

A SYSTEM FOR ASSESSING FRUIT BUNCH QUALITY

FIELD OF INVENTION

- 5 The present invention relates to a system for assessing fruit bunch quality. More particularly, the present invention relates a system for assessing oil palm fresh fruit bunch quality utilizing computer vision, image processing and deep learning techniques to automate the evaluation of the fruit bunch characteristics.

10 BACKGROUND OF THE INVENTION

- Assessing fruit bunch qualities in the agriculture industry posts multifaceted challenges. The variability in fruit characteristics within a single bunch, influenced by factors such as ripeness, size, and health complicates standardized evaluations. Traditional methods
- 15 rely on visual inspection lack precision and are time-consuming. Moreover, the subjective nature of human assessment can lead to inconsistencies in grading. Incorporating technology, such as automated sorting systems or machine learning algorithms, aims to enhance accuracy, yet implementing these technologies widely remains a challenge due to cost constraints and limited accessibility, especially in
- 20 smaller farming operations. Furthermore, factors like environmental conditions, pest infestations, and variations in cultivars add layers of complexity to consistent and accurate assessments. Addressing these challenges necessitates a cohesive approach that combines technological advancements, educational initiatives, and policy support within the agriculture sector to streamline the assessment process, ensuring consistent
- 25 production of high-quality fruits while accommodating the diverse needs of the industry.

Many technologies have been implemented to improve on systems for assessing fruit bunch quality. A publication by Bina Nusantara University titled, “Annotated Datasets

- of Oil Palm Fruit Bunch Piles for Ripeness Grading Using Deep Learning” recites the a new complete dataset obtained directly from palm oil mills in the form of videos and images with different categories in accordance with the real conditions faced by the grading section of the palm oil mill. The video dataset consists of 45 videos with a
5 single category of FFB videos and 56 videos with a collection of FFB with multiple categories for each video. Videos are collected using a smart phone with a size of 1280 × 720 pixels with .mp4 format. In addition, this dataset has also been annotated and labelled based on the maturity level of oil palm fruit with 6 categories, which are unripe, under-ripe, ripe, overripe, empty bunches and abnormal fruit.
- 10 Another publication by Cardiff University in collaboration with Huazhong Agricultural University titled, “Peach Ripeness Classification Based on a New One-Stage Instance Segmentation Model” discloses a large and high-quality peach dataset called NinePeach, and a new one-stage instance segmentation model. The NinePeach dataset
15 aims to reproduce real-world field conditions, encompassing various factors that can significantly influence the accuracy of peach detection, such as varying natural light intensity, instances of multiple fruit adhesion, and occlusion caused by stems and leaves. The proposed one-stage segmentation model does not require Region Proposal Network (RPN) to generate bounding box proposals, it directly identifies object
20 instances by their center locations and sizes and predict their category at the same time. The proposed model also incorporates channel attention and spatial attention mechanisms to enhance object detection capabilities in crucial channels and spatial locations.
- 25 An international publication WO20120239597A2 discloses a fruit ripeness grading system using computer vision application in agricultural quality inspection to ensure ripeness category of fruit. The fruit ripeness grading system comprises a housing having an enclosure for scanning process, an illumination means with optical lens illumination filter provided at the enclosure of the housing, a camera provided at a top portion of the

enclosure of the housing, a feeding device for conveying fruit samples to the housing, a conveyer speed inverter, a processing unit for processing and analyzing the fruit sample image, a data acquisition interface provided between the camera and the processing unit, wherein the feeding device is further provided with a separator
5 controlled by a USB controller and supported by a compressor for controlling and separating the fruit samples, and wherein the processing unit is further provided with a disk top computational unit which serves to transfer data to a computer.

Further, a Chinese patent with publication number CN111753794A discloses a fruit
10 quality classification method, an electronic equipment, and a readable storage medium, and relates to the technical field of data processing. The method comprises the steps of collecting IR images of target fruits at all angles; carrying out scar detection on the target fruit based on the IR image, and determining a suspected scar area of the target fruit; identifying the suspected scar area through a neural network model, and
15 determining the type of the suspected scar; and performing quality classification on the target fruit based on the flaw classification and other quality classification parameter information aiming at the target fruit.

20 SUMMARY OF INVENTION

An object of the present invention is to provide a system for assessing fruit bunch quality that ensures consistent and objective evaluations while minimizing the risk of human error due to subjectivity.

25

Another object of the present invention is to provide real-time analysis of the fruit bunch quality, allowing fruit bunches to be accepted and rejected promptly, enhancing overall operational efficiency and product quality.

It is also an object of the present invention to provide a system that can accurately label each instance of the fruit bunch within a group of fruit bunches in a captured image and the associating characteristics for each fruit bunch.

- 5 In one aspect of the present invention, there is provided a system for assessing fruit bunch quality, the system comprising an image capturing module, for capturing images of one or more fruit bunches; a processing module, for processing the captured images, the processing module comprising: a calibration module, for performing pixel calibration within a field of view of the captured images; an object detection module,
10 for generating at least one bounding box on the captured images, such that each bounding box contains one or more regions of interest on the fruit bunch; and a segmentation module, for generating at least one polygon on the captured images, such that each polygon contains one or more regions of interest on the fruit bunch;; wherein the region of interests within the bounding boxes corresponds to further characteristics
15 of the fruit bunch, including ripeness, damage type, or cleanliness of the fruit bunch, whereas the regions of interests within the polygons corresponds to physical characteristics of the fruit bunch including a size of the fruit bunch and a type of the fruit bunch, in which information from both of the object detection module and the segmentation module are collectively assessed and evaluated to be output thereafter.

