



# Deep Learning based Image Processing Solutions in Food Engineering: A Review

Ninja Begum, Manuj Kumar Hazarika

10.18805/ag.R-2182

## ABSTRACT

Image based assessment of food quality for wholesomeness, nutritional composition, suitability as raw material for processing, degree of processing, product aesthetics, consumer attractiveness etc., are some of the aspirations for applying machine learning in food technology. The initial contributions made by machine learning in the field of artificial intelligence are now more prominent through the techniques of deep learning. This review presents the contributions of machine learning in obtaining image processing based solutions in food technology and the relative advantages of deep learning over machine learning as the technique for solving complex problems like image recognition and image classification. The deep learning based solutions to the problems of image processing are highlighted as the enablers of disruptions in the design and development of different sorting, grading and dietary assessment tools.

**Key words:** Classification, Deep learning, Food, Identification, Machine learning.

Computer vision or image processing-based systems find many applications in the food and agricultural industries. It has its application in fresh produce classification, quality assessment, sorting, automated grading, disease recognition and robotic harvesting. Grading, sorting and disease recognition has helped to evaluate the quality of fresh produce during post-harvest (Tripathi and Maktedar, 2020). Hence, researchers have been motivated to use these applications for evaluation of quality level in fruits and vegetables during post-harvest.

Machine learning has made several contributions in the field of food technology making several breakthroughs. Machine learning in tea industry has helped in quality monitoring emphasizing on grading (Borah, 2005), in disease recognition of tea leaves (Karmokar *et al.*, 2015), in chlorophyll content detection of tea leaf (Sonobe and Hirono, 2020), in categorization of tea leaf (Kamrul *et al.*, 2020). In the beverage industry machine learning found its application in predicting wine quality (Gupta, 2018), Fruit Juice-Alcohol Mixture Analysis (Ordukaya and Karlik, 2016), beer quality assessment (Viejo *et al.*, 2018). In bakery industry it found its way in prediction of baking quality (Isleroglu and Beyhan, 2020), in determination of the optimal state of dough fermentation in bread production (Giefer *et al.*, 2019). Also machine learning has paved its way in first lactation milk yield prediction in buffaloes (Gandhi *et al.*, 2012) and in decision making on fertility of eggs (Onler *et al.*, 2017). Thus machine learning has gained much attention the field of food solving different problems in (a) dietary assessment by making prediction on antioxidant capacity of food (Yordi *et al.*, 2016), by estimating nutritional value of food (Yunus *et al.*, 2018), by classifying food nutrients composition (Aziz and Habeeb, 2019), (b) image recognition and (c) image classification.

This review starts with how machine learning has made breakthrough in the field of artificial intelligence. Emphasizing

Department of Food Engineering and Technology, Tezpur University, Sonitpur-784 028, Assam, India.

**Corresponding Author:** Ninja Begum, Department of Food Engineering and Technology, Tezpur University, Sonitpur-784 028, Assam, India. Email: ninzasworld@gmail.com

**How to cite this article:** Begum, N. and Hazarika, M.K. (2021). Deep Learning Based Image Processing Solutions in Food Engineering: A Review. Agricultural Reviews. DOI: 10.18805/ag.R-2182.

**Submitted:** 26-02-2021    **Accepted:** 12-07-2021    **Online:** 11-09-2021

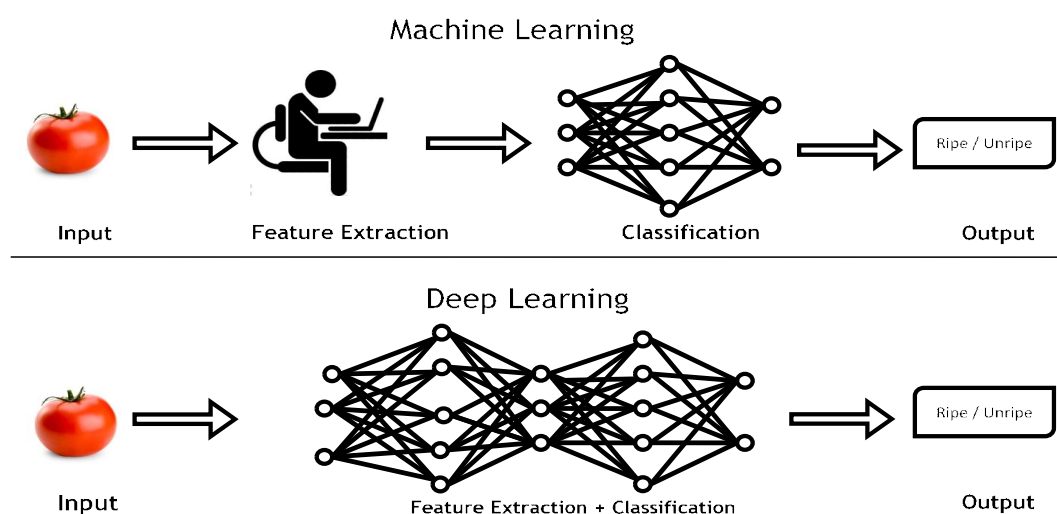
more on deep learning, it is discussed in details. It also aims at presenting a comparative study between machine learning and deep learning as a technique for solving complex problems like image recognition and image classification which enables in the design and development different sorting, grading and dietary assessment techniques.

## Review framework

### Machine learning vs deep learning

In recent years, machine learning along with computer vision techniques has been widely used in the field of agriculture. It has gained popularity as a common tool in the field of information extraction from large data sets. But one of the disadvantages of traditional machine learning is that its lack of ability to analyze data by extracting features automatically. This disadvantage is rectified by using Deep Learning which on the other hand possesses strong feature learning ability wherein solving many complex problems rapidly and effectively. The difference between traditional Machine Learning (ML) and Deep Learning (DL) is depicted in Fig 1.

