Computers and Electronics in Agriculture 230 (2025) 109938

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Review Article

Artificial Neural Networks in Agriculture, the core of artificial intelligence: What, When, and Why

Salvador Castillo-Girones [a], Sandra Munera [b], Marcelino Martínez-Sober [c], José Blasco [a], Sergio Cubero [a], Juan Gómez-Sanchis [c,*]

[a] Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), CV-315, km 10,7,46113, Moncada, (Valencia), Spain
[b] Departamento de Ingeniería Gráfica. Universitat Politècnica de València. Camino de Vera, s/n, 46022, Valencia, Spain
[c] IDAL, Departamento de Ingeniería Electrónica, Universidad de Valencia, Av. de la Universidad, S/N, 46100, Burjassot (Valencia), Spain

ARTICLE INFO

Keywords:
Deep Learning
Neural Networks
Applications
Agricultural Products
Quality

ABSTRACT

Artificial Neural Networks (ANNs) based models have emerged as a powerful tool for solving complex nonlinear problems in agriculture. These models simulate the human nervous system's structure, allowing them to learn hierarchical features from the data and solve nonlinear problems efficiently. Despite requiring a large amount of training data, ANNs with shallow architectures demonstrate superior performance in extracting relevant features and establishing accurate models, instilling confidence in their effectiveness compared to conventional machine learning methods. The versatility of ANNs enables their application in various agricultural domains, including precision agriculture, species classification, phenotyping, and food quality and safety assessment. ANNs combined with image analysis have proven valuable in disease detection, plant phenotyping, and fruit quality evaluation. The use of deep learning in agriculture has experienced exponential growth, as evident from the increasing number of publications in recent years. This article overviews recent advancements in applying ANNs in agriculture. It delves into the fundamental principles behind various types of agricultural data and ANN models, discussing their benefits and challenges. The article offers valuable insights into the proper use and functioning of each neural network, data processing for improved model outcomes, and the diverse applications of ANNs in the agricultural sector. It aims to equip readers with practical information on data utilisation, model selection based on data type, functionality, and current research applications.

1. Introduction

Artificial Intelligence (AI) has transformed numerous industries, providing advanced solutions to complex problems through various techniques and algorithms. Among these techniques, artificial neural networks stand out for their ability to learn and improve, inspired by the structure and functioning of the human brain. The idea behind ANN-based models is to create algorithms that simulate the structure of the human nervous system, which is made up of connected neurons (Nayak et al., 2020). These systems use nodes or artificial neurons interconnected in a layered structure that resembles the human brain, obtaining non-linear relationships from the data more efficiently than other methods (Arul, 2021) to solve complex problems in artificial intelligence. Despite the increased interest in AI by academia, industry, and public institutions, there is no universally accepted definition of AI. Various perspectives have attempted to define AI in terms of human intelligence. Nevertheless, several fundamental elements consistently emerge as essential components of AI, encompassing environmental perception, acquisition and comprehension of input as data, decision-making, execution of actions, autonomous task performance, and attaining predefined goals (S et al., 2020). Among the AI definitions highlights the proposed by the High-Level Expert Group on Artificial Intelligence (HLEG) that states, "AI systems are systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best actions to take to achieve the given goal" (HLEG, 2019).

Deep learning (DL) models are highlighted among the supervised AI-

* Corresponding author.
E-mail addresses: castillo_salgirb@gva.es (S. Castillo-Girones), sanmupi@upv.es (S. Munera), marcelino.martinez@uv.es (M. Martínez-Sober), blasco_josiva@gva.es (J. Blasco), cubero_ser@gva.es (S. Cubero), Juan.Gomez-Sanchis@uv.es (J. Gómez-Sanchis).

https://doi.org/10.1016/j.compag.2025.109938
Received 26 October 2024; Received in revised form 2 January 2025; Accepted 6 January 2025
Available online 16 January 2025

based algorithms (Chollet, 2021). DL is a machine learning subfield based on learning features related to different abstraction levels in a particular problem (Hinton et al., 2006). The DL paradigm uses artificial neural network (ANN)–based models to solve problems. It is expected that DL is associated with ANNs with many hidden layers, but it is not the number of layers that define DL. The underlying concept of deep learning is the hierarchical processing of data (Bengio, 2009), where ANNs are used to obtain increasingly meaningful representations of the data through layered learning.

AI is being applied across many industrial and social sectors and activities. Among them, agriculture presents enormous potential for the development of AI systems. Agriculture has been driven by empirical knowledge passed down from generation to generation. This traditional character, deeply rooted in agricultural practices, has evolved over recent decades thanks to technological advancements following the Industrial Revolution. The 20th century introduced modern machinery and efficient irrigation systems that boosted productivity through fertilizers and pesticides. However, the current state of agricultural technology reveals that while some crops have embraced digitised and sensor-based solutions, conventional techniques still dominate others. In this sense, the emergence of AI represents a new paradigm that redefines how we interpret data, make decisions, and manage resources. Nevertheless, this transition has obstacles for a traditional and relatively less skilled sector compared to industries such as healthcare or manufacturing (Sharma et al., 2021). In industrial environments, processes tend to be predictable.

On the other hand, agriculture operates in a context of high uncertainty, where decisions must be made based on uncontrollable and diverse variables. In this regard, the agricultural sector requires a balance between intuition, grounded in experience with crop management, and data-driven analysis by AI systems that are crucial to ensure social, economical and environmental sustainability in a changing and globalized world. The adaptability of ANNs is particularly relevant to agriculture because of their ability to process large volumes of heterogeneous data, model non-linear relationships, and adapt to dynamic conditions. Agriculture generates diverse data types, ranging from soil composition and weather patterns to satellite imagery and crop health statistics that are often complex, high-dimensional, and noisy. ANNs excel at extracting meaningful insights from such data, making them indispensable for solving agricultural problems.

The relevance of ANNs to agriculture is further underscored by the wide range of applications they enable. Among different ANN applications, yield prediction (Gutiérrez et al., 2019), soil management, precision spraying, crop identification (Xiao et al., 2018), plant phenology (Q. Yang et al. 2020), plant phenotyping (Song et al., 2021), or early pest and disease detection (Polder et al., 2019). In postharvest, these technologies are applied to assess fruit quality (Jahanbakhshi et al., 2020) or prevent fraud (Osako et al., 2020) as principal topics.

Depending on the application, different ANN-based models can be used. Among them, based on the scientific literature, the most used ANNs for agricultural applications are Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Radial Basis Function Networks (RBFN), Autoencoders (AE), Deep Belief Networks (DBN), Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN) and Capsule networks (CN) (Elgendy, 2020). However, in many cases, these techniques are not used correctly. Among the main problems, the incorrect use of nomenclature stands out. Sometimes, the definition of ANN is overused. For example, it is common to assimilate an MLP or other specific techniques to an ANN when an MLP is a particular type of ANN. Another example is the interpretation of a CNN as an evolution of ANNs when it is just another family within ANNs. Currently, public libraries offer various ANN-based techniques that are relatively simple to use. This often leads to the libraries using the default hyperparameters as black boxes fed with an Xtrain and returning a model without adjusting them to the type of data or the problem being addressed. However, for correct and optimal use, parameters that influence the model's reliability must be adjusted, such as the network architecture, the type of activation function or the learning rate. The implication of default hyperparameters is that the model's performance may not be optimal, leading to inaccurate predictions. A typical case is the incorrect use of a continuous output activation function in classification problems.

Another typical error is the use of cross-validation (CV) techniques, resembling generically CV to the k-fold method. K-fold is a widely used CV technique in which the set of samples is divided into k subsets, and the model is trained with k-1 subsets and validated with the remaining one. This process is repeated k times, with a different test subset selected in each iteration. Once the iterations are completed, the accuracy or the error are calculated for each of the produced models, and the final accuracy and error are obtained as an average of the k-trained models. Thus, this technique avoids overfitting by selecting the optimal set of hyperparameters in those cases with a relatively reduced set of samples. However, its use is so widespread that the k-fold may seem like the only cross-validation technique when other techniques exist to better suit the problem under study (Seraj et al., 2023).

A common generalisation is using MAE or RMSE as measures of the performance of regression models and the accuracy in classification problems. Their use is correct, but only give one metric. Moreover, in binary problems, metrics such as accuracy, sensitivity, specificity, recall, or F1-score are expected to be found, which provides insights into the model's performance. However, it may lack robustness when dealing with imbalanced datasets and may not be suitable for specific situations where one class holds greater importance. For example, in a task involving the detection of cancer patients, it is preferable to classify a healthy patient as having cancer than to misclassify an ill patient as healthy. The potential consequence of not considering the class imbalance issue is that the model may be biased towards the majority class, leading to poor performance in detecting the minority class. In this case, it would be advisable to use other metrics, such as the area under the ROC curve. It would be advisable to perform an exploratory analysis of the data before feeding it into the algorithm to check for imbalances and, if necessary, employ techniques to correct biases, such as over- or undersampling techniques or adjusting model parameters to adjust the class weights (Naidu et al., 2023).

Furthermore, little attention is paid to feature engineering. One of the significant advantages of neural models is their ability to extract complex relationships in data. ANNs extract the relationships between the data to produce an output regardless of the input data. While this is true, it is not entirely accurate. To obtain the best results, it is essential to perform sound feature engineering, either extracting the most important features, retaining the most information or selecting those that allow the problem to be modelled with greater simplicity. Finally, a common mistake might be using complex techniques without attempting simpler ones first. However, the problem can be solved using more straightforward classical methods that are easier to interpret. For example, if a problem can be solved with a linear model, it is unnecessary to use a neural one. Furthermore, by observing the parameters of the linear model, some information about the data could be obtained, such as which variables are most important.

This article presents an overview of recent advancements in applying ANNs in agriculture. It explores the fundamental principles behind various types of agricultural data and ANN models, discussing their benefits and challenges. Addressing these topics comprehensively provides a solid foundation for understanding why neural networks are particularly suited for this sector compared to others.

2. Sensing in agriculture. Type of data

Artificial detection in agriculture uses sensors to collect and analyse data about the agricultural ecosystem. It is the basis for generating knowledge and creating tools to help farmers make informed decisions about crop management, such as determining when to plant, water, fertilise or harvest. Various types of sensors are used in agriculture.

Some of them provide one-dimensional (1D) signals, such as weather sensors to monitor temperature, humidity, rainfall, wind speed, soil properties like salinity, acidity, moisture content, nutrient concentration, spectra information, etcetera, and others that capture n-dimensional (nD) data such as those that capture images for remote or proximal sensing (colour, multi or hyperspectral, thermal, time of flight, stereo cameras).

The data these sensors capture must be analysed efficiently to obtain information helpful for farmers. The field is an open environment that is constantly evolving, where meteorological events can rapidly alter the data acquisition conditions. The characteristics of these data depend on multiple factors such as the crop, variety, management, soil, or weather, among others. In addition, these data are sometimes derived from heterogeneous sources, which must be assembled, processed, and converted into useful information for the farmer. Therefore, data analysis techniques must effectively combine and process large amounts of data to quickly adapt to changing conditions and make predictions capable of improving operational decisions. The ANNs are presented as valuable technologies for analysing the problems associated with the field due to their flexibility and adaptability. However, the data structure and the result to be obtained must be considered when using one or another type of ANN.

The election of a particular type or family must be based on the intrinsic structure of the data. Each ANN type uses a different input data. Most kinds of ANN can be fed with various data structures and provide good results, but, in our opinion, some algorithms perform better with 1D data due to their internal structure. In contrast, others are specially developed for n-dimensional (nD) data. Hence, only choosing the most appropriate algorithm with the optimal hyperparameters can lead to the best result. For example, in problems in which the information is encoded in a 1D array, the appropriate ANNs could be based on MLP, DBN or RBFN since they have a structure in which, generally, the input neurons of the network correspond to a predictor variable or attribute, sampled at a particular time.

On the other hand, for nD problems (i.e. multichannel images), the most appropriate ANN architectures could be based on CNN, CN or GAN since they have multidimensional input, processing or output layers in their structure. Examples of such multidimensional structural units are flattening input layers aimed at encoding an image as a vector or convolution unit that extracts features from a region of an image. Due to their typology, other architectures, such as AEs or RNNs, are suitable for different data types. However, RNNs have recurring layers that make them appropriate for handling time series of data.

Examples of ANNs that process signals or 1D data are widely found in the literature. Spectrometric data was used together with a DBN by Lu et al. (2018) to determine the rice germination, Xiong et al. (2021) used an RBFN to analyse potassium content in fresh lettuces using an RBFN, while Pourdarbani et al. (2020) estimated apple ripening training an MLP. RNNs are particularly appropriate for processing temporal data, such as climate data. Ghose et al.(2018) forecast groundwater levels, and Le et al. (2017) drought projections using this type of recursive ANN.

Images are one of the primary data types in agriculture. They can have two dimensions if they are grayscale images or three dimensions data if they are colour, multispectral or hyperspectral images. Multispectral images are composed of several electromagnetic spectrum bands. In contrast, Colour images are a specific type of multispectral images which consist of 3 channels corresponding to the red, green and blue (RGB) electromagnetic spectrum bands. Hyperspectral images comprise several, even hundreds, spectral bands in the electromagnetic spectrum. Some examples are using RGB images and a CNN used by Dong et al. (2021) for strawberry disease identification. Also, RGB images and depth information were used by Yang et al. (2020) for citrus detection and by Zhang et al. (2018) for mango detection, in both cases using a CNN. Hyperspectral images comprise monochromatic images captured in consecutive bands, providing spectral and spatial information. Examples of the use of hyperspectral images processed through ANNs are given by Park et al. (2023) for blueberry firmness estimation through a CNN, by Fan et al. (2019) for rice infestation detection by applying an AE, by Agarwal et al. (2020) for detecting Trigoderma Granarium in grains using CNs, or by Wang et al., (2019) for tomato spotted Wilt virus detection applying GANs. Besides, AEs and GANS, due to their structure, are appropriate for generating new images and data augmentation. For instance, Abbas et al. (2021) employed GANs to create new samples as a preprocessing step before training a CNN to classify tomato leaf diseases. Furthermore, AEs have been used for feature extraction in different domains. Xin et al. (2020) utilised AEs to extract meaningful features for predicting cadmium residues in lettuce leaves. Similarly, Huang et al. (2021) harnessed AEs for feature extraction to predict apple-soluble solid content.

3. Data preprocessing

ANNs can extract (in the hidden layers of the net) important features by learning from the input data. This allows, especially in the case of CNNs, to work with raw data or images as input without preprocessing or a feature engineering stage. Nevertheless, this ability has a drawback: the number of samples to train these algorithms is very high, so the training time is also very long compared to other machine learning methods. Thus, pre-processing steps facilitate the training effort and performance of any machine learning method (Chollet, 2021), which is crucial for other traditional Machine Learning methods to obtain reliable results (Rinnan et al., 2009).

Various preprocessing techniques can be utilised depending on the nature of the data and the type of correction required. Several preprocessing steps can be applied with 1D data, such as those captured by sensors providing signals in agriculture. When analysing data collected from various sources, machine learning relies on data normalisation to account for the differing units of measurement. Data normalisation aims to standardise the values in numerical columns to a standard scale, thus enhancing the stability and effectiveness of machine learning models. This is important because data measured in different ranges can disproportionately impact the model's results (Aurélien Géron, 2019; Müller and Guido, 2016). In some data coming from sensor signals, such as spectral information, scatter correction is required since the interaction of the light with the tissue disperses the light in many directions, causing undesired variations in spectra (Dawson and Acton, 2018). In these cases, multiplicative scatter correction (MSC) and standard average variance (SNV) are the common preprocessing techniques. In MSC, each spectrum is regressed by least squares against the mean spectrum, whereas, in SNV, the spectra are mean-centred and divided by their standard deviation. The signal captured by sensors can be affected by several factors, such as electronic noise, mechanical vibrations, sensor ageing or power supply variations. These factors cause high-frequency peaks in the signals that are convenient to smooth, improve the signal-to-noise ratio, and make it easier to identify underlying patterns. Savitzky-Golay (SG) is one of the primary methods for smoothing the signal where successive sub-sets of adjacent data points are fitted with a low-degree polynomial using the linear least squares. Spectral derivatives are widely used to remove both additive and multiplicative effects in signals like spectra, where they are commonly used for separating peaks of overlapping bands (Fearn et al., 2009).

Preprocessing nD data requires techniques different from those used for signals. In the case of multi (including colour) or hyperspectral images, which are probably the most common nD data captured by agricultural sensors, spatial noise due to variations in the sensor or the illumination, geometric corrections, or dimensionality reduction are among the most required preprocessing.

The correction of hyperspectral images is typically done to reduce the influence of the particular sensitivity of the sensor or the illumination. The most used correction involves obtaining a relative reflectance through a dark reference (to remove the impact of electronic noise due to dark currents in the sensor) and a calibrated white reference of a known reflectance. This results in improved image quality, spectrally and spatially corrected, allowing accurate data analysis (Grahn and Geladi, 2007).