20

Preferably, the image capturing module is further configured to capture frames containing the fruit bunch from a video medium.

- Preferably, the object detection module further comprising a ripeness classification
25 module, a damage detection module, and a cleanliness module.

Preferably, the ripeness classification module is configured to determine a classification of the fruit bunch.

Preferably, the damage detection module is configured to detect an extent of damages done to the fruit bunch and classify the fruit bunch as damaged or undamaged thereafter.

5 Preferably, the cleanliness module is configured to identify a cleanliness level of the fruit bunch, such that the fruit bunch is subsequently labelled dirty or not dirty.

Preferably, the segmentation module further comprising a size estimation module and a stalk detection module.

10 Preferably, the size estimation module is configured to estimate a size of the fruit bunch, such that the fruit bunch is subsequently labelled as small or not small.

Preferably, the stalk detection module is configured to detect whether the fruit bunch has a long stalk or not a long stalk.

15

Preferably, the system further comprising an association module which is configured to associate the respective characteristics as determined by the object detection module and the segmentation module to each fruit bunch within the captured images.

20 In another aspect of the present invention, there is provided a method for assessing fruit bunch quality comprising the steps of: capturing, by an image capturing module, images of one or more fruit bunches; processing, by a processing module, the captured images using the steps of: performing, by a calibration module, pixel calibration within a field of view of the captured images; generating, by an object detection module, at least one
25 bounding box on the fruit bunch, such that each bounding box contains one or more regions of interest on the fruit bunch; and generating, by a segmentation module, at least one polygon on the captured images, such that each polygon contains one or more regions of interest on the fruit bunch; wherein the region of interests within the bounding boxes corresponds to further characteristics of the fruit bunch, including

ripeness, damage type, or cleanliness of the fruit bunch, whereas the regions of interests within the polygons corresponds to physical characteristics of the fruit bunch including a size of the fruit bunch and a type of the fruit bunch, in which information from the object detection module and the segmentation module are collectively assessed and
5 evaluated to be output thereafter.

Preferably, the pixel calibration is performed by the steps of capturing a fiducial marker of a predetermined size within the captured image; determining a location and a pixel size of the fiducial marker using a computer vision technique; and calculating a distance
10 per pixel based on the determined location and pixel size of the fiducial marker.

Preferably, the characteristics of the fruit bunch are determined by the steps of: determining, by a ripeness classification module, a classification of ripeness of the fruit bunch; detecting, by a damage detection module, damages on the fruit bunch; and
15 identifying, by a cleanliness module, a cleanliness level of the fruit bunch.

Preferably, the classification of ripeness of the fruit bunch is determined by the steps of: applying an object detection technique to obtain confidence scores on the bounding boxes; and classifying the ripeness of the fruit bunch based on the confidence scores of
20 each bounding box.

Preferably, the damages on the fruit bunch are detected by the steps of: identifying numbers of damaged fruitlets within the fruit bunch; comparing the numbers of damaged fruitlets to a predetermined threshold range; and designating a label
25 identification to the fruit bunch based on the comparison of damaged fruitlets to the predetermined threshold range.

Preferably, the cleanliness level of the fruit bunch is identified by the steps of: identifying one or more areas of contamination on the fruit bunch within the regions of

interest based on the bounding boxes of the captured images; measuring the areas of contamination on the fruit bunch; comparing the measured areas of contamination to a predetermined threshold range; and designating another label identification to the fruit bunch based on the comparison of the measured areas of contamination to the predetermined threshold range.

Preferably, the size of the fruit bunch is determined by the steps of: measuring a contour area from any one segment within the regions of interest in the polygons of the captured images of the fruit bunch; comparing the contour area to a predetermined threshold; and designating a label identification to the fruit bunch based on the comparison of the contour area to the predetermined threshold.

Preferably, the size of the fruit bunch is also determined by the steps of: calibrating a binocular image capturing module to determine intrinsic parameters and extrinsic parameters of said binocular image capturing module; capturing one or more images of the fruit bunch; rectifying the captured images to ensure at least two captured images are disposed on a same scan line; identifying corresponding features or key points within the rectified images; calculating a disparity map between the rectified images; converting the disparity map to a depth map; generating a three-dimensional point cloud of the fruit bunch based on values obtained from the depth map; normalizing the three-dimensional point cloud using points representing a conveyor floor; and estimating a volume of the fruit bunch based on the generated three-dimensional point cloud.