ML is an application of knowledge engineering where a computer system automatically learns from the previous data to make predictions/classifications of future inputs. ML can



**Fig 1:** Distinction between the Machine Learning architecture and Deep Learning architecture.

further be sub-divided into three learning techniques: Supervised Learning, Unsupervised Learning and Reinforced Learning. In supervised learning, the machine learns under the guidance of previous expert knowledge. The previous knowledge is given as input to the machine in order that it learns the features and later the machine is asked to identify new information. There are various supervised ML algorithms such as Logistic Regression, Naive Bayes, Support Vector Machines, Nearest Neighbor, Artificial Neural Networks, Decision Tree, *etc.*, (Soofi and Awan, 2017). Unsupervised learning can be called 'self-learning' where patterns or features from the input data are understood without any help from a teacher, basically by figuring out the dataset given and finding hidden patterns and make predictions of output. Some common types of algorithms used for unsupervised learning are Clustering, Anomaly detection, Neural Networks, *etc.* Reinforced learning allows the computer to make classification/prediction directly on input data. On making correct decisions, the learning algorithm is rewarded and for wrong decisions, the learning algorithm is penalized. In this way, the model learns the features in a reinforced manner.

Deep Learning (DL) in general uses deep layers of Neural Networks (NN) to analyze data and subsequently make classification. Neural network more specifically Artificial Neural Network (ANN) is a connected structure composed of many interconnected computational units, called neurons that is considered to mimics a neuron in human brain. A good definition of ANN is given by Haykin in (Guresan and Kayakutlu, 2011) which states that ANN is a collection of simple computation units called neurons and these neurons are interconnected. It learns an approximation function that maps the input to the output. The learning process involves modifying the weights in each connection in order to minimize errors. Deep learning architectures are trained using more advanced neural networks which are known as Deep Neural Network. The deep neural network has the ability of automatic feature extraction which makes

it unique and outstanding. The training of the deep learning architectures may be supervised, unsupervised, or reinforcement based. If the architecture is trained about the targeted output it is known as supervised learning. If the architecture discovers the targeted output itself by learning from the given input it is known as unsupervised learning. There are two broad categories of DL techniques: Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN). An RNN is one of the deep learning techniques where the output of the previous step is recursively given as input to the next step. Unlike traditional feed-forward NN, RNN is aware of the time sequences hence finding its application in learning features from Time Series data, Text data, Audio data, *etc.* Various models of RNN include LSTMs, GRUs and Bidirectional RNNs. On the other hand, CNN is a subset of DL that is commonly used to analyze visual data and find patterns for classification. It is used as an input layer followed by Convolutional layers which learn the filters that extract hidden characteristics from the image pixels. Other components such as Pooling layer and Full-connected layer are also used along with Convolutional layer. CNN has become very popular because of its ability to learn from a large corpus of data and make superior predictions of images, videos, text, *etc.* CNN models are tremendously used in image and video processing, classification and segmentation projects.

Here,  $s$  is the total number of images,  $x$  is the height of the image,  $y$  is the width of the image and  $d$  is image depth. The Convolutional layers convolve features of the input by applying filters and pass its result to the successive layer. The activation function known as Rectified Linear Unit (ReLU) is used along with the Convolutional layer which replaces the negative-number of the convolution layer with zero (0), which helps for faster and effective training. After that, the pooling layer performs function such as max pooling or average pooling. The maximum value in a certain filter region is retained in max pooling and the average value in a filter region is retained in average pooling. These are typically

used to reduce the dimensionality of the network. After that, the output of the pooling layer is given as input to the fully-connected layer. The fully connected layer along with the softmax activation function performs classifications on the input data. The output of the fully-connected layer is not normalized and thus, the softmax function is used to distribute the probability between 0 and 1 among the classes. In Fig 2. Pictorial representation of the CNN architecture, the pictorial representation of the different layers of CNN is depicted.

### CNN architectures/frameworks

For image classification, CNN architectures such as LeNet, AlexNet, VGG, GoogLeNet, ResNet, *etc.* have become very popular. All the CNN architectures have their strengths and weaknesses. LeNet is the first and simplest CNN architecture consisting of 7 layers (multiple convolution and average pooling) used for black and white object recognition is the LeNet (LeCun *et al.*, 1998). AlexNet is the first successful CNN for big dataset classification. It consists of 5 convolutions, 3 fully connected, max pooling and dropout layers (Shanmugamani, 2018). At Oxford University, Visual Geometry Group (VGG) proposed a CNN architecture that supports up to 19 convolution layers. It is similar to AlexNet but feature extraction increases as the network goes deeper. Google Net, the winner at the ILSVRC 2014 image classification challenge provides an architecture that significantly decreases error rate as compared to AlexNet and VGG. The basic building block of Google Net is an inception module that used variable kernel sizes to create different receptive fields. These receptive fields generate feature maps that have sparse correlation patterns (Alom

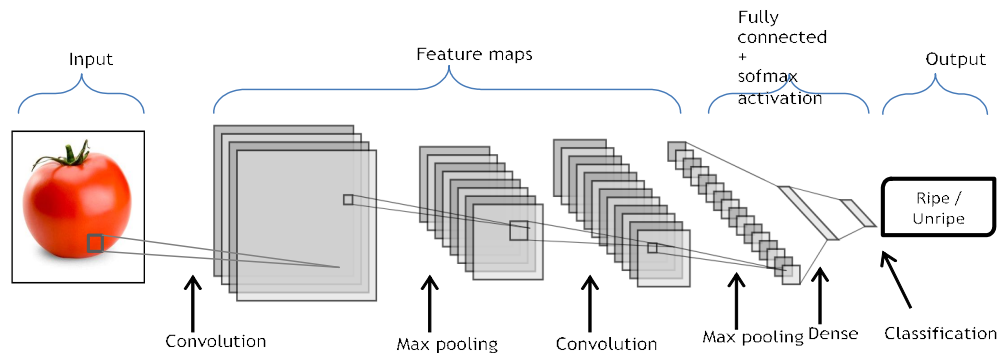
*et al.*, 2018). Further, 1x1 convolutions and global average pooling are also used in this architecture in the intermediate layers.