Among preprocessing steps, dimensionality reduction is aimed at removing redundant features and noisy and irrelevant data to improve learning feature accuracy and reduce the training time (Velliangiri et al., 2019). Improving the learning process and reducing the training computational requirements is critical since this step is usually complex, requires powerful computing units and requires a long time to build the models. Moreover, working with high-dimensional data can lead to overfitting and decrease model interpretability. Reducing dimensionality while retaining the relevant information can reveal hidden patterns, relationships, and underlying structures in the data, which helps to understand the intrinsic characteristics of data better. These techniques can be divided into feature selection and feature extraction (Velliangiri et al., 2019). The choice of dimensionality reduction technique must depend on the data's specific characteristics, the analysis's goals, and the computational resources available. Velasquez et al. (2024) use six dimensionality reduction methods to obtain a subset of wavelengths from hyperspectral images to detect anthracnose early in mango.

Feature selection methods aim to choose a subset of the most influential features from the original dataset concerning the target variable. It helps improve model performance, removing irrelevant features and providing interpretability by simplifying models. These methods are the election in problems aimed at obtaining a reduced number of essential variables, for instance, to select multispectral data from hyperspectral data to build faster models or to construct a camera using only the key wavelengths. Feature selection approaches can be split into filter, wrapper and embedded methods (Jović et al., 2015). Filter methods (Sánchez-Maroño et al., 2007) are used to evaluate features independently of the machine learning model. These methods use different approaches to assess the importance of each feature for prediction. Some standard techniques include correlation-based methods that select features with high correlation with the target variable; mutual information that quantifies the amount of information shared between a feature and the target variable, choosing those with high mutual information; and statistical tests, like chi-squared or ANOVA, that assess the statistical significance of a feature concerning the target variable. Wrapper methods evaluate the performance of a subset of features iteratively using a specific machine learning algorithm. Each iteration adds or removes a variable and compares the performance with the previous iteration, determining if the variable positively or negatively influences the model. Examples include forward selection, backward elimination, and recursive feature elimination (RFE). Forward Selection starts with no features and adds them individually, selecting the one that most improves model performance. Backward Elimination begins with all features and removes them one by one, excluding the one that has the most negligible impact on model performance. Finally, recursive feature elimination (RFE) iteratively trains the model and removes the least essential features until the desired performance level is achieved (El Aboudi and Benhlima, 2016). Generally, wrappers are computationally more demanding than filter methods. Embedded methods embed feature selection within the model-building process, considering the importance of the feature the model is trained (Lal et al., 2006). A popular embedded algorithm is LASSO, which uses a penalty to shrink many regression coefficients to zero, selecting the features with non-zero regression coefficients (R & Rohini, 2016).

Feature extraction methods transform the original data into a lower-dimensional representation. These methods are beneficial for reducing dimensionality and finding underlying patterns or relationships within the data by creating new features containing the most crucial information. Commonly used feature extraction methods are Partial Least Squares (PLS), Linear Discriminant Analysis (LDA), Principal Components Analysis (PCA), autoencoders, and manifold learning (Dai et al., 2020). PLS and LDA are used for supervised learning, whereas PCA, autoencoders, and manifold learning are employed as unsupervised feature extraction methods (Guyon et al., 2006). PLS is used for regression (PLS-R) and classification tasks (PLS-DA) and finds latent variables (linear combinations of the original features) that maximise the covariance between the features and the target variable (Cohen et al., 2013). LDA is a supervised learning algorithm used for classification. It finds linear combinations of features that maximise between-class variance and minimise within-class variance to improve class separation (Haenlein and Kaplan, 2004). PCA identifies principal components, linear combinations of the original features, that maximise the variance. It is an unsupervised technique widely used for data compression and visualisation that reduces data dimensionality while preserving as much variance as possible (Garcia et al., 2022). Autoencoders are ANNs that can be used for non-linear dimensionality reduction by learning compact data representations (Bank et al., 2021). Manifolds help visualise and understand complex data patterns by revealing the lower-dimensional "manifold" where data points reside. It is used for dimensionality reduction, mainly when dealing with non-linear relationships, and is also commonly used for data visualisation. Some manifold learning techniques are Isomap, which builds a graph representing pairwise distances between data points (Balasubramanian and Schwartz, 2002), and Locally Linear Embedding, which preserves the local linear relationships between data points (Roweis and Saul, 2000).

Finally, generating new synthetic samples can be considered a preprocessing step, which is required in cases of limited datasets or when it is not easy to obtain original samples. Sometimes, obtaining actual samples in the field or when complex and expensive laboratory analyses are required is challenging. In many machine learning problems, especially those dealing with images, a large amount of data is needed to train effective models. In practice, obtaining sufficient labelled data is often challenging, particularly in agricultural processes, where collecting field data or conducting the costly laboratory analyses needed for reference labelling remains difficult (Elgendy et al., 2020). Traditionally, chemometric-based methods have been successfully applied in these cases with relatively few samples. Data augmentation techniques allow the size of the training data set to be artificially increased, which can improve model performance by having more samples and introducing variability in the training data. This can help prevent overfitting to the model since the model should not overly depend on specific characteristics in the original data samples. This process can be done in two main ways: generating new data samples from existing ones by applying varied and controlled transformations or using generative ANNs to create new synthetic data from the original ones. Augmentation techniques apply one or several transformations to the original data, such as image rotation, translation, or clipping (Blazhko et al., 2021). In contrast, generative techniques such as Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN) (Lalitha and Latha, 2022) can create new samples based on information provided by the original data set. VAE starts with an encoder layer that transforms the input into a latent feature space. Subsequently, a decoder processes each point in the latent space, which generates a reconstruction of the original data from this point. The VAEs are trained to minimise a global loss function, which has two terms: one that measures the similarity between the original input and the reconstruction and another known as Kullback-Leibler divergence, which measures the difference between the latent distribution generated by the encoder and a standard Gaussian distribution (Huang et al., 2021). GANs start with random samples, generate new data that resembles the actual data, evaluate their similarity, and learn to distinguish them. This process is repeated iteratively, seeking to generate data that is increasingly similar to the actual data until they reach a threshold of similarity (Abbas et al., 2021).

4. Artificial neural networks

Neural networks consist of layers that contain interconnected neurons. A layer of an ANN is a set of neurons whose inputs come from a previous layer (or from the input data in the case of the first layer) and whose outputs are the input of a later layer. ANNs are formed of at least three layers, as shown in Fig. 1: an input layer which receives the input data, one or more hidden layers that serve as the network's learning and abstraction mechanism, extracting meaningful features, and an output layer that makes the prediction. Depending on the complexity and data structure of the problem, the user determines the number of layers, the number of neurons in each layer, and the type of layers based on the data and the expected outcome. ANNs with more layers result in higher-level and more complex features.

There are three main types of layers: fully connected, recurrent, and convolutional. Fully connected layers feature dense structures where each neuron connects to every neuron in the next layer. They are used to learn complex patterns from the data. Recurrent layers of neurons are designed for sequential or time series data, incorporating information from past data to influence the current input and output. Convolutional layers are specialised for image processing, capturing spatial hierarchies by analysing information within small, overlapping regions (Elgendy et al., 2020).

ANNs use forward propagation to predict data, as shown in Fig. 1. This means that data entering the input layer is processed through the entire network, passing from one layer and neuron to the next until reaching the output layer, where the prediction is made. The input of each neuron (node) is weighed, and then a bias value may be added to compute the output of a neuron, allowing the network to adapt and learn patterns. The value of the following neuron is obtained by summing the output of neurons connected to it. This value is then passed to a non-linear activation function (Chollet, 2021). Activation functions determine whether the neuron should be activated based on the input, thus introducing non-linearity into the neural network. This nonlinearity is essential for the network to learn and capture complex patterns within the data effectively. Several activation functions are commonly used, including the sigmoid function, hyperbolic tangent (Tanh), rectified linear activation (ReLU), and leaky ReLU function. The sigmoid function produces outputs between 0 and 1, while Tanh produces outputs between −1 and 1. ReLU is the most used and outputs a positive value for positive inputs and zero for non-positive values. Like ReLU, Leaky ReLU allows for negative values, providing additional flexibility (Arul, 2021).

During the training process, the weights of the nodes in the network are adjusted through an iterative process called backpropagation until the error between the expected and obtained values is minimal. The error is measured through a loss or objective function. There are several types of loss functions, which are chosen based on the type of data being predicted. For instance, Mean Squared Error (MSE) is commonly used for regression problems, Binary Crossentropy is used for two-class classification, and Categorical Crossentropy is used for multiclass classification. Gradient descent is used to optimise this process, where the loss function guides the optimisation process by quantifying the model's performance. During each iteration, the gradient/slope of the loss is calculated. The aim is to find the point where the slope of the loss function is 0, the bottom of the function, and where the error between the desired value and the obtained value is 0. The slope of each point in the loss function is calculated with the chain rule (the cost function derivative). If the slope is positive, it gets multiplied by the learning rate, leading to an update in the weight, as shown in Equation (1) where w is the weight of the neuron, w' is the new weight of the neuron, λ is the learning rate (defined by the user), and ∇L(w) is the gradient of the loss respect to the weights. The user defines the learning rate and the step size taken during optimisation. It determines how much the model's parameters should be adjusted based on the calculated gradient. A higher learning rate means larger steps but could overshoot the optimal values, while a lower learning rate takes smaller steps but might slow down convergence. Finding the right balance is crucial for effective model training (Elgendy et al., 2020).

$$w' = w - \lambda^* \nabla L(w) \tag{1}$$

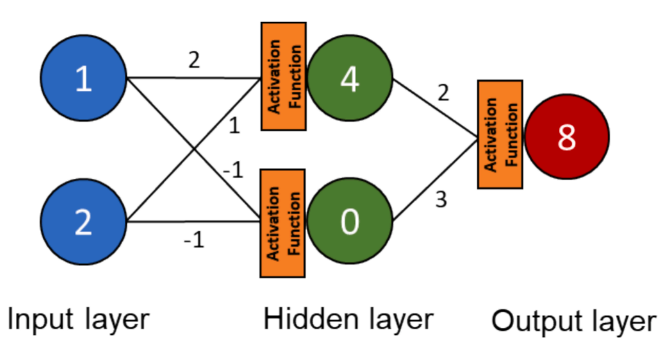

Fig. 1. Forward propagation. Example using the ReLU activation function.

ANNs need to be trained. Training is the process of teaching a neural network to perform a task. In principle, neural networks learn by processing several large sets of labelled or unlabeled data. The training strategy depends on publicly labelled data availability, the task's uniqueness, and the task's similarity. There are three ways to train ANNs: from scratch, transfer learning, and feature extraction. Training a neural network from scratch implies starting the training process without prior information. This approach can be advantageous when working on a task without a pre-existing model. It is necessary for scenarios where the data is distinctive or the problem is new. Using this method, the model can learn task-specific patterns customised to the specific dataset. Transfer learning is a technique that uses a pre-trained model on a related task and applies this knowledge to a similar task. It proves helpful when there is a small number of labelled data for the target task. Transfer learning saves time and resources by leveraging the information gained from a related task. It is effective when the source and target tasks share standard features. Finally, feature extraction uses a pre-trained model to extract essential features from data. It is convenient when the focus is on the high-level representation of data rather than the detailed information. This method allows for using a pre-trained model's ability to capture complex features of the problem without retraining the whole model (Chollet, 2021).

4.1. *Multi-Layer Perceptron (MLP)*

An MLP is one of the most basic forms of ANN. It comprises three or more layers, including an input layer, at least one fully connected hidden layer and an output layer. Each layer has different nodes or neurons; every node is connected to another node in the next layer (Zhu, 2021). MLPs are commonly used in agrifood applications due to their simplicity yet effectiveness. For instance, they have been implemented to predict kiwifruit firmness (Torkashvand et al., 2017), sugarcane yield prediction (P. Wang et al., 2021) or in nutritional status assessment of olive crops (Noguera et al., 2021). Velasquez et al. (2024) used multiple methods, including MLP, to classify asymptomatic mangoes as either healthy or diseased (infected by anthracnose). Among the techniques evaluated, MLP demonstrated superior performance in accurately distinguishing between the two categories.

4.2. *Deep Belief networks (DBNs)*

Geoffrey Hinton proposed DBNs in 2006, which are commonly used for image generation and recognition. They combine the power of deep architectures and probabilistic graphical models in tasks with limited labelled data because they excel in unsupervised learning, automatically capturing complex patterns in data through probabilistic modelling. In

DBNs, hidden layers consist of stacks of Restricted Boltzmann Machines (RBMs) that transform input data into a latent space. RBMs have fully connected connections between their layers, but nodes within each RBM are conditionally independent, as shown in Fig. 2. This allows RBMs to learn complex patterns in the data hierarchy efficiently (Hinton et al., 2006).

RBMs are suitable for tasks like image and speech recognition, image generation and feature extraction due to their hierarchical feature learning and generative capabilities (Hinton et al., 2006). DBNs are not commonly used in the agriculture and food industries. However, some studies have applied them to eliminate variable correlation before using SVM to predict the transpiration rate of strawberry leaves (Shuaishuai et al., 2018), for feature extraction for pesticide residue measurement (Wu et al., 2019) or for detecting rice germination rates (Lu et al., 2018). Besides, some studies have shown that DBNs can produce better results than traditional artificial neural networks (ANNs), such as MLPs, in modelling the spatiotemporal distribution of soil moisture (Song et al., 2016).

4.3. Recurrent neural networks (RNNs)

Recurrent Neural Networks (RNNs) are specifically designed for tasks that involve sequences and time dependencies. They can retain past information, adjust to sequences of varying lengths, and effectively capture contextual information. This makes them ideal for natural language processing and predicting time series applications. On the other hand, other ANNs like MLPs are structured to approximate functions and face difficulties when modelling dynamic systems, as well as systems that involve temporal components with sequential data (Hochreiter and Schmidhuber, 1997).

They work by sequentially processing data and maintaining the memory of past information through the hidden layers. RNNs consist of three main layers: the Input layer, which receives data; the hidden layers, which capture information from their input and retain the memory of past layers; and the output layer, which produces the final output based on the data processed by the hidden layers as shown in Fig. 3 (C. Wang et al. 2021).

However, in basic RNNs, repeated multiplication of weights with a value less than one can cause values to shrink exponentially, leading to very small weights over many time steps. This phenomenon is known as

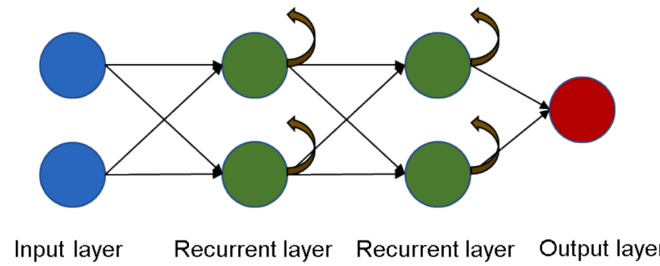

Fig. 3. Recurrent Neural Network.

"vanishing gradients, which pose challenges in training. Therefore, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures were developed to solve that problem and are the most common RNN models. LSTMs are designed to remember important information from past layers, making them great for understanding sequences. They use a gate to decide what to store, update, or forget in their memory. This helps the network capture and retain meaningful patterns over extended sequences (Hochreiter and Schmidhuber, 1997).

GRUs decide what information to remember or forget while processing sequences. They have a simplified system with only two gates: an update gate and a reset gate. This makes GRUs computationally less intensive compared to LSTMs (Jing et al., 2019).

Recurrent Neural Networks (RNNs) are ideal for tasks that involve sequences. RNNs can effectively remember and utilise information from previous steps, making them well-suited for applications like natural language processing and time series prediction. It was challenging to find articles related to agriculture that used RNNs. However, one such study was conducted by Ghose et al. (2018) for groundwater level forecasting, where they used different types of sequential data such as rainfall, temperature, humidity, evapotranspiration and runoff as inputs, and groundwater depth as the output. Another study by Le et al., 2017) projected droughts in California using sequential data such as soil moisture, precipitation, evapotranspiration, moisture recharge, and runoff from 1896 to 2006, which was later validated with the period from 2006 to 2015.

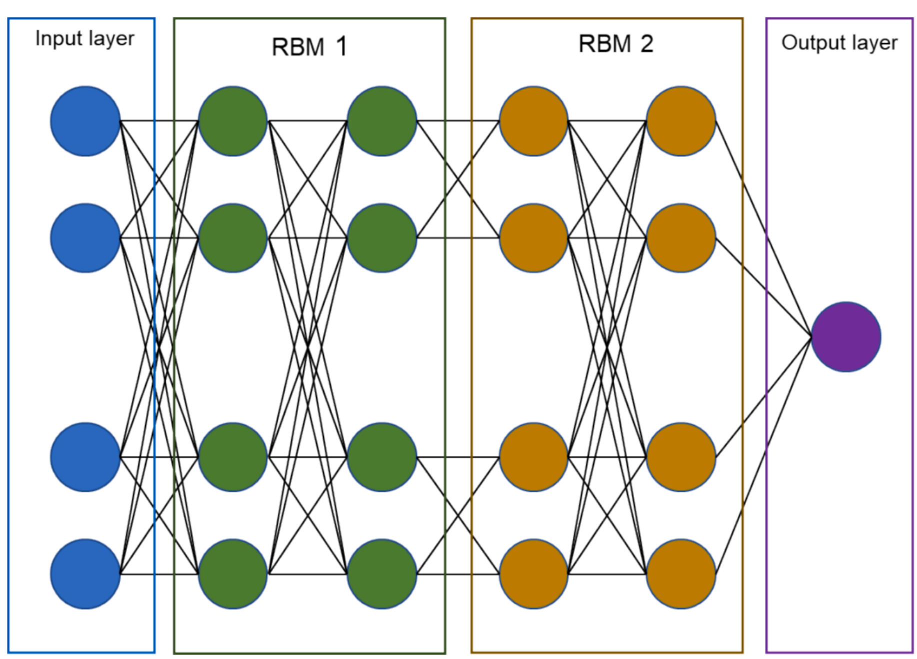

Fig. 2. Deep Belief Network.

