Image and Vision Computing 80 (2018) 24–44

Contents lists available at ScienceDirect

Image and Vision Computing

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

A comprehensive review of fruit and vegetable classification techniques☆

Khurram Hameed*,[1], Douglas Chai[2], Alexander Rassau[3]

School of Engineering, Edith Cowan University, 270 Joondalup Drive, Joondalup WA 6027, Perth, Australia

ARTICLE INFO

Article history:
Received 10 August 2018
Accepted 24 September 2018
Available online 28 September 2018

Keywords:
Recognition
Classification
Fruit
Vegetable
Produce classification
Machine learning
Computer vision

ABSTRACT

Recent advancements in computer vision have enabled wide-ranging applications in every field of life. One such application area is fresh produce classification, but the classification of fruit and vegetable has proven to be a complex problem and needs to be further developed. Fruit and vegetable classification presents significant challenges due to interclass similarities and irregular intraclass characteristics. Selection of appropriate data acquisition sensors and feature representation approach is also crucial due to the huge diversity of the field. Fruit and vegetable classification methods have been developed for quality assessment and robotic harvesting but the current state-of-the-art has been developed for limited classes and small datasets. The problem is of a multi-dimensional nature and offers significantly hyperdimensional features, which is one of the major challenges with current machine learning approaches. Substantial research has been conducted for the design and analysis of classifiers for hyperdimensional features which require significant computational power to optimise with such features. In recent years numerous machine learning techniques for example, Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Decision Trees, Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) have been exploited with many different feature description methods for fruit and vegetable classification in many real-life applications. This paper presents a critical comparison of different state-of-the-art computer vision methods proposed by researchers for classifying fruit and vegetable.

1. Introduction

Many real-life applications such as face recognition, autonomous vehicles, object recognition and robotics rely on attempting to mimic the capabilities of the human brain in order to understand images. In the food industry, fruit and vegetable are a major part of fresh produce and their classification is an extension of object recognition. Conventionally, fruit and vegetable are inspected visually by trained personnel for quality assessment as a produce or a crop. However, manual classification poses many human-related constraints for example, an individual needing to be acquainted with the many characteristics of fruit and vegetable. Manual classification requires a continual and consistent aspect recognition technique to maintain consistency. The agriculture industry now applies mechanized methods of classification and often relies upon computer vision for pre and post-harvesting analysis of crops [1]. Computer vision is a field of mathematical analysis of visual data in terms of images of all kinds and this can be a challenging task when applied to the food industry. Visual data of fruit and vegetable expands from binary to hyperspectral images [2-6]. Advances in imaging techniques have resulted in more sophisticated computer vision leading to its use as an emerging standard for many agricultural applications [7]. In the agriculture industry, one of the most important requirements of computer vision is as a non-destructive technique for quality assessment, sorting, automated grading and robotic harvesting unlike many other techniques [8-11]. Classification of fruit and vegetable is a relatively more complex problem due to the huge variety, for example, irregular intraclass shape, colour and texture, and similar interclass shape, colour and texture. These constraints have caused a lack of multi-class automated fruit and vegetable classification systems. An automated fruit and vegetable classification system with more complex information of fruit and vegetable may prove to be helpful for picking the right fruit and vegetable with the right nutrition. It may also help children and visually impaired people, and improve supermarket

☆ This paper has been recommended for acceptance by Yannis Panagakis.

* Corresponding author.
E-mail addresses: k.hameed@ecu.edu.au (K. Hameed), d.chai@ecu.edu.au (D. Chai), a.rassau@ecu.edu.au (A. Rassau).

[1] PhD Scholar ECU Australia.
[2] Senior Lecturer ECU Australia.
[3] Assoc. Prof. ECU Australia.

https://doi.org/10.1016/j.imavis.2018.09.016
0262-8856/

grocery self-checkouts. A summary of recent fruit and vegetable classification performed in different real-life applications is presented in Table 1. Recent state-of-the-art for fruit and vegetable classification and recognition are a combination of feature description and machine learning algorithms on visual data [1,12-15]. Significant research has been reported for representation of different characteristics of fruit and vegetable as feature vectors [6,16]. Despite much research, many challenges need to be overcome to build an effective fruit and vegetable classification system. Thus, this paper provides a comparative survey of associated limitation for classification of fruit and vegetable and the state-of-the-art computer vision techniques used for this task.

The rest of the paper is organised as follows: major challenges for fruit and vegetable classification are described in Section 2. Recent significant efforts for fruit and vegetable classification are discussed in Section 3. Selection of optimal sensors for data acquisition in this task is analysed in Section 4. Considering the complex applications of fruit and vegetable classification essential pre-processing to avoid noise and occlusion due to the environment is discussed in Section 5. After significant pre-processing the data is processed for distinct features extraction and the techniques for this process are discussed in Section 6. A comparison of the state-of-the-art classification techniques using extracted features is presented in Section 7. Finally, a more precise discussion on deficiencies of current techniques and future directions is presented in Section 8.

2. Key challenges

Recognition and classification of fruit and vegetable as a subset of object classification is an inherently more complex task than other subsets of object classification. Fruit and vegetable present crucial sensory and feature characteristics which are also dependent upon the wide spread applications of it. The key challenges involved in fruit and vegetable classification are categorised as:

- *Appropriate sensor* The selection of a sensor for data acquisition is a key challenge for classification. Sensors ranging from black and white (B/W) cameras to non-visual sensors such as acoustic and tactile sensors have been used for classification of fruit and vegetable, but not all sensors are equally suitable for all applications. As evident from Refs. [9-11,17-19,82], both acoustic and tactile sensors are less suitable for non-destructive classification and recognition. These sensors either need physical contact or excitation of the fruit or vegetable for data acquisition. Additionally, visual sensors are highly sensitive to many factors i.e. illumination condition and background environment. These basic factors are a combination of many complex factors including reflection, refraction, scale, rotation and translation, which need to be considered in depth.
- *Feature selection and representation for classification* Features are the physical characteristics of an object that can distinguish it from other objects. Fruit and vegetable have many physical characteristics i.e. colour, texture, shape and size, which can be used as features for effective classification. Fruit and vegetable have numerous inter and intraclass variations and similarities. The interclass variation are major changes i.e. changes in colour, texture and shape whereas the intraclass variations are generally much more subtle and hard to differentiate i.e. different kinds of mangos or apples have only slight variations in features. An ideal selection of features will allow the system to deal with inter and intraclass classification. The computer-based representation of feature is the other dimension of this challenge. Significant research has been reported related to the representation of features. Investigations have indicated that a single feature cannot be considered sufficient for effective classification of fruit and vegetable, or objects in general [2].
- *Machine vision approach* Machine vision approaches are a set of machine learning algorithms used for classification and recognition of images. Extensive research has been performed since the early 1980s. The algorithms designed can be categorised in many ways, a usual categorisation is neural network (NN) based and hand-crafted. The selection of an appropriate algorithm in any machine learning application is always a critical task but it is even more crucial in the case of fruit and vegetable.

Table 1
Identified applications of fruit and vegetable classification.

Industry	Application	Literature
Food industry	Quality assessment	[1, 2, 6, 7, 13, 17-44]
Agriculture	Robotic harvesting	[4, 14, 15, 45-59]
Retail	Supermarkets, Inventory	[12, 16, 60-81]

3. The state-of-the-art

Significant evidence of efforts made toward the realisation of an automated fruit and vegetable classification system are available [60-66], but no examples of commercial applications of such systems are available to date. Approximately, all previous efforts have a core idea of using one or more kind of the sensor along with a machine learning technique for identification of the features associated with the produce items for example, shape, colour, texture and size to perform the classification. Identification of fruit and vegetable has a large number of challenges associated with it due to irregular shape, size and variable colour. Much research has been performed to identify methods to address these challenges. Practically, all physical aspects of fruit and vegetable have been considered as feasible features for effective classification. Initial efforts were made by using global features i.e. shape and colour for classification and local features like texture were analysed in more advanced approaches. Sensors ranging from the modest black and white cameras through to the most advanced hyperspectral camera have been used to capture the features of fruit and vegetable [3]. Both empirical and Neural Network (NN) based approaches of machine learning have been studied and are continually being improved for this task [12,67,68]. Many factors have been identified in the case of real-world systems that impose constraints on achieving high performance in terms of time and accuracy. Variable background environments, illumination inconsistency, specular reflection and recognition inconsistency are key constraints [66].

Significant challenges have been imposed by fruit and vegetable classification, recognition and detection as sub-fields of object recognition. The task of classification of fruit and vegetable has also been advanced by adopting methods in related fields i.e. leaf classification which can be adopted for classification of vegetable with green leaves [7,45]. Most of the efforts made in this regard are a combination of image analysis as feature description and the machine learning algorithms for classification/recognition [4,7,20,46,69,83,84]. These efforts consider a physical characteristic and represent them in a machine vision based representation called feature description. These features are then given as an input to the classification algorithm to converge on a qualitative output. Numerous techniques have been studied for feature description and classification but there is room for significant redesign and improvement to perform effective classification. An effective fruit and vegetable classification system requires a complete rethinking of all related issues of features, sensors and classification algorithms to implement them as a unified system. An example of this rethinking is selection of a robust feature descriptor w.r.t. affine transforms. To present a detailed comparison

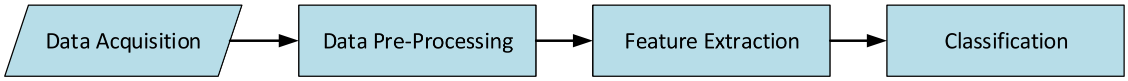

Fig. 1. Process distribution of fruit and vegetable classification into sub-processes.

of the efforts made for fruit and vegetable classification the whole process is divided into sub-processes which are described in Fig. 1. This paper is organised in a sequence of these constituent processes providing a general introduction of the process and then describing the specific variants used in fruit and vegetable classification. A comparable description of the state-of-the-art methods adopted in each of these parts is also presented in their description.

4. Data acquisition

Sampled images, which consist of real-world information are called a dataset and the process of collecting such images in a digital form is called data acquisition. A variety of sensors have been used for this purpose, both passive and active sensors have been exploited for their potential usage. These sensors can be further classified as visual and non-visual sensors. Selection of sensor is highly sensitive to many factor e.g. environment of the application, features sensed, illumination condition, colour camouflage, and occlusion with the environment. Early experiments were performed using B/W cameras as a sensor [3,47]. Light Detection and Ranging (LiDAR) is also used widely for classification of fruit and vegetable in agricultural environments [48]. Significant research has been reported upon the utilisation of Light Structured Sensors (LSS), which exploits the depth data along with colour, shape and texture details [49,69,70]. Classification of fruit and vegetable was initially studied for autonomous harvesting with robots [21]. Numerous research efforts have been reported and are being performed in this direction [46,47,50-52,69]. Colour, thermal, spectral, acoustic, tactile and depth sensors have been used for data acquisition for classification and recognition in the fields of agriculture and food processing. Each sensor has some limitations for example, colour (RGB) images are highly sensitive to the lighting condition and background colour [2,85]. A detailed investigation of literature illustrates that the reflectance properties of objects can be represented by wavelength and hyperspectral cameras can be used for this purpose. This technique has an inherent property of detecting different objects with similar colour or background and is less sensitive to many factors. A recent research has concluded that hyperspectral information combined with other characteristics of fruit can result in an improved performance [22]. This technique has been used in many different classification problems for quality assessment in the food industry [23,71]. Conversely, it is identified that high dimensionality of hyperspectral data is itself a limitation of its use in efficient systems, i.e it requires a large computational power to perform classification with hyperspectral images [23,24,71].

Objects which are above 0 K temperature emits some radiations, which are a function of the emissivity and the surface temperature. This property can also be used for classification of fruit and vegetable. Fruit and vegetable absorb more heat than leaves and background environment, which can be used as a characteristic for classification. However, the classification of a green fruit and leafy vegetable is a challenging task due to approximately similar thermal properties of vegetable and the background [53]. Thermal analysis has recently been employed in many fields i.e. plant disease detection, chilling damage to the fruit in storage, crop maturity estimation and crop yield estimation [54]. This technique is also prone to canopy effect and sensitivity to temperature change [25]. Moreover, no thermal signatures are visible until notable damage has occurred to fruit in some cases [54]. The basic properties of absorption, reflection and refraction of acoustic signals have been used for classification of fruit and vegetable. Acoustic signals have been used for quality assessment of fruit and vegetable by measuring their elasticity as a function of hydration content in their tissues [19,26,27,82]. In the acoustic analysis, a fruit is excited by a physical impact to produce an acoustic wave used to measure the elastic modulus to confirm the firmness and hence freshness. The other method uses an ultrasonic beam targeted on the fruit to measure the co-relation of reference and backscattered beams as a property for classification. However, acoustic analysis is limited to use in fruit and vegetable classification due to its limitations of physical excitation and distortion of the acoustic beam from fruit peel. Acoustic sensors have been used to measure the internal texture of fruit pulp for classification and quality assessment, which depends upon, juiciness, Solid Soluble Content (SSC) and hardness. Both contact and non-contact acoustic sensor are highly sensitive to ambient environmental conditions, which make them less suitable for particular environments i.e. non-destructive, supermarkets and robotic harvesting [17,18,28].

Tactile sensors have been used for measuring fine spatial patterns, roughness and surface friction for classification of fruit and vegetable. These sensors have been used for many intelligent applications i.e. object recognition, robotic grasp, and pose estimation. Tactile sensors have a capability to identify objects which are visually similar but consists of different tactile properties e.g. fruit and vegetable at different levels of maturity [10,86]. Significant results have been reported by analysing a combination of tactile properties and visual properties of objects. More emphasis is evident upon the combination of information from multiple sensors more analogous to the human brain recognition method, which uses a combination of multiple senses for recognition of objects [8]. The state-of-the-art studies have introduced a combination of tactile and visual information as visual-tactile object recognition. However, there are inherent limitations for combining global and local information originated by visual and tactile sensors respectively[87]. A weak paring based approach has been mentioned in Refs. [11,88] for combining inherently different pieces of information. Moreover, tactile sensors are contact based and slow as compared to visual sensors. These limitations make visual-tactile concept less suitable for non-destructive and faster automated classification systems.

The Light Structured Sensors (LSS) has added a new dimension to the machine vision. RGB information combined with the depth information has generated a new set of feature descriptors for classification, segmentation, identification and recognition of objects. The depth is treated as fifth dimension along with colour, shape, size and texture. The RGB data combined with depth (D) data is collectively denoted as RGBD data. There are various applications of RGBD data for example classification, object tracking, surface matching, 3D modelling and pose recognition [89,90]. Numerous commodity sensors are commercially available for sensing the RGDB data [3,51,70] and are being studied. A detailed comparison of sensors in terms of features exploited for fruit and vegetable classification is presented in Table 2.

5. Data pre-processing

The images acquired by visual sensors include some level of noise and distortions. These raw images are generally unsuitable for

Table 2
A comparison of sensors for fruit and vegetable classification.

Sensors	Visual/Non-visual	Sensor type	Features exploited	Advantages	Disadvantages
B/W	Visual	Passive	Geometry and texture	Negligible effect of variable light source	Lack of colour characteristic of object
RGB			Geometry, texture, and colour	Exploits all basic characteristics of object	Highly sensitive to the lighting conditions.
Spectral			Colour and spectral information	Provides more information about reflectance	Computationally expensive for complete spectrum analysis
Thermal			Thermal signatures	Colour invariant	Dependency on minute thermal difference.
RGBD, LSS		Active	RGB image and depth	Complete scene characteristics	Lack of feature descriptors
Acoustic	Non-visual	Both	Elasticity, Cross correlation	Freshness and firmness analysis	Huge distortion at medium boundaries
Tactile		Passive	Roughness, texture, friction, and spatial curves	Non-visual differences works well with same colour and shape	Limited to specific areas and contact to the object, fusion of different informations

extraction of appropriate features for computer vision and image processing applications. To reduce the distortions and noise, a significant pre-processing is essential which is described in this section.