After AlexNet, VGG and Google Net, another groundbreaking result in image classification is achieved by deep Residual Network (ResNet) (He *et al.*, 2016). The core idea of ResNet is called “identity shortcut connection” that jumps over one or more layers. It achieves outstanding performance while training large number of layers. In Table 1, comparison of different CNN architecture is presented.

### Popular food datasets

This section discusses the different datasets prepared for classification of food products.

UNICT-FD889 is the first dataset of food image containing 889 different plates of food (Farinella *et al.*, 2014). Introducing photometric and geometric variation, image of each dish is captured multiple times using a smart phone. The overall dataset contains 3583 images. Another dataset termed as Food 101 containing images of 101 food categories has been used to train different classifiers (Bossard *et al.*, 2014). It consists of 750 training images and 250 manually reviewed test images for each class. Fruit 360 is a huge dataset of total 90483 images of size 100x100 pixel. There are total 131 classes of fruit and vegetables (Muresan and Oltean, 2018). The dataset “UEC FOOD 256” contains photos of 256 different kinds of food items (Kawano and Yanai, 2014). The demarcation of the food item is done by a rectangular box in the photo. Authors in (Singla *et al.*, 2016) prepare two datasets, termed Food-5K and Food-11, in order to recognize food item from non-food items. Further, each datasets is subdivided into training, testing and



Convolutional Neural Network (CNN) Architecture

Fig 2: Pictorial representation of the CNN architecture.

Table 1: Comparison of different CNN architecture.

Architecture	Developed by	Year	No of parameters
LeNet	YannLeCun <i>et al.</i>	1998	60K
AlexNet	Alex Krizhevsky <i>et al.</i>	2012	61M
VGG	Simonyan and Zisserman	2014	-
GoogLeNet	Google	2014	4M
ResNet	Kaiming He	2015	25 M

validation set for learning model. The Food-5K dataset consists of 5000 images selected from publicly available dataset that already exists. Out of 5000 images, 2500 images are of food items and remaining 2500 images contain pictures other than food items. On the other hand, Food-11 uses 16,643 images of 11 categories from publicly available dataset. The major food items that are daily consumed are categories in the dataset. In Table 2, we present the comparison of the different food dataset.

### Tools

In this section, we discuss about the different image processing and deep learning tools used in the food image classification. Various tools and platforms for image classification using deep learning have been developed, allowing the researchers to conduct experiments. Some of the popular deep learning tools are PyTorch, TensorFlow, Deep learning Matlab toolbox, Pylearn2, Caffe, TFLearn, etc.

**Table 2** Comparison of different food datasets

Name	Class	Images	Year
UNICT-FD889	889	3583	2014
UEC FOOD 256	256	28375	2014
Food 101	101	101000	2015
Food-5K	2	5000	2016
Food 11	11	16643	2016
Fruit 360	60	90483	2018

Highlights	Description
Agricultural product	Tomato
Image dataset	Dataset of images of randomly picked fresh fruits and vegetables from markets at Menofia city, Egypt. Images from four faces (Top, Side1, Side2 and End) of 177 tomatoes were used to create the dataset.
Tool	MATLAB
Model/Classifier	SVM classifier
Classes	2
Feature extraction	Color and texture (Color moments, GLCM and Wavelets energy and entropy) extraction and then fused.
Accuracy	Achieved accuracy of 92%.

In (Ileri *et al.*, 2019), authors propose a defect detection system of tomatoes based on image color, texture and shape features has been proposed. Defects such as calyx and stalk were detected by an RBF-SVM classifier with LAB color-space pixel values. Further details are listed below:

Highlights	Description
Agricultural Product	Tomato
Image dataset	Dataset of images from 200 tomatoes having different degrees of defects and red color intensity from Jiangsu, China. Cherry tomatoes and Heirloom tomatoes were used.
Tool	Not mentioned
Model/Classifier	ANN recognition models and SVM classifier
Classes	4
Feature extraction	Lab color features are extracted as color features. Also, contrast, correlation, energy, homogeneity and entropy are extracted as texture features using GLCM. Regularity measure is used as shape features.
Accuracy	Overall accuracy of 97.09%.

Theano is an open source Python library developed by a Montreal Institute for Learning Algorithms (MILA) at the University of Montreal in the year 2007. It uses Python's Numpy library to perform complex mathematical operations. Also, it runs the experiment in multiple GPUs and CPUs using the BLAS library (Al-Rfou *et al.*, 2016). Tensorflow (Abadi *et al.*, 2016) which was created by Google in the year 2015 in order to replace Theano is almost similar to that of theano. Like most deep-learning frameworks, TensorFlow is written with a Python API over a C/C++ engine that makes it run faster. Keras, created by Francois Chollet, a software engineer at Google in the year 2015 is a deep-learning library that sits atop Tensorflow and Theano, providing an intuitive API written in Python. In 2017, Berkeley Vision and Learning Center (BVLC) developed Caffe (Jia *et al.*, 2014). Caffe is written in C++ along with bindings from Python and MATLAB. It is used for training convolutional neural networks and other deep models.