4.4. Radial Basis function network (RBFN)

Broomhead and Lowe (1988) introduced Radial Basis Function Networks with the primary objective of curve fitting. RBFNs offer two key advantages over MLPs: firstly, they are connected to a well-established fitting technique while still leveraging the advantages of neural networks. Secondly, RBFNs can yield good results with a straightforward model in practical applications due to their ability to blend linear dependence on variable weights and capture nonlinear relationships. RBFNs have a simple structure (Fig. 4) and a fast training process with tolerance to input noise (Hashemi Fath et al., 2020). A radial basis function uses a Gaussian function (Fernández-Ruiz et al., 2010), transforming the input into another form that feeds the network to get linear separability. This property allows the network to separate data points with a straight line or plane. This is achieved by transforming the input data into a higher-dimensional space where the data points become linearly separable (Nayak et al., 2020).

These ANNs consist of three layers: an input layer, a hidden layer known as the RBF layer, and an output layer consisting of neurons that perform linear combinations. Each RBF neuron retains a prototype (an instance of the training set). Each neuron computes the Euclidean distance between the input and its respective prototype during classification. The hidden layer relies on a radial basis function as the activation function for approximation. They are used in function approximation, classification and prediction. In a study focused on detecting powdery mildew stages in squash, an RBFN was employed for classification. The study aimed to develop a system that accurately identifies the different stages of powdery mildew in squash plants, a common fungal disease affecting squash production. The RBFN achieved accuracy exceeding 92 % across seven classes and distinguishing between diseased and healthy samples (Abdulridha et al., 2020).

Additionally, RBFNs have been used as an alternative to MLPs for regression tasks, such as predicting lycopene content in tomato juice (Fernández-Ruiz et al., 2010). Another application involving RBFNs for regression was to estimate potassium levels in fresh lettuces, demonstrating their ability to outperform widely used chemometric models like PLS regression (Xiong et al., 2021). The superiority of RBFNs over PLS was further evidenced in a study on rice germination rate by Lu et al. (W. Lu et al., 2018).

4.5. Autoencoders (AEs)

Autoencoders (AEs) are a type of neural network for unsupervised

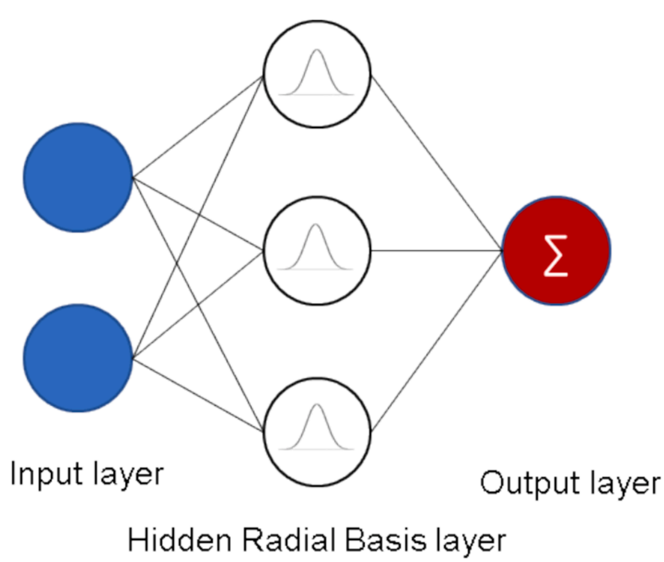

Fig. 4. Radial Basis Function Network.

learning. They can work with data of different dimensions, ranging from 1-D to n-D, but are mainly used for image data. Autoencoders have two parts: an encoder and a decoder, as shown in Fig. 5. The encoder compresses input data by extracting features and creating a compact representation called a bottleneck. The decoder then reconstructs the original data from this bottleneck. Depending on the input data, fully connected layers with 1-D data or convolutional layers in the case of images are used (Bank et al., 2021).

There are three main types of autoencoders: Sparse Autoencoders (SpAEs), Stacked Autoencoders (StAEs) and Variational Autoencoders (VAEs). SpAEs learn compact representations of data while encouraging sparsity in the learned features. Sparsity means that only a few neurons are active at a time in the hidden layer, which can lead to more efficient and interpretable representations as the network focuses on capturing the most important features of the input data. StAEs involve combining multiple autoencoders, one on top of the other, so the output of one autoencoder becomes the input for the next. Each autoencoder in the stack learns to capture different data abstraction levels, allowing the network to capture complex patterns. In VAEs, the bottleneck vector is replaced by two vectors: one represents the standard deviation and the other the mean of the distribution. Apart from learning to represent data, they also learn the uncertainty in their representations, letting them generate new, diverse samples (Bank et al., 2023).

Autoencoders can be used for various tasks, such as data compression, image denoising, generating new images, or extracting features that can be utilised for classification, regression, or prediction. Many studies have employed autoencoders due to their capability of feature extraction, which can be later used for classification, such as for predicting soluble solids in apples (Huang et al., 2021), where a StAE was used for feature extraction before soluble solid prediction using an MLP. The same approach was used to extract the most essential spectral features to predict the nitrogen concentration in oilseed rape leaf (Yu et al., 2018). Other authors have AEs in combination with traditional machine learning methods, such as PLS or SVMs, to extract features. For instance, AEs were used to predict cadmium residues in lettuce leaves. (Xin et al., 2020).

4.6. Convolutional neural networks (CNNs)

CNNs are a specific type of ANN developed to overcome the limitations of traditional neural networks, such as MLPs, in processing and recognising visual data. The idea behind CNNs is to mimic the visual processing in the human brain and take advantage of the spatial hierarchies present in visual data. CNNs excel at capturing local patterns and hierarchical features in images, making them well-suited for tasks such as image classification, object detection, and image segmentation. However, CNNs require a large amount of data for training (O'Shea & Nash, 2015). A CNN consists of several layers, including an input layer that receives the input data, convolutional and max-pooling layers that extract and learn features from the input, fully connected layers, and an output layer where the final prediction is made. This is illustrated in Fig. 6. The concept started with the Neocognitron model developed by Fukushima (1980), who laid the foundation for capturing hierarchical features in visual data using a form of convolutional layers.

The convolutional layer is responsible for extracting features from images and consists of a series of filters or kernels that can be trained. Each filter is applied to the raw pixel values, resulting in a product between the filter and the input pixel, as shown in Fig. 7. This process produces an activation map known as a feature map. Each layer captures different features from the input images, starting from basic ones like lines and progressively becoming more complex and specific in the final layers (Elgendy et al., 2020). CNN models can employ 1D, 2D or 3D convolution filters, where 2D filters are the most used, especially in grayscale and RGB images, which are reshaped into two dimensions before adding them to a 2D CNN. In contrast, 1D is used for 1D signals and 3D filters are usually used for 3D structures such as spectral images

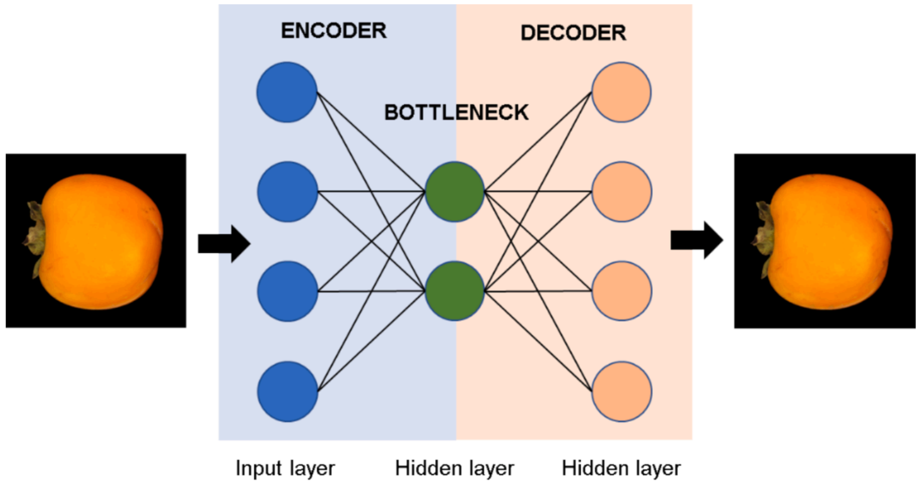

Fig. 5. Autoencoder.

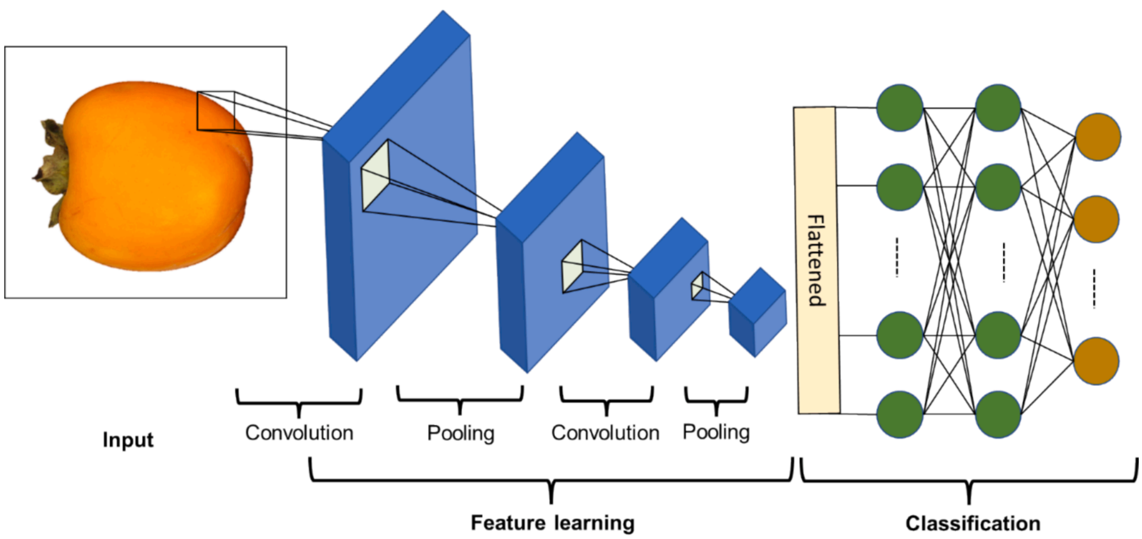

Fig. 6. Convolutional Neural Network.

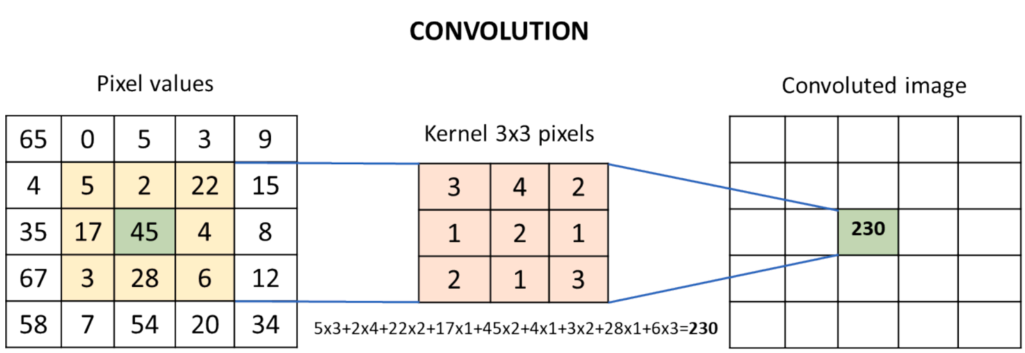

Fig. 7. Example of the convolution process in a CNN.

(Liu et al., 2021).

Pooling layers serve as sub-sampling layers within CNNs. They are crucial in reducing convolutional maps' size while ensuring the model's invariance to rotations and translations. A nonlinear activation function is applied, resulting in the pooling layer's output being determined by selecting the maximum activation value within each feature map of the input layer. Fig. 8 illustrates this process with the ReLU activation function. By preserving the spatial relationships between features, convolution and pooling layers enable the network to learn local patterns effectively (Chollet, 2021). After feature extraction, the fully connected layers use the learned features to classify or predict. However, they require a 1D input. Therefore, flattening is performed to reshape the output of the last convolutional or pooling layer into a vector (Elgendy et al., 2020).

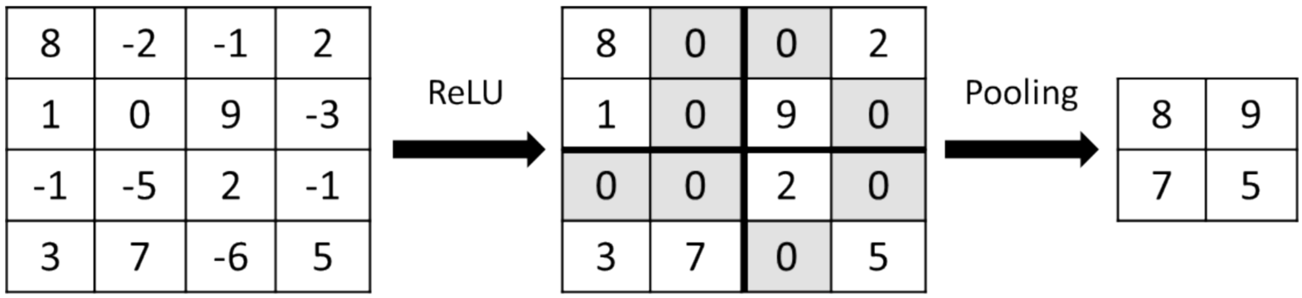

Fig. 8. Example of the pooling process in a CNN.

CNNs have significantly increased in computer vision and image processing applications. Different CNN architectures are employed depending on the objective. Pre-trained architectures like LeNet5, AlexNet, VGGNet, GoogLeNet, and ResNet are commonly used in classification. Region-based CNN, Fast-RCNN, Faster-RCNN, Mask R-CNN, and You Only Look Once (YOLO) are among the most popular for object detection. For image segmentation, SegNet, U-Net, and SegFast are some of the most used architectures (Santosh et al., 2022). The flexibility of CNN to learn and adapt to changing scenarios makes them extensively employed in agriculture, as in grapevine phenotyping (Milella et al., 2019), pest and disease identification as in Banana (Amara et al., 2017) or strawberry (Dong et al., 2021), plant cultivar and species identification (Javanmardi et al., 2021), fruit detection (Gao et al., 2022), or agricultural product quality and safety assessment (Bonora et al., 2022). CNNs proved to be better than other ANNs at processing images for detecting apple defects (Fan et al., 2020), subtle bruises on winter jujube (Feng et al., 2019), grading dragon fruits based on shape, size, weight, colour and diseases (Patil et al., 2021), sorting cherries based on regular and irregular shape (Momeny et al., 2020).

4.7. Generative models

4.7.1. Generative Adversarial networks (GANS)

(Goodfellow et al. (2014) proposed Generative Adversarial Networks (GANs) to address the challenges of approximating intractable probabilistic computations and leveraging the benefits of piecewise linear units in generative models. In GANs, data is formed from random noise using traditional convolutional layers. They are composed of a generator and a discriminator model (Fig. 9). The generator produces synthetic data from noise, and the discriminator distinguishes between real and synthetic data, ensuring that the generated data is close to the real one. Both the generator and discriminator are trained simultaneously and try to surpass each other iteratively.

There are several different GAN implementations. CycleGAN, Conditional GANs, PGGA, or deep convolutional GANs are the most common. The idea of CycleGAN is a type of neural network that can transform images from one domain to another without paired examples for training. It uses two generators and discriminators. For instance, it can turn pictures of cars into pictures of motorbikes, even if it has never seen or learned that transformation before. It learns by trying to transform and then turning it back into a car, making sure it still looks like the original (Wang et al., 2017). Conditional GANs can be used to process input images. In some Conditional GANs, it is possible to provide them with sentences to generate specific outputs. PGGAs are GANs with a progressively growing generator and a discriminator to generate from low-resolution images to high-resolution images. Finally, deep convolutional GANs are a robust variant of GANs that generate high-quality images (Navidan et al., 2021).