Preferably, the type of fruit bunch is determined by the steps of: applying one or more label identification on segments in the regions of interest of the captured images correlating to at least one stalk of the fruit bunch and fruitlets of the fruit bunch; converting the segments containing the fruitlets of the fruit bunch in the captured images into a convex hull; identifying an intersection area between the convex hull and the segments correlating to the stalk of the fruit bunch; measuring a difference between

- the segment correlating to the stalk of the fruit bunch and the intersection area, in which the difference corresponds to an excessive stalk area; generating a minimum bounding box to the excessive stalk area; converting a height of the excessive stalk to metric dimensions using distance per pixel as calibrated by the calibration module based on
- 5 the dimensions of the minimum bounding box; comparing the height of the excessive stalk to a predetermined threshold; and designating a label identification to the fruit bunch based on the comparison of the height of the excessive stalk to the predetermined threshold.
- 10 Preferably, the information of the characteristics obtained from the object detection module and the segmentation module are associated by the steps of matching characteristics of the fruit bunch with at most one fruit bunch in the captured images, the matching includes the steps of defining at least two lists of nodes representing the characteristics of the fruit bunches and individual fruit bunches; iterating through the
- 15 list of nodes using a suitable association algorithm; identifying and designating a label identification of any one of the respective characteristics of the individual fruit bunch to the individual fruit bunch having an intersection percentage value than a minimum intersection percentage value.
- 20 One skilled in the art will readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The embodiment described herein is not intended as limitations on the scope of the invention.

25 **BRIEF DESCRIPTION OF THE DRAWINGS**

For the purpose of facilitating an understanding of the invention, there is illustrated in the accompanying drawing the preferred embodiments from an inspection of which when considered in connection with the following description, the invention, its

construction and operation and many of its advantages would be readily understood and appreciated.

FIG. 1 is a block diagram illustrating a preferred embodiment of a system for assessing
5 fruit quality.

FIG. 2 is a flowchart illustrating an exemplary embodiment of a method for assessing the fruit quality based on its system.

10 **FIG. 3** is a flowchart illustrating an exemplary embodiment of an object detection module for performing an inference on the fruit bunch for detecting characteristics such as ripeness classification, damage detection, and cleanliness of the fruit bunch.

FIG. 4 is a flowchart illustrating an exemplary embodiment of a stalk detection module
15 of a segmentation module for determining a type of said fruit bunch.

FIG. 5 is a flowchart illustrating an exemplary embodiment of a size estimation module of the segmentation module for determining the size of the fruit bunch.

20 **FIG. 6** is a flowchart illustrating an exemplary embodiment of a method for associating the fruit bunch characteristics and individual fruit bunch.

DETAILED DESCRIPTION OF THE INVENTION

25 Hereinafter, the invention shall be described according to the preferred embodiments of the present invention and by referring to the accompanying description and drawings. However, it is to be understood that limiting the description to the preferred embodiments of the invention is merely to facilitate discussion of the present invention and it is envisioned that those skilled in the art may devise various modifications

without departing from the scope of the appended claim.

Hereinafter, the invention shall be described according to the preferred embodiments of the present invention and by referring to the accompanying description and drawings.

5 However, it is to be understood that limiting the description to the preferred embodiments of the invention is merely to facilitate discussion of the present invention and it is envisioned that those skilled in the art may devise various modifications without departing from the scope of the appended claim.

10 It will be understood that each block of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general-purpose computer, special purpose computer, or other programmable data processing apparatus to produce
15 such a machine, such that the instructions, that execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks.

20 These computer program instructions may also be stored in a computer-readable memory that can direct a computer or other programmable data processing apparatus to function in a particular manner, such that the instructions stored in the computer-readable memory produce an article of manufacture including instruction means that implement the function/act specified in the flowchart and/or block diagram block or
25 blocks.

The computer program instructions may also be loaded onto a computer or other programmable data processing apparatus to cause a series of operational steps to be performed on the computer or other programmable apparatus to produce a computer

implemented process such that the instructions that execute on the computer or other programmable apparatus provide steps for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks.

- 5 The invention will now be described in greater detail, by way of example, with reference to the drawings.

FIG. 1 illustrates a preferred embodiment of a system for assessing fruit bunch quality, the system comprising an image capturing module **100**, for capturing images of one or
10 more fruit bunches, in which the image capturing module may be any type of camera such as, for example, a type of charge-coupled device (CCD) camera or a complementary metal oxide semiconductor (CMOS) camera, which has advantages of high noise immunity, very low static power consumption and high efficiency in capturing images in a low light environment. Additionally, the image capturing module
15 **100** may also be configured to record videos of the fruit bunches moving along the conveyor in real-time, in which frames that are detected to contain the fruit bunches are recorded as captured images to be processed by a processing module **200** thereafter. Referring to **FIG. 2**, an imaging chamber may be arranged along a conveyor in which the fruit bunches are loaded onto at Step **201**, such that the image capturing module **100**
20 may be disposed at a position above the conveyor, with its lens facing downwards to the conveyor to capture a top view of the fruit bunches as the conveyor moves said fruit bunches along its path, as illustrated in Step **202**. Preferably, a light source may be implemented with the image capturing module **100** to illuminate on the fruit bunches, in which a broad spectrum of visible light encompassing wavelengths across the visible
25 light, from blue to red, may provide a well balanced and natural-looking illumination to accurately capture colors, textures, and details of the fruit bunches. Besides that, curtains may be applied on any sides of the imaging chamber to ensure consistent lighting may be applied across all images captured by the image capturing module **100**. Additional parameters such as white balance, shutter speed, and aperture of the image

capturing module **100** may be adjusted accordingly in tandem with the illumination of the light source as well as conveyor speed, such as illustrated in Step **203**. In turn, the image capturing module **100** may be able to capture images of the fruit bunches when said fruit bunches are static or in a fixed position, or when the fruit bunches are moving
 5 along the conveyor.