### Applications of Machine Learning in Food

Authors in (Semary *et al.*, 2015) proposed a grading system of tomato fruits based on surface defects using Machine learning as well as computer vision techniques. Classification is done using surface characteristics of the tomatoes using SVM. Color and texture features are extracted and fused using Feature fusion method. Methods used in the process are listed below:

Authors in (Tu *et al.*, 2018) propose a method to identify passion fruit and detect their 175 maturity by developing RGB-D images. Detection of passion fruits is done using faster R-CNN with an accuracy of 92.71%. Detection of maturity was done based on RGB-DSIFT-LLC into SVM classifier with an accuracy of 91.52%.

Highlights	Description
Agricultural product	Passion fruit
Image dataset	4000 images were acquired in total. For image acquisition, a Kinect V2.0 device with infrared (IR) depth and a digital RGB camera was used.
Tool	OpenCV3.0, Caffe, MATLAB 2012b.
Model/Classifier	R-CNN recognition model and SVM classifier
Classes	Five (5) maturity classes: after-mature (AM), mature (M), near-mature (NM), near-young (NY)
Feature extraction	Dense SIFT features extraction using LLC and spatial pyramid matching (SPM)
Accuracy	Detection accuracy of 92.71% and classification accuracy of 91.52%.

Authors in (Kilic *et al.*, 2007) propose a bean classification system using computer vision and artificial neural network. In this work the beans are classified based on their quality.

Highlights	Description
Agricultural product	Beans
Image dataset	Self prepared dataset containing 140 samples of beans.
Tool	MATLAB
Model/Classifier	ANN
Classes	5
Feature extraction	For each channel (R, G, B), four moment values were calculated as features.
Accuracy	Highest classification accuracy of 90.6%

Author in (Militante, 2019) discuss an automated system for grading of fruits. Their study specifically deals with detection and classification of Batuan fruits.

Highlights	Description
Agricultural product	Batuan fruit
Image dataset	Self prepared dataset containing 120 images.
Tool	Not mentioned
Model/Classifier	ANN, SVM, KNN and rule-based classifier used in this work
Classes	3
Feature extraction	Features such as color, texture, or surface information, shape and size are extracted.
Accuracy	Highest classification accuracy of 98%

Authors in (Wan *et al.*, 2018) developed a method for detection of maturity levels (green, orange and red) of fresh market tomatoes by combining the feature color value 185 with the back propagation neural network (BPNN) classification technique. Two varieties of tomatoes Roma and Pear was used.

Highlights	Description
Agricultural product	Tomato
Image dataset	Out of 200 samples collected images, 150 tomato samples for each variety were retained. Thus, a dataset containing 300 images is prepared.
Tool	MATLAB 2016a
Model/Classifier	Back Propagation based Neural Network (BPNN) classifier
Classes	3
Feature extraction	Maximum inscribed circles on the surface of the tomatoes were used for the feature color area (FCA) extraction region.
Accuracy	99.31%

In (Rafiq *et al.*, 2016), authors used Artificial Neural Network (ANN) for quality evaluation of tomatoes based on color and size features. Authors also design an image acquisition system to serve the purpose.

Highlights	Description
Agricultural product	Tomato
Image dataset	Dataset of images for a wide variety of hybrid tomatoes cultivated in Assam, India and categorized into unripe, ripe and overripe. Images from 30 samples from each category were used for the development of ANN model for classification.
Tool	MATLAB 2014a
Model/Classifier	ANN
Classes	3
Feature extraction	Color and size features
Accuracy	Highest classification accuracy reported is 99%.

Authors in (Bendary *et al.* 2015) described an automated approach for multi-class classification of ripeness of tomatoes and evaluation, investigating and classifying its different maturity/ripening stages. PCA was used for feature selection in this work. The selected features are then used to train two different classifiers: SVM and LDA.

Highlights	Description
Agricultural product	Tomato
Image dataset	Self prepared dataset containing 250 images of tomato at different maturity stage.
Tool	MATLAB
Model/Classifier	SVM, LDA
Classes	2
Feature extraction	Color moments and color histogram is used as features.
Accuracy	SVM provides maximum accuracy of 90.80% has been achieved.

### Applications of deep learning in food

In (Kagaya and Aizawa, 2015), classification of food/non-food using CNN as a feature extractor and classifier with three different datasets has been performed. The combinations of training and testing using the all three of them were compared.

Highlights	Description
Agricultural product	Food items
Image dataset	Images from Instagram with search key #Food Dataset (IFD), Food-101 and Caltech-256 Dataset (FCD) is used to prepare dataset.
Tool	Caffe
Model/Classifier	CNN classifier
Classes	2
Accuracy	IFD -96%, FCD -95% and prepared dataset-99%

Authors in (Zhu *et al.*, 2018) used deep learning for classification of vegetables using a large number of image dataset using AlexNet model. Also, a comparison of the experiment was made with the traditional BP neural network and the SVM classifier where AlexNet shows a higher classification accuracy rate of 92.1%.

Highlights	Description
Agricultural product	Vegetables: Broccoli, pumpkin, cauliflower, mushroom and cucumber.
Image dataset	Popular dataset ImageNet used.
Tool	Caffe (a deep learning tool developed in C++)
Model/Classifier	CNN based on AlexNet, SVM and ANN
Classes	5
Accuracy	Highest classification accuracy using CNN is 92.1%

In (Patil and Gaikwad, 2018) deep learning technique has been used for vegetable classification. The authors perform transfer learning using Inception v3 model for extracting 205 features and finally classifying the vegetables.

