[Artificial Intelligence in the Life Sciences 1 (2021) 100010](https://doi.org/10.1016/j.ailsci.2021.100010)

Contents lists available at [ScienceDirect](http://www.ScienceDirect.com/)

Artificial Intelligence in the Life Sciences

journal homepage: [www.elsevier.com/locate/ailsci](http://www.elsevier.com/locate/ailsci)

Machine learning in agriculture domain: A state-of-art survey

Vishal Meshram [a](#_bookmark0), Kailas Patil[a](#_bookmark0),[∗](#_bookmark3), Vidula Meshram [a](#_bookmark0), Dinesh Hanchate [b](#_bookmark1), S.D. Ramkteke [c](#_bookmark2)

a *Department of Computer Engineering, Vishwakarma University, Pune, India 411048*

b *Department of Computer Engineering, Pratishthan’s Kamalnayan Bajaj Institute of Engineering & Technology, Baramati, Pune, Maharashtra, India*

c *ICAR-National Research Center for Grapes, Manjri Farm P.O., Pune-Solapur Road, Pune, India 412307*

a r t i c l e i n f o a b s t r a c t

*Keywords:*

Deep learning Harvesting Machine learning Post-harvesting Pre-harvesting

Precision agriculture

Food is considered as a basic need of human being which can be satisfied through farming. Agriculture not only fulfills humans’ basic needs, but also considered as source of employment worldwide. Agriculture is considered as a backbone of economy and source of employment in the developing countries like India. Agriculture contributes 15.4% in the GDP of India. Agriculture activities are broadly categorized into three major areas: pre-harvesting, harvesting and post harvesting. Advancement in area of machine learning has helped improving gains in agricul- ture. Machine learning is the current technology which is benefiting farmers to minimize the losses in the farming by providing rich recommendations and insights about the crops. This paper presents an extensive survey of lat- est machine learning application in agriculture to alleviate the problems in the three areas of pre-harvesting, harvesting and post-harvesting. Application of machine learning in agriculture allows more eﬃcient and precise farming with less human manpower with high quality production.

# Introduction

Agriculture is considered an important pillar of the world’s econ- omy and also satisfies one of the basic need of human being i.e. food. In most of the countries it is considered the major source of employ- ment. Many countries like India still use the traditional way of farming, farmers are reluctant to use advanced technologies while farming be- cause of either the lack of knowledge, heavy cost or because they are unaware about the advantages of these technologies. Lack of knowl- edge of soil types, yields, crops, weather, and improper use of pesti- cides, problems in irrigation, erroneous harvesting and lack of informa- tion about market trend led to the loss of farmers or adds to additional cost. Lack of knowledge in each stage of agriculture leads to new prob- lems or increases the old problems and add the cost to farming. Growth in the population day by day also increases the pressure on the agri- culture sector. Overall losses in the agriculture processes starting from crop selection to selling of products are very high. As per the famous saying “Information is the Power”, keeping track of information about the crops, environment, and market, may help farmers to take better decisions and alleviate problems related to agriculture. Technologies like blockchain, IoT, machine learning, deep learning, cloud comput- ing, edge computing can be used to get information and process it. Ap- plications of computer vision, machine learning, IoT will help to raise the production, improves the quality, and ultimately increase the prof- itability of the farmers and associated domains. The Precision learning in

the field of agriculture is very important to improve the overall yield of harvesting.

Blockchain technology, cloud computing, internet of things (IoT), machine learning (ML) and deep learning (DL) are the latest emerging trends in the computer field. It has been already used in different do- mains like healthcare, cybercrime, biochemistry, robotics, metrology, banking, medicine, food etc. to solve the complex problems by the re- searchers. Many applications of machine learning, IoT in different do- mains are presented [[1–5]](#_bookmark20). Deep learning algorithms are making ma- chine learning more powerful and accurate. By using automated ma- chine learning (AutoML) one can cut the demand of ML experts, auto- mate the ML pipeline with more accuracy.

While performing agriculture tasks the steps as below is generally followed by farmers.

Step 1: Selection of Crop Step 2: Land Preparation Step 3: Seed Sowing

Step 4: Irrigation & fertilizing

Step 5: Crop Maintenance [use of pesticides, crop pruning etc.] Step 6: Harvesting

Step 7: Post-Harvesting activities

As per the above algorithm, the agriculture related tasks are catego- rized in the for major sub areas. [Fig. 1](#_bookmark4) shows these four sub-domains of agriculture tasks.

∗ Corresponding author: Computer Science & Engineering, Vishwakarma University, Vishwakarma University, Kondhwa (Bk), Pune, Maharashtra 411048, India.

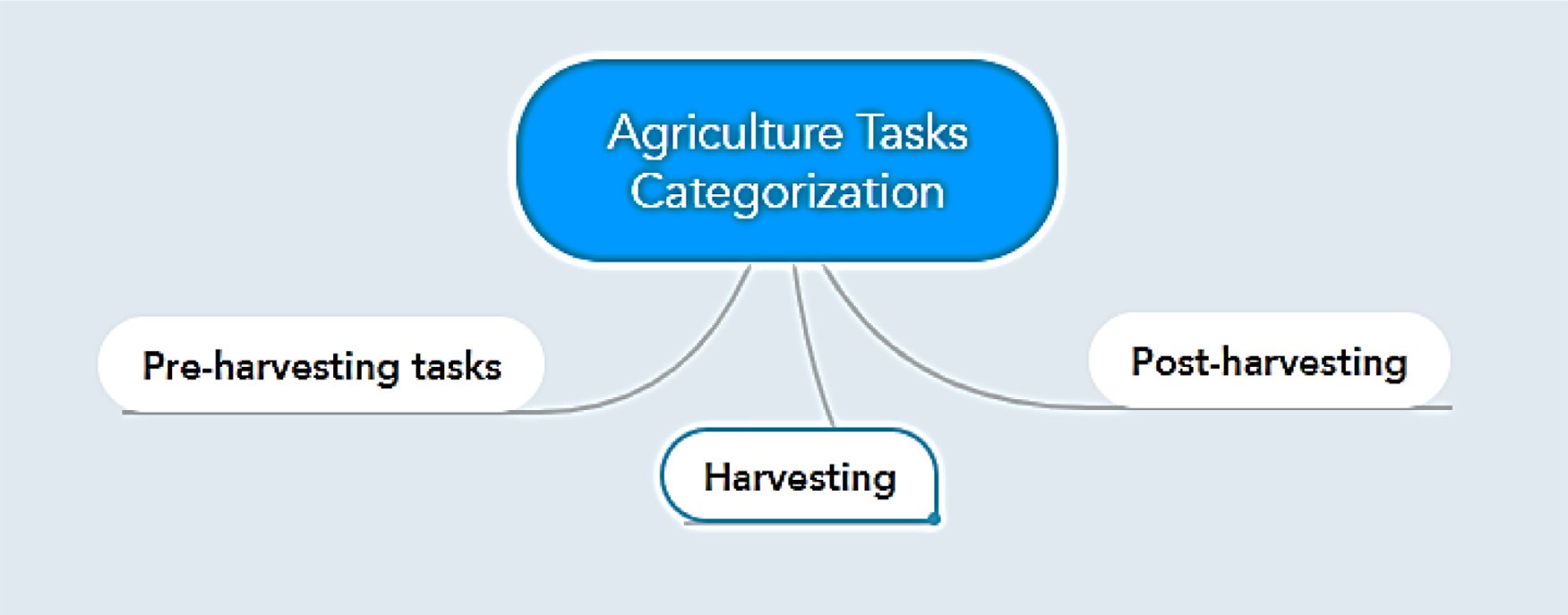
*E-mail address:* [kailas.patil@vupune.ac.in](mailto:kailas.patil@vupune.ac.in) (K. Patil).

<https://doi.org/10.1016/j.ailsci.2021.100010>

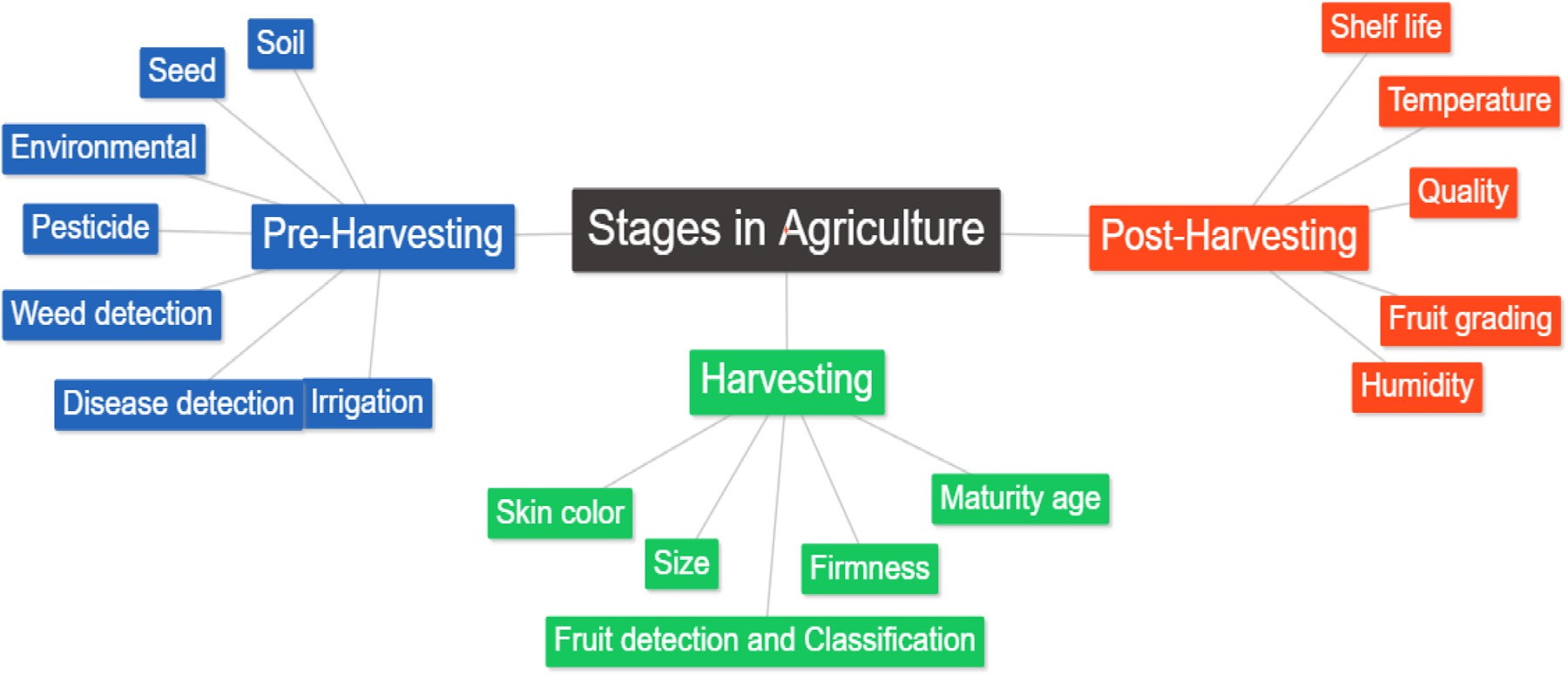
Received 21 August 2021; Received in revised form 28 September 2021; Accepted 30 September 2021