In a preferred embodiment, the image capturing module **100** may be trained and programmed with an object tracking mechanism using tracking algorithms, such that the image capturing module **100** is able to accurately track any of the fruit bunches that
 10 enter the imaging chamber. For example, tracking algorithms such as MeanShift, Kanade-Lucas-Tomasi (KLT) or advanced algorithms such as Simple Online and Realtime Tracking (SORT) or DeepSort may be utilized for tracking the movement of the fruit bunches, whereby the tracking algorithms may be trained with coordinates of the fruit bunches from a starting position of the conveyor to the imaging chamber.
 15 Preferably, the object tracking mechanism operates in conjunction with a line detection mechanism, which is configured to define a virtual line whether horizontally or vertically across a field of view of the image capturing module. Also preferably, the virtual line may be drawn perpendicularly to the movement of the fruit bunches on the conveyor. The image capturing module **100** may then be configured to detect if any of
 20 the fruit bunches crosses the drawn virtual line by means of comparing the coordinates of the fruit bunches with the coordinates of the virtual line. Upon detection of the fruit bunches crossing the virtual line, the processing module **200** may then be triggered to initiate the processing of the captured images by means of a calibration module **300**, an object detection module **400**, and a segmentation module **800**, whereby information
 25 from both modules is assessed and evaluated for various characteristics of the fruit bunch thereafter.

According to **FIG. 2**, the calibration module **300** may be configured to perform pixel calibration within a field of view of the captured images prior to further processing by

the object detection module **400** and the segmentation module **800**. Preferably, the pixel calibration module **300** may be triggered at Step **205** once a first cycle of the entire process has been determined at Step **204** to ensure that the captured images are properly calibrated for accurate processing thereafter. The pixel calibration may be performed

5 by firstly capturing a fiducial marker of a predetermined size within the captured image. Subsequently, a location and a pixel size of the fiducial marker may be determined using a computer vision technique, such as Harris corner detection or Shi-Tomasi corner detection algorithms, whereby these algorithms identify distinctive points or corners within the fiducial marker which aids in determining its location within the image. Once

10 the fiducial marker's corners or distinctive features are identified, further techniques such as perspective transformation or homography may be employed to map the fiducial marker's known dimensions to its corresponding pixels in the captured image, which then facilitates the calculation of pixel size and location of the fiducial marker within the captured image. Upon determining the location and pixel size of the fiducial

15 marker, a distance per pixel may then be calculated at Step **206** to complete the pixel calibration process. Based on **FIG. 2**, upon completion of pixel calibration by the calibration module **300**, the processing module **200** may then trigger the object detection module **400** and the segmentation module **800** to perform an inference each on the captured images at Steps **207** and **208** respectively, such that the assessment of

20 the fruit bunch characteristics from the inference may be performed at Step **209** thereafter. The inference processes may be performed both sequentially or in parallel with one another. At Step **210**, the fruit bunches that were assessed may be accepted or rejected depending on results obtained from the assessment.

25 **FIG. 3** illustrates an exemplary method for performing the inference of the fruit bunch by the object detection module **400**, as illustrated in Step **207** of **FIG. 2**. Preferably, the object detection module **400** may be configured to generate at least one bounding box at Step **301** on the detected fruit bunch within the captured images, such that each bounding box may contain one or more regions of interest on the fruit bunch. Preferably,

the region of interest within the bounding boxes may correspond to characteristics of the fruit bunch, which includes, but is not limited to, ripeness, damage type, or cleanliness of the fruit bunch. The object detection module **400** may employ deep learning techniques or models alongside computer vision techniques to accurately

5 determine the various characteristics of the fruit bunch as mentioned above. In a preferred embodiment, the object detection module **400** may be initially trained with a plurality of images containing fruit bunches having corresponding bounding boxes with regions of interest encompassing the fruit bunches. Once sufficient training is conducted, the object detection module **400** may then be deployed for operation.

10 Additionally, the object detection module **400** may not generate any bounding boxes if the captured images do not contain the fruit bunches. In a preferred embodiment, the object detection module **400** may comprise of a ripeness classification module **500** to determine a classification of ripeness of the fruit bunches, a damage detection module **600** to detect any type of damages on the fruit bunches, and a cleanliness module **700**

15 to identify a cleanliness level of the fruit bunches, all of which will be discussed further herein.

Referring to **FIG. 3**, the ripeness classification module **500** may determine the ripeness classification of any one of the fruit bunches within the captured image by firstly

20 applying an object detection technique, such as non-maximum suppression, on each bounding box generated on the captured images at Step **302**. This technique of non-maximum suppression may eliminate redundant or overlapping bounding boxes or regions of interest generated by the object detection module **400**, keeping only the most confident and accurate bounding box to be used in the classification of ripeness of the

25 fruit bunch. The ripeness classification module **500** may utilize model architectures such as You Only Look Once (YOLO), or Faster Region-based Convolutional Neural Network (R-CNN) to provide an “objectness” score as the confidence score, which reflects how likely the region of interest contains an object. A Vision Transformer (ViT) architecture, which is a non-convolutional layer structure may also be utilized, in which

it divides the captured images into fixed-size patches rather than processing an entire captured image all at once, and treating the patches as “tokens” similar to words in natural language processing (NLP), such that the patches may then be linearly embedded and processed through transformer layers to capture spatial relationships

5 between them. Optionally, additional computer vision techniques such as color-based segmentation may supplement the “objectness” score to further improve on the confidence score for the bounding boxes by defining a color range or color model that represents a specific color corresponding to the different variations of ripeness of the fruit bunches. Once a confidence score has been identified for each bounding box

10 within the captured image, the highest confidence score will determine the classification of ripeness of the fruit bunch, such that a label identification may then be designated for the fruit bunch which includes, but is not limited to labels such as, ripe **303a**, unripe **303b**, overripe **303c**, empty bunch **303d**, and abnormal **303e**, such as illustrated in Step **303**.