5.1. Pre-processing

RGB matrices capture redundant raw information that needs to be processed statistically to cut out unintended information and determine the missing information due to noise, distortion and variable sensor sensitivity to the same physical input from the environment. The raw images are processed at either holistic or elementary level considering the pixel as the lowest level of abstraction for pre-processing. Spatial and non-spatial constraints apply for pixel estimation where each method have its advantages for example, non-spatial estimation is used for contrast enhancement. Representation of three-dimensional objects in two-dimensional images causes geometric distortion which is subject to the relative position of the camera and the object in the case of still images and the speed, stability and angle of the camera for mobile robotic applications. A group of two or more pixels can be used for geometric pre-processing. Significant elementary pre-processing is applied on adjacent pixels to enhance the differences among them common examples of which are image smoothing and gradient used for edge detection. Many signal processing filters have been designed for this purpose. A set of filters at sub-holistic level is applied as convolution for estimation of missing information. The constraint of prior knowledge describes the convolution as statistical or stochastic function. A detailed description of pre-processing techniques with their applications and constraints has been presented in Table 3.

5.2. Segmentation

To extract the distinct section of an image as a Region of Interest (ROI) image segmentation is performed. Image segmentation is a crucial challenge in computer vision systems that determines the overall effectiveness of higher level image analysis [91]. Many segmentation techniques based on brightness, colour, grey scale values, texture and edges have been reported in the literature. However, as the computational capabilities are improving more effective segmentation techniques are evolving [7,92]. A preliminary segmentation can be achieved by detecting the edges and subtracting the unwanted objects or background from the image. Pixel intensity and direction have been used widely for eliminating the local discontinuities at each pixel of a filtered image [93]. Lower and upper thresholds selection to find a discontinuity is crucial for extraction of edge pixels in complex images and different edge detection techniques have a tendency to detect a false edge in pre-processed images. Hence, edge-based segmentation less suitable for images with similar background, occlusion and mixed edges [91]. Pixel level threshold for generating regions in the images has been used for *threshold based segmentation*. Most of the grey scale techniques have been altered for RGB images by applying threshold on three channels separately. Estimation of the threshold is again crucial where many methods use hit and trial for this purpose, but computer vision tasks require a fully automated threshold value convergence for segmentation. An *adaptive threshold* selection based segmentation has been presented in Ref. [94]. A mean grey scale value has been used for finding the optimal threshold with the iterative convergence of mean value. Intraclass variance has been converged and used as threshold in the *Otsu method*, an extension of this method to the RGB images has been presented in Ref. [95]. Thresholding is among the most significantly used techniques for both binary and multi-segmentation in complex images [96]. *Colour histograms* have been used for multilevel segmentation in RGB images while using the Otsu method as an objective function to be maximised for effective segmentation. Meta-heuristic and swarm algorithms have been used for optimal intraclass variance convergence [96,97]. Entropy as RGB histogram function has been used for multilevel RGB segmentation kapur and

Table 3
A description of image pre-processing techniques.

Technique	Description	Applications	Constraints
Intensity estimation	Missing pixel value estimation by spatial and non-spatial analysis	Noisy pixel value determination in grey scale and RGB images	Prior knowledge, likelihood of non-uniform object illumination
Geometric estimation	Estimation of geometric distortion by relative motion, angle, speed and 2D to 3D representation	Determination of geometric details in mobile robotic and remote sensing applications	Knowledge of angle, position and relative speed for sensor and object
Elementary processing	Processing group of neighbouring pixels by signal processing filters	Smoothing and gradient analysis for better edge detection	Complex and non-linear signal processing filters
Holistic processing	Set of filters applied as convolution for image restoration	Determining the holistic image characteristics	Complex stochastic analysis and priori knowledge

minimum cross entropy minimisation has been widely studied for optimal threshold estimation [98], where better segmentation is reported for higher dimensional RGB histograms.

Pixel intensity and spatial connectivity have been used as similarity measures for grouping pixels in *region based methods*. Substantial similarity criteria have been used for instance, pixel intensity differential, running mean and standard deviation among multiple neighbourhoods of candidate pixels. Larger neighbourhood can use colour, texture and spatial information for more complex criteria [99]. This method is effective for images with small numbers of regions, however, more computational power can work with multi-segment images. Comparable performance of significant variants in terms of time and computational requirements has been reported in Ref. [100]. Pixels with similar features are clustered to form *feature based segments*. Both hard and soft clustering are evident from the literature. Fuzzy c-means is among the most widely used soft clustering techniques i.e. a pixel is associated to multiple clusters based on connectivity weight estimation [91]. Variants of fuzzy c-mean with improved performance use spatial information of pixels for weight estimation however, significant performance constraints have been reported while working with the noisy data [101]. The parallel nature of *Neural Networks (NN)* has been widely used for image segmentation [102,103]. A common example of NN based segmentation is the used of spatial information with Self-Organising Maps (SOM) [104]. An inherent limitation of this method is the unavailability of prior information of the number of clusters. SOM has been used to find the optimal number of clusters to perform the segmentation automatically. Significant fuzzy c-mean based variants of NN segmentation are also evident from the literature [105]. The concept of *multi-feature fusion* as a combination of rotation-invariants Local Binary Patterns (LBP), RGB histogram distribution, weighted histograms, region connection statistics and multi-label k-nearest neighbour fusion has been analysed with the existing techniques of automated annotation in Ref. [106]. This concept has been used for segmentation of images using Histogram of Oriented Gradients (HOG) and LBP as feature fusion on RGB and polarised images separately, and improved segmentation results has been presented in [107]. This concept can be used with other significant classifiers for better segmentation.

When considering fruit and vegetable classification, a feature based segmentation has been applied on a pre-segmented Region of Interest (ROI) of apple images for defect detection in Ref. [30]. An experimental setup with intentional lower background intensity is exploited with a low pass filter to find the ROI along with the morphological filling to reduce the effect of false russet removal in artificially defected fruit. Average and standard deviation of intensity has been used to define a global feature on ROI with variable neighbourhood size. A set of supervised and unsupervised classifiers has been applied and significant segmentation effectiveness is presented with super-pixel and supervised classifiers w.r.t. unsupervised classifiers. It is concluded that more accurate results can be achieved for larger neighbourhoods at the cost of computation time. Texture as feature HSI and Colour Co-occurrence(CCM) is used for segmentation based quality assessment of citrus. The texture of citrus leaves with greasy spots, melanose, and scabs is analysed where more effective results are reported for reduced dependency on intensity in texture features [31]. Distance Transform (DT) based watershed segmentation is used with statistical features in RGB images in Ref. [77]. Euclidean, city-block and cross-board based DTs are used for segmentation of fruit and vegetable in binary images with significant effectiveness. A Gabor kernel based global segmentation with eight different orientation of Gabor wavelet is used with Principal Component Analysis (PCA) for automated classification of apple fruit. It is concluded that with Gabor based global segmentation of near infrared (NIR) apple images there is no need of local feature segmentation. The Gabor filter used can extract specific frequency components that can be used for segmentation [32,108]. Recently, Otsu based segmentation has been used for fruit and vegetable defect detection and a common limitation of holes generation for similar intensity level as background has been identified [39,40,57,81]. A combination of LBP, HOG, global colour and shape feature has been used with Otsu thresholding for optimal ROI selection in a multi-class fruit recognition and identified to be improved for effective results [7]. Damage detection in papaya has been performed by k-means clustering after contrast enhancement of colour images where classification has been performed with SVM, decision tree and Naive Bayes with a maximum accuracy of 90.5%. The study reported that the experiments were not performed on a uniform dataset and the result are not comparable with the state-of-the-art. To detect the green apple, a graph based manifold saliency was used with k-mean and Fuzzy-C-Mean (FCM) clustering, where the study reported on imperfect segmentation that needs to be integrated via an area loss function [14]. A more related research has been presented for quality evaluation of packed lettuce, where a patch based segmentation has been performed with CNN. The CNN has been trained with both packed and unpacked lettuce datasets and a 3×3 sliding window is used to estimate the likelihood of each patch of 3×3. The estimated likelihood is then used for threshold-based segmentation with significantly high values of the threshold. A comparison of packed and unpacked lettuce segmentation accuracy has been reported as 83% and 86% respectively [109]. A description of recent segmentation techniques used in various applications of the food industry has been presented in Table 4.

6. Feature extraction

A piece of information related to some particular dynamic property of object in a digital image with higher level of perspective i.e. recognition, classification, retrieval and reconstruction is called a feature descriptor. Fruit and vegetable have several distinct visual characteristics associated with them called features. The most commonly used features for classification and recognition of fruit and vegetable are colour, shape, size and texture. A feature descriptor is either global or partial depending upon their comprehensive or partial representation ability. In particular to the object recognition, a global feature describes the object as a whole in the form of a generalised descriptor for example shape, and a local feature describes many interest points in the form of patches of an image. The interest points are not consistent and can vary from sample to sample in a recognition task [112]. Moreover, usual practices include a combination of local and global features for superior classification effectiveness [85,113,114]. Availability of whole object details is another inherent limitation in image acquisition due to the poor acquisition, noise, partial information, and data loss during conversion (e.g. RGB to grey scale). These limitations pose some constraints on the performance of feature descriptors. Properties of features descriptors for significant representations of features are described in Ref. [115]. A global to global and partial to global recognition based categorisation of feature descriptors is described in Table 5. A non-exhaustive description of shape, texture and colour feature descriptors has been described in this section.

6.1. Shape feature descriptors

The shape of fruit and vegetable has been frequently used for classification. In the food industry, shape and size (morphology) of fruit and vegetable play a critical role in price estimation. This feature is also significant for automatic sorting in the food industry. Spherical or quasi-spherical shapes are easier to describe as feature vectors as compared to natural and more complex shapes of fruit and vegetable. The shape feature vector can be used for quantifying the fruit and vegetable for example, estimating size by projection area, perimeter,

Table 4
A description of segmentation techniques used for fruit and vegetable analysis in the food industry.

Year	Fruit/Veg	Application	Segmentation technique	Ref.
1996	Mixed	Classification	Threshold-based pixel level image subtraction	[66]
2006	Apple	Quality assessment	Feature-based with variable neighbourhood size	[30]
2006	Citrus	Quality assessment	Texture based HSI and Colour Co-occurrence (CCM)	[31]
2007	Apple	Quality assessment	Gabor kernel and PCS avoided local features segmentation	[32,108]
2012	Mixed fruit	Fruit harvesting	Spatial-local adaptive threshold based	[52]
2012	Mixed	Classification	Distance Transform (DT) and watershed	[77]
2013	Vege	Detection	Texture and edge fusion segmentation	[78]
2015	Mixed fruit	Detection	K-mean split and graph-based merge with area threshold	[110]
2016	Apple	Recognition	Dynamic threshold Otsu method	[111]
2016	Mixed fruit	Classification	Square window split and merge segmentation	[72]
2016	Tomato	Quality assessment	Otsu method	[39]
2017	Apple	Bruise detection	HSI based Otsu method	[40]
2017	Eggplant	Grading	Intensity adaptive threshold based Otsu	[57]
2018	Apple	Detection	Graph based k-mean FCM clustering	[14]
2018	Litchi	Robotic harvesting	One dimensional random signal histogram with FCM	[59]
2018	Mixed fruit	Detection	Fusion of LBP, HOG, global colour and shape with Otsu	[7]
2018	Packed food	Quality assessment	3×3 patch likelihood threshold with CNN	[109]
2018	Papaya	Disease detection	K-mean clustering based segmentation	[44]
2018	Pomegranate	Clustering	Threshold Otsu	[81]

Table 5
Properties of feature descriptor.

Property	Global/Local	Description
Description strength		Ability of differentiation among similar and dissimilar characteristics of an image.
Robustness	Global information to the global recognition	Resistant to distortion, noise and small changes during storage and conversion
Resistance		Resistant to affine, projective and colour space transformations
Conciseness and indexing		Ability to reduce the memory size and searching complexity.
Partial matching	Ability of partial to global, recognition among above-mentioned properties.	Ability to recognise and retrieve from partial information.

length, width, major, and minor diagonal for size estimation in the food industry. A shape feature descriptor is a mathematical model that tries to model the shape of an object in a human intuition based method for example shape described as a set of contours. A preliminary technique of a shape descriptor considers the important interest points based on the boundary and the interior of the shapes, various categories of shape interest points are spectral features, curvatures, shape contents, shape matrix, moments and shape signatures [116].

One of the most intuitive categorisations of shape feature descriptors is contour based and region based considering the inherent geometry of shapes. This categorisation is dependent upon whether the shape feature vector is extracted by boundary only or from both boundary and interior as well. A more elementary form of categorisation can be spatial and transform domain, where use of a particular kind of descriptor is dependent upon the application. Representing shape in one or other form can guarantee performance improvement for example, shape description data in the spatial domain can be better handled in the transform domain for lossless conversion and compression [117,118]. The basic geometrical parameters such as curvatures, corners, regions, centre of gravity, convexity, circularity ratio, and Eccentricity associated with the shape of fruit or vegetable can only differentiate the shapes with large differences however, a combination of them can comprehend more fine details. Basic definitions of essential geometrical parameters for shape description are described in Table 6.

Chain codes is a complex mathematical model of basic geometric parameters for describing any geometry in a standardized way. Line segments of a shape geometry are described as a chain of orientation in terms of connectivity [119]. However, the chain codes are prone to noise and deformations [120]. A histogram of surrounding details of an identified key point at an object boundary is maintained in *shape context*, where a combination of all histograms describes the shape of an object as depicted in Fig. 2. However, interest points may vary from sample to sample in a class and need to be fixed manually. Also, the histogram based representation has the capability of representing any spatial information however prone it is to noise and distortion [121]. A point distribution in a shape is represented by *moments based descriptors*. This statistical method of shape representation requires less computational power and shows significant

Table 6
Essential geometric parameters for shape descriptors.

Definition	Geometric parameter
Centre of gravity	$g = \left(\frac{1}{n}\sum_{i=1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n} y_i\right)$
Radial distance	$\rho_i = \parallel p_i - g \parallel_2$
Average bending energy	$E_b = \frac{1}{n}\sum_{s=0}^{n-1} k(S)^2$
Circularity area ratio	$\zeta_A = \frac{A_{shape}}{A_{circle}} = \frac{4\pi A_{shape}}{P_{sphere}}$
Circularity perimeter	$\zeta_P = \frac{A_{shape}}{P^2}$
Circle variance	$\zeta_\rho = \frac{\phi_p}{\mu_p}$
Rectangularity	$\zeta_R = \frac{Ashape}{A_{box}}$
Convexity	$\zeta_C = \frac{P_{hull}}{P_{shape}}$
Solidity	$\zeta_S = \frac{A_{shape}}{A_{hull}}$
Hole area ratio	$\zeta_H = \frac{A_{hole}}{A_{shape}}$
Eccentricity	$\zeta_\epsilon = \frac{\lambda_1}{\lambda_2}$
Ellipse variance	$d = \sqrt{\rho_i^T M^{-1} \rho_i}$
Profile	$\phi_x(i) = \forall_{x=i} : y_{\max} - y_{\min}$

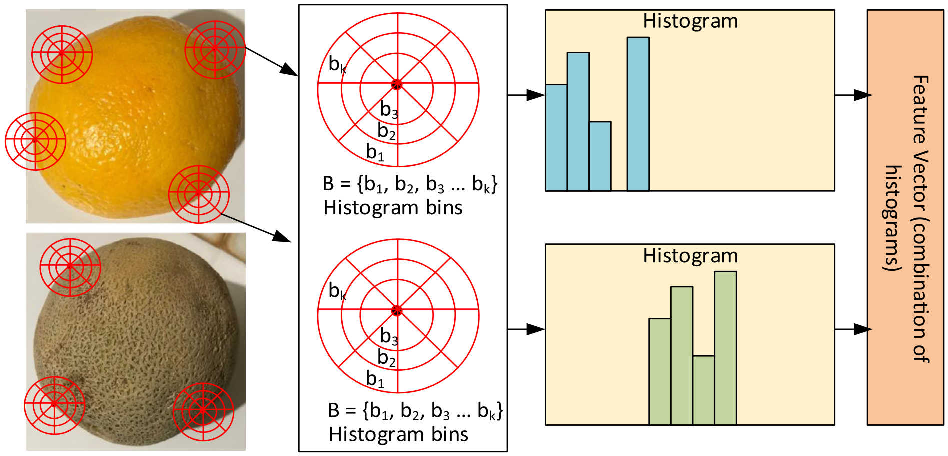

Fig. 2. Shape context feature description vector.

robustness against noise and data redundancy [116] however, it is less efficient for classification of approximately similar shapes due to loss of redundant information in statistical computation [122,116]. Fragmented and simplified details of a shape such as changes in curvatures are called *scale space methods*. This method can work well with the small translation, scale and rotation but is not robust for noisy data an analysis w.rt. to rotation and noise is described in Ref. [120]. Numerous variants of this method are evident from literature [123-126]. *Spatial partitioning* uses local properties to represents the shape globally common examples of local properties are principal axis and axis of least inertia [121,127,128]. A detailed categorisation of shape mathematical models in Ref. [116] is depicted in Table 7. A more detail on various shape representation methods can be found in Refs. [120,129,130], where most of the methods are based on the low dimension geometric parameter. A more recent direction in shape description is use of Bag-of-Curvature (BoC) and Bag-of-Shape-Vocabulary (BoSV) [131] as a variant of Bag-of-Words(BoW) [132]. Different features have been tested for describing the shape vocabulary for example region based visual vocabulary is defined in Ref. [133] based on different local shape primitives. A detailed discussion on shape matching with local shape primitives is presented in Ref. [121,134]. Currently, Convolutional Neural Networks (CNN) are also being used for shape feature representation.