When working with agrifood products, obtaining the vast samples required to train ANNs or performing large amounts of expensive, destructive analysis is complex. Therefore, creating synthetic samples mimicking the real ones is a solution. GANs have been used for this purpose, as in Abbas et al. (2021), who generated new samples for data augmentation to train a CNN for tomato leave disease classification. They are also used for semantic segmentation, as shown by Chen et al., (2021), who segmented occluded tree branches using Pix2Pix GAN and

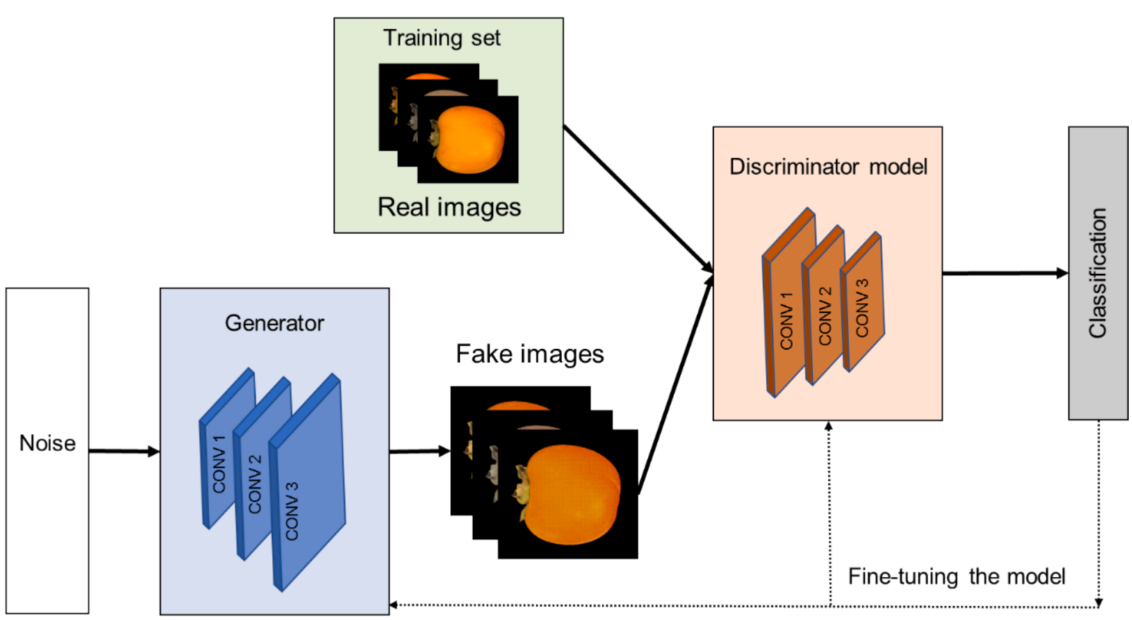

Fig. 9. Generative adversarial Network.

compared the results with other CNNs, obtaining the best results using the GAN. Wang et al., (2019) detected the tomato spotted wilt virus using network analysis based on GANs, integrating the plant segmentation and spectrum classification.

4.7.2. Diffusion models

Diffusion models were introduced in 2015 at the 32nd International Conference on Machine Learning. They are a generative model that models the gradual transformation of a simple probability distribution into the desired target distribution. A simple initial distribution (often Gaussian noise) is transformed step by step into complex data distribution (images) (Sohl-Dickstein et al., 2015). During the training process, the model learns to iteratively transform the initial distribution into samples that resemble the true data distribution. Once trained, it can generate new samples (images) by applying the learned diffusion process to a starting point, such as Gaussian noise. Diffusion models can also perform an inverse process, which means it can go from the complex data distribution back (images) to the simpler distribution. This property allows the model to be used for tasks such as denoising and inpainting. (Ho et al., 2020; Sohl-Dickstein et al., 2015). The application of these networks is relatively new in agriculture, and only one research study has been conducted on them. Chen et al., (2022) used a diffusion model for data augmentation applied to weed recognition enhancement using various pre-trained CNNs (VGG, Inception, DenseNet and ResNet).

4.8. Miscellaneous models

4.8.1. Transformers

Transformers were introduced by Vaswani et al. (2017). Its architecture was initially designed for text translation and natural language processing. However, in agriculture applications, they are mainly used to process images and are called vision transformers.

Vision transformers transform images by dividing them into smaller patches, flattening these patches into 1D vectors, and applying a self-attention mechanism. This mechanism helps the model capture local and global relationships between patches, learning spatial dependencies for practical image understanding. The process involves multiple layers and attention heads, and positional encodings convey information about the spatial arrangement of patches. If applied for image prediction, it is performed with an MLP at the end of the transformer for classification (Dosovitskiy et al., 2021). The basic structure of a vision transformer is shown in Fig. 10.

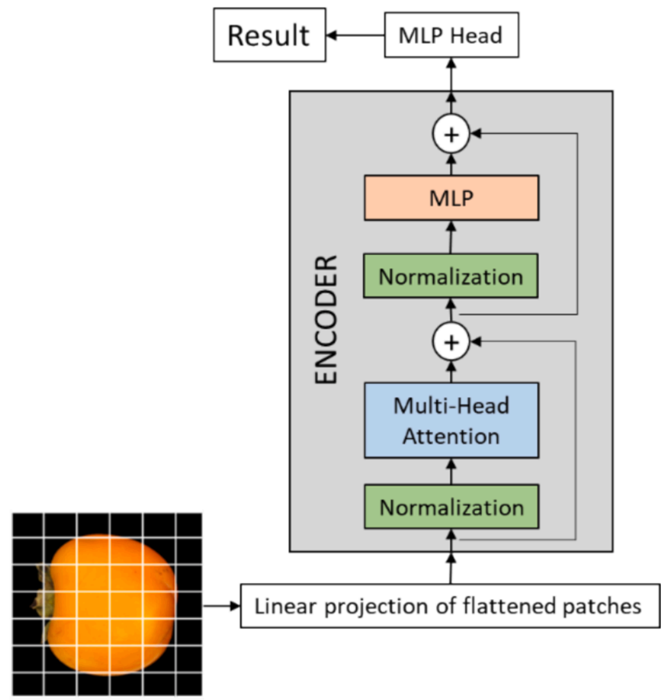

Fig. 10. Vision transformer.

These models are widely used in agriculture for image classification, as in L.-H. Li & Tanone. (2022), who classified several vegetables using a public dataset and obtained a 98 % accuracy. In addition, transformers have been shown to outperform other types of CNNs in image classification, such as in the classification of Cassava leaf diseases (Thai et al., 2021), diagnosing sugarcane leaf diseases (X. Li et al., 2022) and identifying plant diseases (Yu et al., 2023). Furthermore, transformers have been effective in segmentation tasks, such as ripeness estimation in tomatoes, where they outperformed a ResNet CNN (Shinoda et al., 2023). Transformers are handy for large-sized images, as shown in the land cover classification of multispectral LIDAR data (Yu et al., 2022), crop type mapping (K. Li et al., 2022), and satellite images in deforestation monitoring (Kaselimi et al., 2022).

4.8.2. Capsule neural networks (CNs)

Capsule networks were proposed by Geoffrey Hinton in a paper called Learning to Parse Images to solve the loss of location information in neural networks (Hinton et al., 1999). These networks use capsules as the basic unit, which consists of a group of neurons that perform a specific task to combine and generate results. In the context of images, CNNs cannot identify the position of an object as they are only invariant to translation. On the contrary, capsule networks can identify translation and precisely predict the position of an object. Additionally, when reducing the size of convolutional maps, the pooling layers of CNNs lose valuable information required for prediction, while capsule networks do not have this disadvantage. Consequently, capsule networks require less data for training and offer more accurate predictions (mainly used in classification, object recognition, and segmentation) as they know the object's position. (Hinton et al., 1999; Phaye et al., 2018).

CN's structure consists of two main modules: the encoder and the decoder. In the encoder module, there are three essential layers. First is the convolution layer, which processes the input image. Next is the primary caps layer, housing capsules designed to capture fundamental features such as edges and corners. Each capsule within this layer employs convolutional kernels to enhance feature representation. The third layer, called digit caps, is specialised for recognising and encoding information about specific input digits. The decoder module's role is to reconstruct the original input image using fully connected layers that receive the encoder output as input, as shown in Fig. 11. This reconstruction process aims to produce an image that closely resembles the input (Hinton et al., 1999).

CNs are typically used in image classification, object recognition, and segmentation. Some authors have compared capsule networks to other methods, such as CNNs, and have obtained the highest accuracies when using capsule networks, as in identifying carrot appearances (Zhu et al., 2021), in rice recognition (Yu Li et al. (2019)), identifying grassland coverage (Sun et al., 2021) or identifying and diagnosing *Trogoderma granarium* and *Trogoderma variabile* pests in grains (Agarwal et al., 2020). However, the technique is not fully developed. Some experts are sceptical about its advantages, and several changes have been suggested to address its limitations (Kwabena Patrick et al., 2022). Classical capsule networks use shallow convolution, which makes them less effective in handling complex data for classification tasks. To overcome this issue, Dense and Diverse Capsule Networks have been developed (Phaye et al., 2018).

4.8.3. Self-Organising maps (SOMs)

Self-organising maps (SOMs) were introduced by Teuvo Kohonen in 1982 and are designed explicitly for clustering and dimensionality reduction tasks. There are two types of SOMs: supervised and unsupervised (Kohonen, 1982).

SOMs are primarily used as an unsupervised learning technique for

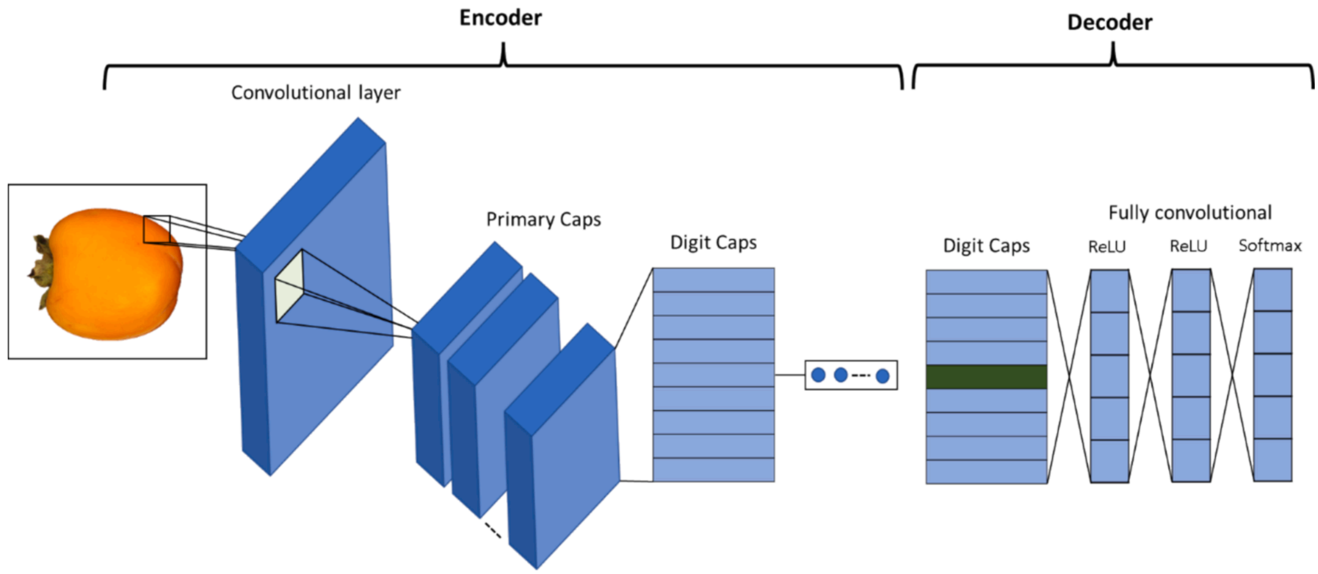

Fig. 11. Capsule Neural Networks.

clustering, visualisation, and dimensionality reduction. They project high-dimensional data onto a lower-dimensional space, typically a 2D grid, while preserving the topological relationships of the data. The architecture of SOMs includes an input layer and an output layer organised as a grid of neurons. Unlike conventional neural network architectures like MLPs or CNNs, SOMs lack hidden layers and utilise competitive learning instead of error-correction-based methods. In contrast, supervised SOMs are tailored to handle labelled data for tasks such as classification or regression. They adjust the training process to incorporate label information, enabling them to perform supervised learning tasks effectively (Miljkovic, 2017).

Several authors have employed SOMs for cluster classification, demonstrating excellent results across various studies. For example, Yufeng Li et al. (2019) applied SOMs to analyse land use, irrigation patterns, and the impact of rice–wheat rotation on pond water quality. Similarly, Theanjumpol et al. (2019) used SOMs to estimate granulation levels in 'Sai Num Pung' tangerine fruit through near-infrared spectroscopy. In this study, SOMs outperformed other methods such as PLS, quadratic discriminant analysis (QDA), K-Nearest Neighbors (KNN), and LDA, achieving a 95 % accuracy in correct classification. Also, using near-infrared spectroscopy, Sringarm et al. (2024) tested LDA, QDA, and SOM for quantitative and qualitative evaluation of maltodextrin products in the industry. However, in this case, better results were obtained using PLS-DA.

5. Applications of ANNs in agriculture

5.1. Pests and disease identification

Early detection of pests and diseases is crucial for maintaining crop health and applying effective treatments. This reduces the need for extensive pest control measures, minimises environmental impact, and leads to significant cost savings for farmers by reducing pesticide use and preventing yield losses. Traditionally, pest and disease detection relied on human inspection, which poses significant challenges, particularly for early detection. Human eyes can miss subtle or early-stage signs, leading to delayed identification and treatment. As a result, the emergence of new techniques which use images and ANNs has significantly increased detection capabilities. Neural networks, especially CNNs, are powerful tools because they can analyse complex patterns and large image datasets. Their role in image analysis makes them ideal for detecting pests and diseases, often present visually on leaves, branches, and other plant parts. While traditional machine learning models can be used for this purpose, ANNs, particularly CNNs, have shown superior performance because they can automatically extract and learn relevant features from images (Alford, 2008; Sawicka and Egbuna, 2020).

One of the main challenges in pest and disease detection is identifying diseases before they become visually apparent or pests when they are still in the early stages. For instance, pests like aphids, thrips, and whiteflies can be challenging to detect when few are visible, and diseases like powdery mildew and downy mildew often start asymptomatically. Detecting these issues early, especially in the field, is critical to preventing their spread and minimising damage (Alford., 2008). Hyperspectral imaging (HSI), which captures and processes information across the electromagnetic spectrum, is particularly useful in disease identification. It allows detecting signs of stress in plants that are not yet visible to the naked eye. For example, Fazari et al. (2021) detected anthracnose in olive fruits. Using HSI increased detection accuracy from 83 % to 87 % on the third day after inoculation, compared to RGB images. Similarly, in detecting Valsa Canker fungus in apple branches, a CNN applied to HSI data achieved a 98 % detection accuracy fifteen days post-inoculation, well before visual symptoms appeared (Cui et al., 2022).

However, the use of ANNs in these applications poses challenges. Creating general-use models is particularly difficult due to the diverse field conditions and crop types. Different living stages of the same pest, as well as the ability of the same pests to affect different plant varieties, pose significant challenges. Similarly, when it comes to diseases, the same disease can impact different varieties in various ways, and the symptoms may change over time. Therefore, it is essential to have large, labelled datasets to train these models. While public datasets like PlantVillage offer valuable resources, they may not always encompass the specific characteristics required for accurate detection in all scenarios. GANs are being explored to create synthetic data for training models, helping to overcome data availability limitations. In one study carried out by Wang et al. (2019), a GAN-based model outperformed a traditional CNN model in detecting Tomato Spotted Wilt Virus using HSI, achieving a higher accuracy of 96 % before symptoms appeared. This demonstrates the potential for advanced neural network techniques to improve early detection capabilities. Another main challenge lies in implementing the models for real-time classification using devices like smartphones. Existing literature primarily focuses on model creation, with only a few authors testing their models for real-time on-site detection, which is the primary purpose of using ANNs for detecting pests and diseases in agriculture due to their ability to make fast and precise decisions. In a study by Barman et al. (2020)SSCNN and MobileNet CNNs were used to classify citrus leaf diseases in real time using smartphone images. SSCNN performed the best, achieving 99 % accuracy during validation while requiring less computation time.

5.2. Object detection and yield prediction

Understanding and predicting agricultural production is necessary for farmers who want to manage their resources efficiently and sustainably. Accurate predictions of crop yields can help farmers make better use of labour, machinery, fertilisers, and pesticides, reducing costs and minimising environmental impact. Additionally, farmers can develop more effective marketing strategies by adjusting prices and distribution based on anticipated production levels and market demand. Recent developments in sensor technologies, including drones and ground-based imaging, are revolutionising the field of agriculture. These technologies are making it possible, together with ANNs, to automate tasks like pruning, monitoring crop health, and evaluating the structural condition of plants.

However, although object detection and yield prediction offer significant potential, several challenges exist. Variability in weather conditions, plant disease outbreaks, and different crop types create difficulties in maintaining consistent accuracy across diverse agricultural settings. Additionally, object detection algorithms must contend with occlusions (where parts of plants or fruits are blocked from view), varying lighting conditions, and differences in plant morphology, all of which can affect detection performance.

CNNs are highly effective at analysing images to identify, segment, and count objects. They play a critical role in modern agriculture by monitoring crop health, estimating yield, and automating processes such as automatic pruning or harvesting. Using different CNN architectures depends on whether the goal is object segmentation, detection, or counting (Chollet, 2021).