Highlights	Description
Agricultural product	Vegetables: Tomato, onion, carrot and cucumber.
Image dataset	ImageNet dataset
Tool	Tensorflow platform
Model/Classifier	Inception-v3 model
Classes	4
Accuracy	Highest classification accuracy achieved is 99%



Authors in (Femling *et al.*, 2018) propose an approach for creating a system which can identify fruit and vegetables in the retail market using images captured with a video camera attached to the system.

Highlights	Description
Agricultural product	Fruits and Vegetables: apple, avocado, banana, bell pepper, clementine, kiwi, orange, pear, potato and tomato.
Image dataset	ImageNet dataset was used to extract 400 images per class. In parallel, camera captured images, in 30 numbers per class were collected.
Tool	Tensorflow framework
Model/Classifier	CNN based on Inception and MobileNet architectures.
Classes	10
Accuracy	Highest classification accuracy of 97% achieved.

In (Muresan and Oltean, 2018), authors prepared a dataset of fruit images called Fruit-360. Further, they also implement a deep neural network to classify different fruits.

Highlights	Description
Agricultural product	Fruits
Image dataset	Self prepared dataset named Fruit-360 was used. It contains 38409 images of 60 fruits at the time of reporting by researchers.
Tool	Tensorflow framework
Model/Classifier	CNN classifier
Classes	60
Accuracy	Highest classification accuracy of 96.3%.

Authors in (Rodriguez *et al.*, 2018) dealt with automated system of identifying different plum varieties at early maturity stages, which is even a difficult task for the human expert. Different varieties of plum are used in this work.

Highlights	Description
Agricultural product	Plum fruit
Image dataset	Self prepared dataset containing 525 images of plum captured in different weeks.
Tool	Tensorflow framework
Model/Classifier	CNN based on AlexNet
Classes	3
Accuracy	Classification accuracy ranges from 91 to 97%.

In (Mergez and Seljak, 2017), an automatic food recognition and classification systems has been proposed. The authors expect that it will help in the process of estimation of food intake and dietary assessment. A newly-introduced deep CNN architecture called NutriNet has been used for recognition and classification.

Highlights	Description
Agricultural product	Food and drinks
Image dataset	Prepared dataset from images gather from google image search based on class label. A dataset of 520 food and drink image classes of 100 images each was prepared.
Tool	Caffe and PyTorch
Model/Classifier	Four deep convolutional neural network architectures: NutriNet, AlexNet (2012), GoogLeNet (2014) and Deep Residual Networks (ResNet, 2015) used as classifier.
Classes	3
Accuracy	Highest classification accuracy of 94.47% achieved.

Authors in (Pan *et al.*, 2017), proposed a new framework called DeepFood, which extracts high quality effective features from food ingredient images using deep learning. Use of deep learning improved its accuracy of classification proving it an advanced machine learning techniques.

Highlights	Description
Agricultural product	Food items
Image dataset	100 images of each class collected from public dataset of food ingredients called MLC dataset.
Tool	Caffe
Model/Classifier	Sequential Minimal Optimization (SMO) based SVM classifier, Random Forest, Bayes Net, Bagging
Classes	41
Feature extraction	The features are extracted using transfer learning algorithms based on CNN
Accuracy	Highest classification accuracy of 87.78% achieved

Authors in (Goswami and Liu, 2017) show that CNNs are quite suitable for classifying food dishes and outperform traditional machine learning approaches in classifying food dishes. Here deep learning is used on images of food dishes: there are multiple cuisines around the world.

Highlights	Description
Agricultural product	Food dishes
Image dataset	Dataset consists of images of 20 popular food dishes from around the world.
Tool	Tensorflowframework
Model/Classifier	SVM, CNN (used VGG-16 [1] model pre-trained on ImageNet for transfer learning)
Classes	20
Feature extraction	Histogram of Oriented Gradients (HOG) as well as a color histogram using the hue channel in HSV color space is extracted as features.
Accuracy	Highest classification accuracy around 45% using CNN

Authors in (Panday *et al.*, 2017) propose a food classification system which can recognize the contents of the meal automatically from the images of the food using deep CNN architecture. They also prepare an Indian food dataset having 50 food dishes of Indian origin.

Highlights	Description
Agricultural product	Food dishes
Image dataset	Trained on Food-101 dataset. Also, self prepared dataset is used in this work. There are 100 images per class for 50 classes.
Tool	MATLAB 2014a
Model/Classifier	Deep CNN fine-tuned sub-networks: AlexNet, GoogLeNet and ResNet architecture.
Classes	50
Accuracy	Highest classification accuracy achieved on the Indian food dataset is 97.60%

Authors in (Fan *et al.*, 2020) use deep CNN to detect defective apples from their images. They have developed a system that can detect defects at the rate of 5 fruit per sec.

Highlights	Description
Agricultural product	Apple
Image dataset	Self prepared dataset with 300 images of apple. Where, 150 correspond to defective apple and 150 normal apples.
Tool	OpenCV, Tensorflow
Model/Classifier	CNN
Classes	2
Accuracy	Highest classification accuracy of 96.5%

Authors in (Vaviya *et al.*, 2019) apply CNN to detect artificially ripen fruits from naturally ripe fruits. The system uses Android Smartphone to capture images of fruit and give it as input to the classifier.