The lower layers of Neural Networks (NN) have been investigated for edge detection due to their capability of learning convolutional kernels. Deeper and complex edge relations can be identified by the deeper layer in the CNN [135]. Considering this ability of CNN, the tedious task of feature descriptor crafting can be performed by the CNN. Generation of effective shape features is however limited to the availability of huge amount of data to train the CNN. Feature descriptor extraction is although an attractive idea but development of such convolutional layers is also a complex task. Approximately, all of the above-described feature descriptors have been used for classification of fruit and vegetable. An energy function minimisation has been used with a model based image interpretation using ACM algorithm to classify the defected apple. It has been identified that using an active contour model (sake) as energy minimisation parameter every component of the shape contour will take approximately $n - 1$ iteration which makes it less feasible for complex images however, significant performance can be achieved in simple images [35]. To reduce the dependency on segmentation Edge Orientation Autocorrelogram (EOAC) has been used in Ref. [136] for produce classification. EOAC can estimate the edges orientation and the spatial correlation among the pixels which are used with a combination of classifiers while an accuracy of 99% has been reported. An erosion based shape representation is used for representing the shape of leafy vegetable and fruit for grading [36]. To detect the immature peach in the orchard a window based scanning of grey scale images has been used in Ref. [2], where the window size was pre-defined and is dataset dependent. The circular disk radius is then estimated by dimensions of the fitted window, which make the complete study highly dependent upon the dataset considered. A Feed-forward Neural Network (FNN) is used in Ref. [68] for classification of fruit and vegetable where the shape is represented as a convex hull covering the complete fruit using graham scan method. Shape combined with other features has been used and an accuracy of 89.1% is reported while using FNN with genetic algorithm (GA). A global shape representation has been used for grading the mangoes in Ref. [42]. Initially, the centroid of mango is estimated using first-order geometric moments (green theorem) and all boundary pixels are then identified with a provision of making this system applicable in a real-world application. A Fourier transform is then used to convert a mango image to the feature vector using lower harmonics. However, lower harmonics are usually distinct for spherical and quasi-spherical shape but can significantly distort the result for complex shapes of fruit and vegetable. A machine vision based fruit counting systems has been designed in Ref. [137] where the mango shape is identified based on the colour and smoothness of pixels while using blob connecting algorithm for mango shape segmentation. The shape of green apple has been represented as perimeter, area and centrifugation on texture-based segmented image of the sample. A maximum area threshold base domain connection has been used for marking the multiple object areas in the image [138]. A heuristic modelling based arc grouping is used to model elliptical mango shape. More recently a shape based tomato maturity system is introduced in Ref. [13]. An experimental setup is carefully designed to capture a single tomato at a time with a dark background. The tomato image is initially centroid and a minimum distance base contour is drawn to describe the tomato shape. The tomato shape is then measured to estimate the maturity level while a performance of approximately 100% is reported. The average of red region of the strawberry fruit is used to find out the main diagonal of the fruit region used to describe the strawberry shape as kite geometry in Ref. [80]. Four boundary points on fruit region are considered to make two sets of equal-length sides selected in a way to make an inscribed rhombus. The size of the rhombus is used to estimate the ripeness of the fruit. A comparison of recent shape based fruit and vegetable analysis is presented in Table 8.

6.2. Texture feature descriptors

Digital images always contain some texture in them, examples of which ranges from spatial patterns in satellite images to arrangement of tissues in microscopic images. The texture is one of the most commonly used properties of fruit and vegetable among colour and shape for classification. Texture is the spatial arrangement of primitives called textons which are fundamental structures at the microscopic level that is pixels in images and the atoms in the human visual perception system. Texture in digital images follows some statistical property of periodic recursion with some degree of variance. This variance can range from statistical to stochastic functions. Texture as a property for classification, recognition, segmentation, synthesis and shape analysis from texture has been studied widely [9]. Significant applications of texture analysis include medical image analysis [140], analysis of satellite images [141], segmentation, content-based image retrieval [142], face recognition [143], object recognition [144], image compression, robotic vision and unmanned aerial vehicles [145], a more broader categorisation is presented in Ref. [9]. Texture description is the core of the texture analysis for any of its application. Much research has been reported in this field while texture representation methods have been divided into five broad categories i.e. statistical, geometrical, structural, model-based, filter based and feature descriptors [9]. The progress in the field of texture analysis is evident from a study of human visual perception according to which the most complex texture can be modelled as an arbitrary order statistics [146]. Most of the early work in the texture

Table 7
Categorisation of mathematical models for shape representation.

Models	Description	Methods	Sub-methods
One dimensional	A perceptual feature of shape derived from boundaries	Complex coordinate Centroid distance Tangent angle Contour curvature Area function Triangle area Chord length	
Polynomial approximation	Neglect discrete pixelisation, by considering the whole shape	Polynomial merging	Distance threshold Tunnelling Polygon evolution
		Polynomial splitting	
Multivariate interpolation/Spatial interpolation	Considering relative orientation i.e. length, curvature and exploiting boundary relation for shape representation	Adaptive grid Bounding box Convex Hull	
		Chain Code	Basic chain code Differential chain code Re-sampling chain Vertex chain Chain code histogram
		Smooth curvature decomposing ALI Method Beam Angle	
		Shape Matrix	Square model shape Polar model shape
		Shape context Chord distribution Shock graph	
Weighted averages (Moments)	The weighted average of pixels, boundaries and function of moments	Boundary moment	
		Region moment	Invariant moment Algebraic moment Zernike moment Radial moment Homocentric moment Orthogonal fourier moment Pseudo-Zernike moment
Scale-space representation	Shape representation as simplified curvatures	Curvature Intersection point map	
Shape transforms	Representation by transform orthogonal or non-orthogonal constituent function	Fourier descriptor	One dimensional fourier Region based fourier
		Wavelet transforms Angular radial transform Shape signature R - Transform Shapelets	

Table 8
Comparison of shape features for fruit and vegetable analysis.

Year	Fruit/Veg	Feature description	Accuracy	Ref.
2011	Apple	Model based ACM shape	91.00%	[35]
2012	Mixed	Edge Orientation Autocorrelogram (EOAC)	98.80%	[136]
2013	Mixed	Erosion morphology based shape	–	[36]
2014	Apple	Fourier descriptor of shape size and Euler number	88.33%	[108]
2014	Peach	circular disk radius estimation for shape	85.00%	[2]
2014	Tomato	Graham scan based convex hull	89.10%	[68]
2014	Tomato	Texture-based blob size h and w ratio	–	[4]
2016	Cucumber	Ellipse fitted contour and ellipsoid mask	100.00%	[23]
2016	Mango	Global shape by centroid and boundary	87.80%	[42]
2016	Tomato	Ratio of equitorial and polar diameter	–	[139]
2017	Green Apple	Perimeter, roundness, centrifugation based shape	90.08%	[138]
2017	Mango	Intensity based blob connection	$R^2 = 0.91$	[137]
2018	Strawberry	Kite geometry based shape	90.00%	[80]
2018	Tomato	Centroid based circular contour estimation	100.00%	[13]

feature description is based on this concept examples of which is Grey Level Co-occurrence Matrix (GLCM) [146,147]. Despite the significant research in this direction, approximately majority of feature descriptors are less feasible for daily life applications in terms of computational requirements and complexity to be implemented as computer vision application. Based on these limitations, the texture descriptors are divided into two categories [148] i.e. high-quality based and high-efficiency based described in Table 9 with identified solutions to the complexities involved. An illustration of complexities of texture in the food images at different illumination, scale and viewpoint conditions is depicted in Fig. 3. The improvements in texture descriptors described in Fig. 4 can be divided in miles stones in a progressive way as filter-based, statistical, Bag-of-Textons (BOT), invariants and Convolutional Neural Networks (CNN) based descriptors.

A bank of filters is used for image convolution to extract the major frequency components in *filter based methods* [149]. Common examples of this method are Gabor filters [150], Gabor wavelet [151], Linear filters [152], and pyramidal wavelets [153]. However, texture cannot be described always in a deterministic way. *Statistical methods* describe the texture as a non-deterministic relationship distribution among the pixels [154]. Examples of the statistical method are Markov Random Field (MRF) and fractal methods. Renaissance of texture as textons is called *BOT*, which is a new dimension in texture representation [155]. A comprehensive mathematical model of textons is described in Refs. [156,157] and a detailed description of operation involved in BOT are described in Fig. 5. Moreover, significant techniques used for each subsequent operation of BOT are listed in Table 10. Although, BOT has shown a significant progress in the semantic representation of texture, it is significantly sensitive to rotation and scale variation; an analysis is presented in Ref. [158]. To reduce the sensitivity of texture descriptors on scale, viewpoint and illumination *scale invariant features* were introduced. Scale Invariant Feature Transforms (SIFT) and LBP are groundbreaking examples of this era. Recently, more deep convolution has been performed with the help of CNN to extract more complex spatial relation among the pixels CNN has shown significant performance in object recognition and texture analysis [159-161]. A key to success and excellent survey on CNN based texture representation is presented in Ref. [162].

Considering the case of texture based fruit and vegetable classification, significant results have been reported. Exploiting the capability of filter-based methods of low computational cost and spatial representation in transform domain a Gabor filter based PCA kernel has been proposed in Ref. [32] for apple quality grading. In this study, the segmentation part has been eliminated by taking advantage of extracting specific frequency components for texture representation while a classification rate of 90.60% is achieved. Scale invariant property of fractal has been used for quantifying the food skin morphological changes as an effect of storage damage and cooking [175]. An average intensity difference has been used for forming fractal image and average Fourier spectrum horizontal and vertical power has been used for frequency domain analysis. It has been identified that fractal changes also correlates to the visual changes. A Spatial Gray-level Dependence Matrix (SGDM) based statistical analysis is used to find 13 statistical features defined to describe the texture of grapefruit peel. The classification has been performed by clustering samples on generalised square distance, where an accuracy of 98% has been reported in Ref. [34]. A co-occurrence matrix based texture has been defined on grey level for 15 classes of fruit and vegetable and eight statistical features has been used for describing the features in Ref. [74]. It is assumed that the same statistical properties will exist due to the iterative nature of texture in fruit and vegetable peel. An accuracy of 89% has been achieved while using texture as a feature. Correspondingly, texture has been represented as Local Activity Spectrum (LAS) in horizontal, vertical and diagonal directions for fruit classification in Ref. [136]. The LAS has been quantised to make a histogram based feature vector of 256 bins, where this method as reported an accuracy of 99%.

In the current state-of-the-art methods of texture representation, a Local Relative Phase Binary Pattern (LRPBP) has been used in Ref. [176], where the texture is used with the LRPBP and an approximated accuracy of 96% has been reported. To reduce the burden of computational power on single board computers or embedded systems, a colour and texture based fruit classification approach has been presented in Ref. [177]. A grey level co-occurrence matrix (GLCM) has been used for texture representation where an accuracy of 83% has been achieved however, no specific details have been indicated explicitly to lower computational cost except running the proposed method on a Field Programmable Gate Array (FPGA). More recently, GLCM has been used with statistical features for classification of diseased papaya fruit [44]. In this research, statistical feature descriptors have been assumed for better discriminatory power for defect detection. Five GLCM features have been used for texture description

Table 9
Categorisation of texture descriptors based on computational constraints for optimal texture representation.

Computational constraints	Properties of descriptor	Complexity involved	Identified solutions
High-quality descriptors	Dealing with significant Intraclass texture irregularities and interclass similarity	Rotation, Variable viewpoint, Variable illumination, Noise	Development of large training datasets for better learning
High-efficiency descriptors	Hyperdimensional texture representation on resource limited hardware i.e. embedded systems	Complex and high dimensional representation of texture	Development of compact and less complex feature descriptor

Fig. 3. Illustration of complexities in the texture of food images (a) scale variation in orange peel (b) scale and viewpoint variation in brown bread (c) scale variation in cracker (d) illumination variation in candy fruit. Images by RawFooT and KTHTIPS texture datasets.

to achieve an accuracy of 90.5% which can be considered as promising. Similarly, a Texture Homogeneity Measuring Technique (THMT) has developed for classification of olives. A homogeneity is measured by considering an adaptive threshold based on defect area pixel intensity variance. Significant accuracy rate has been presented but it can be identified that the approach has a high class dependency [43]. An ROI based multi-feature fusion has been performed by fusing HOG, LBP and GaborLBP for texture representation in Ref. [7]. SVM is then used for classification among the multiple classes of fruit, it is identified that the optimal region selection has improved the overall results by a significant factor. A non-exhaustive comparison of recent research studies has been presented in Table 11.

6.3. Colour feature descriptors

Colour is an important cue for selection or rejection of fruit and vegetable for customers in supermarket or quality assessment personnel [180,181]. The colour is most frequently used feature for image retrieval and recognition. Colour has significant advantages over other features like high frequency ease of extraction, invariant to size, shape and orientation and independent to background complication. Colours are represented in different colour spaces which are designed for specific purpose. A commonly known colour space is RGB, which represents the image in red, green and blue planes. An image generated by the same pixels in an RGB space can have different RGB values for different devices which need to be transformed for standardization. This non-linear nature of RGB makes it less suitable for human visual inspection. To overcome this limitation of RGB space significant other colour spaces have been developed a detailed description of different colour space and their comparative analysis is presented in Ref. [91]. In general, machine learning based colour representation is used for classification of objects from images or videos [182]. Colour of a fruit or vegetable is governed by physical, biochemical and microbial changes during ripening and growth. However, the photometric changes i.e. orientation, scale and illumination can cause a significant effect on the colour of fruit

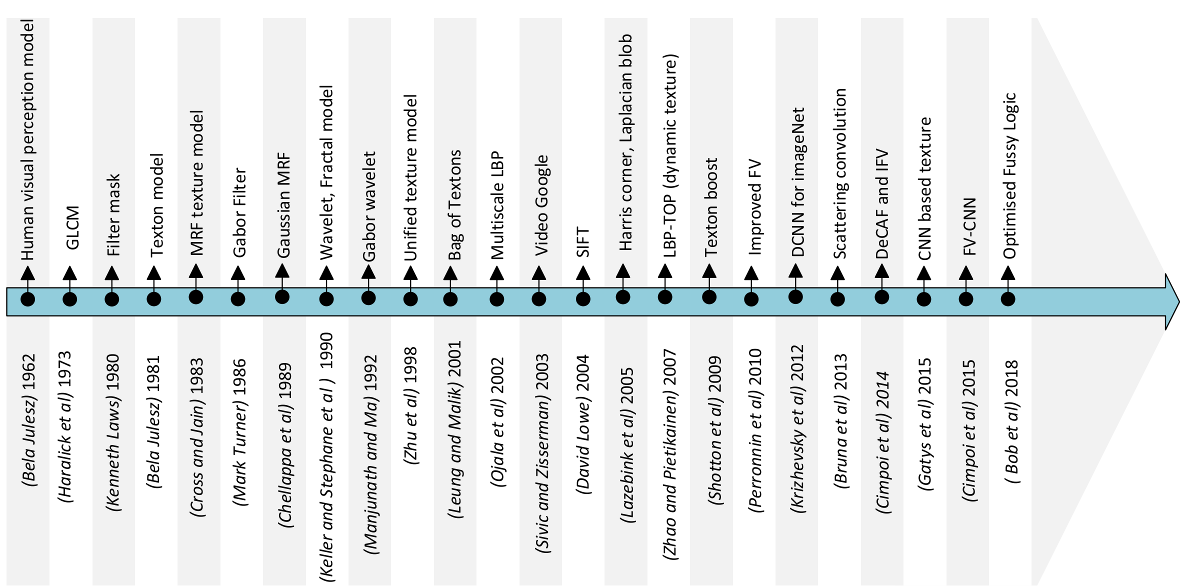

Fig. 4. A time-line of texture representation methods.