Object detection, which identifies the location and presence of objects, relies on architectures like Fast R-CNN, Faster R-CNN, Mask R-CNN, or YOLO (You Only Look Once). Faster R-CNN, for instance, has been widely used to detect apples in RGB images. T. Li et al. (2023) used this model with 800 images across two seasons, achieving an 88 % average precision and a detection speed of 0.32 s per image. Similarly, Mirbod et al. (2023) applied Faster R-CNN for apple detection and Mask R-CNN for segmentation, obtaining a fruit size estimation error of 4.80 %. In the case of YOLO, first shown at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (Redmon et al., 2016), it is an open-source model which surpasses other models in speed, detection precision and generalisation. YOLO models come in various scales for application needs, from small scales for resource-constrained environments to larger versions that provide maximum accuracy and performance. For instance, Hespeler et al. (2021) detected chilli peppers using RGB and thermal imaging, achieving 100 % precision with YOLO v3, outperforming other models in terms of speed with thermal images. However, the dataset was small and lacked real-world complexity. At the same time, MacEachern et al. (2023) found that YOLOv4 small was the best for blueberry yield estimation with an F1 score of 0.82.

Image segmentation, which involves separating different objects or regions in an image, typically uses architectures like U-Net and SegNet. These models divide complex scenes into distinct areas, such as fruit hidden by leaves or branches. U-Net, for instance, has been widely used in agricultural tasks requiring segmentation under challenging conditions, such as occlusion. Z. Chen et al. (2021) compared U-Net with DeepLab for semantic segmentation of apple trees, finding U-Net achieved superior results, with a branch recall of 83.5 %. Similarly, Carneiro et al. (2022) applied U-Net to grapevines, obtaining an intersection over union (IoU) of 91 %, demonstrating its robustness in occluded conditions.

Object counting is crucial in quantifying specific objects like fruits in yield prediction, building on the results from detection and segmentation models. Faster R-CNN and YOLO have been successfully adapted for counting applications. For example, YOLOv5 X was used by C. Gao et al. (2024) to detect and count kiwifruit, achieving 99 % detection accuracy at a fast rate of 21.5 ms per detection, though this model was not tested in real-time conditions.

Despite these advances, occlusion significantly hampers detection and segmentation accuracy. To address this, using depth information from LiDAR and other sensors and images has become more common, as it provides three-dimensional data that enhances detection precision, especially in occluded environments. For example, Ferrer-Ferrer et al. (2023) addressed this by using a Mask R-CNN adapted for RGB + depth data to estimate the size of occluded apples, reducing errors to 5.09 mm for apples with 65 % visibility. Field variability, such as lighting and plant morphology changes, also impacts model performance.

5.3. Species and cultivar discrimination

Species and cultivar classification using satellite or UAV images are crucial for farmers, authorities, and agricultural policies as they enable precise crop identification and monitoring at large scales. For farmers, this technology supports better resource management, early detection of issues, and optimised decision-making. Authorities benefit from efficient crop mapping and regulatory oversight, ensuring compliance with agricultural standards. It also informs policies and helps ensure accurate subsidy allocation by providing detailed data on crop types and land use, promoting more sustainable and equitable farming practices. For example, in European agricultural policies, satellite-based crop classification is vital for verifying the farmer's claims and preventing subsidy fraud. Authorities use this technology to monitor land use, ensuring compliance and fair distribution of subsidies.

CNNs are mainly used to identify species of varieties due to their ability to process and classify high-resolution images. For instance, Asadi & Shamsoddini. (2024) used ten Sentinel bands for crop mapping of 2,547 parcels using images from one season. Various machine learning models, including Random Forest (RF), Support Vector Machine (SVM), and a customised CNN, were used, achieving a 90 % accuracy. RF was also used for feature selection, highlighting bands 6 and 7 (NIR) as the most important for crop mapping. Similarly, Y. Yu et al. (2022) utilised multispectral satellite images combined with LIDAR data from a plane for land cover estimation, applying advanced ANN transformers such as CapVit and CapsNet, along with several other ML models, with the CapVit transformer achieving the best results at 99 % accuracy. This demonstrates the efficacy of transformers in conjunction with diverse data types like multispectral and LIDAR.

Some other examples are crop type mapping using multi-source spatiotemporal data (time-series images, SAR data) by K. Li et al. (2022), who used a CNN, LSTM, and a Proposed transformer, which was the best achieving an F1 score of 0.96. Xiao et al. (2018) aimed to solve the challenges of identifying species from images with complex and varying backgrounds using two CNNs: ResNet50 and Inception V3 with and without Attention Cropping (AC). Inception V3 with AC achieved the best results, with a 69 % accuracy for the PlantCLEF dataset and 95 % for the Oxford flower dataset. Keceli et al. (2022) classified different plant species (tomato, pepper, potato, and maise) and fifteen pests using two public datasets and several CNNs. Their proposed CNN model achieved a 99 % accuracy for species detection and 89 % accuracy for disease detection.

Ground on-field applications allow for continuous crop health and development monitoring, aiding farmers in effective crop management and prompt issue response, such as weed recognition. Many weeds closely resemble crops, particularly in the early growth stages. This resemblance complicates visual detection and can result in classification errors for both manual monitoring and automated systems. Saini and Nagesh. (2024) used eleven different ANN models to recognise weeds in cotton. VGG16 proved the best with 92 to 97 % accuracy using 7,578 images made publicly available under the Cottonweeds dataset. Similarly, Genze et al. (2022) shared their dataset (6,300 patches from 19 UAV images) and their code for crop weed segmentation, thus helping the community. Using a U-Net, an F1 of 0.89 was obtained. Besides, using GANs can help to get more variability, as D. Chen et al. 2024, for weed recognition in cotton fields. A GAN from two public datasets called CottonWeedID15, with 5187 images, was used to create new images, and DeepWeeds, with 17,509 labelled images, and classified using a customised CNN, obtaining a 94 % accuracy using 90 % synthetic images for training, thus ensuring enough variability. Multispectral images are also used for weed detection as Kumar Nagothu et al. (2023 achieved a 95 % accuracy using MobileNet CNN.

As crops progress through their growth cycles, their visual characteristics evolve, presenting a challenge in effectively capturing and interpreting these dynamic changes. Adopting time-series analysis and training models on datasets with various growth stages can enhance the model's ability to generalise effectively. RNNs can be, in some cases, valuable in capturing temporal dependencies and improving accuracy across developmental phases (C. Wang et al. 2021). However, this approach has not been significantly studied. Besides, carrying out field trials and controlled experiments in diverse conditions allows for developing comprehensive datasets with the complete range of environmental variability encountered in real-world situations. This enriches model training and improves its adaptability across different agricultural contexts.

In postharvest applications, accurately identifying fruit and vegetable varieties is crucial for maintaining their authenticity and quality in the market, avoiding fraud and building consumer trust. To achieve this, advanced techniques are needed to prevent misclassification. For instance, Osako et al. (2020) utilised RGB images and a VGG16 CNN model to distinguish four litchi fruit cultivars, achieving 98 % accuracy. Similarly, Jeyaraj et al. (2022) classified nine different varieties of rice using AlexNet CNN, achieving a 98 % accuracy. C. Chen et al. 2024 used RGB images to classify four apricot varieties from five origins harvested from 2018 to 2020. Using a U-Net for segmentation and a VGG16 CNN for classifying, an F1 score of 99 % was obtained. Corn, one of the most important crops in the world, was also studied by other authors. Javanmardi et al. (2021) used RGB images to predict nine corn seed varieties using several machine-learning models: CNN, NN, SVM, KNN, LDA, and CNN-ANN. The CNN-ANN model was the best, achieving a 98.1 % accuracy. Wang and Song. (2023) used hyperspectral images to classify five corn seed varieties with different ML models (SVM, KNN, extreme learning machine, MLP, several CNNs and CNN-LSTM. CNN-LSTM). The CNN-LSTM model achieved a 95 % accuracy. Rong et al. (2020) used spectroscopy and a 1D-CNN to identify five peach varieties, achieving a 94 % accuracy.

5.4. Quality assessment

Quality assessment is crucial in the agricultural industry, especially when measuring and predicting the internal and external quality of fruits and agricultural products. Ensuring that products meet consumer expectations and comply with market regulations is essential. High-quality products appeal to consumers, extend shelf life, reduce waste, and increase profitability. To meet industry standards, the agricultural sector must ensure that products comply with requirements for attributes like size, colour, firmness, sweetness, and the absence of defects such as bruises or rot. Quality standards ensure uniformity in what reaches the market, ultimately protecting consumers and sellers. For instance, Regulation (EU) 2023/2429 sets rules to standardise fruit and vegetable quality within the European Union. This law includes criteria for freshness, physical characteristics, and even permissible imperfections. Besides, internal quality (like sugar content or firmness) is just as important as external quality (appearance) since it directly influences taste, texture, and storability.

ANNs excel in quality prediction by learning complex relationships between external features (colour, size) and internal attributes (sugar content, firmness), providing a complete quality profile. 2D and 3D CNNs are commonly employed for quality assessment using image data due to their effectiveness in image analysis and feature extraction. Similarly, 1D CNNs and MLPs are utilised with spectroscopic or non-image data.

Maturity estimation is essential for harvesting fruit at optimal conditions since accurate estimation helps pick the product at the right time, reducing waste and improving market value, and at postharvest to ensure optimal quality, flavour, and shelf life. Due to the complexity of crop maturation and fruit ripening, the ability of ANNs to handle complex data makes them more suitable than other machine learning models. Some studies estimated maturity in the field, such as Wendel et al. (2018), who estimated the maturity of three different mango cultivars on various dates and climate conditions. A robot platform with HSI cameras was used to predict dry matter. Dry matter is closely related to mango maturity and obtained an R^2 value of 0.64 using a CNN, showing higher results than a PLS model. Alternatively, Gao et al. (2020) estimated the ripeness of strawberries using 244 hyperspectral images from the greenhouse and the laboratory. AlexNet CNN proved to be more accurate than an SVM model, with a 98 % accuracy.

In postharvest, it is important to classify the maturity of the product. This helps determine whether the product needs to be rejected, sold immediately because it is fully ripened, or if it can be stored for some time. In winter jujube, Z. Lu et al. (2021) created a grading robot to classify jujubes into three categories based on their maturity using different CNNs: Modified YOLOv3, SSD, and Faster-R-CNN. The modified YOLOv3 achieved 97 % accuracy using 147 sets of images, each containing 35 jujube fruits.

Besides, estimating quality parameters such as proteins, oil, moisture, or starch content helps farmers and companies ensure that their products meet market and legal standards. Moisture content was predicted by F. Zeng et al. (2022), who used hyperspectral images and low-field magnetic resonance to predict salted sea cucumber moisture on 510 samples to detect fraud. Its proposed CNN obtained an R^2 of 0.99. Liang et al. (2024) used a portable NIR spectrometer, allowing in-site inspection on 122 samples to estimate crude protein in corn feed. He trained with PLS, SVM and a CNN: CropNet model, with which he obtained the most accurate results, an R^2 of 0.89. Other authors, such as Zhang et al. (2022), predicted other components, such as oil content. In this case, oil content in the maise kernel using his and compared different models and model combinations: PLS, SVM, GAN + PLS and GAN + SVM. The most precise results were obtained using a GAN (Nongda616) for data augmentation and an SVM for classification, obtaining an R^2 of 0.80. These results emphasise the usefulness of GANs for data augmentation in model creation.

Alternatively, in fruits, factors such as soluble solids, acidity, and firmness are essential for consumers since they play a key role in ensuring consumer acceptance based on sensory quality. In Fuji apples, Bai et al. (2019) used spectroscopy and predicted soluble solids with an MLP, PLS, and an RF model, obtaining an R^2 of 0.99 with the MLP. 208 samples from three geographical regions were used, allowing the model to eliminate the effects of geographical origin. In Kyoho grape, Xu et al. (2022) used 240 HSI and PLS, SVM, AE + PLS, and AE + SVM models. The AE + SVM model obtained an R^2 of 0.92.

Furthermore, classifying agricultural products based on internal and external defects is essential for ensuring product quality. Defects like blemishes or discolouration can make a product visually unappealing, leading to rejection by consumers or markets. However, defects such as mechanical bruising, internal rotting, or tissue damage are often hidden and only become visible later, potentially affecting the fruit's shelf life and quality. Consumers quickly reject fruits and vegetables with external imperfections, even if the internal quality remains intact. ANNs can identify subtle defects and bruises that may be overlooked by human inspectors or other machine-learning models, leading to more accurate grading and sorting.

Mechanical bruises, for example, are often invisible in the early stages but can cause internal tissue damage, which may eventually lead to rotting. As the internal damage worsens, it can spread to the outer layers, becoming an external defect. When this happens, the product may no longer be suitable for sale, leading to waste. The key to preventing this is early detection. HSI can be essential for this purpose, as demonstrated by Castillo-Girones et al. (2024), in the case of detecting three different levels of invisible bruising in plums. Three different CNN models were used: ResNet50, HSCNN, and a customised CNN. ResNet was the most reliable, with an overall F1 = 90 %. J. Lu et al. (2024) classified the fruit with different defects, including oil spots, mechanical damage, dry scars and black spots. A customised model based on YOLOv5 achieved a 99 % average precision. The model was trained using 330 images of fruits with standard surfaces and surfaces with defects. In addition, different CNNs were used, such as SqueezeNet, ShuffleNet, AlexNet, GoogLeNet, MobileNetV2, ResNet50, and VGG1, to detect and sort surface defects in soybean seeds into six classes. Among these models, MobileNetV2 performed the best, achieving 98 % accuracy and processing 222 seeds per minute, making it the fastest option. Similarly, Cao et al. (2021) automatically graded zizania fruit based on its quality (high and defective quality based on size and defects) using different CNNs: LightNet, AlexNet, VGG11, ResNet18, MobileNet, MobileNetV2 and ShuffleNet V2. LightNet, a light model, was the best and fastest, with 99 % accuracy. The dataset of 4,900 RGB images used in this experiment was publicly released, which can aid other researchers and instil confidence in their work.

Unlike HSI, which only penetrates a few millimetres (Sun., 2010), X-ray images can provide a deeper look inside the fruit, making them more effective for identifying deeper internal issues. For example, in the study of mango seed spoilage locations, Ansah et al. (2023) analysed 98 X-ray images of mangoes using a VGG16 CNN model, achieving an accuracy of 97 %. A Grad-CAM was performed to identify the regions in the images where the model was directing its attention, thus confirming its effectiveness and focus on the seeds. Similarly, in detecting pear disorders, Tempelaere et al. (2023) analysed 80 X-ray images to find internal disorders and segment them using a U-Net, obtaining an IOU of 0.88. In avocado, Matsui et al. (2023) detected internal fruit rot with a 98 % detection accuracy and 3.15 RMSE for segmentation using 286 images.

The quality of agricultural products is highly influenced by crop management. Farmers must assess the nutritional status of plants at the field level to improve crop health, increase yield and product quality, and significantly decrease chemical contamination in the soil. Non-destructive systems, along with ANNs, are becoming a promising solution to the challenge of avoiding time-consuming and costly ionomic analyses. For example, Noguera et al. (2021) in olive leaves, N, P, and K were predicted using multispectral UAV images (5 wavelengths) in two orchards covering 150.7 ha and 6177 images. MLP, PLS, SVM and Gaussian process regression were used, and MLP showed the best results. Similarly, Yu et al. (2018) predicted N content in oilseed rape leaves using hyperspectral images on 192 samples from plants with four different fertilisation levels. A stack AE was used to extract the most important features and an MLP to predict, achieving an R^2 of 0.90. Xiong et al. (2021) used transmission spectroscopy in 211 lettuce leaves from 2 different varieties for K prediction. RBFN and PLS models were trained, and the best results were obtained using RBFN with an R^2 of 0.86.

