 34th International Conference on Machine Learning*, 3145–3153. https://proceedings.mlr.press/v70/shrikumar17a.html

Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). *Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps* (arXiv:1312.6034). arXiv. https://doi.org/10.48550/arXiv.1312.6034

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv Preprint arXiv:1409.1556*.

Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S., & Batra, N. (2020). PlantDoc: A Dataset for Visual Plant Disease Detection. *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, 249–253. https://doi.org/10.1145/3371158.3371196

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, *12*(19), 3136.

Skovsen, S., Dyrmann, M., Mortensen, A. K., Laursen, M. S., Gislum, R., Eriksen, J., Farkhani, S., Karstoft, H., & Jorgensen, R. N. (2019, June). The GrassClover Image Dataset for Semantic and Hierarchical Species Understanding in Agriculture. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Soltani, N., Dille, J. A., Burke, I. C., Everman, W. J., VanGessel, M. J., Davis, V. M., & Sikkema, P. H. (2016). Potential Corn Yield Losses from Weeds in North America. *Weed Technology*, *30*(4), 979–984. https://doi.org/10.1614/WT-D-16-00046.1

Song, R., Zhang, Z., & Liu, H. (2017). Edge connection based Canny edge detection algorithm. *Pattern Recognition and Image Analysis*, *27*(4), 740–747. https://doi.org/10.1134/S1054661817040162

Sothearith, Y., Appiah, K. S., Mardani, H., Motobayashi, T., Yoko, S., Eang Hourt, K., Sugiyama, A., & Fujii, Y. (2021). Determination of the Allelopathic Potential of Cambodia's Medicinal Plants Using the Dish Pack Method. *Sustainability*, *13*(16), Article 16. https://doi.org/10.3390/su13169062

Sripada, R. P., Heiniger, R. W., White, J. G., & Meijer, A. D. (2006). Aerial Color Infrared Photography for Determining Early In-Season Nitrogen Requirements in Corn. *Agronomy Journal*, *98*(4), 968–977. https://doi.org/10.2134/agronj2005.0200

Srivastava, A. K., Safaei, N., Khaki, S., Lopez, G., Zeng, W., Ewert, F., Gaiser, T., & Rahimi, J. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. *Scientific Reports*, *12*(1), 3215.

Subeesh, A., Bhole, S., Singh, K., Chandel, N. S., Rajwade, Y. A., Rao, K., Kumar, S., & Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture*, *6*, 47–54.

Sudars, K., Jasko, J., Namatevs, I., Ozola, L., & Badaukis, N. (2020). Dataset of annotated food crops and weed images for robotic computer vision control. *Data in Brief*, *31*, 105833.

Suh, H. K., Ijsselmuiden, J., Hofstee, J. W., & van Henten, E. J. (2018). Transfer learning for the classification of sugar beet and volunteer potato under field conditions. *Biosystems Engineering*, *174*, 50–65.

Sultana, N., Jahan, M., & Uddin, M. S. (2022). An extensive dataset for successful recognition of fresh and rotten fruits. *Data in Brief*, *44*, 108552. https://doi.org/10.1016/j.dib.2022.108552

Sun, H., Liu, H., Ma, Y., & Xia, Q. (2021). Optical remote sensing indexes of soil moisture: Evaluation and improvement based on aircraft experiment observations. *Remote Sensing*, *13*(22), 4638.

Sun, Z., Di, L., Fang, H., & Burgess, A. (2020). Deep learning classification for crop types in north dakota. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2200–2213.

Suresha, M., Shreekanth, K. N., & Thirumalesh, B. V. (2017). Recognition of diseases in paddy leaves using knn classifier. *2017 2nd International Conference for Convergence in Technology (I2CT)*, 663–666. https://doi.org/10.1109/I2CT.2017.8226213

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. *Thirty-First AAAI Conference on Artificial Intelligence*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818–2826.

Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, *15*(1), 1–28.

Thompson, C. N., Guo, W., Sharma, B., & Ritchie, G. L. (2019). Using normalized difference red edge index to assess maturity in cotton. *Crop Science*, *59*(5), 2167–2177.

Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture*, *7*(1), 1–19. https://doi.org/10.1016/j.inpa.2019.09.006

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Türkoğlu, M., & Hanbay, D. (2019). Plant disease and pest detection using deep learning-based features. *Turkish Journal of Electrical Engineering and Computer Sciences*, *27*(3), 1636–1651.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (arXiv:1706.03762). arXiv. https://doi.org/10.48550/arXiv.1706.03762

Veeranampalayam Sivakumar, A. N., Li, J., Scott, S., Psota, E., J. Jhala, A., Luck, J. D., & Shi, Y. (2020). Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. *Remote Sensing*, *12*(13), 2136.

Vincent, B., & Dardenne, P. (2021). Application of NIR in Agriculture. In Y. Ozaki, C. Huck, S. Tsuchikawa, & S. B. Engelsen (Eds.), *Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and Applications* (pp. 331–345). Springer. https://doi.org/10.1007/978-981-15-8648-4_14

Wang, L., Qu, J. J., & Hao, X. (2008). Forest fire detection using the normalized multi-band drought index (NMDI) with satellite measurements. *Agricultural and Forest Meteorology*, *148*(11), 1767–1776. https://doi.org/10.1016/j.agrformet.2008.06.005

Weng, L., Kang, Y., Jiang, K., & Chen, C. (2022). Time Gated Convolutional Neural Networks for Crop Classification. *arXiv Preprint arXiv:2206.09756*.

Wiesner-Hanks, T., & Brahimi, M. (2018). *Image set for deep learning: Field images of maize annotated with disease symptoms*. https://osf.io/p67rz/

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, *38*(1), 259–269.