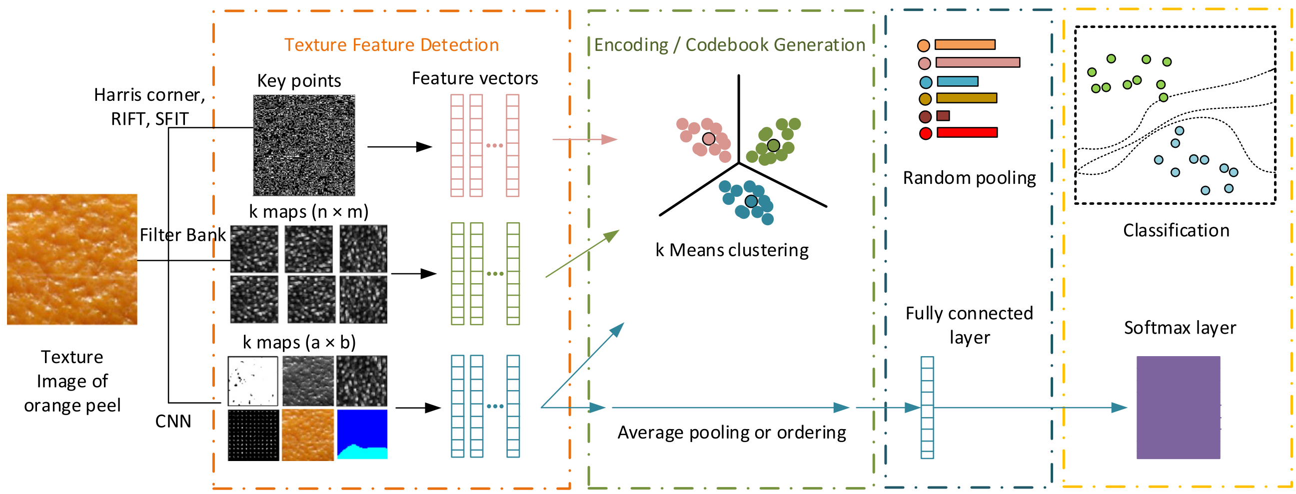

Fig. 5. Generic representation of BOT variants [148].

as illustrated in Fig. 6. To reduce the photometric effects a colour descriptor must has a significant invariance property [5]. A diagonal model-based representation and effect of photometric changes on a digital image is studied in Ref. [183]. Based on the diagonal model five different invariance properties of colour feature descriptor are presented in Table 12.

Significant colour descriptors have been used for fruit and vegetable colour representation. Colour of fruit and vegetable can be described as a whole or in terms of regions of homogeneous colour i.e. globally or locally. Histograms, moment invariants, SIFT and coherence vectors have been significantly used for the description of colours in fruit and vegetable classification. Histogram of each colour channel is combined to make an *RGB histogram*. However, an RGB image consists of numerous RGB levels which need to be normalised after quantisation in particular histogram level [184]. Moreover, pixel level histograms are invariant w.r.t. photometric changes. Also, histograms do not contain semantic information of the image. Considering the image as a function of RGB triplets, *colour moments* are decried in Refs. [185,186]. *Scale Invariant Feature Transform* (SIFT) [187] is invariant to common photometric changes as the gradients of an image are invariant to photometric changes [188]. Significant variants of SIFT are HSV SIFT, HUE SIFT, opponent SIFT, C-SIFT, *rg*SIFT and RGB-SIFT [189-191]. *Colour Coherence Vector* (CCV) describes holistic colour distribution with spatial pixel relevance by dividing the image into connected components. Much research has been reported using CCV as scene recognition and object recognition with variable viewpoints [192,193]. A summary of invariance properties of discussed colour feature descriptors w.r.t. to the diagonal model is presented in Table 13.

An HSI based colour histogram representation in Ref. [66] is among the most initial efforts for classification of produce, where one-dimensional histogram of H, S and I channels are fused to represent the colour. Similarly, an HSI based CCV has been used for disease detection in citrus peel in Ref. [34]. An RGB based histogram

Table 10
The state-of-the-art techniques of BOT as represented in Fig. 5.

Steps	List of approaches	The state-of-the-art
Texture feature descriptor	Sparse methods Fractal methods Dense methods	Harris Laplacian (RIFT, SIFT and SPIN) [163, 164] Multi-Fractal Spectrum [165] Gabor wavelet LM filters [155] Schmid Filters Maximum response (8 filters) [166] Local Binary Pattern (LBP) Basic Image Features (BIF)[167] Weber Local Descriptor (WLD) [167]
Codebook generation	Predefined method [167] k-means clustering [155] Gaussian Mixture Model (GMM) [168] Spare code learning [169]	
Encoding	Voting based methods	Hard voting [155] Soft voting [170]
	Reconstruction based methods	Sparse coding [171] Local constraint Linear Coding (LCC) [172]
	Fisher Vector (FV) based	Fisher Vector (FV) [173] Improved Fisher Vector (IFV) [168] Vector of Locally Aggregated Descriptor (VLAD) [174]
Feature pooling	Average Pooling Max Pooling Spatial Pyramid Pooling (SPP)	
Classifier	Nearest Neighbour Classifier (NNC) [166] Kernel Support Vector Machine (Kernel- SVM) [164] Linear Support Vector Machine (Linear-SVM) [160]	

Table 11
Comparison of recent texture feature description methods for fruit and vegetable analysis.

Year	Fruit/Veg	Application	Feature vector	Accuracy	Ref.
2009	Grapefruit	Quality assessment	SGDM based 13 statistical	98.30%	[34]
2010	Mixed fruit	Recognition	Five statistical features	94.00%	[74]
2012	Pomegranates	Quality assessment	Statistical features	98.80%	[178]
2012	Vegetable	Classification	kurtosis and skewness	95.00%	[77]
2013	Mixed fruit	Quality assessment	Curvelet-based statistical feature	91.42%	[179]
2015	Mixed fruit	Classification	Wavelet Entropy	89.50%	[29]
2016	Mixed fruit	Classification	Local relative phase binary patterns (LRPBP)	95.83%	[176]
2017	Apple	Recognition	Grey scale difference with statistical feature	98.08%	[138]
2017	Grapevine bud	Detection	SFIT with BOF	96.50%	[56]
2017	Mango	Segmentation	Dense SIFT-based histogram visual word	88.00%	[137]
2017	Mixed fruit	Recognition	Mean, Standard Deviation, Skewness and Kurtosis	83.30%	[177]
2018	Mixed fruit	Detection	Fused HOG, LBP, and GaborLBP	98.50%	[7]
2018	Olive fruit	Quality assessment	THMT based threshold comparison	100.00%	[43]

of banana is feed into a neural network for fruit quality assessment where the proposed method is evaluated manually by measuring the quality of banana with classification [194]. A fusion of CCV, Global Colour Histogram(GCH) and unser's descriptor has been compared with SVM and LDA for classification of fruit in Ref. [73]. It is also reported that the colour based approach has out-performed as compared to more complex appearance-based approaches. Five statistical properties of each colour channel in HSV space has been extracted for detection of citrus fruit in Ref. [180]. Experiments have been performed with different combination of statistical features, where the fusion of more features results in better detection rate it can also be identified that the results are significantly better as compare to RGB space. The quantisation of RGB images in histograms is assumed placement in less bins for more colour levels. A more precise digitisation of the histogram is performed for estimation and a comparison of different machine learning methods has been presented in Ref. [68]. Visible optical fibre sensor with RGB Light Emitting Diode (LED) has been used for fruit quality assessment in [195], various ripening stages were recorded to generate a dataset. The optical instruments used in this study have reported a significant result, while the coefficient of determination R^2 was recorded as 0.879. The sRGB conversion to $L^*a^*b^*$ space is performed to determine the red area share on the peel of mango for mango ripeness estimation. The experimental setup is designed to capture the mango images without background and a pixel count based red area share has been estimated. However, the method is limited to the particular fruit pattern and cannot work with the complex fruit peel properties [196]. A citrus crop estimation is performed by using water shedding segmentation and distance transform and marker controller. The colour as HSV feature is used for counting the citrus fruit [197]. More recently, different colour models have been analysed for recognition of litchi fruit during day and night time where, statistical features of YIQ, RGB, HSV and YUV colour space have been used for representation of a litchi bunch. A CCD camera was fitted with an illumination system on a mobile robot for image acquisition. It is identified that the overlapping of background and pixel at night time is significantly less as compared to daytime [59]. A region based division of tomato has been performed to estimate the ripeness, the colour in terms of RGB is considered from every five regions and are converted to HSI. To estimate the ripeness level, the region based colour variations of different samples have been identified by selecting the most significant effect [13]. A performance compression of colour based fruit and vegetable analysis is provided in Table 14.

Fig. 6. Illustration of changes in digital images (a) Scale variation in Orange peel (b) Illumination change up to 75% in Chilli paper (c) Viewpoint change for Acerolas. Images by RawFooT and KTHTIPS food datasets.

Table 12
Colour invariance properties of colour feature descriptors w.r.t. diagonal model.

Property	Diagonal representation	Description
Scale-invariant	$\begin{bmatrix} R^r \\ G^r \\ B^r \end{bmatrix} = \begin{bmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{bmatrix} \begin{bmatrix} R^u \\ G^u \\ B^u \end{bmatrix}$	The equivalent change in RGB channels w.r.t. to intensity change, where x is scaling factor.
Intensity shift invariant	$\begin{bmatrix} R^r \\ G^r \\ B^r \end{bmatrix} = \begin{bmatrix} R^u \\ G^u \\ B^u \end{bmatrix} + \begin{bmatrix} o_1 \\ o_1 \\ o_1 \end{bmatrix}$	The equal shift in intensity values in all RGB channels i.e. ($o_1 = o_2 = o_3$), where o is shifting factor.
Scale and shift invariant	$\begin{bmatrix} R^r \\ G^r \\ B^r \end{bmatrix} = \begin{bmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{bmatrix} \begin{bmatrix} R^u \\ G^u \\ B^u \end{bmatrix} + \begin{bmatrix} o_1 \\ o_1 \\ o_1 \end{bmatrix}$	The descriptor is invariant to the changes of scale and shift w.r.t light intensity.
Light colour invariant	$\begin{bmatrix} R^r \\ G^r \\ B^r \end{bmatrix} = \begin{bmatrix} r & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & b \end{bmatrix} \begin{bmatrix} R^u \\ G^u \\ B^u \end{bmatrix}$	The images channels scale independently i.e. ($r \neq g \neq b$).
Light colour and shift invariant	$\begin{bmatrix} R^r \\ G^r \\ B^r \end{bmatrix} = \begin{bmatrix} r & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & b \end{bmatrix} \begin{bmatrix} R^u \\ G^u \\ B^u \end{bmatrix} + \begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}$	The model changes arbitrarily for both shift and colour i.e. ($r \neq g \neq b$) and ($o_1 \neq o_2 \neq o_3$).

7. Classification

The initial introduction to computer vision traces back to 1960s. It is now an essential part of the state-of-the-art systems in industrial automation, intelligent security, autonomous vehicles, food industry, robotics and medical imaging [201]. Fruit and vegetable classification is a problem of assigning a qualitative fruit or vegetable class $c_i \in \{1, 2, \ldots C\}$ to an observed input i_o. RGB images have been studied extensively to exploit significant characteristics of fruit and vegetable like colour, shape, texture and size for conventional computer vision systems. Robotic harvesting [50], quality analysis [1], disease identification [44] and damage analysis [40] are among the leading applications of vision based fruit and vegetable classification. Recent research has used a variety of machine learning models for example, KNN, SVM, decision trees and Neural Networks (NN) and their variants [6,77,108,202] for this purpose. Linear and non-linear hyperdimensional data can be classified with the SVM which is a non-linear mapping of data with the help of kernel functions. KNN is an instance based non-parametric similarity measure learning for data of infinite dimensions and a decision tree is a probability based graph for multi-class classification. SVM and KNN have been widely used for fruit and vegetable classification and a comparable classification effectiveness w.r.t. Multi-layer Perceptron (MLP) and Radial Bias Functions (RBF) has been reported [51]. However, hyperdimensional approximation for multi-class fruit and vegetable classification using SVM poses significant performance constraints which have been addressed by combining the SVM with the meta-heuristic optimisation for optimal parameter estimation in Ref. [7]. The capability of holistic feature extraction of CNN has reported significant object classification effectiveness. Currently, Neural Networks (NN) has gained a significant importance in the food industry [2,13,29,45,72,180,203]. One of the constraints of CNN is the scarcity of substantial dataset for training the CNN. The development of pre-trained networks for general objects classification is a serious attempt to address this issue [45,204-206]. These pre-trained networks can also deal with the scarcity of fruit and vegetable datasets as they exploit the more essential features of images. Significant variants techniques of using pre-trained CNN for object classification have been presented in Table 15. A more recent comparison of available classification algorithms on different computer vision datasets has been presented in Ref. [207].

Robotic harvesting, quality assessment and the produce classification are among the most evident applications of fruit and vegetable from literature. A prototype system for the produce classification has been introduced in Ref. [66], where colour and texture is used for classification using KNN as machine learning technique. Illumination has been considered as a key driving factor for the colour variance, hence lighting and relative position of produce item have been considered carefully. This experiment is the most initial effort in the produce classification however significant performance has been reported. Similarly, a mobile platform has been designed for robotic harvesting using a far vision (FAR) and near (NEAR) vision system developed in Ref. [47]. A four camera integrated image is then analysed for intensity-based roundness and smoothness for watermelon detection while ignoring too small and big object, a significant performance has been reported. Recently, much research has been reported for fruit and vegetable classification a non-exhaustive detail of which is discussed here. A defect segmentation for apple fruit has been performed in Ref. [30] by threshold-based segmentation with multiple supervised classification models to segment the defected area. The comparison considered all pixel of the image as noise free however, this assumption has led to a significant performance lack, although larger neighbourhood analysis has reported reasonable performance. A colour and Infra-red (IR) fusion has been

Table 13
Colour descriptor invariance w.r.t diagonal model, where ✓ indicates invariance and ✗ represents lack of invariance.

Colour feature descriptor	Invariances				
	Scale	Intensity Shift	Scale and shift	Light colour	Light colour and intensity shift
Histogram (RGB)	✗	✗	✗	✗	✗
Hue histogram	✓	✓	✓	✗	✗
rg histogram	✓	✗	✗	✗	✗
Colour moments	✗	✓	✗	✗	✗
SIFT	✓	✓	✓	✗	✗
HSV-SIFT	✗	✗	✗	✗	✗
Hue-SIFT	✓	✓	✓	✗	✗
rg-SIFT	✓	✗	✗	✗	✗
Opponent SIFT	✓	✓	✓	✗	✗
C-SIFT	✓	✗	✗	✗	✗
RGB-SIFT	✓	✓	✓	✓	✓

Table 14
Comparison of colour features for fruit and vegetable analysis.

year	Fruit/Veg	Colour feature vector	Colour space	Accuracy	Ref.
2009	Grapefruit	HSI based CCV	HSI	98.30%	[34]
2009	Banana	RGB Histogram	RGB	–	[194]
2009	Pomegranate	Threshold on R/G ration and RGDB LDA	RGB	90.00%	[198]
2010	Strawberry	Dominant colour in a^* channel	CIE Lab	88.80%	[199]
2012	Mixed	Red and green component in HSI domain	HSI	95.00%	[77]
2013	Citrus	LUT based CCI in $L^*a^*b^*$	$L^*a^*b^*$	95.00%	[200]
2014	Tomato	Colour histogram	RGB	89.10%	[68]
2014	Mango	Optical RGB with fine LED light	RGB	87.90%	[195]
2015	Olive	Histogram of gradients of R, G and B channel	RGB	100.00%	[79]
2016	Mango	R/G ratio by pixel count	RGB	94.00%	[196]
2016	Tomato	Mean, deviation and skewness on R, G and B channel	RGB	100.00%	[39]
2017	Citrus	H channel thresholding in HSV	HSV	93.00%	[197]
2018	Litchi	Statistical features of Y, I and Q	YIQ	93.75%	[59]
2018	Citrus	Watershed on RGB	RGB	93.13%	[15]
2018	Tomato	HSI based colour matching	HSI	100.00%	[13]

used for counting apple fruit on the tree [53]. A Haar filter based fruit detection has been used with the Adaboost algorithm on a mobile robot. The analysis has taken an advantage of colour-IR fusion for dealing with occlusions however, Haar filters are not robust enough w.r.t. noise and distortion in data. Grapefruit peel condition has been analysed in Ref. [34] for five diseases by texture analysis. An LDA based texture features selection has been performed for spatial intensity level comparison, but the reduction in features size has reported in performance lack. It can be identified that the experiments have been performed in a constraints environment for better performance, where the colour space chosen has a limitation of low lighting condition. A supermarket produce classification system has been presented in Ref. [74] for 15 classes and 2633 images. Statistical colour and texture features have been used for classification, however significant over-fitting is evident from classification results due to a small number of training samples moreover introduction to more significant features i.e. morphology can also improve the performance. Another similar approach of produce classification is presented in Ref. [73], where packed fruit and vegetable are also considered for classification of items. Statistical features of colour and vocabulary based texture have been used along with the fusion of classifiers in this study. It has been reported that the experiments can achieve a more better result using more complex features e.g. appearance-based features.