Available online 2 October 2021

2667-3185/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)



**Fig. 1.** General categorization of agriculture tasks.



**Fig. 2.** Important parameters considered in each stage of farming.

**Table 1**

Important factors to be considered in each stage.

S. No. Stage Activities / Factors References

1. Pre-harvesting Soil, seeds quality, fertiliser/pesticide application, pruning, cultivar selection, genetic and environmental conditions, irrigation, crop load, weed detection, disease detection.

[[6](#_bookmark26), [7](#_bookmark29), [9](#_bookmark33)]

1. Harvesting Fruit/crop size, skin color, firmness, taste, quality, maturity stage, market window, fruit detection and classification. [[7]](#_bookmark29)
2. Post-harvesting Factors affecting the fruit shelf-life such as temperature, humidity, gasses used in fruit containers, usage of chemicals in postharvest and fruit handling processes to retain the quality, fruit grading as per quality.

[[7]](#_bookmark29)

During pre-harvesting tasks farmers focuses on selection of crops, land preparation, seed sowing, irrigation, and crop maintenance which includes use pesticides, pruning etc. In yield estimation the farmers do the activities like yield mapping and counting the number of fruits so that they can predict the production and make the necessary arrange- ments required at the time of harvesting or post-harvesting. While har- vesting farmers are focused on maturity of crops or fruits market need quality. Whereas in post-harvesting farmers are focused on post-harvest storage and processing systems. [Fig. 2](#_bookmark5) shows the important factors that should be considered in each stage of farming. [Table 1](#_bookmark6) summarizes few works in each stage of agriculture tasks.

The major branches of the agriculture are Agronomy, Horticulture, Forestry, Livestock, Fisheries, Agriculture Engineering and Economics. The scope of the paper is confined to use of machine learning in agri- culture, specifically on fruits.

In the following sections, the review of the most recent techniques of machine vision systems used for classification and object detection in

each stage of farming is presented. [Section 2](#_bookmark7) explains the use of ML in the pre-harvesting stage. In [Section 3](#_bookmark9), usage of ML in the stage of harvesting is explained and in [Section 4](#_bookmark12) usage of ML in the post-harvesting stage is explained. [Sections 5](#_bookmark15) and [6](#_bookmark16) focuses on discussion and challenges in use of the Artificial Intelligence (AI), ML, and DL.

# Pre-harvesting

Pre-harvesting parameters play a key role in overall growth of crop/fruits. In pre-harvesting machine learning is used to capture the parameters of soil, seeds quality, fertilizer application, prun- ing, genetic and environmental conditions and irrigation. Focus- ing on each component it is important to minimize the overall losses in production. Here few important components in the pre- harvesting are considered and how neural networks and machine learning are used to capture the parameters of each component.

**Table 2**

Analysis of pre-harvesting parameter: Soil.

Important

S. No. Property features

Classes defined in the work

Dataset used (Public / Own)

Total number of images used for training

Models / Method / Algorithms compared

Best model / method/

algorithm Results Reference

1. Soil Village wise soil fertility indices of available Soil Reaction (pH), Organic Carbon (OC) and Boron (B), Phosphorus (P), and Potassium (K)

For P, K and OC three classes: Low, Medium, and High. For B six classes: Very Low, Low, Medium, Moderately High, High, and very High. For pH Four classes: Strongly Acidic (SA), Highly Acidic (HA), Moderately Acidic (MA), and Slightly Acidic (SLA).

public (reports available during the years

2014 to 2017)

NA Extreme Learning Machine (ELM) with different activation functions like

sine-squared, Gaussian radial basis, triangular basis, hyperbolic tangent, and hard limit

ELMs with Gaussian radial basis function

80% of accuracy [[10]](#_bookmark35)

1. Soil Organic

matter (SOM) and pH parameter

SOM and pH parameters Own 523 soil samples four Machine

Learning models Cubist regression model (Cubist), extreme learning machines (ELM), least squares-support vector machines

(LS-SVM), and

partial least squares regression (PLSR)

ELM R2 = 0.81 [[11]](#_bookmark36)

1. Moisture

content (MC),

Estimating moisture content (MC), organic

Own 140 set Cubist, partial least squares regression

LS-SVM is best for MC and OC

MC - RMSEP:0.457%,

[[12]](#_bookmark38)

organic carbon

carbon (OC), and

(PLSR), least squares and TN is best

RPD:2.24 TN -

(OC), and nitrogen (TN)

nitrogen (TN)

support vector machines (LS-SVM), and principal component regression (PCR)

by the Cubist

RMSEP: 0.071

and RPD :1.96

1. soil moisture Auto-regressive error function (AREF) combined with computational models

Own The soil moisture and density were determined by volumetric rings with 100 cm3 collected in

eight positions along the plots, at

depths from 25 mm to

75 mm

NA One Neuro-Fuzzy model (ANFIS) and two artificial neural networks (a

Multi-Layer Perceptron (MLP) and a Radial Basis Function (RBF)). Multiple linear regression (MLR) models with two and six independent variables

Neural Network with AREF

RMSE between 1.27% and

1.30%, R2

around 0.80, and APE between 3.77%

and 3.75%

[[13]](#_bookmark40)

1. Soil

Temperature

soil temperature (ST) at 6 different depths of 5,

10, 20, 30, 50 and

100 cm

Public (For Bandar Abbas, 10 years measured data sets for the period

of 1996–2005

and for Kerman, 7 years measured data sets for

the period of 1998–2004)

NA ELM, SaE-ELM,

genetic programming (GP) and artificial neural network (ANN)

SaE-ELM MABE - 0.8660–1.5338 C R - 0.9084–0.9893

[[14]](#_bookmark43)

* 1. *Soil*

Liakos, et al. [[8]](#_bookmark31) and Sharma, et al. [[9]](#_bookmark33) presented a soil manage- ment survey with the application of ML techniques for prediction or identification of soil properties (estimation of soil temperature, soil dry- ing, and moisture content). The categorization and estimation of the soil attributes help farmers in minimizing extra cost on fertilizers, cut the demand of soil analysis experts, increase profitability, and improve health of soil, whereas Suchithra and Pai [[10]](#_bookmark35) presented pH values and soil fertility indices classification and predication model. Yang, et al.

[[11]](#_bookmark36) observed that important indicators of soil fertility are pH values and

Soil Organic matter (SOM) and thus the authors have done prediction of SOM and pH parameters in paddy soil. Morellos, et al. [[12]](#_bookmark38) has predicted organic carbon (OC), nitrogen (TN), and moisture content (MC) param- eters of the soil. The aim of study was to compare machine learning algorithms and linear multivariate algorithms on basis of their perfor- mance of prediction. As soil moisture is frequently associated with vari- ability in yield, Johann, et al. [[13]](#_bookmark40) have estimated the moisture content of soil using with Auto-regressive error function (AREF) along with ma- chine learning algorithms. Nahvi, et al. [[14]](#_bookmark43) developed a new model by employing Self-adaptive evolutionary (SaE) agent in extreme machine learning (ELM) architecture. This new model is used for the assessment

of daily soil temperature (ST) at 6 different depths of 5, 10, 20, 30, 50 and 100 cm. The detail summary of work done by different authors on soli parameter is mentioned in [Table 2](#_bookmark8).