15

In a preferred embodiment, the damage detection module **600** detects an extent of the damages done to the fruit bunches by firstly identifying numbers of damaged fruitlets within the fruit bunch based on the captured images at Step **304**. In this instance, damage to the fruit bunch may usually be caused by pest infestation such as rats, and

20 clear identification of damaged or undamaged fruit bunches are essential in the assessments of fruit bunch quality in the agriculture industry to ensure quality fruit bunches are distributed to the consumers. At Step **305**, the numbers of damaged fruitlets in the fruit bunch may then be compared to a predetermined threshold range which may be configured beforehand based on any type of model training or the likes. Similar to

25 the Step **303**, a label identification may be designated to the fruit bunch based on the comparison of damaged fruitlets to the predetermined threshold range. If the numbers of damaged fruitlets exceed the predetermined threshold range, the fruit bunch may then be designated with the label “damaged” **305a**. If the numbers of damaged fruitlets do not exceed the predetermined threshold range, the fruit bunch may then be

designated with the label “not damaged” **305b**.

In a preferred embodiment, the cleanliness module **700** may be configured to identify the cleanliness level of the fruit bunches in the captured images as mentioned above.

- 5 Contamination such as mud, leaves, dirt, debris, animal droppings or the likes that appear on a surface of the fruit bunch are key factors in identifying the cleanliness level of the fruit bunches. At Step **306**, one or more areas of contamination on the fruit bunch within the regions of interest based on the bounding boxes in the captured images are identified. Subsequently, the areas of contamination on said fruit bunches are then
10 measured accordingly at Step **307**, using the following equation:

$$\text{Contamination Area Percentage} = \frac{\text{Contamination area}}{\text{Fruit bunch area}}$$

- Accordingly, the measured contamination area percentage may then be compared to a
15 predetermined threshold range as illustrated in Step **308**. Another label identification may then be designated accordingly based on the comparison between the measured areas of contamination and the predetermined threshold range, whereby a label “dirty” **308a** may be designated to the fruit bunch if the measured areas of contamination exceed the predetermined threshold range. Accordingly, a label “not dirty” **308b** may
20 then be designated to the fruit bunch if the measured areas of contamination do not exceed the predetermined threshold range.

- In a preferred embodiment, the segmentation module **800** may perform its inference on the fruit bunches as illustrated in **FIG. 4** and **FIG. 5**. Preferably, the segmentation
25 module **800** may be configured to generate polygons within the captured images, in which the polygons represent one or more regions of interest which corresponds to physical characteristics of the fruit bunch, which includes, but is not limited to, a type and a size of the fruit bunch by means of a stalk detection module **900** and a size estimation module **1000** respectively. Also preferably, the segmentation module **800**

utilizes a computer vision technique known as pixel-wise segmentation, also known as semantic segmentation, which involves dividing an image into distinct regions or segments, assigning each pixel in the image to a specific class or category, thereby creating a pixel-level classification map. The objective of pixel-wise segmentation is to

5 label each pixel in the captured image with a particular class, such as identifying objects, shapes, or regions within the image. In this instance, it is to identify the size of the fruit bunch and the type of the fruit bunch, particularly whether the fruit bunch has an excessive stalk or not. There are various approaches to produce pixel-wise segmentation of the fruit bunch, which includes fully convolutional networks (FCN),

10 U-Net, Mask R-CNN, SegFormer or the likes. Particularly, the size of the fruit bunch to be determined is whether the fruit bunch is small or not small, whereas the type of fruit bunch to be determined is whether the fruit bunch has a long stalk or no long stalk, which will be discussed further herein. Additionally, the segmentation module **800** may operate either in sequence or in tandem with the object detection module to determine

15 the different characteristics of the fruit bunch as mentioned above.

Based on **FIG. 4**, the type of fruit bunch may be determined by the stalk detection module **900** by firstly applying one or more label identifications on segments in the regions of interest of the captured images that correspond to at least one stalk of the

20 fruit bunch and to the fruitlets of the fruit bunch, as illustrated in Step **401**. At Step **402**, the segments containing the fruitlets of the fruit bunch in the captured images may be converted into a convex hull area, which refers to a smallest convex shape or polygon that encloses a given set of points within a two-dimensional or three-dimensional space. Subsequently at Step **403**, an intersection area between the convex hull area and the

25 segments correlating to the stalk of the fruit bunch may be identified, such that a difference between the segments correlating to the stalk of the fruit bunch and the intersection area may be measured at Step **404**, in which the difference may correspond to an area containing excessive stalk. At Step **405**, a minimum bounding box may be generated to entirely enclose the excessive stalk area, in which a height of the excessive

stalk may then be converted into metric dimensions thereafter by using the distance per pixel as calibrated by the calibration module **300**, such as illustrated in Step **406**. Preferably, the height of the excessive stalk to be converted may be based on dimensions of the minimum bounding box. The height of the excessive stalk may then

5 be compared to a predetermined threshold range at Step **407**, such that a label identification may be designated to the fruit bunch based on the comparison of the height of the excessive stalk to the predetermined threshold range. If the height of the excessive stalk exceeds the predetermined threshold range, the fruit bunch may be designated with a “long stalk” **407a** label. Otherwise, the fruit bunch may be designated

10 with a “not long stalk” **407b** label instead if the height of the excessive stalk does not exceed the predetermined threshold range.