Many examples of vision-based quality grading are also evident from the literature. Colour-based statistical features have been described for quality assessment of citrus fruit describing colour in HSV space. A distance-based classification i.e. KNN and CNN have been used for classification of defected citrus [180]. A statistical histogram-based apple quality assessment has been presented in a non-destructive way. A soft clustering has been performed for classification however, the ACM energy minimisation used for segmentation of apple shape poses significant performance constraints. Moreover, the invariance of colour space to illumination has not been considered carefully while performing the experiment in a controlled environment [35]. More recently, olive quality assessment has been performed for oil extraction, olive image histogram has been used with Fisher discriminant analysis for linear classification. However, the training has been performed with a very small number of samples and more complex feature vector can be used for this purpose keeping in view the limitations of global histograms used [79]. As the current advancement in computer vision has presented an emphasis on image representation as its elementary characteristics, a BOF based image representation has been used for this purpose. A machine vision based mango crop estimation is performed by detecting mango fruit in the canopy of the mango tree, a manual counting is performed on segmented images for estimation. Dense SIFT has been used for constructing a Bag-of-Visual

Table 15
Significant variants of CNN based approaches.

CNN Variants	Description	Literature	Application
Pre-trained CNN model	Basic filter bank and feature encoding and pooling techniques.	AlexNet [161]	Introduction of CNN based feature encoding for image classification challenge by image net.
		VGGM [208]	Texture classification performance than AlexNet at similar complexity.
		VGGVD [209]	Deeper layer set for better classification performance.
		GoogleNet [204-206]	Smaller filter banks and deeper convolution layers for image classification.
		ResNet [84]	Significantly deeper than previous CNN based pre-trained models.
Fine-tuned model	Conversion of the fully connected layer to n nodes specific to classes in the dataset for classification.	TCNN [210]	Global average pooling the output from multiple CONV layers.
		BCNN [211 ,212]	Introduction of orderless bilinear pooling methods for high dimensional feature representation.
		Compact BCNN [213]	Dimensionality reduction of features in BCNN for better performance.
Hand-crafted CNN model	Using traditional hand-crafted feature descriptor methods for convolution layers.	ScatNet [214]	Using Gabor wavelet as a function for convolution layer.
		PCANet [215]	Using PCA filters as convolution layer along with LBP and histogram for feature pooling.

Table 16
Comparison of machine vision techniques for fruit and vegetable classification.

year	Fruit/Veg	Dataset size	Classifier	Accuracy	Ref.
2006	Apple	526	KNN, LDC, QDC, LR, SVM, FNN, K-means, SOM, NN	99.30%	[30]
2006	Citrus	–	Specialised	95.00%	[31]
2007	Apple	166	Gabor wavelet PCA	90.50%	[32]
2008	Apple	46	PCA	-	[33]
2009	Grapefruit	180	Squared distance	98.30%	[34]
2010	Mixed	2633	Hyperdimensional SVM	86.00%	[73]
2011	Apple	–	Histogram based FCM	96.00%	[35]
2012	Apple	210	N-neighbouring	96.00%	[76]
2012	Mixed	2633	K-mean clustering	98.80%	[136]
2012	Pomegranates	–	Hyperdimensional SVM	99.88%	[178]
2012	Vege	296	Decision trees	95.00%	[77]
2013	Apple	92	BPNN networks	88.00%	[36]
2013	Jatropha	–	K-means, fuzzy c-means (FCM)	87.20%	[216]
2013	Mixed	–	Network based	96.55%	[78]
2014	Apple	–	K-means fuzzy c-means (FCM)	60.00%	[217]
2014	18 fruit	1653	FSCBC+FNN	89.10%	[68]
2015	18 fruit	1653	WE, PCA,BBO, FNN	89.50%	[29]
2015	Fruit	(5 classes)	Transfer Learning	50.00%	[218]
2015	Olive	77	Fisher Discriminant Analysis (FDA)	100.00%	[79]
2016	18 fruit	1653	FNN and Deep Learning(DL)	99.88%	[72]
2016	Figs	120	SVM, LDA, LOGLC	100.00%	[38]
2016	Tomato	520	Three layer FNN	100.00%	[39]
2017	Apple	–	Artificial Neural Network (NN)	94.94%	[40]
2017	Almond	2000	KNN, L-SVM, Chi-SVM	91.00%	[6]
2017	Eggplant	50	KNN	88.00%	[57]
2017	Grapevine	760	SVM	97.70	[56]
2017	Mango	200	SVM	87.00%	[42]
2017	Mango	2464	SVM and dense segmentation	98.00%	[137]
2017	Vege	(26 classes)	KNN, SVM, ABC-FNN, FSCABC-FNN	95.60%	[202]
2017	Vege	(5 classes)	SVM	90.79%	[41]
2017	Tomato	–	ANN	98.50%	[203]
2018	Apple	55	K-means, FCM	91.84%	[14]
2018	20 cultivars	–	-	100.00%	[81]
2018	Dates	8000	Caffee Net	99.24%	[219]
2018	Fruit	1778	SVM	98.50%	[7]
2018	Litchi	480	FCM	97.50%	[59]
2018	Lettuce	320	CNN	86.00%	[109]
2018	Maize	910	PLS-DA	100.00%	[220]
2018	Orange	335	Naive Bayes, ANN, Decision Tree	93.45%	[15]
2018	Papaya	114	Decision tree	95.98%	[58]
2018	Papaya	129	SVM, Decision Tree, Naive bayes	90.15%	[44]
2018	Strawberry	337	Histogram Comparison	94.00%	[80]
2018	Strawberry	2969	KNN, FCM, K-means	100.00%	[43]
2018	Tomato	150	BPNN	100.00%	[13]

(BOV) words for super-pixel classification using KNN [137]. Sweet and bitter almond visual classification with the key points based BOF is performed in Ref. [6], where each almond image can be represented as a frequency histogram of BOF in the codebook. Corners, regions and blobs have been used to represent the almonds and an accuracy of 91% is reported. However, a complete analysis of the invariance of features w.r.t. different transforms need to be performed and a more large dataset per class should be used for more reliable results. An accuracy of 99.24% has been reported with an SVM is used with LBP, HOG and CNN based feature for generating image patches used as an input for classification [7]. These patches are then analysed with CafeNet for classification, the overlapping among the multiple patches detecting windows has been used for patch selection as final feature vector for decision making. This window based method has been used for classifying fruit with occlusions however, some instances with occlusion have been detected falsely. It is also identified that complex background poses significant performance and computation constraints. Morphology of fruit and vegetable has also been considered for different food industry applications. The approximation of the elliptical shape of strawberry fruit has been represented with Elliptical Fourier Descriptor (EFD) while using SVM and decision tree for shape-based classification. Length of contour, area and major axis of the estimated ellipse has been used for shape representation. Chain codes difference of optimal ellipse area ration and optimal boundary length ratio has been used for finding the elliptical similarity for classification, where an accuracy of 91% has been reported [16]. Another morphology based strawberry classification is performed in Ref. [80], shape and size have been estimated by kite analysis for classification of strawberry. However, morphological analysis is limited to automated sorting only but no quality assessment can be performed due to lack other features description i.e. colour or texture. A more precise morphological analysis is performed by combining the morphology and contour based colour information for tomato ripeness estimation [13]. Dark image background is used to segment and centroid estimation of tomato, where colour information is considered on equidistant contour regions in the tomato boundary for ripeness estimation. More considerable utilisation of this technique can be in classification of multiple types of same fruit or vegetable with a slight visual difference at the global level e.g. classification of different types of apple or mangoes.

More significant efforts for classification of fruit and vegetable have utilised approximately all possible feature and have tested machine vision boundaries. A Fitness Scaled Chaotic Artificial Bee Colony (FSCBC) algorithm has been tested with Feed-forward Neural Network (FNN) as a hybrid classification techniques [68]. Selected windows on the fruit images are used for feature extraction and classification with FNN-FSCBC where an accuracy of 89.10% has been achieved. Another FNN based on wavelet entropy PCA has been presented in Ref. [29]. The FNN has been trained by FSCBC and biogeography-based optimisation is applied for classification. SVM and fuzzy algorithm have been used for grading of mangoes in Ref. [42] with an accuracy of 87%. An apple bruise detection has been performed for automated quality assessment and disease detection in Ref. [40] using a thermal camera and Artificial Neural Network (ANN). A packed fresh-cut lettuce analysis has been performed in Ref. [109] for supermarket produce. A Deep Learning (DL) based classification of on CIELAB colour space has been performed with super-pixel segmentation in this study. A more detailed comparison of the state-of-the-art fruit and vegetable classification methods has been presented in Table 16.

8. Summary

A comprehensive review of the fruit and vegetable classification process has been presented. A detailed comparative study is presented to consider significant characteristics of sensors, feature description and classification algorithms. A comparison of the techniques used in the field of fruit and vegetable classification is established to comprehended the current key challenges in this field. The study explores the major constraints of utilisation of currently available sensors and the combination of multiple sensors for data acquisition in different applications of food industry. A brief description of difficulties in multi-sensory data fusion is also discussed in the paper. Significant points have been made on the importance of pre-processing and segmentation required for computer vision based analysis in the food industry. The feature description of pre-processed and segmented images is discussed in detail with an emphasis on fruit and vegetable characteristics. Finally, an overview of classification techniques used with various features and their combination in different applications of food industry has been presented.

8.1. Conclusion

Based on the literature, an up-to-date review of fruit and vegetable classification and constituent processes is presented in this paper and the previous efforts made have been recorded well. Significant challenges in terms of data acquisition devices, feature representation and classification algorithms have been identified to overcome. The sensors used for the data acquisition in the food industry are found constrained due to substantial limitations in various applications for example, some of the applications are non-destructive in nature, have environmental occlusions, presents inter and intraclass similarities and complex features. Other significant limitation on the use of multiple sensors in the same application of fruit and vegetable analysis is different nature of data produced by them. This different nature of data is also limited for providing significant multisensory data fusion. The feature descriptors developed and used in the state-of-the-art are also insufficient in such a capability. Moreover, no sufficient feature descriptors are available for the most recent kind of sensors i.e. RGBD sensors. Other significant limitations of feature descriptors are due to their sensitivity to many natural pheromones of image capturing. These limitations are significantly evident from the relevant literature and are presented in the paper. The machine vision algorithms evident from literature are insignificant to cope with multi-feature hyperdimensional information for classification. The fruit and vegetable have numerous classes and each of them presents a multi-feature nature. The classification algorithms identified are constrained by the scarcity of substantial datasets available. It has been identified that most of the experiments performed in the literature are either limited in terms of classes or the size of the dataset. The current research for the development of pre-trained CNN is a step toward developing a capability of providing off-the-shelf components for computer vision applications. However, these pre-trained CNN are data dependent and availability of significantly large dataset of fruit and vegetable is scarce. Considering the detailed discussion on the fruit and vegetable classification a suggestion can be raised that a complete rethinking is required for more effective use of computer vision in the food industry.

8.2. Future directions

Significant limitations of the state-of-the-art techniques in different application areas have been identified. Most of the emerging new sensors have not been exploited for the applications of fruit and vegetable. The major reasons for their scarce utilisation in fruit and vegetable classification is the unavailability of substantial datasets. The data needs to be collected and augmented to build new datasets to take advantage of RGBD sensors for more effective results. Among the numerous applications of this area, some have not been studies well e.g. supermarket self-checkout and use of recent RGBD sensors for this task. Significant evidence of automated self-checkout and utilisation of visual data in intelligent self-checkout are presented as future technology [60-66]. The constraints: lighting condition, timeliness, large dataset, effectiveness and accuracy are there to introduce this new technology in supermarket. Approximately, 150 classes of fruit and vegetable have been identified in a rough internet survey in Australian supermarkets, none of the previous studies have discussed such a number of classes. Recent advanced commodity RGBD sensors are being used for object classification [70,221-227], which can also be used for more effective classification of fruit and vegetable.

Detailed survey of the fruit and vegetable classification techniques has been presented to investigate the intuitive use of recent techniques in computer vision based automated self-checkout. The technologies explored were specifically chosen to meet the predefined goals. Based on the knowledge developed from this study, our future areas of research will be:

- The utilisation of RGBD data for fruit and vegetable classification
- System level design of RGBD sensor based supermarket self-checkout
- Optimal ways of dealing with scarcity of large RGBD datasets
- Optimisation of the state-of-the-art machine learning techniques with RGBD data

Acknowledgements

This work was supported by Higher Education Commission (HEC) Pakistan, The Islamia University of Bahawalpur (IUB) Pakistan (5-1/HRD/UECTP(Batch-V)/1182/2017/HEC) and Edith Cowan University (ECU) Australia. The authors would like to thank HEC, IUB Pakistan and ECU Australia for PhD grant of first corresponding author of this paper.

References

[1] A. Bhargava, A. Bansal, Fruits and vegetables quality evaluation using computer vision: a review, J. King Saud Univ. - Comput. Inf. Sci. (2018) 1–15. https://doi.org/10.1016/j.jksuci.2018.06.002.
[2] F. Kurtulmus, W.S. Lee, A. Vardar, Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network, Precis. Agric. 15 (1) (2014) 57–79.
[3] A. Gongal, S. Amatya, M. Karkee, Q. Zhang, K. Lewis, Sensors and systems for fruit detection and localization: a review, Comput. Electron. Agric. 116 (2015) 8–19.
[4] K. Yamamoto, W. Guo, Y. Yoshioka, S. Ninomiya, On plant detection of intact tomato fruits using image analysis and machine learning methods, Sensors 14 (7) (2014) 12191–12206.

[5] A. Gijsenij, T. Gevers, J. Van De Weijer, Computational color constancy: survey and experiments, IEEE Trans. Image Process. 20 (9) (2011) 2475–2489. https://doi.org/10.1109/TIP.2011.2118224.

[6] A. Nasirahmadi, S.H. Miraei Ashtiani, Bag-of-Feature model for sweet and bitter almond classification, Biosyst. Eng. 156 (2017) 51–60. https://doi.org/10.1016/j.biosystemseng.2017.01.008.

[7] H. Kuang, C. Liu, L.L.H. Chan, H. Yan, Multi-class fruit detection based on image region selection and improved object proposals, Neurocomputing 283 (2018) 241–255. .

[8] S. Lacey, C. Campbell, K. Sathian, Vision and touch: multiple or multisensory representations of objects? Perception 36 (10) (2007) 1513–1521.

[9] M. Tuceryan, A.K. Jain, Texture Analysis, 1998. (January). https://doi.org/10.1142/9789814343138.

[10] A. Schmitz, Y. Bansho, K. Noda, H. Iwata, T. Ogata, S. Sugano, Tactile object recognition using deep learning and dropout, Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on, IEEE. 2014, pp. 1044–1050.

[11] H. Liu, Y. Yu, F. Sun, J. Gu, Visual tactile fusion for object recognition, IEEE Trans. Autom. Sci. Eng. 14 (2) (2017) 996–1008.

[12] Y.-D. Zhang, Z. Dong, X. Chen, W. Jia, S. Du, K. Muhammad, S.-H. Wang, Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation, Multimedia Tools and Applications (2017) 1–20.

[13] P. Wan, A. Toudeshki, H. Tan, R. Ehsani, A methodology for fresh tomato maturity detection using computer vision, Comput. Electron. Agric. 146 (February 2017) (2018) 43–50. https://doi.org/10.1016/j.compag.2018.01.011.

[14] B. Li, Y. Long, H. Song, Detection of green apples in natural scenes based on saliency theory and Gaussian curve fitting, Int. J. Agric, Biol. Eng. 11 (1) (2018) 192–198. https://doi.org/10.25165/j.ijabe.20181101.2899.