* 1. *Seeds*

Seed germination is a vital factor for quality of seed, which is an important determining factor of yield and quality of production. Seed germination rate calculation is still done manually with the help of trained persons which is not only a tiresome process but also prone to error. Thus, various machine leaning and image recognition tech- niques have been proposed by different authors to automate the process of seed sorting and calculation. Various computer vision, machine learn- ing techniques, Convolution Neural Network (CNN) methods have been presented in D. Sivakumar, et al. [[15]](#_bookmark45), Huang, et al. [[16]](#_bookmark47), Zhu, et al. [[17]](#_bookmark49). Image recognition technique for seed sorting with high accuracy is developed by Young, et al. [[18]](#_bookmark51). Ke-ling, et al. [[19]](#_bookmark53) used a multi- layer perceptron neural network model for improving the accuracy of the classification method to separate pepper seeds of high-quality from low-quality. Uzal, et al. [[20]](#_bookmark55) and Veeramani et al. [[21]](#_bookmark56) used the deep neural network (DNN) model using CNN for the assessment of the quan- tity of seeds per pod in soybean and for sorting of haploid seeds on basis of shape, phenotypic expression, and the embryo pose. Nkemelu, et al. [[22]](#_bookmark57), built a model using CNN for plant seedlings classification into 12 species. Medeiros, et al. [[23]](#_bookmark59) assessed the proficiency of computer vision as an alternative to routine vigor tests to expedite the process of accurate evolution of seed physiological potential. Amiryousefi, et al. [[24]](#_bookmark60) used image analysis technique, principal component analysis (PCA), to save time and cost of placing seeds in different clusters by reducing the fea- tures to be considered for clustering. Vlasov, et al. [[25]](#_bookmark62), Kurtulmuş, et al.

[[26]](#_bookmark64) used machine learning (ML) techniques for eﬃcient seed classifica-

tion. The detail summary of work done by different authors is mentioned in [Table 3](#_bookmark10).

* 1. *Pesticides and disease detection*

In-time disease detection is the most important task to save crops from major loss. Some farmers regularly analyze leaf or branches of tree while growing and identify the diseases or many times to avoid the diseases, they apply the pesticides on all the crops equally. Both the activities are based on human experience which is prone to errors and risky. Decision of which pesticide, when to apply and where to apply is totally dependent on type of disease, its stage and affected area. Appli- cation of unnecessary pesticide on all the crops may harm crops as well as farmer’s health. Precision agriculture helps farmers for application of the right pesticide at right time at right place. Many works combined pesticides prediction with the detection of disease on plants. This section discusses bout disease detection using machine learning.

Alagumariappan, et al. [[27]](#_bookmark66), developed a real-time decision support system integrated with a camera sensor module for plant disease iden- tification. In this work authors evaluated the performance of three ma- chine learning algorithms namely, Extreme Learning Machine (ELM) and Support Vector Machine (SVM) with linear and polynomial kernels and observed that the performance of ELM is better when compared to other algorithms. Savary, et al. [[28]](#_bookmark18) studied how diseases cause the crop losses and their implications for global food production losses and food security. The objective of this work is to show that crop loss research is vital and should be consider as full branch of plant science.

Sujatha, et al. [[29]](#_bookmark19), compared the ML algorithms (SVM, RF, SGD) with DL algorithms (Inception-v3, VGG-16, VGG-19) in terms of citrus plant disease detection and observed that DL methods performed much better. Karada˘g, et al. [[30]](#_bookmark21) studied detection of healthy and fusarium diseased peppers (capsicum annuum) from the reflections obtained from the pepper leaves with the help of spectroradiometer. Artificial Neural Networks (ANN), Naive Bayes (NB) and K-nearest Neighbor (KNN) ma- chine learning algorithms were used for classification. Authors claimed

that leaf reflections can be used in disease detection. Pandya [[31]](#_bookmark22), pre- sented data about different types of pesticides, their applications and impact on environment. Arsenovic, et al. [[32]](#_bookmark23) discussed the shortcom- ings of available DL models used for plant disease detections. A novel model is built which consist of two-stage architecture Disease Net, for classification of plant disease, which achieved 93.67% training accu- racy. Barbedo [[33]](#_bookmark24), explored the new approach by using DL to identify plant diseases from individual lesions and spots instead of considering entire leaf. This approach helps to detect multiple diseases on the same leaf with 12% higher accuracy. Saleem, et al. [[34]](#_bookmark25), presented a detail review of DL models used to envision different diseases of plant. Many research gaps have been enlisted in the plant disease detection and sug- gested that advanced DL algorithms should be used to increase the ac- curacy.

Liu, et al. [[36]](#_bookmark27), Kour, et al. [[37]](#_bookmark30) studied the apple leaf diseases and apple fruit diseases respectively. A CNN model was proposed to classify apple leaf diseases into Brown spot, Rust, Mosaic, and Al- ternaria leaf spot. A new dataset was created consisting of 13,689 im- ages of diseased leaves which was used to train the novel architecture based on AlexNet in [[34]](#_bookmark25). For apple disease detection and classification in Kashmir Valley, another model called Fuzzy Rule-Based Approach for Disease Detection (FRADD) was proposed in [[35]](#_bookmark28). Though the ac- curacy of the model is good, it takes into account only one disease known as scab and limited numbers of fruit types. Xing, et al. [[38]](#_bookmark32) pro- posed a new model called Weakly DenseNet-16, to overcome the limi- tations of pre-trained models which are trained on ImageNet dataset. A dataset consisting of 17 species of citrus pests and seven types of cit- rus diseases (9051 images of citrus pests and 3510 images of citrus diseases) was created. Weakly DenseNet-16 performed well with the accuracy 93.33% as compared to MobileNet-v1 (85.04%), MobileNet- v2 (87.82%), ShuﬄeNet-v1 (83.44%), ShuﬄeNet-v2 (83.21%), NIN-

16 (91.66%), SENet-16 (88.36%), and VGG-16 (92.93%). Doh, et al.

[[39]](#_bookmark34) proposed a solution to detect the citrus fruit diseases using their physical attributes such as the texture, color, structure of holes on the fruit and morphology. The proposed solution composed of K-Means clus- tering technique, ANN and SVM algorithms. Results show that the use of SVM with ANN helps in increasing disease detection and classifica- tion rate. The detailed summary of the published works is presented in [Table 4](#_bookmark11).

# Harvesting

After taking care of parameters in pre-harvesting stage like soil, seeds, weeds etc., when the fruits/vegetables are ready then harvest- ing is the most important stage. The important parameters should be focused in this stage are fruit/crop size, skin color, firmness, taste, qual- ity, maturity stage, market window, fruit detection and classification for harvesting. Careful and right harvesting of fruit is directly correlated with the profit. In the survey, we observed that auto-harvesting robots, machine learning, deep learning techniques are achieving better results and helping farmers in reducing the losses in harvesting stage. This sec- tion presents the application of ML, DL algorithms in the harvesting.

Hua, et al. [[40]](#_bookmark37) presented a detail survey on automated fruit harvest- ing systems for sweet pepper, tomato, apple and kiwifruit as an example to demonstrate the recent advances in intelligent automatic harvesting robots in horticulture. The use automatic robots in field helps to increase the production, saves the harvesting time which ultimately increase the profits of the farmers. Kushtrim, et al. [[41]](#_bookmark39) developed a CNN model based on single shot detector (YOLO) algorithm for on-tree fruit detec- tion. A dataset consisting of real and synthetic images of apple and pear trees was created. For labeling the images, open-source labeling tool called as BBox-Label-Tool was used. More than 5000 images of pear and apple fruits were used while training the model. Amazon cloud platform was used to train the model. The model achieved more than 90% accu- racy for on-tree fruit detection. Two deep neural network models were investigated in the proposed work, a small CNN model and a VGG-16

**Table 3**

Analysis of pre-harvesting parameter: Seed.

Important Sr. No. Property features

Classes defined in the work

Dataset used (Public / Own)

Total no of images used for training

Models / Method / Algorithms compared

Best model / method /

algorithm Results

Model evaluation

technique Reference

1. Seed color, shape, and texture

maze seed Own 4000 ensemble learning,

K-nearest neighbor (KNN), logistic regression, support vector machine (SVM), and Speeded Up Robust Features (SURF)

algorithm to classify the extracted features, GoogLeNet, VGG19

GoogleNet 95% Confusion Table, [[16]](#_bookmark47) Training loss.

Testing loss. Training accuracy.