Application	Data type	Topic	Models used	Results	Reference
Pests and disease identification	RGB images	Real-time classification of citrus leaf disease	CNN: SSCNN and MobileNet	SSCNN the best	(Barman et al., 2020)
	HSI images	Apple valsa canker detection	CNN: proposed model	98 % accuracy	(Cui et al., 2022)
	HSI images	Tomato Spotted Wilt Virus detection	New proposed GAN, CNN	Proposed GAN 98 % accuracy and 96 % before symptoms appear	(Wang et al., 2019)
	HSI images	Anthracnose detection in olives	CNN: ResNet101, VGG-16, ResNet50 and Densenet121	92 % accuracy with ResNet101	(Fazari et al., 2021)
	RGB images	Plant species and pest detection	CNN: diverse CNN models and proposed model	99 % accuracy for species detection, 89 % accuracy for disease detection	(Keceli et al., 2022)
Object detection and yield prediction	RGB images	Kiwifruit detection	CNN: YOLO5	99 % kiwifruit detection	(Gao et al., 2024)
	RGB images	Apple fruit detection	CNN: Faster R-CNN	88 % average precision	(Li et al., 2023)
	RGB + Depth images	Fruit detection and size estimation	CNN: Mask R-CNN	Apple detection F1 = 88 %. Mean Absolute Error (MAE) diameter estimation of 5.6 mm	(Ferrer-Ferrer et al., 2023)
	RGB images	Grapevine segmentation	CNN: U-Net	Intersection Over Union of 91 %	(Carneiro et al., 2022)
	Thermal imaging	Chili peppers detection	CNN: YOLOv3 and Mask-RCNN	YOLOv3 has better detection and ten times faster	(Hespeler et al., 2021)
	RGB images	Yield and maturity estimation of wild blueberry	CNN: six networks used	YoloV4 F1 = 0.82	(MacEachern et al., 2023)
	RGB images	On-tree apple fruit size estimation	CNN: Faster R-CNN and Mask R-CNN	R^2 = 0.67	(Mirbod et al., 2023)
	RGB images	Semantic segmentation of apple trees	GAN, CNN (U-Net and DeepLabv3)	U-Net CNN the best	(Chen et al., 2021)
Species and cultivar discrimination	RGB images	Apricot variety classification	CNN: U-Net	F1 = 99 %	(C. Chen et al., 2024)
	Multispectral images	Crop mapping	RF, SVM, CNN: customised model	88 % accuracy	(Asadi and Shamsoddini, 2024)
	RGB images	Weed detection	CNN: 11 models	VGG16 97 % accuracy	(Saini and Nagesh, 2024)
	RGB images	Weed recognition	Diffusion model + CNN	94 % accuracy using 10 % of real images	(D. Chen et al., 2024)
	RGB images	Crop weed segmentation	CNN: U-Net	F1 = 0.89	(Genze et al., 2022)
	RGB images	Corn seed variety classification	CNN, NN, SVM, KNN, LDA, CNN-ANN	CNN-ANN 98.1 % accuracy	(Javanmardi et al., 2021)

(continued on next page)

(*continued*)

Application	Data type	Topic	Models used	Results	Reference
	Multispectral images	Crop type mapping	CNN, LSTM, Transformer, Proposed transformer	Proposed transformer F1 = 0.96	(K. Li et al., 2022)
	RGB images	Crop weed detection	CNN: MobileNet	95 % accuracy	(Kumar Nagothu et al., 2023)
	RGB images	Litchi fruit cultivar discrimination	CNN: VGG-16	98 % accuracy	(Osako et al., 2020)
	Spectroscopy	Peach variety detection	CNN	94 % accuracy	(Rong et al., 2020)
	RGB images	Plant species identification	CNN: ResNet50 and Inception V3 with and without Attention Cropping (AC)	Inception V3 with AC: 69 % accuracy for PlantCLEF and 95 % for the Oxford flower dataset.	(Xiao et al., 2018)
	HSI images	Sweet maise seed variety identification	SVM, KNN, extreme learning machine, MLP, several CNNs and CNN-LSTM	CNN-LSTM 95 % accuracy	(Wang and Song, 2023)
	Multispectral images + LIDAR	Land cover classification	Transformer: CapVit and CapsNet and several other ML models	CapVit 99 % accuracy	(Yu et al., 2022)
	RGB images	Real-time rice variety classification	CNN: AlexNet	98 % accuracy	(Jeyaraj et al., 2022)
Quality assessment	X-Ray	Mango seed spoilage location	CNN: VGG16	97 % accuracy	(Ansah et al., 2023)
	Spectroscopy	Estimation of crude protein in corn feed	PLS, SVM, CNN: CropNet, CropResNet	CropResNet R^2 = 0.89	(Liang et al., 2024)
	Spectroscopy	Apple soluble solid prediction	ANN, PLS, RF	Proposed ANN R^2 = 0.99	(Bai et al., 2019)
	HSI images	Maise kernel oil content prediction	CNN and PLSR	R^2 = 0.92	(Zhang et al., 2022a)
	HSI images	Kyoho grape firmness and pH prediction	PLS, SVM, AE + PLS, AE + SVM	AE + SVM R^2 = 0.92	(M. Xu et al., 2022)
	HSI images	Salted sea cucumber moisture prediction	CNN: proposed CNN	R^2 = 0.99	(Zeng et al., 2022)
	HSI images	Bruise detection in Plums	CNN: ResNet50, HSCNN, customised CNN	ResNet is the most reliable. F1 = 90 %	(Castillo-Girones et al., 2024)
	X-Ray images	Internal disorder segmentation in pear	CNN: U-Net	Intersection Over Union (IOU) of 0.88	(Tempelaere et al., 2023)
	X-Ray	Hass avocado internal fruit rot	CNN: U-Net	98 % Classification accuracy. RMSE segmentation = 3.15	(Matsui et al., 2023)
	HSI images	In-field strawberry ripeness estimation	CNN and SVM	CNN (AlexNet) 98 % accuracy	(Gao et al., 2020)
	HSI images	Mango maturity estimation	CNN and PLS	CNN R^2 = 0.64	(Wendel et al., 2018)
	RGB images	Automated Zizania quality grading	CNN: LightNet, AlexNet, VGG11, ResNet18, MobileNet, MobileNetV2 and ShuffleNet V2	LightNet is the best and fastest, with 99 % accuracy	(Cao et al., 2021)
	RGB images	Design of a winter jujube grading robot	CNN: YOLOv3, SSD and a Faster-R-CNN	YOLOv3 97 % accuracy	(Lu et al., 2021)
	RGB images	Citrus peel fruit defect detection	CNN: proposed model based on YOLOv5, Yolo5s, Yolo7tiny, Yolo8n	99 % average precision	(Lu et al., 2024)
	Multispectral images	Olive crop nutritional status assessment	MLP, PLS, SVM and Gaussian Process Regression (GPR)	MLP better results	(Noguera et al., 2021)
	HSI images	Oilseed rape nitrogen prediction	AE + MLP	R^2 = 0.90	(Yu et al., 2018)
	Spectroscopy	Fresh lettuce potassium prediction	RBFN, PLS	RBFN R^2 = 0.86	(Xiong et al., 2021)
Toxic compounds and residues detection	HSI images	Peanut aflatoxin detection	CNN	99 % accuracy	(Zhu et al., 2022)
	HSI images	Apple pesticide residue detection	KNN, SVM, CNN: AlexNet	99 % accuracy	(Jiang et al., 2019)
	Raman spectroscopy	Rice chemical residue identification	CNN: Squeezenet	0.98 coefficient of determination	(Weng et al., 2023)
	Spectroscopy	Lettuce leaves pesticide residue identification	DBN for data extraction and SVM, PLS-DA, and KNN for classification	DBM-SVM 95 % accuracy	(Wu et al., 2019)

5.5. *Toxic compounds and residues detection*

Food safety is vital for consumers, who increasingly demand products free of pesticides. Additionally, toxic compounds in agricultural products at postharvest are prohibited due to their health risks, which can lead authorities to withdraw such items from the market. Optical sensors and ANNs are needed to manage these intricate tasks in this context. However, in situations such as identifying components with very low concentrations, ANNs face difficulty providing accurate predictions.

Using HSI or spectroscopy and CNNs has enabled the detection of pesticide residues and toxic biological compounds in agricultural products. For example, Jiang et al. (2019) detected four different pesticides in apples (chlorpyrifos, mycelium and mixtures of those two at 100 ppm) using 72 hyperspectral images per pesticide. They predicted using KNN, SVM, and an AlexNet CNN, which obtained a 99 % accuracy. Zhu et al. (2022) detected aflatoxin B1 in peanuts using the spectra of every pixel (197,265 spectra) and a customised CNN obtaining a 99 % accuracy. Wu et al. (2019) employed NIR spectroscopy to detect two pesticides (fenvalerate and triazoline at 0.24 and 0.25 mg/L) in 240 lettuce leaves. DBN was used for data extraction, and SVM, PLS-DA, and KNN were used for classification. The combination of DBM-SVM was the best, with a 95 % accuracy. Besides, Raman spectroscopy is gaining attention for its ability to detect even small concentrations of toxic components and pesticides, as in rice chemical residues, where Weng et al. (2023) identified chlormequat chloride and acephate at 15.5 and 100 mg/L concentrations in rice with an R^2 of 0.98.

6. Conclusions and future trends

ANNs have become key tools in agriculture, addressing critical challenges and enhancing efficiency across various applications. CNNs are the dominant approach for image-based tasks due to their exceptional image analysis, feature extraction, and pattern recognition capabilities. They are particularly effective in detecting pests and diseases, object detection, yield prediction, species classification, and quality assessment. For example, CNNs enable early detection of pests and diseases by identifying subtle visual patterns in RGB or hyperspectral images. Besides, emerging techniques like Capsule Networks offer promising solutions to overcome CNN limitations by incorporating object position identification and requiring less training data.

Additionally, AEs, GANs and Diffusion Models can be very useful for data augmentation, generating synthetic datasets to address the scarcity of labelled data. These techniques enhance model performance in tasks like pest detection and weed recognition. AEs and DBNs further support image-based applications by extracting features from complex datasets.

For non-image data, such as spectroscopic or time-series inputs, neural networks like MLPs, 1D-CNNs, RNNs, and RBFNs are commonly employed. MLPs are versatile for predicting the nutrient content of crops or quality attributes like sugar levels in fruits, while RNNs excel in capturing temporal dependencies for monitoring crop growth cycles. RBFNs effectively model localised relationships for specific predictions such as nutrient levels. Emerging architectures like Transformers demonstrate potential for large-scale crop mapping and species classification using multispectral and spatiotemporal data. Their ability to process diverse data types positions them as powerful tools for precision agriculture. Meanwhile, SOMs aid in clustering and visualising agricultural datasets, offering insights into underlying patterns.

The application of ANNs in agriculture is expected to expand with the integration of emerging technologies such as the Internet of Things, cloud platforms, and 5G networks, together with the cost reduction of hardware capable of running ANNs. These technologies enable real-time data collection from sensors, drones, and satellites, providing comprehensive datasets for ANN models. Hybrid approaches that combine ANNs with other machine learning techniques or advanced algorithms like reinforcement learning could further enhance predictive accuracy and decision-making capabilities. Additionally, advancements in explainable AI likely improve trust in ANN-driven systems by providing interpretable insights into their decisions.

Despite these advancements, significant challenges remain. Field variability, resulting from differences in soil types, weather conditions, or crop varieties, creates complexities that necessitate robust models capable of generalising across diverse environments. The limited availability of labelled datasets is a critical barrier for training models. Furthermore, rural areas often face infrastructure limitations such as insufficient internet bandwidth or computational resources to deploy ANN-based systems effectively. Another challenge is the cost of implementing ANN solutions at scale. Small farmers may lack access to the technology or expertise required to adopt these innovations. Ethical considerations regarding data privacy and ownership need attention as agriculture becomes increasingly digitised.

To address these challenges, data collection strategies should be implemented to create repositories of diverse agricultural datasets while ensuring privacy and security. Investments in rural connectivity infrastructure are also essential to support solutions for real-time ANN applications. In addition, creating affordable hardware solutions designed for small-scale farmers promotes equitable access to these technologies. Implementing training programs that educate farmers on effectively using tools based on ANNsis also important.

CRediT authorship contribution statement

Salvador Castillo-Girones: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Sandra Munera:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Marcelino Martínez-Sober:** Writing – review & editing, Supervision. **José Blasco:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization. **Sergio Cubero:** Writing – review & editing, Validation, Supervision. **Juan Gómez-Sanchis:** Writing – review & editing, Validation, Supervision, Resources, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was partially funded through the projects MICIU AEI PID2023-150192OR-C31, C32 and C-33 with the support of FEDER funds, project AEI TED2021-130117B-C31 and C33, by the project GVA-PROMETEO CIPROM/2021/014. Salvador Castillo thanks INIA for the FPI-INIA grant number PRE2020-094491, partially supported by European Union FSE funds. Sandra Munera expresses her thanks for the postdoctoral contract Juan de la Cierva - Formación (FJC2021-047786-I) co-funded by MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR.

Data availability

No data was used for the research described in the article.

References

Abbas, A., Jain, S., Gour, M., Vankudothu, S., 2021. Tomato plant disease detection using transfer learning with C-GAN synthetic images. Comput Electron Agric 187, 106279. https://doi.org/10.1016/j.compag.2021.106279.

Agarwal, M., Al-Shuwaili, T., Nugaliyadde, A., Wang, P., Wong, K.W., Ren, Y., 2020. Identification and diagnosis of whole body and fragments of Trogoderma granarium and Trogoderma variabile using visible near infrared hyperspectral imaging technique coupled with deep learning. Comput Electron Agric 173, 105438. https://doi.org/10.1016/j.compag.2020.105438.

Alford, D.V., 2008. Pest and Disease Management Handbook. Pest and Disease Management Handbook 1–615. https://doi.org/10.1002/9780470690475.

Amara, J., Bouaziz, B., Algergawy, A., 2017. A Deep Learning-based Approach for Banana Leaf Diseases Classification, in: BTW.

Ansah, F.A., Amo-Boateng, M., Siabi, E.K., Bordoh, P.K., 2023. Location of seed spoilage in mango fruit using X-ray imaging and convolutional neural networks. Sci Afr 20, e01649. doi: 10.1016/j.sciaf.2023.e01649.

Arul, V.H., 2021. Deep learning methods for data classification. In: Binu, D., Rajakumar, B.R. (Eds.), Artificial Intelligence in Data Mining, 5. Academic Press, pp. 87–108. https://doi.org/10.1016/B978-0-12-820601-0.00001-X.

Asadi, B., Shamsoddini, A., 2024. Crop mapping through a hybrid machine learning and deep learning method. Remote Sens Appl 33, 101090. https://doi.org/10.1016/j.rsase.2023.101090.

Bai, Y., Xiong, Y., Huang, J., Zhou, J., Zhang, B., 2019. Accurate prediction of soluble solid content of apples from multiple geographical regions by combining deep learning with spectral fingerprint features. Postharvest Biol Technol 156, 110943. https://doi.org/10.1016/j.postharvbio.2019.110943.

Balasubramanian, M., Schwartz, E., 2002. Schwartz, E.L.: The Isomap Algorithm and Topological Stability. Science 295, 5552. Science 295, 7. doi: 10.1126/science.295.5552.7a.

Bank, D., Koenigstein, N., Giryes, R., 2021. Autoencoders.

Bank, D., Koenigstein, N., Giryes, R., 2023. Autoencoders. Machine Learning for Data Science Handbook 353–374. https://doi.org/10.1007/978-3-031-24628-9_16.

Barman, U., Choudhury, R.D., Sahu, D., Barman, G.G., 2020. Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf

disease. Comput Electron Agric 177, 105661. https://doi.org/10.1016/j.compag.2020.105661.
Bengio, Y., 2009. Learning Deep Architectures for AI. Foundations and Trends®. Machine Learning 2, 1–127. https://doi.org/10.1561/2200000006.
Blazhko, U., Shapaval, V., Kovalev, V., Kohler, A., 2021. Comparison of augmentation and pre-processing for deep learning and chemometric classification of infrared spectra. Chemometrics and Intelligent Laboratory Systems 215, 104367. https://doi.org/10.1016/j.chemolab.2021.104367.
Bonora, A., Bortolotti, G., Bresilla, K., Grappadelli, L.C., Manfrini, L., 2022. Reprint of: A convolutional neural network approach to detecting fruit physiological disorders and maturity in 'Abbé Fétel' pears. Biosyst Eng 223, 224–232. https://doi.org/10.1016/j.biosystemseng.2022.10.005.
Broomhead, D.S., Lowe, D., 1988. Multivariable Functional Interpolation and Adaptive Networks. Complex Syst. 2.
Cao, J., Sun, T., Zhang, W., Zhong, M., Huang, B., Zhou, G., Chai, X., 2021. An automated zizania quality grading method based on deep classification model. Comput Electron Agric 183, 106004. https://doi.org/10.1016/j.compag.2021.106004.
Carneiro, G.A., Magalhães, R., Neto, A., Sousa, J.J., Cunha, A., 2022. Grapevine Segmentation in RGB Images using Deep Learning. Procedia Comput Sci 196, 101–106. https://doi.org/10.1016/j.procs.2021.11.078.
Castillo-Girones, S., Van Belleghem, R., Wouters, N., Munera, S., Blasco, J., Saeys, W., 2024. Detection of subsurface bruises in plums using spectral imaging and deep learning with wavelength selection. Postharvest Biol Technol 207, 112615. https://doi.org/10.1016/j.postharvbio.2023.112615.
Chen, D., Qi, X., Zheng, Y., Lu, Y., Li, Z., 2022. Deep Data Augmentation for Weed Recognition Enhancement: A Diffusion Probabilistic Model and Transfer Learning Based Approach. https://doi.org/10.48550/arxiv.2210.09509.
Chen, D., Qi, X., Zheng, Y., Lu, Y., Huang, Y., Li, Z., 2024b. Synthetic data augmentation by diffusion probabilistic models to enhance weed recognition. Comput Electron Agric 216, 108517. https://doi.org/10.1016/j.compag.2023.108517.
Chen, Z., Ting, D., Newbury, R., Chen, C., 2021. Semantic segmentation for partially occluded apple trees based on deep learning. Comput Electron Agric 181, 105952. https://doi.org/10.1016/j.compag.2020.105952.
Chen, C., Wang, L., Liu, H., Liu, J., Xu, W., Huang, M., Gou, N., Wang, C., Bai, H., Jia, G., Wuyun, T., 2024a. Construction of apricot variety search engine based on deep learning. Hortic Plant J 10, 387–397. https://doi.org/10.1016/j.hpj.2023.02.007.
Chollet, F., 2021. Deep Learning with Python. Second Edition, Deep Learning with Python.
Cohen, J., Cohen, P., West, S.G., Aiken, L.S., 2013. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, Third Edition. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Third Edition, pp. 1–704.
Cui, R., Li, J., Wang, Y., Fang, S., Yu, K., Zhao, Y., 2022. Hyperspectral imaging coupled with Dual-channel convolutional neural network for early detection of apple valsa canker. Comput Electron Agric 202, 107411. https://doi.org/10.1016/j.compag.2022.107411.
Dai, T., Hao, Z., Wang, Z., -, al, Ali, S.K., Sadoon -, S.A., Chen, H., Liu, Y., Pan, X., 2020. Feature Extraction Methods: A Review. J Phys Conf Ser 1591, 012028. doi: 10.1088/1742-6596/1591/1/012028.
Dawson, P.L., Acton, J.C., 2018. Impact of proteins on food color. In: Yada, R.Y. (Ed.), Proteins in Food Processing (second Edition), Woodhead Publishing Series in Food Science, Technology and Nutrition, 22. Woodhead Publishing, pp. 599–638. https://doi.org/10.1016/B978-0-08-100722-8.00023-1.
Dong, C., Zhang, Z., Yue, J., Zhou, L., 2021. Automatic recognition of strawberry diseases and pests using convolutional neural network. Smart Agricultural Technology 1, 100009. https://doi.org/10.1016/j.atech.2021.100009.
Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2021. AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE. doi: 10.48550/arXiv.2010.11929.
El Aboudi, N., Benhlima, L., 2016. Review on wrapper feature selection approaches. In: In: 2016 International Conference on Engineering & MIS (ICEMIS), pp. 1–5. https://doi.org/10.1109/ICEMIS.2016.7745366.
Elgendy, M., 2020. Deep Learning for Vision Systems. Manning Publications Co. 480.
Elgendy, M., O'Reilly for Higher Education (Firm), Safari, an O.M.Company., 2020. Deep Learning for Vision Systems. Manning Publications Co, Sheler Island, NY.
Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., Zhang, C., Huang, W., 2020. On line detection of defective apples using computer vision system combined with deep learning methods. J Food Eng 286, 110102. https://doi.org/10.1016/j.jfoodeng.2020.110102.
Fan, Y., Zhang, C., Liu, Z., Qiu, Z., He, Y., 2019. Cost-sensitive stacked sparse auto-encoder models to detect striped stem borer infestation on rice based on hyperspectral imaging. Knowl Based Syst 168, 49–58. https://doi.org/10.1016/j.knosys.2019.01.003.
Fazari, A., Pellicer-Valero, O.J., Gómez-Sanchıs, J., Bernardi, B., Cubero, S., Benalia, S., Zimbalatti, G., Blasco, J., 2021. Application of deep convolutional neural networks for the detection of anthracnose in olives using VIS/NIR hyperspectral images. Comput Electron Agric 187, 106252. https://doi.org/10.1016/j.compag.2021.106252.
Fearn, T., Riccioli, C., Garrido-Varo, A., Guerrero-Ginel, J.E., 2009. On the geometry of SNV and MSC. Chemometrics and Intelligent Laboratory Systems 96, 22–26. https://doi.org/10.1016/j.chemolab.2008.11.006.
Feng, L., Zhu, S., Zhou, L., Zhao, Y., Bao, Y., Zhang, C., He, Y., 2019. Detection of Subtle Bruises on Winter Jujube Using Hyperspectral Imaging With Pixel-Wise Deep Learning Method. IEEE Access 7, 64494–64505. https://doi.org/10.1109/ACCESS.2019.2917267.
Fernández-Ruiz, V., Torrecilla, J.S., Cámara, M., Sánchez Mata, M.C., Shoemaker, C., 2010. Radial basis network analysis of color parameters to estimate lycopene content on tomato fruits. Talanta 83, 9–13. https://doi.org/10.1016/j.talanta.2010.08.020.
Ferrer-Ferrer, M., Ruiz-Hidalgo, J., Gregorio, E., Vilaplana, V., Morros, J.-R., Gené-Mola, J., 2023. Simultaneous fruit detection and size estimation using multitask deep neural networks. Biosyst Eng 233, 63–75. https://doi.org/10.1016/j.biosystemseng.2023.07.010.
Fukushima, K., 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36, 193–202. https://doi.org/10.1007/BF00344251.
Gao, F., Fang, W., Sun, X., Wu, Z., Zhao, G., Li, G., Li, R., Fu, L., Zhang, Q., 2022. A novel apple fruit detection and counting methodology based on deep learning and trunk tracking in modern orchard. Comput Electron Agric 197, 107000. https://doi.org/10.1016/j.compag.2022.107000.
Gao, C., Jiang, H., Liu, X., Li, H., Wu, Z., Sun, X., He, L., Mao, W., Majeed, Y., Li, R., Fu, L., 2024. Improved binocular localization of kiwifruit in orchard based on fruit and calyx detection using YOLOv5x for robotic picking. Comput Electron Agric 217, 108621. https://doi.org/10.1016/j.compag.2024.108621.
Gao, Z., Shao, Y., Xuan, G., Wang, Y., Liu, Y., Han, X., 2020. Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. Artificial Intelligence in Agriculture 4, 31–38. https://doi.org/10.1016/j.aiia.2020.04.003.
Garcia, J.B., Gornay, J., Lacoue-Negre, M., Garcia, S.M., Er-Rmyly, J., Bendoula, R., Roger, J.-M., 2022. A novel methodology for determining effectiveness of preprocessing methods in reducing undesired spectral variability in near infrared spectra. J near Infrared Spectrosc 30, 74–88. https://doi.org/10.1177/09670335211047959.
Genze, N., Ajekwe, R., Güreli, Z., Haselbeck, F., Grieb, M., Grimm, D.G., 2022. Deep learning-based early weed segmentation using motion blurred UAV images of sorghum fields. Comput Electron Agric 202, 107388. https://doi.org/10.1016/j.compag.2022.107388.
Aurélien Géron, 2019. Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems. O'Reilly Media 851.
Ghose, D., Das, U., Roy, P., 2018. Modeling response of runoff and evapotranspiration for predicting water table depth in arid region using dynamic recurrent neural network. Groundw Sustain Dev 6, 263–269. https://doi.org/10.1016/j.gsd.2018.01.007.
Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems. Curran Associates Inc.
Grahn, H.F., Geladi, P., 2007. Techniques and Applications of Hyperspectral Image Analysis. Techniques and Applications of Hyperspectral Image Analysis 1–368. https://doi.org/10.1002/9780470010884.
Gutiérrez, S., Wendel, A., Underwood, J., 2019. Ground based hyperspectral imaging for extensive mango yield estimation. Comput Electron Agric 157, 126–135. https://doi.org/10.1016/j.compag.2018.12.041.
Guyon, I., Nikravesh, M., Gunn, S., Zadeh, L.A. (Eds.), 2006. Feature Extraction. Studies in Fuzziness and Soft Computing 207. doi: 10.1007/978-3-540-35488-8.
Haenlein, M., Kaplan, A.M., 2004. A Beginner's Guide to Partial Least Squares Analysis. Understanding Statistics 3, 283–297. https://doi.org/10.1207/S15328031US0304_4.
Hashemi Fath, A., Madanifar, F., Abbasi, M., 2020. Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems. Petroleum 6, 80–91. https://doi.org/10.1016/j.petlm.2018.12.002.
Hespeler, S.C., Nemati, H., Dehghan-Niri, E., 2021. Non-destructive thermal imaging for object detection via advanced deep learning for robotic inspection and harvesting of chili peppers. Artificial Intelligence in Agriculture 5, 102–117. https://doi.org/10.1016/j.aiia.2021.05.003.
Hinton, G.E., Ghahramani, Z., Teh, Y.W., 1999. Learning to Parse Images. In: Solla, S., Leen, T., Müller, K. (Eds.), Advances in Neural Information Processing Systems. MIT Press.
Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A Fast Learning Algorithm for Deep Belief Nets. Neural Comput 18, 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527.
HLEG, 2019. HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES Definition developed for the purpose of the AI HLEG's deliverables.
Ho, J., Jain, A., Abbeel, P., 2020. Denoising Diffusion Probabilistic Models. Adv Neural Inf Process Syst 2020-December. doi: 10.48550/arxiv.2006.11239.
Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Comput 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
Huang, Y., Wang, J., Li, N., Yang, J., Ren, Z., 2021. Predicting soluble solids content in "Fuji" apples of different ripening stages based on multiple information fusion. Pattern Recognit Lett 151, 76–84. https://doi.org/10.1016/j.patrec.2021.08.003.
Jahanbakhshi, A., Momeny, M., Mahmoudi, M., Zhang, Y.-D., 2020. Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks. Sci Hortic 263, 109133. https://doi.org/10.1016/j.scienta.2019.109133.
Javanmardi, S., Miraei Ashtiani, S.-H., Verbeek, F.J., Martynenko, A., 2021. Computer-vision classification of corn seed varieties using deep convolutional neural network. J Stored Prod Res 92, 101800. https://doi.org/10.1016/j.jspr.2021.101800.
Jeyaraj, P.R., Asokan, S.P., Samuel Nadar, E.R., 2022. Computer-Assisted Real-Time Rice Variety Learning Using Deep Learning Network. Rice Sci 29, 489–498. https://doi.org/10.1016/j.rsci.2022.02.003.

Jiang, B., He, J., Yang, S., Fu, H., Li, T., Song, H., He, D., 2019. Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues. Artificial Intelligence in Agriculture 1, 1–8. https://doi.org/10.1016/j.aiia.2019.02.001.

Jing, L., Gulcehre, C., Peurifoy, J., Shen, Y., Tegmark, M., Soljacic, M., Bengio, Y., 2019. Gated Orthogonal Recurrent Units: On Learning to Forget. Neural Comput 31, 765–783. https://doi.org/10.1162/neco_a_01174.

Jović, A., Brkić, K., Bogunović, N., 2015. A review of feature selection methods with applications. In: In: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1200–1205. https://doi.org/10.1109/MIPRO.2015.7160458.

Kaselimi, M., Voulodimos, A., Daskalopoulos, I., Doulamis, N., Doulamis, A., 2022. A Vision Transformer Model for Convolution-Free Multilabel Classification of Satellite Imagery in Deforestation Monitoring. IEEE Trans Neural Netw Learn Syst 1–9. https://doi.org/10.1109/TNNLS.2022.3144791.

Keceli, A.S., Kaya, A., Catal, C., Tekinerdogan, B., 2022. Deep learning-based multi-task prediction system for plant disease and species detection. Ecol Inform 69, 101679. https://doi.org/10.1016/j.ecoinf.2022.101679.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol Cybern 43, 59–69. https://doi.org/10.1007/BF00337288.

Kumar Nagothu, S., Anitha, G., Siranthini, B., Anandi, V., Siva Prasad, P., 2023. Weed detection in agriculture crop using unmanned aerial vehicle and machine learning. Mater Today Proc. https://doi.org/10.1016/j.matpr.2023.03.350.

Kwabena Patrick, M., Felix Adekoya, A., Abra Mighty, A., Edward, B.Y., 2022. Capsule Networks – A survey. Journal of King Saud University - Computer and Information Sciences 34, 1295–1310. https://doi.org/10.1016/j.jksuci.2019.09.014.

Lal, T.N., Chapelle, O., Weston, J., Elisseeff, A., 2006. Embedded Methods, in: Guyon, I., Nikravesh, M., Gunn, S., Zadeh, L.A. (Eds.), Feature Extraction: Foundations and Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 137–165. doi: 10.1007/978-3-540-35488-8_6.

Lalitha, V., Latha, B., 2022. A review on remote sensing imagery augmentation using deep learning. Mater Today Proc. https://doi.org/10.1016/j.matpr.2022.03.341.

Le, J.A., El-Askary, H.M., Allali, M., Struppa, D.C., 2017. Application of recurrent neural networks for drought projections in California. Atmos Res 188, 100–106. https://doi.org/10.1016/j.atmosres.2017.01.002.

Li, T., Fang, W., Zhao, G., Gao, F., Wu, Z., Li, R., Fu, L., Dhupia, J., 2023. An improved binocular localization method for apple based on fruit detection using deep learning. Information Processing in Agriculture 10, 276–287. https://doi.org/10.1016/j.inpa.2021.12.003.

Li, Yu., Qian, M., Liu, P., Cai, Q., Li, X., Guo, J., Yan, H., Yu, F., Yuan, K., Yu, J., Qin, L., Liu, H., Wu, W., Xiao, P., Zhou, Z., 2019a. The recognition of rice images by UAV based on capsule network. Cluster Comput 22, 9515–9524. https://doi.org/10.1007/s10586-018-2482-7.

Li, L.-H., Tanone, R., 2022. Vision Transformer Approach for Vegetables Recognition. In: In: 2022 International Seminar on Application for Technology of Information and Communication (ISemantic), pp. 113–118. https://doi.org/10.1109/iSemantic55962.2022.9920435.

Li, Y., Wright, A., Liu, H., Wang, J., Wang, G., Wu, Y., Dai, L., 2019b. Land use pattern, irrigation, and fertilization effects of rice-wheat rotation on water quality of ponds by using self-organizing map in agricultural watersheds. Agric Ecosyst Environ 272, 155–164. https://doi.org/10.1016/J.AGEE.2018.11.021.

Liang, J., Wang, B., Xu, X., Xu, J., 2024. Integrating portable NIR spectrometry with deep learning for accurate Estimation of crude protein in corn feed. Spectrochim Acta A Mol Biomol Spectrosc 314, 124203. https://doi.org/10.1016/j.saa.2024.124203.

Liu, Y., Pu, H., Sun, D.-W., 2021. Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. Trends Food Sci Technol 113, 193–204. https://doi.org/10.1016/j.tifs.2021.04.042.

Lu, J., Chen, W., Lan, Y., Qiu, X., Huang, J., Luo, H., 2024. Design of citrus peel defect and fruit morphology detection method based on machine vision. Comput Electron Agric 219, 108721. https://doi.org/10.1016/j.compag.2024.108721.

Lu, W., Guo, Y.-M., Dai, D.-J., Zhang, C.-Y., Wang, X.-Y., 2018. Rice Germination Rate Detection Based on Fluorescent Spectrometry and Deep Belief Network. Guang Pu Xue Yu Guang Pu Fen Xi/spectroscopy and Spectral Analysis 38, 1303–1312. https://doi.org/10.3964/j.issn.1000-0593(2018)04-1303-10.

Lu, Z., Zhao, M., Luo, J., Wang, G., Wang, D., 2021. Design of a winter-jujube grading robot based on machine vision. Comput Electron Agric 186, 106170. https://doi.org/10.1016/j.compag.2021.106170.

MacEachern, C.B., Esau, T.J., Schumann, A.W., Hennessy, P.J., Zaman, Q.U., 2023. Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks. Smart Agricultural Technology 3, 100099. https://doi.org/10.1016/j.atech.2022.100099.

Matsui, T., Sugimori, H., Koseki, S., Koyama, K., 2023. Automated detection of internal fruit rot in Hass avocado via deep learning-based semantic segmentation of X-ray images. Postharvest Biol Technol 203, 112390. https://doi.org/10.1016/j.postharvbio.2023.112390.

Milella, A., Marani, R., Petitti, A., Reina, G., 2019. In-field high throughput grapevine phenotyping with a consumer-grade depth camera. Comput Electron Agric 156, 293–306. https://doi.org/10.1016/j.compag.2018.11.026.

Miljkovic, D., 2017. Brief review of self-organizing maps. In: 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics. https://doi.org/10.23919/MIPRO.2017.7973581.

Mirbod, O., Choi, D., Heinemann, P.H., Marini, R.P., He, L., 2023. On-tree apple fruit size estimation using stereo vision with deep learning-based occlusion handling. Biosyst Eng 226, 27–42. https://doi.org/10.1016/j.biosystemseng.2022.12.008.

Momeny, M., Jahanbakhshi, A., Jafarnezhad, K., Zhang, Y.-D., 2020. Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach. Postharvest Biol Technol 166, 111204. https://doi.org/10.1016/j.postharvbio.2020.111204.

Müller, A.C., Guido, S., 2016. Introduction to machine learning with Python : a guide for data scientists 378.

Naidu, G., Zuva, T., Sibanda, E.M., 2023. A Review of Evaluation Metrics in Machine Learning Algorithms. Lecture Notes in Networks and Systems 724 LNNS, 15–25. doi: 10.1007/978-3-031-35314-7_2.

Navidan, H., Moshiri, P.F., Nabati, M., Shahbazian, R., Ghorashi, S.A., Shah-Mansouri, V., Windridge, D., 2021. Generative Adversarial Networks (GANs) in networking: A comprehensive survey & evaluation. Computer Networks 194, 108149. https://doi.org/10.1016/j.comnet.2021.108149.