[15] A. Wajid, N.K. Singh, P. Junjun, M.A. Mughal, Recognition of Ripe, Unripe and Scaled Condition of Orange Citrus Based on Decision Tree Classification, International Conference on Computing, Mathematics and Engineering Technologies, 2018. pp. 2–5.

[16] T. Ishikawa, A. Hayashi, S. Nagamatsu, Y. Kyutoku, I. Dan, T. Wada, K. Oku, Y. Saeki, T. Uto, T. Tanabata, Classification of strawberry fruit shape by machine learning, Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. XLII (June) (2018) 4–7.

[17] H. Zhang, J. Wu, Z. Zhao, Z. Wang, Nondestructive firmness measurement of differently shaped pears with a dual-frequency index based on acoustic vibration, Postharvest Biol. Technol. 138 (December 2017) (2018) 11–18. https://doi.org/10.1016/j.postharvbio.2017.12.002.

[18] M. Lashgari, A. Maleki, J. Amiriparian, Application of acoustic impulse response in discrimination of apple storage time using neural network, 24. 2017, 1075–1080. (June).

[19] J. Foerster, M. Geyer, O. Schlüter, P. Fey, M. Kiefer, Acoustic resonance analysis for quality characterization of fruits and vegetables, Landtechnik 65 (2) (2010) 96–98.

[20] M. Vogl, J.-Y. Kim, S.-D. Kim, A fruit recognition method via image conversion optimized through evolution strategy, Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on, IEEE. 2014, pp. 1497–1502.

[21] a.R Jimenez, a.K Jain, R. Ceres, J. Pons, Automatic fruit recognition : a survey and new results using Range/Attenuation images, Pattern Recogn. 32 (1999) 1719–1736.

[22] Y. Sun, X. Gu, K. Sun, H. Hu, M. Xu, Z. Wang, K. Tu, L. Pan, Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches, LWT- Food Sci. Technol. 75 (2017) 557–564. https://doi.org/10.1016/j.lwt.2016.10.006.

[23] H. Cen, R. Lu, Q. Zhu, F. Mendoza, Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification, Postharvest Biol. Technol. 111 (2016) 352–361. https://doi.org/10.1016/j.postharvbio.2015.09.027.

[24] H. Okamoto, W.S. Lee, Green citrus detection using hyperspectral imaging, Comput. Electron. Agric. 66 (2) (2009) 201–208.

[25] S. Naik, B. Patel, Thermal imaging with fuzzy classifier for maturity and size based non-destructive mango (*Mangifera indica* L.) grading, 2017 International Conference on Emerging Trends and Innovation in ICT, ICEI 2017, 2017, 15–20. https://doi.org/10.1109/ETIICT.2017.7977003.

[26] S. Lee, B.-K. Cho, Evaluation of the firmness measurement of fruit by using a non-contact ultrasonic technique, Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on, IEEE. 2013, pp. 1331–1336.

[27] I. Aboudaoud, B. Faiz, E. Aassif, A. Moudden, D. Izbaim, D. Abassi, M. Malainine, M. Azergui, The maturity characterization of orange fruit by using high frequency ultrasonic echo pulse method, IOP Conference Series: Materials Science and Engineering, vol. 42, IOP Publishing. 2012, pp. 12038.

[28] P. Yoiyod, Complex Relative Permittivity Analysis for Monitoring Watermelon, 2017, 8–9.

[29] S. Wang, Y. Zhang, G. Ji, J. Yang, J. Wu, L. Wei, Fruit classification by wavelet-entropy and feedforward neural network trained by fitness-scaled chaotic abc and biogeography-based optimization, Entropy 17 (8) (2015) 5711–5728. https://doi.org/10.3390/e17085711.

[30] D. Unay, B. Gosselin, Automatic defect segmentation of 'Jonagold' apples on multi-spectral images: a comparative study, Postharvest Biol. Technol. 42 (3) (2006) 271–279. https://doi.org/10.1016/j.postharvbio.2006.06.010.

[31] R. Pydipati, T.F. Burks, W.S. Lee, Identification of citrus disease using color texture features and discriminant analysis, Comput. Electron. Agric. 52 (1-2) (2006) 49–59.

[32] B. Zhu, L. Jiang, Y. Luo, Y. Tao, Gabor feature-based apple quality inspection using kernel principal component analysis, J. Food Eng. 81 (4) (2007) 741–749. https://doi.org/10.1016/j.jfoodeng.2007.01.008.

[33] K. Vijayarekha, Multivariate image analysis for defect identification of apple fruit images, IECON Proceedings (Industrial Electronics Conference), 2008. pp. 1499–1503. https://doi.org/10.1109/IECON.2008.4758175.

[34] D.G. Kim, T.F. Burks, J. Qin, D.M. Bulanon, Classification of grapefruit peel diseases using color texture feature analysis, Int. J. Agric. Biol. Eng. 2 (23) (2009) 41–50. Int J Agric & Biol Eng Open Access at http://www.ijabe.org. https://doi.org/10.3965/j.issn.1934-6344.2009.03.041-050.

[35] G. Moradi, M. Shamsi, M.H. Sedaaghi, S. Moradi, Apple defect detection using statistical histogram based Fuzzy C-means algorithm, IEEE Electr. Eng. (2011) 11–15.

[36] M. Jhuria, A. Kumar, R. Borse, Image Processing for Smart Farming: Detection of Disease and Fruit Grading, 2013 IEEE 2nd International Conference on Image Information Processing, IEEE ICIIP 2013, 2013, 521–526. https://doi.org/10.1109/ICIIP.2013.6707647.

[37] S.R. Dubey, A.S. Jalal, Application of Image Processing in Fruit and Vegetable Analysis: A Review, J. Intell. Syst. 24 (4) (2015) 405–424. https://doi.org/10.1515/jisys-2014-0079.

[38] G. Ortac, A.S. Bilgi, Y.E. Gorgulu, A. Gunes, H. Kalkan, K. Tasdemir, Classification of black mold contaminated figs by hyperspectral imaging, 2015 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2015, 2016. pp. 227–230. https://doi.org/10.1109/ISSPIT.2015.7394332.

[39] M.P. Arakeri, Lakshmana, Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry, Procedia Comput. Sci. 79 (2016) 426–433. https://doi.org/10.1016/j.procs.2016.03.055.

[40] D. Jawale, M. Deshmukh, Real Time Automatic Bruise Detection in (Apple) Fruits Using Thermal Camera, 2017, 1080–1085.

[41] N. Belsha, N. Hariprasad, The enhanced content based image retrieval system and classification of infected vegetables, Proceedings of 2017 3rd IEEE International Conference on Sensing, Signal Processing and Security, ICSSS 2017, 2017. pp. 83–88. https://doi.org/10.1109/SSPS.2017.8071570.

[42] C.S. Nandi, B. Tudu, C. Koley, A machine vision technique for grading of harvested mangoes based on maturity and quality, IEEE Sensors J. 16 (16) (2016) 6387–6396. https://doi.org/10.1109/JSEN.2016.2580221.

[43] N.M. Hussain Hassan, A.A. Nashat, New effective techniques for automatic detection and classification of external olive fruits defects based on image processing techniques, Multidim. Syst. Sign. Process. (2018) 1–19. https://doi.org/10.1007/s11045-018-0573-5.

[44] M.T. Habib, A. Majumder, A.Z.M. Jakaria, M. Akter, M.S. Uddin, F. Ahmed, Machine vision based papaya disease recognition, J. King Saud Univ. - Comput. Inf. Sci. (2018) 0–9. https://doi.org/10.1016/j.jksuci.2018.06.006.

[45] W.-S. Jeon, S.-Y. Rhee, Plant leaf recognition using a convolution neural network, Int. J. Fuzzy Logic Intelligent Syst. 17 (1) (2017) 26–34. http://www.ijfis.org/journal/view.html?doi=10.5391/IJFIS.2017.17.1.26. https://doi.org/10.5391/IJFIS.2017.17.1.26.

[46] E. Barnea, R. Mairon, O. Ben-Shahar, Colour-agnostic shape-based 3D fruit detection for crop harvesting robots, Biosyst. Eng. 146 (2016) 57–70.

[47] Y. Edan, D. Rogozin, T. Flash, G.E. Miles, Robotic melon harvesting, IEEE Trans. Robot. Autom. 16 (6) (2000) 831–835.

[48] D. Andújar, V. Rueda-Ayala, H. Moreno, J.R. Rosell-Polo, C. Valero, R. Gerhards, C. Fernández-Quintanilla, J. Dorado, H.-W. Griepentrog, Discriminating crop, weeds and soil surface with a terrestrial LIDAR sensor, Sensors 13 (11) (2013) 14662–14675.

[49] E.E. Aksoy, A. Abramov, F. Wörgötter, H. Scharr, A. Fischbach, B. Dellen, Modeling leaf growth of rosette plants using infrared stereo image sequences, Comput. Electron. Agric. 110 (2015) 78–90.

[50] Y. Zhao, L. Gong, Y. Huang, C. Liu, A review of key techniques of vision-based control for harvesting robot, Comput. Electron. Agric. 127 (2016) 311–323.

[51] Y. Tao, J. Zhou, Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking, Comput. Electron. Agric. 142 (2017) 388–396.

[52] K. Kapach, E. Barnea, R. Mairon, Y. Edan, O. Ben-Shahar, Computer vision for fruit harvesting robots state of the art and challenges ahead, Int. J. Comput. Vis. Robotics 3 (1-2) (2012) 4–34.

[53] J. Wachs, H.I. Stern, T. Burks, V. Alchanatis, I. Bet-Dagan, Apple detection in natural tree canopies from multimodal images, Proceedings of the 7th European Conference on Precision Agriculture, Wageningen, The Netherlands, vol. 68, 2009. pp. 293302.

[54] S. Khanal, J. Fulton, S. Shearer, An overview of current and potential applications of thermal remote sensing in precision agriculture, Comput. Electron. Agric. 139 (2017) 22–32. https://doi.org/10.1016/j.compag.2017.05.001.

[55] A. Danti, M. Madgi, B.S. Anami, Mean and Range Color Features Based Identification of Common Indian Leafy Vegetables, 5. 2012, 151–160. (3).

[56] D.S. Pérez, F. Bromberg, C.A. Diaz, Image classification for detection of winter grapevine buds in natural conditions using scale-invariant features transform, bag of features and support vector machines, Comput. Electron. Agric. 135 (2017) 81–95.

[57] Rahman, A. Yeasmin, O. Mohammad, Development of a Computer Vision based Eggplant Grading System, International Conference on Advances in Electrical Engineering, IEEE Press, Dhaka, 2017, pp. 285–290.

[58] L.F.S. Pereira, S. Barbon, N.A. Valous, D.F. Barbin, Predicting the ripening of papaya fruit with digital imaging and random forests, Comput. Electron. Agric. 145 (2018) 76–82.

[59] J. Xiong, R. Lin, Z. Liu, Z. He, L. Tang, Z. Yang, X. Zou, The recognition of litchi clusters and the calculation of picking point in a nocturnal natural environment, Biosyst. Eng. 166 (2018) e10. https://doi.org/10.1016/j.biosystemseng.2017.11.005.

[60] M. Dhankhar, And Object Recognition Kiosk for Retail Checkouts, Automated Object Recognition Kiosk for Retail Checkouts, 2015, https://www.google.com/patents/US20150109451.

[61] J. IIZUKA, Information Processing Apparatus and Program, 2017, https://www.google.com.au/patents/EP3232388A1?cl=en.

[62] T.C. Evans, E. Gavrilovich, R.C. Mihai, I. Isbasescu, (12) Patent Application Publication (10) Pub. No.: US 2006/0222585 A1 Figure 1, 2015, 10.1037/t24245-000, arXiv:0403007.

[63] T. Tashiro, S.-c. Terminal, Self-checkout Terminal, 2009.

[64] P.T. Catoe, And Self-checkout System, Automated Self-checkout System, 2014, https://www.google.tl/patents/US8825531.

[65] N.C. Herwig, Method and Apparatus for Reducing Recognition Times in an Image-based Product Recognition System, 2014, https://www.google.com/patents/US20140036630.

[66] R.M. Bolle, J.H. Connell, N. Haas, R. Mohan, G. Taubin, Veggievision: A produce recognition system, Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on, IEEE. 1996, pp. 244–251.

[67] Y. Sakai, T. Oda, M. Ikeda, L. Barolli, A vegetable category recognition system using deep neural network, Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2016 10th International Conference on, IEEE. 2016, pp. 189–192.

[68] Y. Zhang, S. Wang, G. Ji, P. Phillips, Fruit classification using computer vision and feedforward neural network, J. Food Eng. 143 (2014) 167–177. https://doi.org/10.1016/j.jfoodeng.2014.07.001.

[69] L. Jiang, A. Koch, S.A. Scherer, A. Zell, Multi-class fruit classification using RGB-D data for indoor robots, Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on, IEEE. 2013, pp. 587–592.

[70] E. Rachmawati, I. Supriana, M.L. Khodra, Toward a new approach in fruit recognition using hybrid RGBD features and fruit hierarchy property, 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2017, 1–6. http://ieeexplore.ieee.org/document/8239110/. (September). https://doi.org/10.1109/EECSI.2017.8239110.

[71] J.H. Cheng, B. Nicolai, D.W. Sun, Hyperspectral imaging with multivariate analysis for technological parameters prediction and classification of muscle foods: a review, Meat Sci. 123 (2017) 182–191. https://doi.org/10.1016/j.meatsci.2016.09.017.

[72] Y. Zhang, P. Phillips, S. Wang, G. Ji, J. Yang, J. Wu, Fruit classification by biogeography-based optimization and feedforward neural network, Expert. Syst. 33 (3) (2016) 239–253. https://doi.org/10.1111/exsy.12146.

[73] A. Rocha, D.C. Hauagge, J. Wainer, S. Goldenstein, Automatic fruit and vegetable classification from images, Comput. Electron. Agric. 70 (1) (2010) 96–104. http://www.sciencedirect.com/science/article/pii/S016816990900180X. https://doi.org/10.1016/j.compag.2009.09.002.

[74] S. Arivazhagan, R.N. Shebiah, S.S. Nidhyanandhan, L. Ganesan, Fruit recognition using color and texture features, Inf. Sci. 1 (2) (2010) 90–94.

[75] R.L. Radojevi, D.V. Petrovi, V.B. Pavlovi, Z.M. Nikoli, M.P. Uroševi, Digital Parameterization of Apple Fruit Size, Shape and Surface Spottiness, 6. 2011, 3131–3142. (13). https://doi.org/10.5897/AJAR11.291.

[76] S.R. Arlimatti, Window based method for automatic classification of apple fruit, Int. J. Eng. Res. Appl. (IJERA) 2 (4) (2012) 1010–1013.

[77] M. Suresha, Texture features and decision trees based vegetables classification, Int. J. Comput. Appl. (April) (2012) 21–26.

[78] T. Chowdhury, S. Alam, M.A. Hasan, I. Khan, Vegetables detection from the glossary shop for the blind., IOSR J. Electr. Electron. Eng. (IOSR-JEEE) 8 (3) (2013) 43–53.

[79] D. Martinez Gila, D. Aguilera Puerto, J. Gamez Garcia, J. Gomez Ortega, Automatic classification of olives for oil production using computer vision, International Conference on Industrial Technology (ICIT), 2015. pp. 1651–1656. http://ieeexplore.ieee.org/ielx7/7108493/7125066/07125334.pdf?tp=&arnumber=7125334&isnumber=7125066%5Cnhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7125334. https://doi.org/10.1109/ICIT.2015.7125334.

[80] L.M. Oo, N.Z. Aung, A simple and efficient method for automatic strawberry shape and size estimation and classification, Biosyst. Eng. 170 (2018) 96–107. https://doi.org/10.1016/j.biosystemseng.2018.04.004.

[81] M.R. Amiryousefi, M. Mohebbi, A. Tehranifar, Pomegranate seed clustering by machine vision, Food Sci. Nutr. 6 (1) (2018) 18–26. https://doi.org/10.1002/fsn3.475.

[82] T.H. de Groot, E. Woudenberg, A.G. Yarovoy, Urban objects classification with an experimental acoustic sensor network, IEEE Sensors J. 15 (5) (2015) 3068–3075.

[83] S. Arora, A. Bhaskara, R. Ge, T. Ma, Provable bounds for learning some deep representations, International Conference on Machine Learning, 2014. pp. 584–592.

[84] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. pp. 770–778.