Testing accuracy

1. Cotton

Jinxin5, Jinxi7,

own, dataset

13,160 SVM, PLS-DA,

self-design CNN 80% classification

[[17]](#_bookmark49)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Seed |  | Shennongmian1, collected from |  | and LR models |  |  | accuracy |  |
|  |  | Xinjiangzaomian1S, hihezi, Xinjiang |  | based on deep |  |  |  |
|  |  | Xinluzao- Uyghur |  | features |  |  |  |
|  |  | mian29, Autonomous |  | extracted by |  |  |  |
|  |  | Xinluzhong52 Region, China |  | self-design CNN |  |  |  |
|  |  | and |  | and ResNet |  |  |  |
|  |  | Xinluzhong42 |  | models |  |  |  |
| 3 | pepper | 15 features (ten | germinated seed Own | 400 seeds | multilayer | multilayer | 90% | classification | [[19]](#_bookmark53) |
|  | seeds | R, G, B, L∗, a∗, color features:  b∗, hue, | (1) and  un-germinated seed (0) |  | perceptron (MLP); BLR  binary logistic | perceptron and binary logistic  regression |  | accuracy |  |
|  |  | saturation, |  |  | regression, |  |  |  |  |
|  |  | brightness, and |  |  | single feature |  |  |  |  |
|  |  | Gray, three |  |  | models |  |  |  |  |
|  |  | geometric |  |  |  |  |  |  |  |
|  |  | features: width, |  |  |  |  |  |  |  |
|  |  | length, and |  |  |  |  |  |  |  |
|  |  | projected area, |  |  |  |  |  |  |  |
|  |  | seed weight and |  |  |  |  |  |  |  |
|  |  | density) |  |  |  |  |  |  |  |
| 4 | soybean | 38 tailored | 2-SPP, 3-SPP, Own | 18,178 | tailored features | CNN | 86.20% | accuracy | [[20]](#_bookmark55) |
|  | pods | features, | and 4-SPP |  | extraction (FE) |  |  |  |  |
|  |  | geometrical |  |  | followed by a |  |  |  |  |
|  |  | characteristics |  |  | Support Vector |  |  |  |  |
|  |  | (area, |  |  | Machines |  |  |  |  |
|  |  | perimeter, major |  |  | (SVM), CNN |  |  |  |  |
|  |  | and minor axis |  |  |  |  |  |  |  |
|  |  | length), shape |  |  |  |  |  |  |  |
|  |  | features |  |  |  |  |  |  |  |
|  |  | (density, |  |  |  |  |  |  |  |
|  |  | elongation, |  |  |  |  |  |  |  |
|  |  | ompactness, |  |  |  |  |  |  |  |
|  |  | rugosity and axis |  |  |  |  |  |  |  |
|  |  | ratio), first 4 |  |  |  |  |  |  |  |
|  |  | Hu moments, |  |  |  |  |  |  |  |
|  |  | and finally a 25 |  |  |  |  |  |  |  |
|  |  | bins histogram |  |  |  |  |  |  |  |
|  |  | of the profile of |  |  |  |  |  |  |  |
|  |  | the pod |  |  |  |  |  |  |  |
|  |  | straighten mask |  |  |  |  |  |  |  |
|  |  | added along the |  |  |  |  |  |  |  |
|  |  | short axis |  |  |  |  |  |  |  |
| 5 | haploid | texture, | True-Diploid, Own | 4021 | DeepSort, | DeepSort | 0.961 | 5-fold | [[21]](#_bookmark56) |
|  | maize | morphology, | True-Haploid |  | Support Vector |  |  | cross-validation |  |
|  | seeds | color and shape |  |  | Machine (SVM), |  |  |  |  |
|  |  |  |  |  | Random Forest |  |  |  |  |
|  |  |  |  |  | (RF), and |  |  |  |  |
|  |  |  |  |  | Logistic |  |  |  |  |
|  |  |  |  |  | Regression (LR) |  |  |  |  |

Analysis of pre-harvesting parameter: Pesticides and disease detection.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sr. No. | Property | Important features | Classes defined in the work | Dataset used Total no of (Public / images used Own) for training | Models / Method / Algorithms compared | Best model / method  / algorithm | Results | Model evaluation technique | Reference |
| 1 | Disease | color, shape, | 12 different species | Own 79,265 | AlexNet, VGG 19, | PlantDiseaseNet | 94% | TOP-1 | [[32]](#_bookmark23) |
|  | detection | and texture | and 42 different | (PlantDisease) | Inception, DenseNet, |  |  | Accuracy |  |
|  |  |  | classes (both healthy |  | ResNet, PlantDiseaseNet |  |  |  |  |
|  |  |  | and diseased) |  | Object Detection: |  |  |  |  |
|  |  |  |  |  | Two-Stage Methods - |  |  |  |  |
|  |  |  |  |  | Faster R-CNN, Faster |  |  |  |  |
|  |  |  |  |  | R-CNN with TDM, Faster |  |  |  |  |
|  |  |  |  |  | R-CNN with FPN, |  |  |  |  |
|  |  |  |  |  | One-Stage Methods - |  |  |  |  |
|  |  |  |  |  | YOLOv3, SSD513, |  |  |  |  |
|  |  |  |  |  | RetinaNet |  |  |  |  |
| 2 | Plant | individual | Healthy, | Own (Plant- PDDB - | GoogLeNet CNN | GoogLeNet CNN | 12% higher | Confusion | [[33]](#_bookmark24) |
|  | disease | lesions and | Mildly diseased, | Disease) 1575 XDB - |  |  |  | matrices |  |
|  |  | spots | Moderately diseased, | 46,409 |  |  |  |  |  |
|  |  |  | Severely diseased |  |  |  |  |  |  |
| 3 | Plant | deep | 8 classes : 5 disease | Own 1965 | extreme learning | ResNet50 model and | 98% | accuracy, | [[35]](#_bookmark28) |
|  | disease and | features | (Coryneum |  | machine (ELM), support | SVM |  | sensitivity, |  |
|  | pest |  | beijerinckii, Apricot |  | vector machine (SVM), | classifier |  | specificity, |  |
|  | detection |  | monilia laxa, Peach |  | and K-nearest neighbor |  |  | and |  |
|  |  |  | monilia laxa, Cherry |  | (KNN), VGG16, VGG19, |  |  | F1-score, |  |
|  |  |  | myzus cerasi, |  | and AlexNet |  |  | confusion |  |
|  |  |  | Xanthomonas |  |  |  |  | matrix |  |
|  |  |  | arboricola); 3 pest |  |  |  |  |  |  |
|  |  |  | (Walnut leaf mite ga, |  |  |  |  |  |  |
|  |  |  | Peach |  |  |  |  |  |  |
|  |  |  | sphaerolecanium |  |  |  |  |  |  |
|  |  |  | prunastri, Erwinia |  |  |  |  |  |  |
|  |  |  | amylovora) |  |  |  |  |  |  |
| 4 | Apple Leaf | edge, | 4 classes: Brown | Own 13,689 | AlexNet Precursor, VGG | AlexNet Precursor | 97.62% | confusion | [[36]](#_bookmark27) |
|  | Diseases | corner, | spot, Rust, Mosaic, |  | 19, Inception, DenseNet, |  |  | matrix |  |
|  |  | color, shape | and Alternaria leaf |  | ResNet, PlantDiseaseNet, |  |  |  |  |
|  |  | and object, | spot |  | SVM BP AlexNet |  |  |  |  |
|  |  |  |  |  | GoogLeNet ResNet-20 |  |  |  |  |
|  |  |  |  |  | VggNet-16 Our Work |  |  |  |  |
| 5 | Apple Fruit | background | 4 classes: Poor, | Own (Two NA | Fuzzy Rule-Based | FRADD | 91.66 | accuracy | [[37]](#_bookmark30) |
|  | Disease | and | Average, Good, | datasets) | Approach for Disease |  |  |  |  |
|  |  | foreground | Excellent |  | Detection (FRADD) |  |  |  |  |
|  |  | pixels |  |  |  |  |  |  |  |

fine-tuned model to classify the fruits by Hossain, et al. [[42]](#_bookmark41). The first model was built with six layers while the second was fine-tuned visual geometry group-16 pre-trained DL model. Two datasets were used to evaluate the performance of the proposed models. Dataset-1 is publicly available and it consists of 2633 color images whereas dataset-2 con- sists of total 5946 images, distributed among 10 classes. It was claimed VGG-16 fine-tuned model achieved excellent accuracy on both datasets. Kirk, et al. [[43]](#_bookmark42) studied on improving network performance on unseen data through a structured approach and analysis of the network input. Instead of modifying network architecture and increasing depth of neu- ral network, the fusion of features was chosen. Result shows that the model complexity for more accuracy and generalization capabilities can be avoided by using bio-inspired features. It is claimed that for the color centric data classes this approach shows more promising results with the robust DL model in real world. For this the work author created dataset consists of 6189 images over 2 months, August and September 2018, and manually annotated 150 of them. Altaheri, et al. [[44]](#_bookmark44) proposed a machine vision system to categorize date fruit images according their maturity stages which help in harvesting decision. A dataset of 8072 images were created consisting of five date types: Naboot Saif, Kha- las, Barhi, Meneifi, and Sullaj with different pre-maturity and maturity stages. The images were captured in various angles, scales, illumination conditions, and there were few occluded images. Transfer learning from two famous CNN models AlexNet and VGGNet were used to build the three classification models to classify date fruit according to their matu- rity stage, type, and whether they are harvestable or not. Result shows that VGG-16 model outperformed with the accuracy of 99.01% in 20.6

msec. Bauer, et al. [[45]](#_bookmark46) developed a platform that chains up-to-date ML techniques, modern computer vision, and integrated software engineer- ing practices to measure yield-related phenotypes from ultra-large aerial imagery named as AirSurf. Author claims that this platform help to in- crease the yield and crop marketability before the harvest. Zhang, et al.

[[46]](#_bookmark48) developed a harvesting robot for autonomous harvesting which consists of low priced gripper and ML technique for detection of cutting- point. The purpose of the study was to develop an autonomous harvester system which can harvest any crop with peduncle rather than damag- ing to its flesh. Onishi, et al. [[47]](#_bookmark50) proposed a new system (robot arm) consisting of Single Shot MultiBox Detector (SSD) and stereo camera for autonomous detection and harvesting of fruits. The system was tested on apple tee called “Fuji”. Robot arm detects the harvestable fruit po- sition and harvest it by twisting the hand axis. An experimental result shows that system was able to detect 90% fruits and took only 16 s for harvesting. Liu, et al. [[48]](#_bookmark52) proposed a novel pipeline consisting of seg- mentation, 3D localization and frame to frame tracking for accurately counting the fruits from order of images. This model was evaluated on orange and apple fruits dataset. [Table 5](#_bookmark13), presented the detail summary of harvesting techniques.