In a preferred embodiment, as illustrated in **FIG. 5**, the size of the fruit bunch may be determined by the size estimation module **1000** by firstly measuring a contour area from

15 any one segment within the region of interests in the polygons generated by the segmentation module **800** of the captured images of the fruit bunch, as illustrated in Step **501**. Preferably, the segment may refer to an outline of a shape of the fruit bunch within the bounding boxes of the captured images. Upon measuring the contour area of said fruit bunch, the contour area may then be compared to a predetermined threshold

20 range at Step **502**, and subsequently have a label identification assigned to the fruit bunch based on the comparison between the measured contour area and the predetermined threshold range. If the measured contour area exceeds the predetermined threshold range, the fruit bunch may be labelled as “not small” **502b**, otherwise the fruit bunch may then be labelled as “small” **502a**, if the measured contour area does not

25 exceed the predetermined threshold range.

In another embodiment, the size of the fruit bunch may also be determined by obtaining depth information of said fruit bunch. Preferably, the image capturing module **100** may be a binocular image capturing module **100**, also known as a stereo vision camera that

can produce an RGBD image, which is an image that combines color information from Red, Green, Blue (RGB) channels with depth information (D). At Step **503**, the binocular image capturing module **100** may be firstly calibrated to determine its intrinsic parameters such as, focal length and optical center, and extrinsic parameters such as relative position and translation, of each image capturing module used. A plurality of images of the fruit bunches may be subsequently captured at Step **504**, with at least one image captured from each camera of the binocular image capturing module **100**. At Step **505**, the captured images may then be rectified to ensure corresponding points within at least two captured images are disposed on a same scan line, as configured by the binocular image capturing module **100**. Subsequently at Step **506**, corresponding features or key points within the captured images that were rectified may be identified for feature matching, such that a disparity map may be calculated at Step **507** thereafter. Preferably, the disparity map may represent a horizontal pixel-wise difference between the corresponding key points in the captured images used for feature matching. At Step **508**, the disparity map may then be converted to a depth map using a known baseline value, such as a distance between at least two image capturing modules and the focal length of said image capturing modules as calibrated in Step **503**. The depth may be calculated as follows:

$$Depth = \frac{Baseline \times Focal\ Length}{Disparity}$$

At Step **509**, values obtained from the depth calculation may then be used to generate a three-dimensional point cloud, in which each pixel in the captured images may correspond to a three-dimensional point in space. (x, y) coordinates from the captured images and corresponding depth values may then be used to calculate the (x, y, z) coordinates in said three-dimensional point cloud. At Step **510**, the three-dimensional point cloud may be normalized by using points representing the conveyor floor, such that each point constitutes the height of the point above said conveyor floor. The following is the equation to calculate the height of each point:

$$\text{Height at each pixel} = \text{Depth of ground} - \text{depth measured in each point}$$

Once the three-dimensional point cloud of the fruit bunch has been fully generated at
 5 Step **509**, the volume of said fruit bunch may then be estimated by means of
 voxelization, which is dividing the three-dimensional space into small cubes called
 voxels and counting occupied voxels, or by surface reconstruction of the three-
 dimensional point cloud, such as illustrated in Step **511**. Further, the points in the three-
 dimensional point cloud that correspond to the fruit bunch may be cropped or selected
 10 by assigning individual points to said fruit bunch through the segmentation module
1000. Upon estimating the volume of the fruit bunch at Step **511**, the size estimation
 module **1000** may then compare said estimated volume to a predetermined threshold at
 Step **512**, such that the label identification of “small” **512a** may be assigned accordingly
 if said estimated fruit bunch volume does not exceed said predetermined threshold.
 15 Otherwise, the label identification of “not small” **512b** may be assigned if the estimated
 fruit bunch volume exceeds the predetermined threshold.

Upon determining all the individual characteristics pertaining to the ripeness, the
 damage type, the cleanliness level, the size, and the type of the fruit bunches in the
 20 captured images, the processing module **200** further assesses the information gathered
 from the object detection module **400** and the segmentation module **800** to associate
 the features to the fruit bunches accordingly by means of spatial distribution of the
 different characteristics over the captured images. For example, when there are two fruit
 bunches within a single captured image, the system is required to differentiate the
 25 characteristics of one fruit bunch against the characteristics of another fruit bunch. For
 the ripeness classification, the bounding box covers the entirety of the fruit bunch.
 However, for the other characteristics such as the damage type, the cleanliness level,
 the size, and the type of the fruit bunch, each of those characteristics cover a partial and
 intersecting area over the fruit bunch in the captured images.