Wu, X., Zhan, C., Lai, Y., Cheng, M.-M., & Yang, J. (2019). IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. *IEEE CVPR*, 8787–8796.

Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. *Sensors*, *21*(11), 3647.

Yaloveha, V., Podorozhniak, A., & Kuchuk, H. (2022). Convolutional neural network hyperparameter optimization applied to land cover classification. *Radioelectronic and Computer Systems*, *1*, 115–128.

Yin, Y., Li, H., & Fu, W. (2020). Faster-YOLO: An accurate and faster object detection method. *Digital Signal Processing*, *102*, 102756.

Yoo, H.-J. (2015). Deep convolution neural networks in computer vision: A review. *IEIE Transactions on Smart Processing and Computing*, *4*(1), 35–43.

Yu, J., Schumann, A. W., Cao, Z., Sharpe, S. M., & Boyd, N. S. (2019). Weed detection in perennial ryegrass with deep learning convolutional neural network. *Frontiers in Plant Science*, *10*, 1422.

Yu, J., Sharpe, S. M., Schumann, A. W., & Boyd, N. S. (2019). Deep learning for image-based weed detection in turfgrass. *European Journal of Agronomy*, *104*, 78–84.

Yue, J., Tian, J., Tian, Q., Xu, K., & Xu, N. (2019). Development of soil moisture indices from differences in water absorption between shortwave-infrared bands. *ISPRS Journal of Photogrammetry and Remote Sensing*, *154*, 216–230. https://doi.org/10.1016/j.isprsjprs.2019.06.012

Zhang, K., Wu, Q., Liu, A., & Meng, X. (2018). Can deep learning identify tomato leaf disease? *Advances in Multimedia*, *2018*.

Zhang, R., Wang, C., Hu, X., Liu, Y., Chen, S., & others. (2020). Weed location and recognition based on UAV imaging and deep learning. *International Journal of Precision Agricultural Aviation*, *3*(1).

Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019). Graph convolutional networks: A comprehensive review. *Computational Social Networks*, *6*(1), 11. https://doi.org/10.1186/s40649-019-0069-y

Zhong, Y., & Zhao, M. (2020). Research on deep learning in apple leaf disease recognition. *Computers and Electronics in Agriculture*, *168*, 105146.

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). *Learning Deep Features for Discriminative Localization*. 2921–2929. https://openaccess.thecvf.com/content_cvpr_2016/html/Zhou_Learning_Deep_Features_CVPR_2016_paper.html

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, *1*, 57–81. https://doi.org/10.1016/j.aiopen.2021.01.001

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, *109*(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555

Appendices

VI	Definition	Formula
Normalized Difference Vegetation Index (Rouse Jr et al., 1973)	This index is a measure of healthy, green vegetation. The combination of its normalized difference formulation and use of the highest absorption and reflectance regions of chlorophyll make it robust over a wide range of conditions	$NDVI = \dfrac{p_{800} - p_{680}}{p_{800} + p_{680}}$
Soil Adjusted Vegetation Index (Huete, 1988)	This index is similar to NDVI, but it suppresses the effects of soil pixels.	$SAVI = \dfrac{1.5 * (p_{800} - p_{680})}{p_{800} + p_{680} + 0.5}$
Leaf Area Index (Boegh et al., 2002)	This index is used to estimate foliage cover and to forecast crop growth and yield.	$LAI = (3.618 * EVI - 0.118) > 0$
Difference Vegetation Index (Tucker, 1979)	This index distinguishes between soil and vegetation	$DVI = p_{800} - p_{680}$
Green Difference Vegetation Index (Sripada et al., 2006)	This index was originally designed with color-infrared photography to predict nitrogren requirements for corn	$GDVI = p_{800} - p_{555}$
Enhanced Vegetation Index (A. Huete et al., 2002)	Improve the NDVI by optimizing the vegetation signal in LAI regions	$EVI = \dfrac{2.5 * p_{800} - p_{680}}{p_{800} + 6 * p_{680} - 7.5 * p_{450} + 1}$
Simple Ratio (Birth & McVey, 1968)	This index is a ratio of (1) the wavelength with highest reflectance for vegetation and (2) the wavelength of the deepest chlorophyll absorption.	$SR = \dfrac{p_{800}}{p_{680}}$
Normalized Difference Nitrogen Index (Serrano, Peñuelas, et al., 2002)	This index is designed to estimate the relative amounts of nitrogen contained in vegetation canopies.	$NDNI = \dfrac{log\left(\frac{1}{p_{1510}}\right) - log\left(\frac{1}{p_{1680}}\right)}{log\left(\frac{1}{p_{1510}}\right) + log\left(\frac{1}{p_{1680}}\right)}$
Moisture Stress Index (Ceccato et al., 2001)	This index is a reflectance measurement that is sensitive to increasing leaf water content.	$MSI = \dfrac{p_{1599}}{p_{819}}$
Normalized Difference Water Index (B.-C. Gao, 1996)	This index is sensitive to changes in vegetation canopy water content because reflectance at 857 nm and 1241 nm has similar but slightly different liquid water absorption properties.	$NDWI = \dfrac{p_{857} - p_{1241}}{p_{857} + p_{1241}}$
Normalized Multi-band Drought Index (Wang et al., 2008)	This index takes into account a soil moisture background to monitor potential drought conditions.	$NMDI = \dfrac{p_{860} - (p_{1640} - p_{2130})}{p_{860} + (p_{1640} - p_{2130})}$

Appendix A. A representative table of different vegetation indexes and their respective formula and description.