# Post-harvesting

Post-harvesting is last and most crucial area in agriculture which require more attention. After successfully completing all stages starting from yield-estimation till harvesting, negligence in post-harvesting may spoil all the efforts and cause severe loss to farmers. The subtasks that

**Table 5**

Analysis of harvesting techniques.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No.** | **Property** | **Important features** | **Classes defined in the work** | **Dataset used (Public / Own)** | **Total no of images used for training** | **Models / Method / Algorithms compared** | **Best model**  **/ Method**  **/Algorithm** | **Results** | **Model evaluation technique** | **Reference** |
| 1 | Real-Time | fruit | apple | own | 5000 | Single-Shot Convolution | YOLO | 90% | confusion | [[41]](#_bookmark39) |
|  | Fruit | shapes, | and pear fruits |  |  | Neural |  |  | matrix. |  |
|  | Detection | color and/or |  |  |  | Network (YOLO) |  |  |  |  |
|  | within tree | other |  |  |  |  |  |  |  |  |
|  |  | attributes |  |  |  |  |  |  |  |  |
| 2 | fruit classifi- | NA | 1st dataset: 15 | 1st dataset: | 1st dataset: | 2 deep learning Models : | VGG-16 | 99.75% | Confusion | [[42]](#_bookmark41) |
|  | cation |  | classes, 2nd | Public, 2nd | 2633, 2nd | 1) light model of six CNN | based |  | matrix |  |
|  |  |  | dataset: 10 | Dataset: | dataset:5946 | layers and 2)VGG-16 | architecture |  |  |  |
| 3 | Outdoor | Bio-Inspired | classes  3 classes: Ripe | own  own | 4219 | based architecture  Feature Pyramid | L∗a∗b∗Fruits | performance | F1 score, the | [[43]](#_bookmark42) |
|  | Fruit | Features, | Strawberry, | (DeepFruit) |  | Networks, Residual | system | increase of | harmonic |  |
|  | Detection | fusion of | Unripe |  |  | Neural Networks |  | 6.6 times | mean of |  |
|  |  | features | Strawberry, |  |  | and RetinaNet |  |  | precision |  |
|  |  |  | Both Classes |  |  |  |  |  | and recall |  |
| 4 | Date Fruit | local and | five date types | own | 8000 | VGG-16, AlexNet | VGG-16 | 99.01% | Confusion | [[44]](#_bookmark44) |
|  | Classifica- | spatial | in different |  |  |  |  |  | matrix. |  |
|  | tion | features and | pre-maturity |  |  |  |  |  |  |  |
|  |  | patterns | and maturity |  |  |  |  |  |  |  |
|  |  |  | stages: Naboot |  |  |  |  |  |  |  |
|  |  |  | Saif, Khalas, |  |  |  |  |  |  |  |
|  |  |  | Barhi, Menei, |  |  |  |  |  |  |  |
|  |  |  | and Sullaj |  |  |  |  |  |  |  |
| 5 | fruit | NA | apples Detected, | public | 169 | Single Shot MultiBox | YOLO | 0.9 | precision, | [[47]](#_bookmark50) |
|  | harvesting |  | Undetected |  |  | Detector (YOLO) |  |  | recall |  |
|  | robot |  |  |  |  |  |  |  |  |  |

can be consider in this stage are shelf-life of fruits and vegetables, post- harvest grading and export. Every country has their own standard rules and regulations for grading the fruits [[49–51]](#_bookmark54).

In [[52]](#_bookmark58), an information manual with directions for “Post-harvest management of mango for quality and safety assurance” was presented. This is very insightful for all the stakeholders of horticultural supply chain. Study showed that wrong post-harvest handling methods can af- fect the quality and quantity of fruits which increases the overall losses. 31% losses which are identified at retail level were caused by decay only. The other practices which add losses are poor harvesting, careless handling, and improper packaging and carriage conditions.

The wrong disease management during production causes the de- cay at high-level of pre-harvest infections. The decays in the form of anthracnose and stem end rot are very commonly observed. A train- ing manual for “handling fresh fruits, vegetables and root crops” for Grenada was presented in [[53]](#_bookmark61), as a part of the “Agricultural Market- ing Improvement” Project TCP/GRN/2901 which was implemented by Grenada Government and FAO. The goal of this project was to increase the profits for horticulture products and root crop growers through a well-organized agricultural marketing system. This document provides in detail study about all post-harvest stages with how to minimize the losses in every stage. Ucat, et al. [[54]](#_bookmark63) explored the use of image pro- cessing with deep leaning algorithm to classify Cavendish banana as per their grades. Python, OpenCV and Tensorflow were used to build the model to classify the bananas into different categories such as Class A big-hand or small-hand, Class B big-hand or small-hand and Cluster class (part of hand). Result shows that the model achieved more than 90% classification accuracy. Ireri, et al. [[55]](#_bookmark65) proposed a machine vision system for post-harvest tomato grading. The system works on RGB im- ages given as an input to the system. Dataset was created by manually labeling the tomato images into four categories according to their de- fect, healthy and ripeness parameters. Four different models were built to classify image into one of the category according to the matching fea- tures, total 15 features were considered while taking the decision Result shows that RBF-SVM performed well as compared to others for cate- gory 1 i.e. healthy or defected with 0.9709 detection accuracy. Piedad, et al. [[56]](#_bookmark67) developed a system for banana (Musa acuminata AA Group

’Lakatan’) classification using ML techniques based on tier-based. A non- invasive tier-based technique was used in this study. ANN, SVM and RF classifiers were used to classify bananas into extra class, class I, class II and rejected classes. Result shows that the random-forest algorithm outperformed as compared to others with the 94.2% accuracy. Lia, et al.

[[57]](#_bookmark68) studied and compared two hyper-spectral imaging technologies namely long-wave near infrared (LW-NIR) and short-wave near infrared (SW-NIR) for early identification of Bruise of ‘Pinggu’ peaches. An im- proved watershed segmentation algorithm based on morphological gra- dient reconstruction and marker extraction was developed and tested on multispectral PC images in this study. Experimental result shows that a proposed algorithm accurately classified 96.5% of the bruised and 97.5% of sound peaches respectively. An automated real–time grading system with quality inspection for apple fruit was developed by Sofu, et al. [[58]](#_bookmark69). The developed system comprises a roller, transporter and class conveyors joined with an enclosed cabin with camera, load cell and control panel units. System not only classifies the apples on the basis of color, size and weight parameters but also identifies defective apples. The proposed system took only 0.52 s to capture the apple image and process. Average 15 apples per seconds were sorted by the system. Author claims average sorting accuracy between 73 and 96% and the system can be used to sort different fruits like orange, potatoes and so on. A grading and sorting system based on machine vision for date fruit was developed by Ohali [[59]](#_bookmark70). The system was able to categorize the date fruit into three classes (grade 1, 2 or 3) from the given RGB image as an input. A back-propagation algorithm was tested in the study which showed 80% accuracy. Fruits and vegetables quality depends on their parameters like shape, size, texture, color and defects. Different meth- ods needs to apply in order to classify the fruits and vegetables accord- ing to their quality parameters like data collection, pre-processing of data, image segmentation, feature extraction, and finally classification. Bhargava, et al. [[60]](#_bookmark71) presented a detail survey to compare the various algorithms used in every stage of the fruits and vegetables quality in- spection. Meshram, et al. [[61]](#_bookmark72) proposed a new framework called “MNet: Merged Net” to reduce the fruits misclassification problem. Author cre- ated his own dataset of top Indian fruits consists of 12,000 images with

**Table 6**

Analysis of post-harvesting works.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sr. No. | Property | Classes defined in the work | Dataset used (Public / Own) | Total no of images used for training | Models / Method / Algorithms compared | Results | Model evaluation technique | Reference |
| 1 |  | 4 classes | own | 1116 | Python | 0.9 | accuracy | [[54]](#_bookmark63) |
|  | POSTHARVEST |  |  |  | OpenCV and |  |  |  |
|  | GRADING |  |  |  | Tensorflow |  |  |  |
|  | CLASSIFI- |  |  |  |  |  |  |  |
|  | CATION OF |  |  |  |  |  |  |  |
|  | CAVENDISH |  |  |  |  |  |  |  |
|  | BANANA |  |  |  |  |  |  |  |
| 2 | Defect dis- | 4 classes: | own | 8000 | linear-SVM, | 0.9709 | Confusion | [[55]](#_bookmark65) |
|  | crimination | category 1, |  |  | quadratic- |  | matrix |  |
|  | and grading | 2, 3, and 4. |  |  | SVM, |  |  |  |
|  | in tomatoes | depends |  |  | cubic-SVM, |  |  |  |
|  |  | upon defect, |  |  | and radial |  |  |  |
|  |  | healthy, and |  |  | basis |  |  |  |
|  |  | ripeness |  |  | function |  |  |  |
|  |  | (red color |  |  | (RBF-SVM), |  |  |  |
|  |  | intensity) |  |  | ANN, |  |  |  |
|  |  |  |  |  | decision |  |  |  |
|  |  |  |  |  | tree, and |  |  |  |
|  |  |  |  |  | random |  |  |  |
|  |  |  |  |  | forest |  |  |  |
| 3 | Postharvest | extra class, | own | 1164 | artificial | 0.942 |  | [[56]](#_bookmark67) |
|  | classifica- | class I, class |  |  | neural |  | Classification |  |
|  | tion of | II and reject |  |  | network, |  | Accuracy, |  |
|  | banana | class |  |  | support |  | F-Score, |  |
|  | (Musa |  |  |  | vector |  | Confusion |  |
|  | acuminata) |  |  |  | machines |  | matrix |  |
|  |  |  |  |  | and random |  |  |  |
|  |  |  |  |  | forest |  |  |  |
| 4 | Automatic | small, | own | 183 | K-means, | 0.79 | statistical | [[58]](#_bookmark69) |
|  | apple | normal, |  |  | C4.5 |  | test |  |
|  | sorting | large, light |  |  | decision tree |  |  |  |
|  | system | and dark, |  |  |  |  |  |  |
|  |  | defective |  |  |  |  |  |  |
|  |  | and non- |  |  |  |  |  |  |
|  |  | defective |  |  |  |  |  |  |
| 5 | Date fruit | 3 classes: | own | 1860 | back | 0.8 | Confusion | [[59]](#_bookmark70) |
|  | grading | grades 1, 2 |  |  | propagation |  | matrix |  |
|  |  | and 3 |  |  | neural |  |  |  |
|  |  |  |  |  | network |  |  |  |
|  |  |  |  |  | (BPNN) |  |  |  |

six classes. [Table 6](#_bookmark14), presented the detail summary of post-harvesting works.