Referring to **FIG. 6**, the association of the fruit bunch characteristics obtained from the object detection module **400** and the segmentation module **800** as illustrated at Step **209** in **FIG. 2**, may be conducted by means of a bipartite matching process, which is a

5 process that efficiently associates objects or regions by the object detection module **400** with ground truth annotations or with results obtained from the segmentation module **800**. Preferably, the obtained information of the fruit bunch characteristics may be associated by matching said characteristics of the fruit bunch with at most one fruit bunch in the captured images at Step **601**. At Step **602**, at least two lists of nodes may

10 then be defined as input, in which one node may comprise a list of individual fruit bunches as detected by the object detection module **400**, whereas another node may comprise the different characteristics as determined by the object detection module **400** and the segmentation module **800**.

15 At Step **603**, the list of nodes is iterated using a suitable association algorithm of choice for the respective fruit bunches and different characteristics corresponding to said fruit bunches. For example, a maximum intersection algorithm may be implemented, such that said algorithm may firstly determine if the intersection area percentage meets a minimum intersection area percentage criterion at Step **604**. Step **603** may be repeated

20 until the intersection area percentage of the fruit bunch in the captured images meets said minimum intersection area percentage criterion. Upon meeting the criterion, the algorithm may match any one of the fruit bunch characteristics with the fruit bunch where the intersection area is largest and check if the intersection area percentage exceeds the maximum intersection area percentage as illustrated in Step **605**. The

25 maximum intersection area percentage may remain as 0 if the intersection area percentage does not exceed the maximum intersection area percentage, such as illustrated in Step **606** and repeats Step **603**. If the maximum intersection area exceeds the intersection area percentage, said intersection area percentage will be designated with the value of the maximum intersection area percentage as illustrated in Step **607**.

A label identification of any one of the characteristics of the individual fruit bunch may be designated to according to the list of nodes where the intersection area percentage value is higher than the minimum intersection percentage value through a performed comparison at Step **68**. If the intersection area percentage value is found to be less than
5 the minimum intersection area percentage value, the label identification designated to said fruit bunch may be written as “None” at Step **609**. Otherwise, if the maximum intersection area percentage exceeds the minimum intersection area percentage, the label identification of any one of the characteristics of the individual fruit bunch may then be designated to the bounding box of the fruit bunch, such as illustrated at Step
10 **610**, Based on the label identification designated to each fruit bunch, the fruit bunches may be accepted or rejected thereafter, such as illustrated in Step **210** of **FIG. 2**.

The present disclosure includes as contained in the appended claims, as well as that of the foregoing description. Although this invention has been described in its preferred
15 form with a degree of particularly, it is understood that the present disclosure of the preferred form has been made only by way of example and that numerous changes in the details of construction and the combination and arrangements of parts may be resorted to without departing from the scope of the invention.

CLAIMS

1. A system for assessing fruit bunch quality, the system comprising:
 an image capturing module (100), for capturing images of one or more fruit
 5 bunches;
 a processing module (200), for processing the captured images, the processing
 module (200) comprising:
 a calibration module (300), for performing pixel calibration within a field of
 view of the captured images;
 10 an object detection module (400), for generating at least one bounding box on
 the captured images, such that each bounding box contains one or more regions of
 interest on the fruit bunch; and
 a segmentation module (800), for generating at least one polygon on the
 captured images, such that each polygon contains one or more regions of interest
 15 on the fruit bunch;
 wherein the region of interests within the bounding boxes corresponds to further
 characteristics of the fruit bunch, including ripeness, damage type, or cleanliness
 of the fruit bunch, whereas the regions of interests within the polygons corresponds
 to physical characteristics of the fruit bunch including a size of the fruit bunch and
 20 a type of the fruit bunch, in which information from both of object detection
 module (400) and the segmentation module (800) are collectively assessed and
 evaluated to be output thereafter.
2. The system as according to claim 1, wherein the image capturing module (100) is
 25 further configured to capture frames containing the fruit bunch from a video
 medium.
3. The system as according to any one of the preceding claims, wherein the object
 detection module (400) further comprising a ripeness classification module (500),

a damage detection module (600), and a cleanliness module (700).

4. The system as according to claim 3, wherein the ripeness classification module (500) is configured to determine a classification of the fruit bunch.
- 5 5. The system as according to claim 3, wherein the damage detection module (600) is configured to detect an extent of damages done to the fruit bunch and classify the fruit bunch as damaged (305a) or undamaged (305b) thereafter.
- 10 6. The system as according to claim 3, wherein the cleanliness module (700) is configured to identify a cleanliness level of the fruit bunch, such that the fruit bunch is subsequently labelled dirty (308a) or not dirty (308b).
- 15 7. The system as according to any one of the preceding claims, wherein the segmentation module (800) further comprising a stalk detection module (900) and a size estimation module (1000).
- 20 8. The system as according to claim 7, wherein the stalk detection module (900) is configured to detect whether the fruit bunch has a long stalk (407a) or not a long stalk (407b).
- 25 9. The system as according to claim 7, wherein the size estimation module (1000) is configured to estimate a size of the fruit bunch, such that the fruit bunch is subsequently labelled as small (502a) or not small (502b).
10. The system as according to any one of the preceding claims, further comprising an association module (1100) which is configured to associate the respective characteristics as determined by the object detection module (400) and the segmentation module (800) to each fruit bunch within the captured images.