[85] L. Kabbai, M. Abdellaoui, A. Douik, Hybrid local and global descriptor enhanced with colour information, IET Image Process. 11 (2) (2016) 109–117.

[86] N. Gorges, S.E. Navarro, D. Göger, H. Wörn, Haptic object recognition using passive joints and haptic key features, Robotics and Automation (ICRA), 2010 IEEE International Conference on, IEEE. 2010, pp. 2349–2355.

[87] H. Liu, Y. Wu, F. Sun, B. Fang, D. Guo, Weakly paired multimodal fusion for object recognition, IEEE Trans. Autom. Sci. Eng. 15 (2) (2017) 784–795. https://doi.org/10.1109/TASE.2017.2692271.

[88] N. Pendneault, A.-M. Cretu, 3D Object recognition from tactile data acquired at salient points, 2017.

[89] A.E. Johnson, M. Hebert, Using spin images for efficient object recognition in cluttered 3D scenes, IEEE Trans. Pattern Anal. Mach. Intell. 21 (5) (1999) 433–449. https://doi.org/10.1109/34.765655.

[90] R. Méndez Perez, F. Auat Cheein, J.R. Rosell Polo, Flexible system of multiple RGB-D sensors for measuring and classifying fruits in agri-food Industry, Comput. Electron. Agric. 139 (2017) 231–342.

[91] F. Garcia-Lamont, J. Cervantes, A. López, L. Rodriguez, Segmentation of images by color features: a survey, Neurocomputing 292 (2018) 1–27. https://doi.org/10.1016/j.neucom.2018.01.091.

[92] R. Gothwal, S. Gupta, D. Gupta, A.K. Dahiya, Color Image Segmentation Algorithm Based on RGB Channels, 2014, 1–5.

[93] D.A. Mély, J. Kim, M. McGill, Y. Guo, T. Serre, A systematic comparison between visual cues for boundary detection, Vis. Res. 120 (2016) 93–107. https://doi.org/10.1016/j.visres.2015.11.007.

[94] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis, and Machine Vision, Thomson Learn. (2008) 812. https://doi.org/10.1007/978-1-4899-3216-7.

[95] H. Fan, F. Xie, Y. Li, Z. Jiang, J. Liu, Automatic segmentation of dermoscopy images using saliency combined with Otsu threshold, Comput. Biol. Med. 85 (March) (2017) 75–85. https://doi.org/10.1016/j.compbiomed.2017.03.025.

[96] L. He, S. Huang, Modified firefly algorithm based multilevel thresholding for color image segmentation, Neurocomputing 240 (2017) 152–174. https://doi.org/10.1016/j.neucom.2017.02.040.

[97] A.K. Bhandari, A. Kumar, S. Chaudhary, G.K. Singh, A novel color image multilevel thresholding based segmentation using nature inspired optimization algorithms, Expert Systems with Applications 63 (2016) 112–133. https://doi.org/10.1016/j.eswa.2016.06.044.

[98] A. Dirami, K. Hammouche, M. Diaf, P. Siarry, Fast multilevel thresholding for image segmentation through a multiphase level set method, Signal Process. 93 (1) (2013) 139–153. https://doi.org/10.1016/j.sigpro.2012.07.010.

[99] E. Rashedi, H. Nezamabadi-Pour, A stochastic gravitational approach to feature based color image segmentation, Eng. Appl. Artif. Intell. 26 (4) (2013) 1322–1332. https://doi.org/10.1016/j.engappai.2012.10.002.

[100] C. Rueden, Statistical region merging, ImageJ 26 (11) (2017) 1452–1458. https://imagej.net/Statistical_Region_Merging.

[101] D. Mújica-Vargas, F.J. Gallegos-Funes, A.J. Rosales-Silva, A fuzzy clustering algorithm with spatial robust estimation constraint for noisy color image segmentation, Pattern Recogn. Lett. 34 (4) (2013) 400–413. https://doi.org/10.1016/j.patrec.2012.10.004.

[102] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2018) 834–848. arXiv:1412.7062. http://arxiv.org/abs/1412.7062. https://doi.org/10.1109/TPAMI.2017.2699184.

[103] Y. Guo, Y. Liu, T. Georgiou, M.S. Lew, A review of semantic segmentation using deep neural networks, Int. J. Multimedia Information Retrieval 7 (2) (2018) 87–93. https://doi.org/10.1007/s13735-017-0141-z.

[104] S.H. Ong, N.C. Yeo, K.H. Lee, Y.V. Venkatesh, D.M. Cao, Segmentation of color images using a two-stage self-organizing network, Image Vis. Comput. 20 (4) (2002) 279–289.

[105] E. Rajaby, S.M. Ahadi, H. Aghaeinia, Robust color image segmentation using fuzzy c-means with weighted hue and intensity, Digital Signal Process. A Rev. J. 51 (2016) 170–183. https://doi.org/10.1016/j.dsp.2016.01.010.

[106] S. Xia, P. Chen, J. Zhang, X. Li, B. Wang, Utilization of rotation-invariant uniform LBP histogram distribution and statistics of connected regions in automatic image annotation based on multi-label learning, Neurocomputing 228 (November 2016) (2017) 11–18. https://doi.org/10.1016/j.neucom.2016.09.087.

[107] F. Wang, S. Ainouz, C. Lian, A. Bensrhair, Multimodality semantic segmentation based on polarization and color images, Neurocomputing 253 (2017) 193–200. https://doi.org/10.1016/j.neucom.2016.10.090.

[108] V. Ashok, D. Vinod, Automatic quality evaluation of fruits using Probabilistic Neural Network approach, Proceedings of 2014 International Conference on Contemporary Computing and Informatics, IC3I 2014, 2014. pp. 308–311. https://doi.org/10.1109/IC3I.2014.7019807.

[109] D.P. Cavallo, M. Cefola, B. Pace, A.F. Logrieco, G. Attolico, Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material, J. Food Eng. 223 (2018) 46–52. https://doi.org/10.1016/j.jfoodeng.2017.11.042.

[110] V.H. Pham, B.R. Lee, An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm, Vietnam J. Comput. Sci. 2 (1) (2015) 25–33.

[111] L. Jidong, Z. De-An, J. Wei, D. Shihong, Recognition of apple fruit in natural environment, Optik-Int. J. Light Electron Opt. 127 (3) (2016) 1354–1362.

[112] D.A. Lisin, M.A. Mattar, M.B. Blaschko, E.G. Learned-Miller, M.C. Benfield, Combining local and global image features for object class recognition, Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on, IEEE. 2005, pp. 47.

[113] K. Yu, Y. Lin, J. Lafferty, Learning image representations from the pixel level via hierarchical sparse coding, Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE. 2011, pp. 1713–1720.

[114] C. Singh, P. Sharma, Performance analysis of various local and global shape descriptors for image retrieval, Multimedia Systems 19 (4) (2013) 339–357.

[115] B. Drost, M. Ulrich, N. Navab, S. Ilic, Model globally, match locally: efficient and robust 3D object recognition, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, Ieee. 2010, pp. 998–1005.
[116] M. Yang, K. Kpalma, J. Ronsin, A. Survey, S. Feature, A survey of shape feature extraction techniques to cite this version :, Pattern Recogn. (2010)
[117] E. Salahat, M. Qasaimeh, Recent advances in features extraction and description algorithms: a comprehensive survey, Proce. IEEE Int. Conference Ind. Technol. (2017) 1059–1063. arXiv:1703.06376. https://doi.org/10.1109/ICIT.2017.7915508.
[118] S. Loncaric, A survey of shape analysis techniques, Pattern Recogn. 31 (8) (1998) 983–1001.
[119] H. Freeman, On the encoding of arbitrary geometric configurations, IRE Trans. Electron. Comput. 2 (1961) 260–268.
[120] L. Kurnianggoro, Wahyono, K.H. Jo, A survey of 2D shape representation: methods, evaluations, and future research directions, Neurocomputing 0 (2018) 1–16. https://doi.org/10.1016/j.neucom.2018.02.093.
[121] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Trans. Pattern Anal. Mach. Intell. 24 (4) (2002) 509–522. arXiv:arXiv:1011.1669v3. https://doi.org/10.1109/34.993558.
[122] M.-K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Inf. Theory 8 (2) (1962) 179–187.
[123] F. Berrada, D. Aboutajdine, S.E. Ouatik, A. Lachkar, Review of 2D shape descriptors based on the curvature scale space approach, Int. Conf. Multimedia Comput. Syst. -Proc. (1). (2011) https://doi.org/10.1109/ICMCS.2011.5945600.
[124] H. Asada, The curvature primal sketch, IEEE Trans. Pattern Anal. Mach. Intell. (1) (1986) 2–14.
[125] S. Kopf, T. Haenselmann, W. Effelsberg, Shape-based posture and gesture recognition in videos, Proc. IS&T/SPIE Electron. Imaging (EI) (2005) 114–124. https://doi.org/10.1117/12.587946.
[126] W. Maghrebi, A. Borchani, M.A. Khabou, A.M. Alimi, A system for historic document image indexing and retrieval based on xml database conforming to mpeg7 standard, 2007, 114–125.
[127] L. Kurnianggoro, Y. Yang, K.H. Jo, A Similarity-based Approach for Shape Classification Using Region Decomposition, International Conference on Intelligent Computing, vol. 12, 2015. pp. 51–60. https://doi.org/10.1007/978-3-319-42294-7.
[128] L.J. Latecki, R. Lakämper, Convexity rule for shape decomposition based on discrete contour evolution, Comput. Vis. Image Underst. 73 (3) (1999) 441–454. https://doi.org/10.1006/cviu.1998.0738.
[129] D. Zhang, G. Lu, Review of shape representation and description techniques, Pattern Recogn. 37 (1) (2004) 1–19. https://doi.org/10.1016/j.patcog.2003.07.008.
[130] O. Ramos Terrades, S. Tabbone, E. Valveny, A review of shape descriptors for document analysis, Proceedings of the International Conference on Document Analysis and Recognition, ICDAR 1 (Icdar) (2007) 227–231. https://doi.org/10.1109/ICDAR.2007.4378709.
[131] X. Wang, B. Feng, X. Bai, W. Liu, L. Jan Latecki, Bag of contour fragments for robust shape classification, Pattern Recogn. 47 (6) (2014) 2116–2125. https://doi.org/10.1016/j.patcog.2013.12.008.
[132] J. Sivic, A. Zisserman, Video Google: a text retrieval approach to object matching in videos, IEEE Int. Conf. Comput. Vis. (2003) 1470–1477. arXiv:1504.06897. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1238663. https://doi.org/10.1109/ICCV.2003.1238663.
[133] X. Bai, C. Rao, X. Wang, Shape vocabulary: a robust and efficient shape representation for shape matching, IEEE Trans. Image Process. 23 (9) (2014) 3935–3949. https://doi.org/10.1109/TIP.2014.2336542.
[134] J.J. De Mesquita Sa Junior, A.R. Backes, Shape classification using line segment statistics Jarbas, Inf. Sci. 305 (2015) 349–356. https://doi.org/10.1016/j.ins.2015.01.027.
[135] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks arXiv:1311.2901v3 [cs.CV] 28 Nov 2013, Computer VisionECCV 2014 8689 (2014) 818–833. arXiv:1311.2901. http://link.springer.com/10.1007/978-3-319-10590-1_53%5Cnhttp://arxiv.org/abs/1311.2901%5Cnpapers3://publication/uuid/44feb4b1-873a-4443-8baa-1730ecd16291. https://doi.org/10.1007/978-3-319-10590-1_53.
[136] F.A. Faria, J.A. Dos Santos, A. Rocha, R.D.a.S. Torres, Automatic classifier fusion for produce recognition, Brazilian Symposium of Computer Graphic and Image Processing (2012) 252–259. https://doi.org/10.1109/SIBGRAPI.2012.42.
[137] W.S. Qureshi, A. Payne, K.B. Walsh, R. Linker, O. Cohen, M.N. Dailey, Machine vision for counting fruit on mango tree canopies, Precis. Agric. 18 (2) (2017) 224–244.
[138] D. Li, M. Shen, D. Li, X. Yu, Green apple recognition method based on the combination of texture and shape features, 2017 IEEE International Conference on Mechatronics and Automation, ICMA 2017, 2017. pp. 264–269. https://doi.org/10.1109/ICMA.2017.8015825.
[139] J.B.U. Dimatira, E.P. Dadios, F. Culibrina, J.A. Magsumbol, J.D. Cruz, K. Sumage, M.T. Tan, M. Gomez, Application of fuzzy logic in recognition of tomato fruit maturity in smart farming, IEEE Region 10 Annual International Conference, Proceedings/TENCON, 2016. pp. 2031–2035. https://doi.org/10.1109/TENCON.2016.7848382.
[140] L. Nanni, A. Lumini, S. Brahnam, Local binary patterns variants as texture descriptors for medical image analysis, Artif. Intell. Med. 49 (2) (2010) 117–125. https://doi.org/10.1016/j.artmed.2010.02.006.
[141] C. He, S. Li, Z. Liao, M. Liao, Texture classification of PolSAR data based on sparse coding of wavelet polarization textons, IEEE Trans. Geosci. Remote Sens. 51 (8) (2013) 4576–4590. https://doi.org/10.1109/TGRS.2012.2236338.
[142] L. Zheng, Y. Yang, Q. Tian, SIFT meets CNN: a decade survey of instance retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 40 (5) (2018) 1224–1244. arXiv:1608.01807. https://doi.org/10.1109/TPAMI.2017.2709749.
[143] C. Ding, J. Choi, D. Tao, L.S. Davis, Multi-directional multi-level dual-cross patterns for robust face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 38 (3) (2016) 518–531.
[144] E. Oyallon, S. Mallat, Deep roto-translation scattering for object classification, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn. 07-12-June (2015) 2865–2873. arXiv:1412.8659. https://doi.org/10.1109/CVPR.2015.7298904.
[145] L.A. Gatys, A.S. Ecker, M. Bethge, Image style transfer using convolutional neural networks, IEEE Conf. Comput. Vis. Pattern Recogn. (2016) 2414–2423. arXiv:1505.07376. https://doi.org/10.1109/CVPR.2016.265.
[146] B. Julesz, Visual pattern discrimination, IRE Trans. Inf. Theory 49 (1962) 41–46. http://dl.acm.org/citation.cfm?id=1972%5Cnhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1057698. https://doi.org/10.1007/BF00344749.
[147] R.M. Haralick, Statistical and structural approach to texture, Proceeding of IEEE vol 67 no 5 67 (5) (1979) 786–804. https://doi.org/10.1109/PROC.1979.11328.
[148] L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa, M. Pietikainen, A Survey of Recent Advances in Texture Representation, 2018, arXiv:1801.10324.. http://arxiv.org/abs/1801.10324.
[149] K.I. Laws, Rapid texture identification, Image Processing for Missile Guidance, International Society for Optics and Photonics. 1980, pp. 376–382.
[150] P. Demyelinated, N. Fiber, Biological cybernetics, Cybernetics 77 (1988) 73–77.
[151] B.S. Manjunath, W.Y. Ma, Texture features for browsing and retrieval of large image data, 18. 1996, 837–842. (8).
[152] J. Malik, P. Perona, Preattentive texture discrimination with early vision mechanisms, J. Opt. Soc. Am. A 7 (5) (1990) 923. https://www.osapublishing.org/abstract.cfm?URI=josaa-7-5-923. https://doi.org/10.1364/JOSAA.7.000923.
[153] W.T. Freeman, E.H. Adelson, The Design and Use of Steerable Filters, 1991. arXiv:arXiv:0912.1845v2,. https://doi.org/10.1109/34.93808.
[154] A. Materka, M. Strzelecki, Texture analysis methods: a review, Methods 11 (1998) 1–33. 10.1.1.97.4968. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.4968&rep=rep1&type=pdf.
[155] T. Leung, J. Malik, Representing and recognizing the visual appearance of materials using three-dimensional textons, Int. J. Comput. Vis. 43 (1) (2001) 29–44.
[156] J. Xie, W. Hu, S.C. Zhu, Y.N. Wu, Learning sparse FRAME models for natural image patterns, Int. J. Comput. Vis. 114 (2-3) (2015) 91–112. https://doi.org/10.1007/s11263-014-0757-x.
[157] S.C. Zhu, X.W. Liu, Y.N. Wu, Exploring texture ensembles by efficient Markov chain monte carlo-toward a "trichromacy" theory of texture, IEEE Trans. Pattern Anal. Mach. Intell. 22 (6) (2000) 554–569. https://doi.org/10.1109/34.862195.
[158] Wang, S. Hussain, R. Mitchell, A. Richard, N. Andrew, D. May, An analysis of scale and rotation invariance in the bag-of-features method for histopathological image classification, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Raza2011, 2011, pp. 66–74. arXiv:15334406. https://doi.org/10.1016/j.cogdev.2010.08.003.Personal.
[159] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, A. Vedaldi, Describing textures in the wild, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn. (2014) 3606–3613. ISSN 10636919. arXiv:1311.3618. https://doi.org/10.1109/CVPR.2014.461.
[160] M. Cimpoi, S. Maji, I. Kokkinos, A. Vedaldi, Deep filter banks for texture recognition, description, and segmentation, Int. J. Comput. Vis. 118 (1) (2016) 65–94. ISSN 15731405. arXiv:1507.02620. https://doi.org/10.1007/s11263-015-0872-3.
[161] A. Krizhevsky, I. Sutskever, H. Geoffrey E., ImageNet Classification with deep convolutional neural networks, Adv. Neural Inf. Proces. Syst. 25 (NIPS2012) (2012) 1–9. arXiv:1102.0183. https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf. https://doi.org/10.1109/5.726791.
[162] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, Pattern Recogn. 77 (2017) 354–377. arXiv:1512.07108. https://doi.org/10.1016/j.patcog.2017.10.013.
[163] C. Schmid, A Sparse texture representation using Af ne-invariant regions, Pattern Recogn. 27 (8) (2003) 1265–1278.
[164] C. Schmid, Local features and kernels for Classification of texture and object categories: a comprehensive study, Pattern Recogn. (2006) 0–7. http://ieeexplore.ieee.org/ielx5/10922/34371/01640452.pdf?tp=&arnumber=1640452&isnumber=34371%5Cnhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1640452. https://doi.org/10.1007/s11263-006-9794-4.
[165] Y. Xu, H. Ji, C. Fermüller, Viewpoint invariant texture description using fractal analysis, Int. J. Comput. Vis. 83 (1) (2009) 85–100. https://doi.org/10.1007/s11263-009-0220-6.
[166] M. Varma, A. Zisserman, A statistical approach to texture classification from single images, Int. J. Comput. Vis. 62 (1-2) (2005) 61–81. arXiv:/dx.doi.org/10.1023%2FA%3A1010933404324,. https://doi.org/10.1023/B:VISI.0000046589.39864.ee.
[167] M. Crosier, L.D. Griffin, Using basic image features for texture classification, Int. J. Comput. Vis. 88 (3) (2010) 447–460. https://doi.org/10.1007/s11263-009-0315-0.