# Discussion

This paper has extensively reviewed the available literature on appli- cation of machine learning and deep learning in agriculture. Different state-of-the-art machine learning and deep learning models in differ- ent stages of agriculture, including pre-harvesting, harvesting and post- harvesting in different domains were reviewed. Deep learning technol- ogy is becoming mature day-by-day. This survey shows that use of CNN in agriculture is huge and it is also getting remarkable results. By ex- ploiting depth, other structure and hardware support, the learning ca- pacity and accuracy of the CNN is significantly improved. Still there are challenges like dataset creation, time required for training and test- ing, hardware support, deployment of big models on small devices like boards or android phones, user awareness etc.

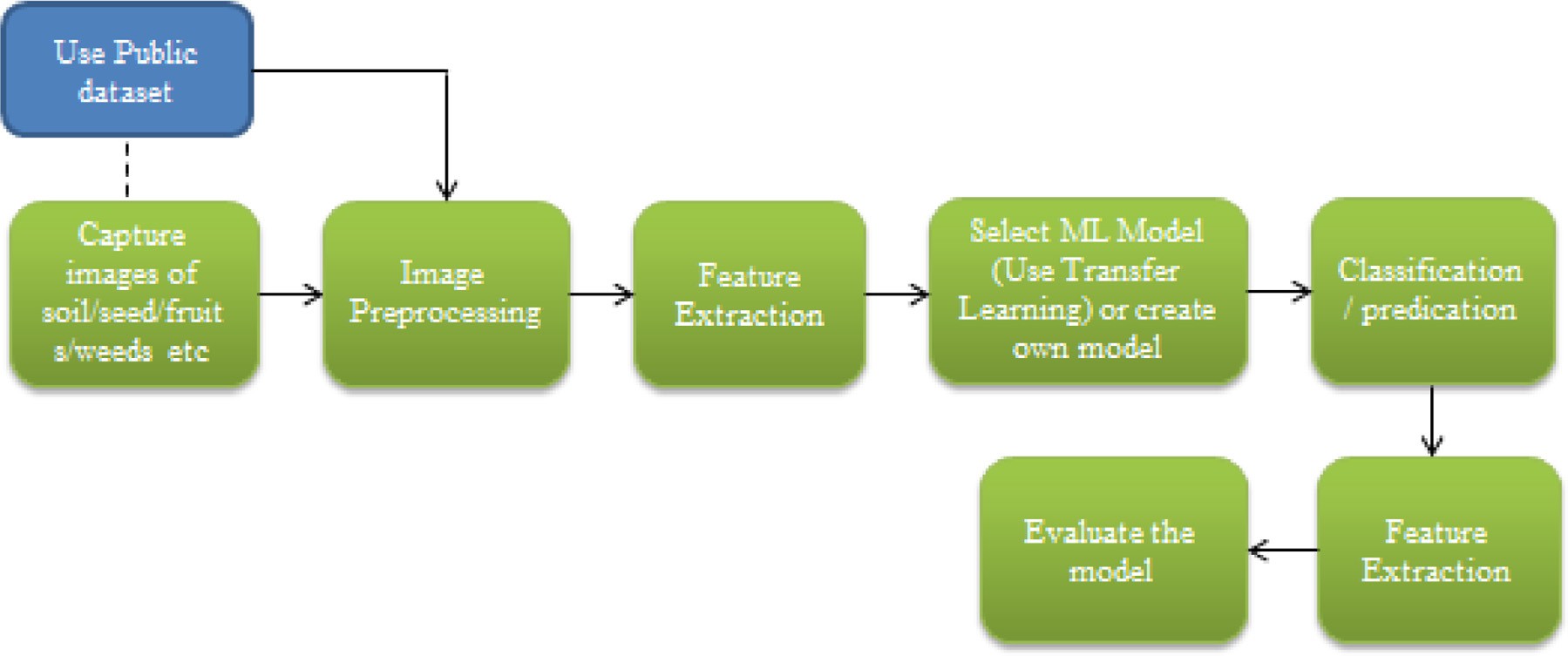
A popular technique called “Transfer Learning” is often used to miti- gate the problems of small dataset, time required for training and to im- prove the accuracy of the model. Internet of Things (IoT) systems com- bined with machine learning provides a beneficial solution to improve farming gains. Real time parameters of the farms are gathered using IoT, and the collected data is used by machine learning algorithms either to predict or for recommendations to farmers for improvements in farm- ing. From the survey it is also observed that Single-Shot Convolution

Neural YOLO (You only look once) is a state-of-the-art, real-time object detection system which must be used for detection and localization to increase the classification accuracy.

Automated machine learning (AutoML) is the latest approach which can be used to build highly eﬃcient, more accurate, high quality ML models in a less time [[62](#_bookmark73),[63](#_bookmark74)]. AutoML is used to automate the entire ML pipeline shown in the [Fig. 3](#_bookmark17), starting from data cleaning to model selection and hyperparameters tuning. These are time-consuming and iterative tasks of machine learning model development. As compared to traditional ML model development which is time-consuming, resource- intensive, need domain expertize, AutoML can accelerate the complete process to get production-ready model in less time without requiring domain expertise. In depth surveys on automated machine learning (Au- toML) is presented in [[64–69]](#_bookmark75).

# Challenges and recommendations

From this survey one can comprehend the importance of machine learning in the agriculture domain. In each phase of agriculture starting from pre-harvesting to post-harvesting, researchers have applied ma- chine learning algorithms to solve the complex problems. Today’s need is to develop precise and customized machine learning models which can perform fast, automatically analyze bigger, more complex data and help to optimize the agriculture processes like classification, recommen- dations or predications.



**Fig. 3.** Steps of Machine Learning used in literature.

The benefits of machine learning in agriculture domain are enor- mous. However, the benefits come with its challenges. Few such chal- lenges while implementing machine learning algorithms in agriculture domain are listed as follows:

1. *Data:* Data is the most fundamental requirement to build the machine learning models. Many researchers faced the challenges regarding data like lack of data, unavailability of data in required format, poor quality of data, data may contain extraneous features etc. From this survey it is observed that, many researchers use data source sites like Kaggel, Meandly, IEEE Dataport etc. to get the data to build models. If the required data is not available then researchers need to build their own dataset [[70–75]](#_bookmark70).
2. *Pre-processing of the data:* As there are lot of problems associated with data, one has to apply the different pre-processing techniques to make the data suitable for training, testing, and validation testing the model. This might be time consuming process.
3. *Selection of machine learning algorithms:* Wide list of machine learn- ing algorithm is available which make it diﬃcult to find out more suitable algorithm to build the customize machine learning model. Many times, it is required to do random selection or after compar- ing results of multiple algorithms one can come to conclusion for best suitable algorithm. This trial-and error technique may delay the model deployment process.
4. *Training and testing of the machine learning model:* Building the accu- rate model needs huge data for training. Testing and validation are also important to check the accuracy of the model before its deploy- ment. Building a model from scratch for best desired and possible outcomes needs long training and multiple time testing which are very time-consuming tasks. It needs high configuration hardware re- sources; domain knowledge programmers, testing tools etc. Overfit- ting and underfitting are the common challenges faced while build- ing the models.
5. *Deployment of models:* This is the most challenging phase to bring the models in the production as there is absence of deployment skills, third party library dependencies, size of models, complex real-world scenarios, deployment platform hardware limitations, (like android phones, embedded boards) etc.

Some more challenges are important to make a note of:

1. Understanding the business need and identification of problem.
2. Understanding user and their interaction with technology
3. User friendly application design.
4. Performance of models in the real-word scenarios.
5. Power consumption by model and battery limitations to run the model on the devices.
6. For computer vision models camera configurations at user end.

The applications of machine learning and deep learning in the field of agriculture are huge with many challenges. After this in-depth sur- vey following are a few recommendations to make the implementation process more fast, accurate, smooth and deployable.

1. Focus to build a machine learning model to solve specific problem like classification or recommendation.
2. For training the model try to create own dataset and make this avail- able to other researchers through open platform like Kaggel, Me- andly, IEEE Dataport etc.
3. For testing and validation of the models use publically available dataset.
4. To reduce the time required for training a model use the “Transfer Learning” techniques.
5. AutoML is the state-of-the-art approach which can be used to build more accurate, high quality ML models in a less time.
6. Deployment of the model in real-time application is recommended to help the intended users in their mundane work.

# Conclusion

In this paper an in-depth survey of applications of machine learning algorithms in agriculture domain is presented. According to this review, agriculture activities are broadly categorized into three major areas as pre-harvesting, harvesting and post harvesting. Important parameters to be considered in each stage are shown in [Fig. 2](#_bookmark5) and [Table 1](#_bookmark6). Ma- chine learning algorithms/techniques used in each stage are reviewed and presented in [Tables 2](#_bookmark8),[3](#_bookmark10),[4](#_bookmark11),[5](#_bookmark13) and [6](#_bookmark14) respectively. Machine learning is the state-of-art technology which is used to solve complex problem in the agriculture and helping farmers to reduce their losses. In this survey it is seen that machine learning algorithms have obtained remarkable outcomes to solve agriculture related problems.

Our study indicated that there is need to follow the machine learning pipeline with standard experimental methods. Researches should create their own dataset and make this available to others through different platforms, so that others can use it for testing and validation of their own models. This comprehensive survey of various machine learning algorithms used in different stages of agriculture will be more helpful to other researches who are working in this field.