Highlights	Description
Agricultural product	Banana
Image dataset	Fruit 360 dataset is used to train the classifier to identify fruits. A dataset is also prepared using 14 bananas artificially ripen and 14 bananas which are normally ripen. Each class has 1000 images which are captured from different angles.
Tool	OpenCV, Tensorflow framework
Model/Classifier	CNN
Classes	2
Accuracy	Highest classification accuracy achieved on the Indian food dataset is 97%

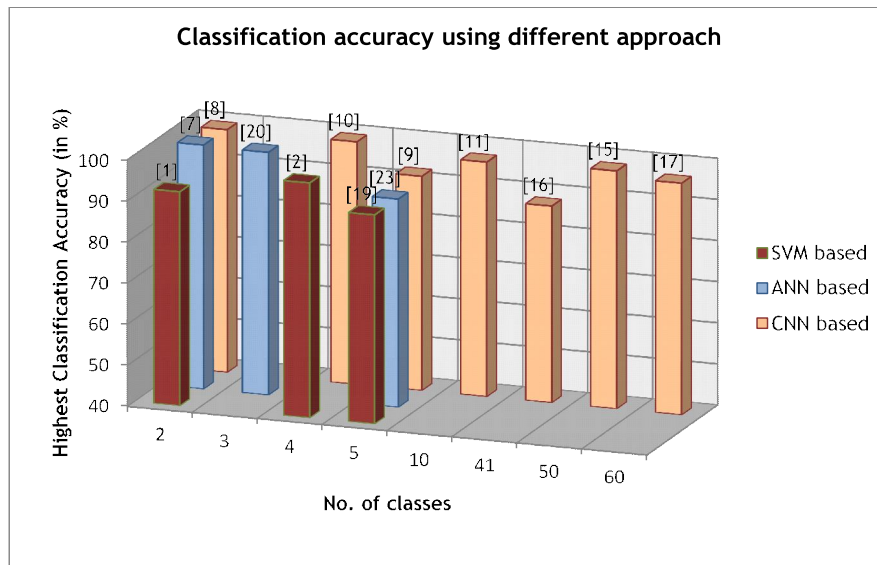


Fig 3: Highest classification accuracy of different machine learning approaches

In this section, we put forward the insight of the review of literature pertaining to classification food or agricultural products using machine learning techniques. Most of the research paper target quality assessment of agricultural products. For quality assessment, detection of maturity or ripeness, artificial ripening, etc. have been taken into account. Some works have also been conducted to automatically identify food from non-food items from images. The literature can be broadly classified in to two categories based on the type of machine learning algorithm used for classification: (a) SVM-based, (b) ANN-based, (c) CNN-based. Out of the different papers reviewed on application of machine learning approaches in food classification, figure 3 presents only the papers capable of achieving highest classification. X-axis represents number of classes, Y-axis represents classification accuracy of each individual method. Classification accuracy of the above machine learning approaches can be identified from the references on top of each bar in the bar diagram. Each bar represents the highest classification accuracy of each individual method in classifying 2, 3, 4, 5, 10, 41, 50 and 60 classes respectively.

In SVM-based works, SVM classifier with different kernels is used to classify the images of agricultural products. Further in ANN-based works, the extracted features from the images are given as input to ANN which

is used to classify different food products. The SVM-based and ANN-based classifier, require features extracted from images as input. Some of the common features extracted are color moments, texture, shape and size. Maximum number of classes is limited to 5 classes for SVM-based classifiers and ANN based works. Among all the approaches, CNN-based approach outperforms all other approaches. In CNN-based works, pre-existing models such as AlexNet, ResNet, GoogLeNet, etc. has been used with modifications. Also, transfer learning approaches have been adopted in case of CNN architectures. The classification accuracy achieved using CNN classifiers approaches 99% for different number of classes. From Fig 3, it is also evident that the CNN-based approach is able to handle large number food classes without much degradation in classification accuracy. In figure 3, it is also evident that CNN-based models are increasingly being used for solving classification problems with highest accuracy in food domain compared to SVM-based and ANN-based methods. This increase in use of CNN is due to its performance in terms of classification accuracy and its ability to handle large number of classes. Further, deep learning architectures such as CNN have greater capability of handling large datasets. CNN based methods outperforms SVM and ANN based methods in the way that the latter two require manual feature extraction

process before training the model while CNN extracts features automatically. Thus input to the CNN architectures can be in raw form.

## CONCLUSION

Computer vision, being a non-destructive quality assessment tool, is replacing the manual and destructive quality assessment methods in the classification, identification and nutritional quality assessment. In the papers reviewed in this work, many have shown applicability of the machine learning techniques for quality assessment, dietary assessment, automatic sorting and grading of agriculture produces. Deep learning is reported to outperform other techniques due to automatic feature extraction capability and the ability to handle large datasets. Continued improvement in the performance of deep learning in image processing will encourage many researchers to apply deep learning in food domain to solve more complex problems non-destructively and with more classification accuracy.

## Declaration of competing interest

The authors declare that they have no conflict of interest.

## ACKNOWLEDGEMENT

The authors gratefully acknowledge the help and support provided by Tezpur University, Assam, India.

## REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G. and Isard, M. (2016). Tensorflow: A system for large scale machine learning. In: 12<sup>th</sup> USENIX symposium on operating systems design and implementation (OSDI' 16). pp. 265-283.
- Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Van Esesn, B.C., Awwal, A.A.S. and Asari, V.K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches, arXiv preprint arXiv: 1803.01164.
- Al-Rfou, R., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J., Belikov, A. and Belopolsky, A. (2016). Theano: A python framework for fast computation of mathematical expressions, arXiv: arXiv:1605.
- Aziz, A.S.B.A. and Habeeb, R.A.A. (2019). Classification of food nutrients composition using deep learning. url: [https://www.researchgate.net/publication/338935647\\_Classification\\_of\\_Food\\_Nutrients\\_Composition\\_using\\_Deep\\_Learning](https://www.researchgate.net/publication/338935647_Classification_of_Food_Nutrients_Composition_using_Deep_Learning)
- Borah, S. (2005). Machine vision for tea quality monitoring with special emphasis on fermentation and grading. Tezpur University, Assam. url: <https://shodhganga.inflibnet.ac.in/handle/10603/100305>.
- Bossard, L., Guillaumin, M. and Gool, L.V. (2014). Food-101-mining discriminative components with random forests. In: European conference on computer vision, Springer. pp. 446-461. [https://dx.doi.org/10.1007/978-3-319-10599-4\\_29](https://dx.doi.org/10.1007/978-3-319-10599-4_29)
- El-Bendary, N., Hariri, E.El, Hassanien, A.E. and Badr, A. (2015). Using machine learning techniques for evaluating tomato ripeness. Expert Systems with Applications. 42(4): 1892-1905. <https://doi.org/10.1016/j.eswa.2014.09.057>.
- Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., Zhang, C. and Huang, W. (2020). On line detection of defective apples using computer vision system combined with deep learning methods. Journal of Food Engineering. 286: 110102. <https://doi.org/10.1016/j.jfoodeng.2020.110102>.
- Farinella, G.M., Allegra, D. and Stanco, F. (2014). A benchmark dataset to study the representation of food images. In: European Conference on Computer Vision, Springer. pp. 584-599. [https://dx.doi.org/10.1007%2F978-3-319-16199-0\\_41](https://dx.doi.org/10.1007%2F978-3-319-16199-0_41).
- Femling, F., Olsson, A. and Alonso-Fernandez, F. (2018). Fruit and vegetable identification using machine learning for retail applications. In: 14<sup>th</sup> International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), IEEE. pp. 9-15. <https://doi.org/10.1109/SITIS.2018.00013>
- Gandhi, R.S., Monalisa, D., Dongre, V.B., Ruhil, A.P., Singh, A. and Sachdeva, G.K. (2012). Prediction of first lactation 305-day milk yield based on monthly test day records using artificial neural networks in Sahiwal cattle. Indian Journal of Dairy Science. 65(3). <https://doi.org/10.5146/IJDS.V65I3.25895.G11927>.
- Giefer, L.A., Lutjen, M., Rohde, A.K., Freitag, M. (2019). Determination of the optimal state of dough fermentation in bread production by using optical sensors and deep learning. Applied Sciences. 9(20): 4266. <https://doi.org/10.3390/app9204266>.
- Goswami, A. and Liu, H. (2017). Deep dish: Deep learning for classifying food dishes, Stanford University Reports. url: <http://cs231n.stanford.edu/reports/2017/pdfs/6.pdf>
- Gupta, Y. (2018). Selection of important features and predicting wine quality using machine learning techniques. Procedia Computer Science. 125: 305-312. <https://doi.org/10.1016/j.procs.2017.12.041>.
- Guresen, E. and Kayakutlu, G. (2011). Definition of artificial neural networks with comparison to other networks. Procedia Computer Science. 3: 426-433. <https://doi.org/10.1016/j.procs.2010.12.071>.
- He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778.
- Ireri, D., Belal, E., Okinda, C., Makange, N. and Ji, C. (2019). A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. Artificial Intelligence in Agriculture. 2: 28-37. <https://doi.org/10.1016/j.aiia.2019.06.001>.
- Isleroglu, H. and Beyhan, S. (2020). Prediction of baking quality using machine learning based intelligent models. Heat and Mass Transfer, Springer. 56: 2045-2055. <https://dx.doi.org/10.1007/s00231-020-02837-6>.
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S. and Darrell, T. (2014). Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on Multimedia. pp. 675-678. <https://doi.org/10.1145/2647868.2654889>
- Kagaya, H. and Aizawa, K. (2015). Highly accurate food/non-food image classification based on a deep convolutional neural network. In: International conference on image analysis and processing, Springer. pp. 350-357. [https://dx.doi.org/10.1007/978-3-319-23222-5\\_43](https://dx.doi.org/10.1007/978-3-319-23222-5_43).