11. A method for assessing fruit bunch quality comprising the steps of:
 - capturing, by an image capturing module (100), images of one or more fruit bunches;
 - 5 processing, by a processing module (200), the captured images using the steps of:
 - performing, by a calibration module (300), pixel calibration within a field of view of the captured images;
 - generating, by an object detection module (400), at least one bounding box on the fruit bunch, such that each bounding box contains one or more regions of interest on the fruit bunch; and
 - 10 generating, by a segmentation module (800), at least one polygon on the captured images, such that each polygon contains one or more regions of interest on the fruit bunch;
 - wherein the region of interests within the bounding boxes corresponds to further characteristics of the fruit bunch, including ripeness, damage type, or cleanliness of the fruit bunch, whereas the regions of interests within the polygons corresponds to physical characteristics of the fruit bunch including a size of the fruit bunch and a type of the fruit bunch, in which information from both of object detection module (400) and the segmentation module (800) are collectively assessed and
 - 20 evaluated to be output thereafter.
12. The method as according to claim 11, wherein the pixel calibration is performed by the steps of:
 - capturing a fiducial marker of a predetermined size within the captured image;
 - 25 determining a location and a pixel size of the fiducial marker using a computer vision technique; and
 - calculating a distance per pixel based on the determined location and pixel size of the fiducial marker.

13. The method as according to claim 11, wherein the characteristics of the fruit bunch quality are determined by the steps of:
 determining, by a ripeness classification module (500), a classification of ripeness of the fruit bunch;
 5 detecting, by a damage detection module (600), damages on the fruit bunch; and
 identifying, by a cleanliness module (700), a cleanliness level of the fruit bunch.

14. The method as according to claim 13, wherein the classification of ripeness of the fruit bunch is determined by the steps of:
 10 applying an object detection technique to obtain confidence scores on the bounding boxes; and
 classifying the ripeness of the fruit bunch with a label identification based on the confidence scores of each bounding box.

- 15 15. The method as according to claim 13, wherein the damages on the fruit bunch are detected by the steps of:
 identifying numbers of damaged fruitlets within the fruit bunch;
 comparing the numbers of damaged fruitlets to a predetermined threshold range;
 and
 20 designating a label identification to the fruit bunch based on the comparison of damaged fruitlets to the predetermined threshold range.

16. The method as according to claim 13, wherein the cleanliness level of the fruit bunch is identified by the steps of:
 25 identifying one or more areas of contamination on the fruit bunch within the regions of interest based on the bounding boxes of the captured images;
 measuring the areas of contamination on the fruit bunch;
 comparing the measured areas of contamination to a predetermined threshold range; and

designating another label identification to the fruit bunch based on the comparison of the measured areas of contamination to the predetermined threshold range.

17. The method as according to claim 11, wherein the size of the fruit bunch is
5 determined by the steps of:
measuring a contour area from any one segment within the region of interests in
the polygons of the captured images of the fruit bunch;
comparing the contour area to a predetermined threshold; and
designating a label identification to the fruit bunch based on the comparison of the
10 contour area to the predetermined threshold.
18. The method as according to claim 11, wherein the size of the fruit bunch is also
determined by the steps of:
calibrating a binocular image capturing module (100) to determine intrinsic
15 parameters and extrinsic parameters of said binocular image capturing module
(100);
capturing one or more images of the fruit bunch;
rectifying the captured images to ensure at least two captured images are disposed
on a same scan line;
20 identifying corresponding features or key points within the rectified images;
calculating a disparity map between the rectified images;
converting the disparity map to a depth map;
generating a three-dimensional point cloud based on values obtained from the
depth map;
25 normalizing the three-dimensional point cloud using points representing a
conveyor floor; and
estimating a volume of the fruit bunch based on the generated three-dimensional
point cloud.

19. The method as according to claim 11, wherein the type of fruit bunch is determined by the steps of:
- applying one or more label identification on segments in the regions of interest of the captured images correlating to at least one stalk of the fruit bunch and fruitlets of the fruit bunch;
 - converting the segments containing the fruitlets of the fruit bunch in the captured images into a convex hull area;
 - identifying an intersection area between the convex hull area and the segments correlating to the stalk of the fruit bunch;
 - measuring a difference between the segments correlating to the stalk of the fruit bunch and the intersection area, in which the difference corresponds to an excessive stalk area;
 - generating a minimum bounding box to the excessive stalk area;
 - converting a height of the excessive stalk to metric dimensions using distance per pixel as calibrated by the calibration module based on the dimensions of the minimum bounding box;
 - comparing the height of the excessive stalk to a predetermined threshold; and
 - designating a label identification to the fruit bunch based on the comparison of the height of the excessive stalk.
20. The method as according to claim 11, wherein the information of the characteristics obtained from the object detection module (100) and the segmentation module (800) are associated by the steps of:
- matching characteristics of the fruit bunch with at most one fruit bunch in the captured images, such that the matching includes the steps of:
 - defining at least two lists of nodes representing the individual fruit bunch and the characteristics of the fruit bunches;
 - iterating through the list of nodes using a suitable association algorithm; and

identifying and designating a label identification of any one of the respective characteristics of the individual fruit bunch to the individual fruit bunch having an intersection percentage value higher than a minimum intersection percentage value.