# Financial and ethical disclosures

This work is not supported fully or partially by any funding organi- zation or agency.

# Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

# References

1. [Siqueira VS, Borges MM, Furtado RG, Dourado CN, MCosta R. Artificial In- telligence applied to support medical decisions for the automatic analysis of echocardiogram images: a Systematic Review. Artif Intell Med 2021;120(5):102165 doi.org/10.1016/j.artmed.2021.102165.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0001)
2. [Karie NM, Kebande VR, Venter HS. Diverging deep learning cognitive comput- ing techniques into cyber forensics. Forensic Sci Int 2019;1:61–7 vol. 1, pp. 61-67doi.org/10.1016/j.fsisyn.2019.03.006.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0002)
3. [Doowon J. Artificial intelligence security threat, crime, and forensics: taxonomy and open issues. IEEE Access 2020;8 184560-184574. 10.1109/ACCESS.2020.3029280.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0003)
4. [Meshram VV, Patil K, Meshram VA, Shu FC. An Astute Assistive Device for Mobility and Object Recognition for Visually Impaired People. IEEE Trans Hum Mach Syst 2019;49(5):449–60 10.1109/THMS.2019.2931745.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0004)
5. [Patil K, Jawadwala Q, Shu FC. Design and construction of electronic aid for visually impaired people. IEEE Trans Hum Mach Syst 2018;48(2):172–82 10.1109/THMS.2018.2799588.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0005)
6. [Arah IK, Amaglo H, Kumah EK, Ofori H. Preharvest and postharvest factors affecting the quality and shelf life of harvested tomatoes: a mini review. Int J Agron 2015;2015 6,Article ID 478041http://dx.doi.org/10.1155/2015/478041.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0006)
7. [Prange RK. Pre-harvest, harvest and post-harvest strategies for or- ganic production of fruits and vegetables. Acta Hortic 2012;933:43–50 DOI10.17660/ActaHortic.2012.933.3.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0007)
8. [Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agricul- ture: a review. *Sensors* (Switzerland) 2018;18(8):1–29 10.3390/s18082674.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0008)
9. [Sharma A, Jain A, Gupta P, Chowdary V. Machine learning applications for precision agriculture: a comprehensive review. IEEE Access 2021;9:4843–73 10.1109/AC- CESS.2020.3048415.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0009)
10. [Suchithra MS, Pai ML. Improving the prediction accuracy of soil nutrient classi- fication by optimizing extreme learning machine parameters. Inf Process Agricul 2019;7(1):72–82.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0010)
11. [Yang M, Xu D, Chen S, Li H, Shi Z. Evaluation of machine learning approaches to predict soil organic matter and pH using vis-NIR spectra. *Sensors* (Switzerland) 2019;19(2):263–77 10.3390/s19020263.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0011)
12. [Morellos A, Pantazi X, Moshou D, Alexandridis T, Whetton R, Tziotzios G, Wieben-](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0012) [sohn J, Bill R, Mouazen AM. Machine learning based prediction of soil total nitro- gen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosyst Eng 2016;152:104–16 http://dx.doi.org/10.1016/j.biosystemseng. 2016.04.018.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0012)
13. [Johann AL, de Araújo AG, Delalibera HC, Hirakawa AR. Soil moisture modeling based on stochastic behavior of forceson a no-till chisel opener. Comput Electron Agricul 2016;121:420–8 http://dx.doi.org/10.1016/j.compag.2015 .12.020, 2016.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0013)
14. [Nahvi B, Habibi J, Mohammadi K, Shamshirband S, Saleh Al Razgan O. Us- ing self-adaptive evolutionary algorithm to improve the performance of an ex- treme learning machine for estimating soil temperature. Comput Electron Agricul 2016;124:0168–1699 http://dx.doi.org/10.1016/j.compag.2016.03.025.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0014)
15. [Sivakumar D, SuriyaKrishnaan K, Akshaya P, Anuja GV, Devadharshini GT. Comput- erized growth analysis of seeds using deep learning method. Int J Recent Technol Eng 2019 Volume-7Issue-6S5.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0015)
16. [Huang S, Fan X, Sun L, Shen Y, Suo X. Research on classification method of maize seed defect based on machine vision. J Sens 2019;2019:9 Article ID 2716975https://doi.org/10.1155/2019/2716975.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0016)
17. [Zhu S, Zhou L, Gao P, Bao Y, He Y, Feng L. Near-infrared hyperspectral imaging com- bined with deep learning to identify cotton seed varieties. Molecules 2019;24:3268 10.3390/molecules24183268.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0017)
18. [Young J, Se JK, Dayeon K, Keondo L, Wan CK. Super-high-purity seed sorter us- ing low-latency image-recognition based on deep learning. IEEE Robot Autom Lett 2018:2377–3766 https://doi.org/10.1186/s12859-018-2267-2.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0018)
19. [Ke-ling TU, Lin-juan LI, Li-ming YANG, Jian-hua WANG, Qun SUN. Selection for high quality pepper seeds by machine vision and classifiers. J Integr Agric 2018;17(9):1999–2006 10.1016/S2095-3119(18)62031-3.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0019)
20. [Uzal LC, et al. Seed-per-pod estimation for plant breeding us- ing deep learning. Comput Electron Agricul 2018;150:196–204 https://doi.org/10.1016/j.compag.2018.04.024.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0020)
21. [Veeramani B, Raymond JW, Chanda P. DeepSort: deep convolutional networks for sorting haploid maize seeds. BMC Bioinformatics 2018;19(9):289 Suppl- https://doi.org/10.1186/s12859-018-2267-2.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0021)
22. D. Nkemelu, D. Omeiza, and N. Lubalo,”Deep convolutional neural network for plant seedlings classification”, 2018, arXiv:1811.08404v1 [cs.CV].
23. [Medeiros Dantas de, Pereira MDias, Fernanda Santos Neri Soares T, Gomes](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0023) [Noronha B, Teixeira Pinheiro D. Computer vision as a complementary method to vigour analysis in maize seeds. J Exp Agricul Int 2018;25(5):1–8 Article no.JEAI.43464.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0023)
24. [Amiryousefi MR, Mohebbi M, Tehranifar A. Pomegranate seed clustering by machine vision. Food Sci Nutr 2017;6(1):18–26 10.1002/fsn3.475.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0024)
25. [Vlasov AV, Fadeev AS. A machine learning approach for grain crop’s seed classification in purifying separation. IOP Conf Ser 2017;803(2017):012177 10.1088/1742-6596/803/1/012177.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0025)
26. [Kurtulmuş F, Alibaş İ, Kavdir I. Classification of pepper seeds using ma- chine vision based on neural network. Int J Agricul Biol Eng 2016;9(1):51–62 10.3965/j.ijabe.20160901.1790.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0026)
27. [Alagumariappan P](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0027), [Dewan NJ](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0027), [Muthukrishnan GN](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0027), [Bojji Raju BK](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0027), [Bilal RAA](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0027), [Sankaran V. Intelligent plant disease identification system usingmachine learning. Eng Proc 2020;2:49 10.3390/ecsa-7-08160.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0027)
28. [Savary S, Ficke A, Aubertot J-N, Hollier C. Crop losses due to diseases and their implications for global food production losses and food security. Food Secur 2012 DOI 10.1007/s12571-012-0200-5.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0028)
29. [Sujatha R, Chatterjee JM, Jhanjhi NZ, Brohi SN. Performance of deep learning vs machine learning in plant leaf disease detection. Microprocess Microsyst 2021;80 https://doi.org/10.1016/j.micpro.2020.103615.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0029)
30. [Karada˘g K, Tenekeci ME, Tasaltın R, Bilgilic A. Detection of pepper fusarium disease using machine learningalgorithms based on spectral reflectance. Sustain Comput 2018:8 https://doi.org/10.1016/j.suscom.2019.01.001.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0030)
31. [Pandya IY. Pesticides and their applications in agriculture. Asian J Appl Sci Technol (AJAST) 2018;2(2):894–900 ISSN: 2456-883X.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0031)
32. [Arsenovic M, Karanovic M, Sladojevic S, Anderla A, Stefanovic D. Solving current limitations of deep learning based approaches for plant disease detection. Symmetry (Basel) 2019;11:939 10.3390/sym11070939.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0032)
33. [Barbedo JGA. Plant disease identification from individual le- sions and spots using deep learning. Biosyst Eng 2019;180:96–107 https://doi.org/10.1016/j.biosystemseng.2019.02.002.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0033)
34. [Saleem MH, Potgieter J, Arif KM. Plant disease detection and classification by deep learning. Plants 2019;8:468 10.3390/plants8110468.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0034)
35. [Türkoğlu M, Hanbay D. Plant disease and pest detection using deep learning-based features. Turk J Electr Eng Comput Sci 2019;27:1636–51 10.3906/elk-1809-181.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0035)
36. [Liu B, Zhang Y, He D, Li Y. Identification of apple leaf diseases based on deep con- volutional neural networks. Symmetry (Basel) 2018;10:11 10.3390/sym10010011.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0036)
37. [Kour V, Arora S. Fruit Disease Detection Using Rule-Based Classification. In: Pro- ceedings of Smart Innovations in Communication and Computational Sciences, Ad- vances in Intelligent Systems and Computing (ICSICCS-2018); 2019. p. 295–312. 10.1007/978-981-13-2414-7\_28.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0037)
38. [Xing S, Lee M, Lee K. Citrus pests and diseases recognition model using weakly dense connected convolution network. Sensors 2019;19:3195 10.3390/s19143195.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0038)
39. [Doh B, Zhang D, Shen Y, Hussain F, Doh RF, Ayepah K. Automatic Citrus Fruit Dis- ease Detection By Phenotyping Using Machine Learning. In: Proceedings of the 25th international conference on automation & computing. Lancaster UK: Lancaster Uni- versity; 2019. 5-7 September.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0039)
40. [Hua Y, Zhang N, Yuan X, Quan L, Yang J, Nagasaka K, Zhou X. Recent advances in intelligent automated fruit harvesting robots. Open Agricul J 2019;13:101–6 10.2174/1874331501913010101.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0040)
41. [Kushtrim B, Demetrio P, Alexandra B, Brunella M, Grappa C. Single-shot convolu- tion neural networks for real-time fruit detection within the tree. Front Plant Sci 2019;10:611 10.3389/fpls.2019.00611.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0041)
42. [Hossain MS, Al-Hammadi M, Muhammad G. Automatic fruit classification us- ing deep learning for industrial applications. IEEE Trans Ind Inf 2019;15(2) 10.1109/TII.2018.2875149.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0042)
43. [Kirk R, Cielniak G, Mangan M. L∗ a∗ b∗ Fruits: a Rapid and Robust Outdoor Fruit De-](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0043)