- Kamrul, M.H., Rahman, M., Robin, M.R.I., Hossain, M.S., Hasan, M.H., Paul, P. (2020). A deep learning based approach on categorization of tea leaf. In: Proceedings of the International Conference on Computing Advancements. pp. 1-8. <https://doi.org/10.1145/3377049.3377122>.
- Karmokar B.C., Ullah, M.S., Siddiquee, M.K. and Alam, K.M.R. (2015). Tea leaf diseases recognition using neural network ensemble. *International Journal of Computer Applications*. 114(17): 975-8887. <http://dx.doi.org/10.5120/20071-1993>.
- Kawano, Y. and Yanai, K. (2014). Automatic expansion of a food image dataset leveraging existing categories with domain adaptation. In: *European Conference on Computer Vision*, Springer, Cham.
- Kılıç, K., Boyacı, I. H., Köksel, H. and Küsmenoğlu, İ. (2007). A classification system for beans using computer vision system and artificial neural networks. *Journal of Food Engineering*. 78(3): 897-904. <http://dx.doi.org/10.1016/j.jfoodeng.2005.11.030>.
- LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. In: *Proceedings of the IEEE*. 86(11): 2278-2324. <https://doi.org/10.1109/5.726791>.
- Mezgec, S. and Koroušić Seljak, B. (2017). NutriNet: a deep learning food and drink image recognition system for dietary assessment. *Nutrients* 9(7): 657. <https://doi.org/10.3390/nu9070657>.
- Militante, S. (2019). Fruit grading of *Garcinia Binucao* (batuan) using image processing. *International Journal of Recent Technology and Engineering (IJRTE)*. 8(2): 1829-1832. <https://dx.doi.org/10.35940/ijrte.b1028.078219>.
- Muresan, H. and Oltean, M. (2018). Fruit recognition from images using deep learning. *Acta Universitatis Sapientiae, Informatica*. 10(1): 26-42. <http://dx.doi.org/10.2478/ausi-2018-0002>.
- Önler, E., Çelen, I.H., Gulhan, T. and Boynukara, B. (2017). A study regarding the fertility discrimination of eggs by using ultrasound. *Indian Journal of Animal Research*. 51(2): 322-326. <http://dx.doi.org/10.18805/ijar.v0iOF.4561>.
- Ordukaya, E. and Karlik, B. (2016). Fruit juice-alcohol mixture analysis using machine learning and electronic nose. *IEEE Transactions on Electrical and Electronic Engineering*. 11(S1): S171-S176. <http://dx.doi.org/10.1002/tee.22250>.
- Pan, L., Pouyanfar, S., Chen, H., Qin J. and Chen, S.C. (2017). Deep Food: Automatic multi-class classification of food ingredients using deep learning. In: 2017 IEEE 3<sup>rd</sup> international conference on collaboration and internet computing (CIC), IEEE. pp. 181-189. <https://doi.org/10.1109/CIC.2017.00033>.
- Pandey, P., Deepthi, A., Mandal, B. and Puhan, N.B. (2017). Foodnet: Recognizing foods using ensemble of deep networks. *IEEE Signal Processing Letters*. 24(12): 1758-1762. <https://doi.org/10.1109/LSP.2017.2758862>.
- Patil, O. and Gaikwad, V. (2018). Classification of vegetables using Tensorflow. *International Journal for Research in Applied Science and Engineering Technology*. 6(4): 2926-2934. <http://dx.doi.org/10.22214/ijraset.2018.4488>.
- Rafiq, A., Makroo, H.A. and Hazarika, M.K. (2016). Artificial neural network-based image analysis for evaluation of quality attributes of agricultural produce. *Journal of Food Processing and Preservation*. 40(5): 1010-1019. <https://doi.org/10.1111/jfpp.12681>.
- Rodríguez, F.J., García, A., Pardo, P.J., Chávez, F. and Luque-Baena, R.M. (2018). Study and classification of plum varieties using image analysis and deep learning techniques. *Progress in Artificial Intelligence*. 7(2): 119-127. <http://dx.doi.org/10.1007/s13748-017-0137-1>.
- Semary, N.A., Tharwat, A., Elhariri, E. and Hassanien, A.E. (2015). Fruit-based tomato grading system using features fusion and support vector machine, in: *Intelligent Systems' 2014*, Springer, 2015, pp. 401-410. [https://dx.doi.org/10.1007/978-3-319-11310-4\\_35](https://dx.doi.org/10.1007/978-3-319-11310-4_35).
- Shanmugamani, R. (2018). *Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras*. Packt Publishing Ltd.
- Singla, A., Yuan, L. and Ebrahimi, T. (2016). Food/non-food image classification and food categorization using pre-trained googlenet model. In: *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*. pp: 3-11. <http://dx.doi.org/10.1145/2986035.2986039>.
- Sonobe, R., Hirono, Y. and Oi, A. (2020). Non-destructive detection of tea leaf chlorophyll content using hyperspectral reflectance and machine learning algorithms. *Plants*. 9(3): 368. <https://doi.org/10.3390/plants9030368>.
- Soofi, A.A. and Awan, A. (2017). Classification techniques in machine learning: applications and issues. *Journal of Basic and Applied Sciences*. 13: 459-465. <http://dx.doi.org/10.6000/1927-5129.2017.13.76>.
- Tripathi, M.K. and Makedar, D.D. (2020). A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey. *Information Processing in Agriculture*. 7(2): 183-203. <https://doi.org/10.1016/j.inpa.2019.07.003>.
- Tu, S., Xue, Y., Zheng, C., Qi, Y., Wan, H. and Mao, L. (2018). Detection of passion fruits and maturity classification using red-green-blue depth images. *Biosystems Engineering*. 175: 156-167. <https://doi.org/10.1016/j.biosystemseng.2018.09.004>.
- Vaviya, H., Yadav, A., Vishwakarma, V. and Shah, N. (2019). Identification of artificially ripened fruits using machine learning. In: *2nd International Conference on Advances in Science and Technology (ICAST)*. url: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3368903](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3368903).
- Viejo, C.G., Fuentes, S., Torrico, D., Howell, K. and Dunshea, F.R. (2018). Assessment of beer quality based on foamability and chemical composition using computer vision algorithms, near infrared spectroscopy and machine learning algorithms. *Journal of the Science of Food and Agriculture*. 98(2): 618-627. <https://doi.org/10.1002/jsfa.8506>.
- Wan, P., Toudeshki, A., Tan, H. and Ehsani, R. (2018). A methodology for fresh tomato maturity detection using computer vision. *Computers and Electronics in Agriculture*. 146: 43-50. <https://doi.org/10.1016/j.compag.2018.01.011>.
- Yordi, E.G., Koelig, R., Matos, M.J., Mota, Y.C., Uriarte, E., Martinez, A.P., Santana, L. and Molina, E. (2016). Machine-learning models to predict the antioxidant capacity of food. In: *International Conference on Multidisciplinary Sciences (MOL2NET 2016)*, 2<sup>nd</sup> edition. <http://dx.doi.org/10.3390/mol2net-02-03829>.
- Yunus, R., Arif, O., Afzal, H., Amjad, M.F., Abbas, H., *et al.* (2018). A framework to estimate the nutritional value of food in real time using deep learning techniques. *IEEE Access*. 7: 2643-2652. <https://doi.org/10.1109/ACCESS.2018.2879117>.
- Zhu, L., Li, Z., Li, C. Wu, J. and Yue, J. (2018). High performance vegetable classification from images based on alexnet deep learning model. *International Journal of Agricultural and Biological Engineering*. 11(4): 217-223.