Nayak, J., Vakula, K., Dinesh, P., Naik, B., Pelusi, D., 2020. Intelligent food processing: Journey from artificial neural network to deep learning. Comput Sci Rev 38, 100297. https://doi.org/10.1016/j.cosrev.2020.100297.

Noguera, M., Aquino, A., Ponce, J.M., Cordeiro, A., Silvestre, J., Calderón, R., Marcelo, M. da E., Pedro Jordão, Andújar, J.M., da Encarnação Marcelo, M., Pedro Jordão, Andújar, J.M., 2021. Nutritional status assessment of olive crops by means of the analysis and modelling of multispectral images taken with UAVs. Biosyst Eng 211, 1–18. doi: 10.1016/j.biosystemseng.2021.08.035.

O'Shea, K., Nash, R., 2015. An Introduction to Convolutional Neural Networks. Int J Res Appl Sci Eng Technol 10, 943–947. https://doi.org/10.22214/ijraset.2022.47789.

Osako, Y., Yamane, H., Lin, S.-Y., Chen, P.-A., Tao, R., 2020. Cultivar discrimination of litchi fruit images using deep learning. Sci Hortic 269, 109360. https://doi.org/10.1016/j.scienta.2020.109360.

Park, B., Shin, T., Cho, J.-S., Lim, J.-H., Park, K.-J., 2023. Improving blueberry firmness classification with spectral and textural features of microstructures using hyperspectral microscope imaging and deep learning. Postharvest Biol Technol 195, 112154. https://doi.org/10.1016/j.postharvbio.2022.112154.

Patil, P.U., Lande, S.B., Nagalkar, V.J., Nikam, S.B., Wakchaure, G.C., 2021. Grading and sorting technique of dragon fruits using machine learning algorithms. J Agric Food Res 4, 100118. https://doi.org/10.1016/j.jafr.2021.100118.

Phaye, S.S.R., Sikka, A., Dhall, A., Bathula, D.R., 2018. Dense and Diverse Capsule Networks: Making the Capsules Learn Better. ArXiv abs/1805.0.

Polder, G., Blok, P.M., de Villiers, H.A.C., van der Wolf, J.M., Kamp, J., 2019. Potato Virus Y Detection in Seed Potatoes Using Deep Learning on Hyperspectral Images. Front Plant Sci 10, 209. https://doi.org/10.3389/fpls.2019.00209.

Pourdarbani, R., Sabzi, S., Kalantari, D., Paliwal, J., Benmouna, B., García-Mateos, G., Molina-Martínez, J.M., 2020. Estimation of different ripening stages of Fuji apples using image processing and spectroscopy based on the majority voting method. Comput Electron Agric 176, 105643. https://doi.org/10.1016/j.compag.2020.105643.

R, M., Rohini, R., 2016. LASSO: A feature selection technique in predictive modeling for machine learning. pp. 18–20. doi: 10.1109/ICACA.2016.7887916.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. In: You Only Look Once: Unified, Real-Time Object Detection. IEEE Computer Society, Los Alamitos, CA, USA, pp. 779–788. https://doi.org/10.1109/CVPR.2016.91.

Rinnan, Å., van den Berg, F., Engelsen, S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends in Analytical Chemistry 28, 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007.

Rong, D., Wang, H., Ying, Y., Zhang, Z., Zhang, Y., 2020. Peach variety detection using VIS-NIR spectroscopy and deep learning. Comput Electron Agric 175, 105553. https://doi.org/10.1016/j.compag.2020.105553.

Roweis, S.T., Saul, L.K., 2000. Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science 1979 (290), 2323–2326. https://doi.org/10.1126/science.290.5500.2323.

S, S., M, L.C., E, G.G., G, D.P., F, M.-P., B, D., 2020. AI WATCH. Defining Artificial Intelligence . doi: 10.2760/382730 (online).

Saini, P., Nagesh, D.S., 2024. CottonWeeds: Empowering precision weed management through deep learning and comprehensive dataset. Crop Protection 181, 106675. https://doi.org/10.1016/j.cropro.2024.106675.

Sánchez-Maroño, N., Alonso-Betanzos, A., Tombilla-Sanromán, M., 2007. Filter methods for feature selection - A comparative study. In: Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-540-77226-2_19/COVER.

Santosh, K.C., Das, N., Ghosh, S., 2022. Chapter 3 - Deep learning models. In: Santosh, K. C., Das, N., Ghosh, S. (Eds.), Deep Learning Models for Medical Imaging, Primers in Biomedical Imaging Devices and Systems. Academic Press, pp. 65–97. https://doi.org/10.1016/B978-0-12-823504-1.00013-1.

Sawicka, B., Egbuna, C., 2020. Chapter 1 - Pests of Agricultural Crops and Control Measures. In: Egbuna, C., Sawicka, B. (Eds.), Natural Remedies for Pest, Disease and Weed Control. Academic Press, pp. 1–16. https://doi.org/10.1016/B978-0-12-819304-4.00001-4.

Seraj, A., Mohammadi-Khanaposhtani, M., Daneshfar, R., Naseri, M., Esmaeili, M., Baghban, A., Habibzadeh, S., Eslamian, S., 2023. Chapter 5 - Cross-validation, in: Eslamian, S., Eslamian, F. (Eds.), Handbook of Hydroinformatics. Elsevier, pp. 89–105. doi: 10.1016/B978-0-12-821285-1.00021-X.

Sharma, N., Sharma, R., Jindal, N., 2021. Machine Learning and Deep Learning Applications-A Vision. Global Transitions Proceedings 2, 24–28. https://doi.org/10.1016/j.gltp.2021.01.004.

Shinoda, R., Kataoka, H., Hara, K., Noguchi, R., 2023. Transformer-based Ripeness Segmentation for Tomatoes. Smart Agricultural Technology 100196. https://doi.org/10.1016/j.atech.2023.100196.

Shuaishuai, L., Li, L., shiwang, C., Fanjia, M., Haihua, W., Zhanzhan, S., Sigrimis, N.A., 2018. Prediction Model of Transpiration Rate of Strawberry in Closed Cultivation

Based on DBN-LSSVM Algorithm. IFAC-PapersOnLine 51, 460–465. doi: 10.1016/j.ifacol.2018.08.171.
Sohl-Dickstein, J., Weiss, E.A., Maheswaranathan, N., Ganguli, S., 2015. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. In: 32nd International Conference on Machine Learning. https://doi.org/10.48550/arxiv.1503.03585.
Song, P., Wang, J., Guo, X., Yang, W., Zhao, C., 2021. High-throughput phenotyping: Breaking through the bottleneck in future crop breeding. Crop J 9, 633–645. https://doi.org/10.1016/j.cj.2021.03.015.
Song, X., Zhang, G., Liu, F., Li, D., Zhao, Y., Yang, J., 2016. Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model. J Arid Land 8, 734–748. https://doi.org/10.1007/s40333-016-0049-0.
Sringarm, C., Numthuam, S., Jiamyangyuen, S., Kittiwachana, S., Kielar, F., Wongsaipun, S., Rungchang, S., 2024. Quantitative and qualitative evaluation of maltodextrin products in the industry using near-infrared spectroscopy. Int J Food Sci Technol 59, 7391–7402. https://doi.org/10.1111/IJFS.17476.
Sun, Y., Liu, H., Guo, Z., 2021. Capsule network-based approach for estimating grassland coverage using time series data from enhanced vegetation index. Artificial Intelligence in Geosciences 2, 26–34. https://doi.org/10.1016/j.aiig.2021.08.001.
Tempelaere, A., Minh Phan, H., Van De Looverbosch, T., Verboven, P., Nicolai, B., 2023. Non-destructive internal disorder segmentation in pear fruit by X-ray radiography and AI. Comput Electron Agric 212, 108142. https://doi.org/10.1016/j.compag.2023.108142.
Thai, H.-T., Tran-Van, N.-Y., Le, K.-H., 2021. Artificial Cognition for Early Leaf Disease Detection using Vision Transformers. In: In: 2021 International Conference on Advanced Technologies for Communications (ATC), pp. 33–38. https://doi.org/10.1109/ATC52653.2021.9598303.
Theanjumpol, P., Wongzeewasakun, K., Muenmanee, N., Wongsaipun, S., Krongchai, C., Changrue, V., Boonyakiat, D., Kittiwachana, S., 2019. Non-destructive identification and estimation of granulation in 'Sai Num Pung' tangerine fruit using near infrared spectroscopy and chemometrics. Postharvest Biol Technol 153, 13–20. https://doi.org/10.1016/J.POSTHARVBIO.2019.03.009.
Torkashvand, A.M., Ahmadi, A., Nikravesh, N.L., 2017. Prediction of kiwifruit firmness using fruit mineral nutrient concentration by artificial neural network (ANN) and multiple linear regressions (MLR). J Integr Agric 16, 1634–1644. https://doi.org/10.1016/S2095-3119(16)61546-0.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention Is All You Need. Adv Neural Inf Process Syst 2017-December, 5999–6009. doi: 10.48550/arxiv.1706.03762.
Velliangiri, S., Alagumuthukrishnan, S., Thankumar joseph, S.I., 2019. A Review of Dimensionality Reduction Techniques for Efficient Computation. Procedia Comput Sci 165, 104–111. doi: 10.1016/j.procs.2020.01.079.
Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., Li, · Xiang, Li, X., Wang, C, 2021b. A review of deep learning used in the hyperspectral image analysis for agriculture. Artif Intell Rev 54, 5205–5253. doi: 10.1007/s10462-021-10018-y.
Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., Wang, F.-Y., 2017. Generative adversarial networks: introduction and outlook. IEEE/CAA Journal of Automatica Sinica 4, 588–598. https://doi.org/10.1109/JAS.2017.7510583.
Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., Li, X., 2021a. A review of deep learning used in the hyperspectral image analysis for agriculture. Artif Intell Rev 54, 5205–5253. https://doi.org/10.1007/s10462-021-10018-y.
Wang, Y., Song, S., 2023. Variety identification of sweet maize seeds based on hyperspectral imaging combined with deep learning. Infrared Phys Technol 130, 104611. https://doi.org/10.1016/j.infrared.2023.104611.
Wang, D., Vinson, R., Holmes, M., Seibel, G., Bechar, A., Nof, S., Tao, Y., 2019. Early Detection of Tomato Spotted Wilt Virus by Hyperspectral Imaging and Outlier Removal Auxiliary Classifier Generative Adversarial Nets (OR-AC-GAN). Sci Rep 9. https://doi.org/10.1038/s41598-019-40066-y.
Wendel, A., Underwood, J., Walsh, K., 2018. Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. Comput Electron Agric 155, 298–313. https://doi.org/10.1016/j.compag.2018.10.021.
Weng, S., Tang, L., Qiu, M., Wang, J., Wu, Y., Zhu, R., Wang, C., Li, P., Sha, W., Liang, D., 2023. Surface-enhanced Raman spectroscopy charged probes under inverted superhydrophobic platform for detection of agricultural chemicals residues in rice combined with lightweight deep learning network. Anal Chim Acta 1262, 341264. https://doi.org/10.1016/j.aca.2023.341264.
Wu, M., Sun, J., Lu, B., Ge, X., Zhou, X., Zou, M., 2019. Application of deep brief network in transmission spectroscopy detection of pesticide residues in lettuce leaves. J Food Process Eng 42, e13005. doi: 10.1111/JFPE.13005.
Xiao, Q., Li, G., Xie, L., Chen, Q., 2018. Real-world plant species identification based on deep convolutional neural networks and visual attention. Ecol Inform 48, 117–124. https://doi.org/10.1016/j.ecoinf.2018.09.001.
Xin, Z., Jun, S., Yan, T., Quansheng, C., Xiaohong, W., Yingying, H., 2020. A deep learning based regression method on hyperspectral data for rapid prediction of cadmium residue in lettuce leaves. Chemometrics and Intelligent Laboratory Systems 200, 103996. https://doi.org/10.1016/j.chemolab.2020.103996.
Xiong, Y., Ohashi, S., Nakano, K., Jiang, W., Takizawa, K., Iijima, K., Maniwara, P., 2021. Application of the radial basis function neural networks to improve the nondestructive Vis/NIR spectrophotometric analysis of potassium in fresh lettuces. J Food Eng 298, 110417. https://doi.org/10.1016/j.jfoodeng.2020.110417.
Xu, M., Sun, J., Yao, K., Cai, Q., Shen, J., Tian, Y., Zhou, X., 2022. Developing deep learning based regression approaches for prediction of firmness and pH in Kyoho grape using Vis/NIR hyperspectral imaging. Infrared Phys Technol 120, 104003. https://doi.org/10.1016/j.infrared.2021.104003.
Yang, Q., Shi, L., Han, J., Yu, J., Huang, K., 2020b. A near real-time deep learning approach for detecting rice phenology based on UAV images. Agric for Meteorol 287, 107938. https://doi.org/10.1016/j.agrformet.2020.107938.
Yang, C.H., Xiong, L.Y., Wang, Z., Wang, Y., Shi, G., Kuremot, T., Zhao, W.H., Yang, Y., 2020a. Integrated detection of citrus fruits and branches using a convolutional neural network. Comput Electron Agric 174, 105469. https://doi.org/10.1016/j.compag.2020.105469.
Yu, Y., Jiang, T., Gao, J., Guan, H., Li, D., Gao, S., Tang, E., Wang, W., Tang, P., Li, J., 2022. CapViT: Cross-context capsule vision transformers for land cover classification with airborne multispectral LiDAR data. International Journal of Applied Earth Observation and Geoinformation 111, 102837. https://doi.org/10.1016/j.jag.2022.102837.
Yu, X., Lu, H., Liu, Q., 2018. Deep-learning-based regression model and hyperspectral imaging for rapid detection of nitrogen concentration in oilseed rape (Brassica napus L.) leaf. Chemometrics and Intelligent Laboratory Systems 172, 188–193. https://doi.org/10.1016/j.chemolab.2017.12.010.
Yu, S., Xie, L., Huang, Q., 2023. Inception convolutional vision transformers for plant disease identification. Internet of Things 21, 100650. https://doi.org/10.1016/j.iot.2022.100650.
Zeng, F., Shao, W., Kang, J., Yang, J., Zhang, X., Liu, Y., Wang, H., 2022. Detection of moisture content in salted sea cucumbers by hyperspectral and low field nuclear magnetic resonance based on deep learning network framework. Food Research International 156, 111174. https://doi.org/10.1016/j.foodres.2022.111174.
Zhang, L., An, D., Wei, Y., Liu, J., Wu, J., 2022a. Prediction of oil content in single maize kernel based on hyperspectral imaging and attention convolution neural network. Food Chem 395, 133563. https://doi.org/10.1016/j.foodchem.2022.133563.
Zhang, J., He, L., Karkee, M., Zhang, Q., Zhang, X., Gao, Z., 2018. Branch detection for apple trees trained in fruiting wall architecture using depth features and Regions-Convolutional Neural Network (R-CNN). Comput Electron Agric 155, 386–393. https://doi.org/10.1016/j.compag.2018.10.029.
Zhang, L., Wang, Y., Wei, Y., An, D., 2022b. Near-infrared hyperspectral imaging technology combined with deep convolutional generative adversarial network to predict oil content of single maize kernel. Food Chem 370, 131047. https://doi.org/10.1016/j.foodchem.2021.131047.
Zhu, H., Yang, L., Sun, Y., Han, Z., 2021. Identifying carrot appearance quality by an improved dense CapNet. J Food Process Eng 44, e13586. doi: 10.1111/jfpe.13586.
Zhu, H., Yang, L., Ding, W., Han, Z., 2022. Pixel-level rapid detection of aflatoxin B1 based on 1D-modified temporal convolutional network and hyperspectral imaging. Microchemical Journal 183, 108020. https://doi.org/10.1016/j.microc.2022.108020.
Zhu, C. (Ed.), 2021. Appendix B - Deep learning basics, in: Machine Reading Comprehension. Elsevier, pp. 213–237. doi: 10.1016/B978-0-323-90118-5.00015-1.

Glossary

AE: Auto Encoder
AUC: Area Under Curve
ANN: Artificial Neural Network
CN: Capsule Network
CNN: Convolutional Neural Network
DBN: Deep Belief Network
DT: Decision Tree
GAN: Generative Adversarial Network
GRU: Gated Recurrent Unit
HIS: hyperspectral imaging
IoU: Intersection Over Union
KNN: K-Nearest Neighbours
LDA: Linear Discriminant Analysis
LR: Linear Regression
LSSVM: Least Squares Support Vector Regression
LSTM: Long Short-Term Memory
ML: Machine Learning
MLP: Multilayer Perceptron
MLR: Multiple Linear Regression
MSC: Multiplicative Scatter Correction
MSE: Mean Square Error
NB: Naïve Bayes
PCA: Principal Components Analysis
RBFN: Radial Basis Function Network
ReLU: Rectified Linear Activation
RGB-D: RGB images with depth information
RNN: Recurrent Neural Network
S-G: Savitzky-Golay
SNV: Standard Normal Variate
SOM: Unsupervised Self-Organizing Map
UAV: Unnamed Aerial Vehicle
VAE: Variational Autoencoder