[tection System Combining Bio-Inspired Features with One-Stage Deep Learning Net- works. Sensors 2020;20:275 10.3390/s20010275.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0043)

1. [Altaheri H, Alsulaiman M, Muhammad G. Date fruit classification for robotic harvest- ing in a natural environment using deep learning. IEEE Access 2019;7:117115–33 10.1109/ACCESS.2019.2936536.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0044)
2. [Bauer A](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Bostrom AG](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Ball J](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Applegate C](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Cheng T](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Laycock S](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Rojas SM](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Kirwan J](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045), [Zhou Ji. Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: a case study of lettuce production. Hortic Res 2019;6:70 https://doi.org/10.1038/s41438-019-0151-5.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0045)
3. [Zhang T, Huang Z, You W, Lin J, Tang X, Huang H. An autonomous fruit and veg- etable harvester with a low-cost gripper using a 3D sensor. Sensors 2020;20:93 10.3390/s20010093.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0046)
4. [Onishi Y, Yoshida T, Kurita H, Fukao T, Arihara H, Iwai A. An auto- mated fruit harvesting robot by using deep learning. Robomech J 2019;6:13 https://doi.org/10.1186/s40648-019-0141-2.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0047)
5. X. Liu, S.W. Chen, S. Aditya, N. Sivakumar, S. Dcunha, C. Qu, C.J. Taylor, J. Das, and

V. Kumar,”Robust Fruit Counting: Combining Deep Learning, Tracking, and Struc- ture from Motion”, arXiv:1804.00307v2 [cs.CV] 2 Aug 2018.

1. United states department of Agriculture (USDA) Grade standards for Fruits, <https://www.ams.usda.gov/grades-standards/fruits> (Accessed: July 2021).
2. EU, Fruit and Vegetables: Marketing Standards, 2011**,**<https://ec.europa.eu/agriculture/fruit-and-vegetables/marketing-standards_en> (Accessed: July 2021).

[https://upload.indiacode.nic.in/showfile?actid=AC\_CEN\_23\_31\_00011\_193701\_](https://upload.indiacode.nic.in/showfile?actid=AC_CEN_23_31_00011_193701_1535099362507\046type%3Drule\046filename%3Dfruits_and_vegetables_grading_and_marking_rules%2C_2004.pdf) [51] Government Of India, AGMRK ("Agricultural Marketing Adviser"), 2004, [1535099362507&type=rule&filename=fruits\_and\_vegetables\_grading\_and\_marking\_](https://upload.indiacode.nic.in/showfile?actid=AC_CEN_23_31_00011_193701_1535099362507\046type%3Drule\046filename%3Dfruits_and_vegetables_grading_and_marking_rules%2C_2004.pdf)

[rules,\_2004.pdf (accessed:July 2021).](https://upload.indiacode.nic.in/showfile?actid=AC_CEN_23_31_00011_193701_1535099362507\046type%3Drule\046filename%3Dfruits_and_vegetables_grading_and_marking_rules%2C_2004.pdf)

1. [Esguerra Elda B. Post-harvest management of mango for quality and safety assurance guidance for horticultural supply chain stakeholders. Rome: Food and Agriculture Organization of The United Nations; 2018.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0052)
2. Gaetano Paltrinieri. Handling of fresh fruits, vegetables and root crops

-a training manual- for grenadaAgricultural marketing improvement TCP/GRN/2901. Food and Agriculture Organization of the United Nations; 2021. <http://www.fao.org/3/au186e/au186e.pdf>. accessed:July.

1. [Ucat RC, Dela Cruz JC. Postharvest grading classification of cavendish banana using deep learning and tensorflow. In: 2019 International Symposium on Mul- timedia and Communication Technology (ISMAC); 2019. p. 1–6. 10.1109/IS- MAC.2019.8836129.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0054)
2. [Ireri D, Belal E, Okinda C, Makange N, Ji C. A computer vision system for defect dis- crimination and grading in tomatoes using machine learning and image processing. Artif Intell Agricul 2019;2:28–37 https://doi.org/10.1016/j.aiia.2019.06.001.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0055)
3. [Jr Piedad E, Larada JI, Pojas GJ, Vithalie L, Ferrer V. Postharvest classification of ba- nana (Musa acuminata) using tier-based machine learning. Postharvest Biol Technol 2018;145:93–100 https://doi.org/10.1016/j.postharvbio.2018.06.004.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0056)
4. [Lia J, Chena L, Huanga W. Detection of early bruises on peaches (Amyg- dalus persica L.) using hyperspectral imaging coupled with improved wa- tershed segmentation Algorithm. Postharvest Biol Technol 2018;135:104–13 http://dx.doi.org/10.1016/j.postharvbio.2017.09.007.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0057)
5. [Sofu MM, Er O, Kayacan MC, Cetisli B. Design of an automatic apple sort- ing system using machine vision. Comput Electron Agric 2016;127:395–405 http://dx.doi.org/10.1016/j.compag.2016.06.030.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0058)
6. [Ohali YA. Computer vision based date fruit grading system: design and implementation. J King Saud Univ – Comput Inf Sci 2011;23:29–36 10.1016/j.jksuci.2010.03.003.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0059)
7. [Bhargava A, Bansal A. Fruits and vegetables quality evaluation using com- puter vision: a review. J King Saud Univ – Comput Inf Sci 2018;33:243–57 https://doi.org/10.1016/j.jksuci.2018.06.002.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0060)
8. [Meshram VA, Patil K, Ramteke SD. MNet: a framework to reduce fruit image misclas- sification. Ing Syst Inf 2021;26(2):159–70 https://doi.org/10.18280/isi.260203.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0061)
9. Microsoft, “What is automated machine learning (AutoML)?” <https://docs.microsoft.com/en-US/azure/machine-learning/concept-automated-ml>, (Accessed: July 2021).
10. [Hibayesian, “awesome-automl-papers”, 2017, https://github.com/hibayesian/ awesome- automl- papers (Accessed: July 2021).](https://github.com/hibayesian/awesome-automl-papers)
11. X. He, K. Zhao, X. Chu, “AutoML: A Survey of the State-of-the-Art”, 2021, arXiv:1908.00709v6 [cs.LG] 16 Apr 2021.
12. Zoller M, Huber MF. Benchmark and survey of automated machine learning frame- works. J Artif Intell Res 2021;70:409–74. arXiv:1904.12054v5 [cs.LG] 26 Jan 2021.
13. R. Elshawi, M. Maher, S. Sakr, “Automated machine learning: state-of-the-art and open challenges”, 2019, arXiv:1906.02287v2 [cs.LG] 11 Jun 2019.
14. Q. Yao, M. Wang, Y. Chen, W. Dai, Y. Li, W. Tu, Q. Yang, Y. Yu, “Taking the hu- man out of learning applications: a survey on automated machine learning”, 2019, arXiv:1810.13306v4 [cs.AI] 16 Dec 2019.
15. L. Yang, and A. Shami, “On hyperparameter optimization of machine learning algo- rithms: theory and practice”, 2020, arXiv:2007.15745v2 [cs.LG] 7 Aug 2020.
16. H.J. Escalante, “Automated Machine Learning - a brief review at the end of the early years”, 2020, arXiv:2008.08516v3 [cs.LG] 24 Aug 2020.
17. [Mihai Oltean. Fruits 360 dataset. Mendeley Data 2018;V1 10.17632/rp73yg93n8.1](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0070).
18. [prabira Kumar sethy. Indian Fruits-40. Mendeley Data 2020;V1 10.17632/bg3js4z2xt.1.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0071)
19. [Meshram VA, Thanomliang K, Ruangkan S, Chumchu P, Patil K. FruitsGB: top indian fruits with quality. IEEE Dataport 2020 https://dx.doi.org/10.21227/gzkn-f379.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0072)
20. Meshram VA, Patil K. FruitNet: Indian Fruits Dataset with quality (Good, Bad & Mixed quality). Mendeley Data 2021;V1. doi:[10.17632/b6fftwbr2v.1](https://doi.org/10.17632/b6fftwbr2v.1).
21. [Tripathi M, Maktedar D, Dhanajay D. Fruits and Vegetables. Mendeley Data 2020;V2 10.17632/73kpfrbcck.2.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0074)
22. [Math M, Kumar R, Dharwadkar V, Nagaraj Dr. Real-world tomato image dataset for deep learning and computer vision applications involving precision agriculture. Mendeley Data 2020;V1 10.17632/9zyvdgp83m.1.](http://refhub.elsevier.com/S2667-3185(21)00010-6/sbref0075)