[168] G. Sharma, F. Jurie, Local Higher-Order Statistics (LHS) describing images with statistics of local non-binarized pixel patterns, Comput. Vis. Image Underst. 142 (2016) 13–22. arXiv:1510.00542. https://doi.org/10.1016/j.cviu.2015.09.007.
[169] K. Skretting, J. Husøy, Texture classification using sparse frame-based representations, EURASIP J. Adv. Signal Process. 2006 (1) (2006) 052561. https://asp-eurasipjournals.springeropen.com/articles/10.1155/ASP/2006/52561. https://doi.org/10.1155/ASP/2006/52561.
[170] J. Ren, X. Jiang, J. Yuan, Noise-resistant local binary pattern with an embedded error-correction mechanism, IEEE Trans. Image Process. 22 (10) (2013) 4049–4060. https://doi.org/10.1109/TIP.2013.2268976.
[171] J. Yang, K. Yu, Y. Gong, T. Huang, Linear spatial pyramid matching using sparse coding for image classification, Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE. 2009, pp. 1794–1801.
[172] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, Y. Gong, Locality-constrained linear coding for image classification, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE. 2010, pp. 3360–3367.
[173] F. Perronnin, C. Dance, Fisher kernels on visual vocabularies for image categorization, Proc. {CVPR} (2006) https://doi.org/10.1109/CVPR.2007.383266.
[174] F. Perronnin, M. Douze, P. Pe, C. Schmid, J. Sa, Aggregating local image descriptors into compact codes, Analysis 34 (9) (2012) 1704–1716.
[175] R. Quevedo, L.G. Carlos, J.M. Aguilera, L. Cadoche, Description of food surfaces and microstructural changes using fractal image texture analysis, J. Food Eng. 53 (4) (2002) 361–371. https://doi.org/10.1016/S0260-8774(01)00177-7.
[176] Y. Pan, L. Liu, L. Yang, Y. Wang, Texture feature extracting method based on local relative phase binary pattern, Computer Science and Network Technology (ICCSNT), 2016 5th International Conference on, IEEE. 2016, pp. 749–753.
[177] S. Jana, S. Basak, R. Parekh, Automatic Fruit Recognition from Natural Images using Color and Texture Features, Devices for Integrated Circuit (DevIC), 2017, IEEE. 2017, pp. 620–624.
[178] M.M. Pawar, Identification of infected pomegranates using color texture feature analysis, Int. J. Comput. Appl. 43 (22) (2012) 30–34.
[179] S.A. Khoje, S.K. Bodhe, A. Adsul, Automated skin defect identification system for fruit grading based on discrete curvelet transform, Int. J. Eng. Technol. 5 (4) (2013) 3251–3256. https://doi.org/10.18520/cs/v112/i08/1704-1711.
[180] J.J. Lopez, M. Cobos, E. Aguilera, Computer-based detection and classification of flaws in citrus fruits, Neural Comput. & Applic. 20 (7) (2011) 975–981. https://doi.org/10.1007/s00521-010-0396-2.
[181] P. Cano Marchal, D. Martínez Gila, J. Gámez García, J. Gómez Ortega, Expert system based on computer vision to estimate the content of impurities in olive oil samples, J. Food Eng. 119 (2) (2013) 220–228. https://doi.org/10.1016/j.jfoodeng.2013.05.032.
[182] J. Vogel, B. Schiele, Semantic modeling of natural scenes for content-based image retrieval, Int. J. Comput. Vis. 72 (2) (2007) 133–157. https://doi.org/10.1007/s11263-006-8614-1.
[183] J. Von Kries, Influence of adaptation on the effects produced by luminous stimuli, handbuch der Physiologie des Menschen 3 (1905) 109–282.
[184] C.L. Novak, S.A. Shafer, Anatomy of a color histogram, Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on, IEEE. 1992,
[185] F. Mindru, T. Tuytelaars, L.V. Gool, T. Moons, Moment invariants for recognition under changing viewpoint and illumination, Comput. Vis. Image Underst. 94 (1-3) (2004) 3–27. http://linkinghub.elsevier.com/retrieve/pii/S1077314203001930. https://doi.org/10.1016/j.cviu.2003.10.011.
[186] J.-L. Shih, L.-H. Chen, Colour image retrieval based on primitives of colour moments, IEE Proc. - Vis., Image, and Signal Process. 149 (6) (2002) 370. http://digital-library.theiet.org/content/journals/10.1049/ip-vis_20020614. https://doi.org/10.1049/ip-vis:20020614.
[187] D.G. Lowe, Distinctive image features from scale invariant keypoints, Int. J. Comput. Vis. 60 (2004) 91–11020042. arXiv:0112017. http://portal.acm.org/citation.cfm?id=996342. https://doi.org/10.1023/B:VISI.0000029664.99615.94.
[188] K. Van De Sande, T. Gevers, C. Snoek, Evaluating color descriptors for object and scene recognition, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1582–1596. https://doi.org/10.1109/TPAMI.2009.154.
[189] a. Bosch, A. Zisserman, X. Munoz, Scene classification using a hybrid generative/discriminative approach, IEEE Trans. Pattern Anal. Mach. Intell. 30 (4) (2008) 712–727. https://doi.org/10.1109/TPAMI.2007.70716.
[190] J.M. Geusebroek, R. Van Den Boomgaard, A.W. Smeulders, H. Geerts, Color invariance, IEEE Trans. Pattern Anal. Mach. Intell. 23 (12) (2001) 1338–1350. https://doi.org/10.1109/34.977559.
[191] G.J. Burghouts, J.M. Geusebroek, Performance evaluation of local colour invariants, Comput. Vis. Image Underst. 113 (1) (2009) 48–62. https://doi.org/10.1016/j.cviu.2008.07.003.
[192] G. Pass, R. Zabih, J. Miller, Comparing images using color coherence vectors, Proceedings of the fourth ACM international conference on Multimedia (MULTIMEDIA '96) (1998) 1–14. http://dl.acm.org/citation.cfm?id=244130.244148. https://doi.org/10.1145/244130.244148.
[193] Z. Kuś, J. Cymerski, J. Radziszewska, A. Nawrat, Applying colour image-based indicator for object tracking, Studies in Systems, Decision and Control 106 (2018) 23–33. https://doi.org/10.1007/978-3-319-64674-9_2.
[194] Y. Wang, Y. Cui, S. Chen, P. Zhang, Huafang Huang, Study on fruit quality measurement and evaluation based on color identification, Proc.SPIE, 2009.7513–7513.
[195] O.K.M. Yahaya, M.Z. Matjafri, A.A. Aziz, A.F. Omar, Non-destructive quality evaluation of fruit by color based on RGB LEDs system, 2014 2nd International Conference on Electronic Design, ICED 2014, 2014. pp. 230–233. no. 1001. https://doi.org/10.1109/ICED.2014.7015804.
[196] Y. Makino, K. Goto, S. Oshita, A. Sato, M. Tsukada, A grading method for mangoes on the basis of peel color measurement using a computer vision system, Agric. Sci. 7 (6) (2016) 327–334. https://doi.org/10.4236/as.2016.76033\ignorespacesA.
[197] U.O. Dorj, M. Lee, S. seok Yun, An yield estimation in citrus orchards via fruit detection and counting using image processing, Comput. Electron. Agric. 140 (2017) 103–112. https://doi.org/10.1016/j.compag.2017.05.019.
[198] J. Blasco, S. Cubero, J. Gómez-Sanchs, P. Mira, E. Molt, Development of a machine for the automatic sorting of pomegranate (*Punica granatum*) arils based on computer vision, J. Food Eng. 90 (1) (2009) 27–34.
[199] X. Liming, Z. Yanchao, Automated strawberry grading system based on image processing, Comput. Electron. Agric. 71 (2010) S32–S39.
[200] A. Vidal, P. Talens, J.M. Prats-Montalbán, S. Cubero, F. Albert, J. Blasco, In-line estimation of the standard colour index of citrus fruits using a computer vision system developed for a mobile platform, Food Bioprocess Technol. 6 (12) (2013) 3412–3419. https://doi.org/10.1007/s11947-012-1015-2.
[201] S. Caldera, A. Rassau, Review of Deep Learning Methods in Robotic Grasp Detection, 2018, 1–22. (May). https://doi.org/10.20944/preprints201805.0484.v1.
[202] G. Zeng, Fruit and Vegetables Classification System Using Image Saliency and Convolutional Neural Network, Information Technology and Mechatronics Engineering Conference (ITOEC), 2017 IEEE 3rd, IEEE Press. 2017, pp. 613–617.
[203] M. Zaborowicz, P. Boniecki, K. Koszela, A. Przybylak, J. Przybył, Application of neural image analysis in evaluating the quality of greenhouse tomatoes, Sci. Hortic. 218 (2017) 222–229. https://doi.org/10.1016/j.scienta.2017.02.001.
[204] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on, IEEE. 2015,
[205] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, AAAI: Association for the Advancement of Artificial Intelligence, vol. 4, 2017. pp. 12.
[206] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. pp. 2818–2826.
[207] C. Zhang, C. Liu, X. Zhang, G. Almpanidis, An up-to-date comparison of state-of-the-art classification algorithms, Expert Systems with Applications 82 (2017) 128–150.
[208] K. Chatfield, K. Simonyan, A. Vedaldi, A. Zisserman, Return of the Devil in the Details: Delving Deep into Convolutional Nets, 2014, 1–11. http://arxiv.org/abs/1405.3531. arXiv:1405.3531. https://doi.org/10.5244/C.28.6.
[209] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-scale Image Recognition, 2014, 1–14. http://arxiv.org/abs/1409.1556. arXiv:1409.1556. https://doi.org/10.1016/j.infsof.2008.09.005.
[210] V. Andrearczyk, P.F. Whelan, Using filter banks in Convolutional Neural Networks for texture classification, Pattern Recogn. Lett. 84 (2016) 63–69. arXiv:1601.02919. https://doi.org/10.1016/j.patrec.2016.08.016.
[211] T.Y. Lin, A. Roychowdhury, S. Maji, Bilinear CNN models for fine-grained visual recognition, Proc. IEEE Int. Conf. Comput. Vis. 2015 Inter (2015) 1449–1457. arXiv:1504.07889. https://doi.org/10.1109/ICCV.2015.170.
[212] T.-Y. Lin, S. Maji, Visualizing and understanding deep texture representations, 2015, arXiv:1511.05197. http://arxiv.org/abs/1511.05197. https://doi.org/10.1109/CVPR.2016.305.
[213] Y. Gao, O. Beijbom, N. Zhang, T. Darrell, Compact Bilinear Pooling, 2015, (2) arXiv:1511.06062. http://arxiv.org/abs/1511.06062. https://doi.org/10.1109/CVPR.2016.41.
[214] J. Bruna, S. Mallat, Invariant scattering convolution networks, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1872–1886. arXiv:1203.1513. https://doi.org/10.1109/TPAMI.2012.230.
[215] T.H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, Y. Ma, PCANet: a simple deep learning baseline for image classification? IEEE Trans. Image Process. 24 (12) (2015) 5017–5032. arXiv:1404.3606. https://doi.org/10.1109/TIP.2015.2475625.
[216] M.M. Devikar, M.K. Jha, Segmentation of Images Using Histogram Based, 6. 2013, 225–231. (1).
[217] VaniAshok, D.D. Vinod, Using K-means cluster and fuzzy C means for defect, Int. J. Comput. Eng. & Technol. (IJCET), ISSN (July) (2014) 11–19.
[218] Imari, A. Kuno, Y. Sato, Evaluating Freshness of Produce Using Transfer Learning, Frontiers of Computer Vision (FCV), 2015 21st Korea-Japan Joint Workshop on, IEEE Press. 2015, pp. 1–4.
[219] M.S. Hossain, G. Muhammad, S. Umar, Improving consumer satisfaction in smart cities using edge computing and caching: a case study of date fruits classification, Futur. Gener. Comput. Syst. 88 (2018) 333–341. S0167739X1733056X. http://www.sciencedirect.com/science/article/pii/S0167739X1733056X.
[220] K. Sendin, M. Manley, P.J. Williams, Classification of white maize defects with multispectral imaging, Food Chem. 243 (June 2017) (2018) 311–318. https://doi.org/10.1016/j.foodchem.2017.09.133.
[221] S. Sun, N. An, X. Zhao, M. Tan, A PCA CCA network for RGB-D object recognition, 2018, 1–12. (February). https://doi.org/10.1177/1729881417752820.
[222] X. Wen, H. Liu, G. Yan, F. Sun, Weakly paired multimodal fusion using multilayer extreme learning machine, Soft. Comput. 22 (11) (2018) 3533–3544. https://doi.org/10.1007/s00500-018-3108-y.

[223] Y. Yin, H. Li, Multi-view CSPMPR-ELM feature learning and classifying for RGB-D object recognition, Clust. Comput. (2018) 1–11. https://doi.org/10.1007/s10586-018-1695-0.
[224] M. Schwarz, A. Milan, A.S. Periyasamy, S. Behnke, RGB-D object detection and semantic segmentation for autonomous manipulation in clutter, 2018. https://doi.org/10.1177/0278364917713117.
[225] L. Li, B. Qian, J. Lian, W. Zheng, Y. Zhou, Short Papers Traffic Scene Segmentation Based on RGB-D Image and Deep Learning, 19. 2018, 1664–1669. (5).
[226] A.G.S.O.J. Garcia-rodriguez, M. Cazorla, M. Cazorla, Interactive 3D object recognition pipeline on mobile GPGPU computing platforms using low-cost RGB-D sensors, J. Real-Time Image Proc. 14 (3) (2018) 585–604. https://doi.org/10.1007/s11554-016-0607-x.
[227] F. Zhou, Y. Hu, X. Shen, MSANet: multimodal self-augmentation and adversarial network for RGB-D object recognition, Vis. Comput. (2018) http://link.springer.com/10.1007/s00371-018-1559-x. https://doi.org/10.1007/s00371-018-1559-x.