

Article

Disease and Defect Detection System for Raspberries Based on Convolutional Neural Networks

José Naranjo-Torres ^{1,†} , Marco Mora ^{1,*,†} , Claudio Fredes ² and Andres Valenzuela ³

¹ Laboratory of Technological Research in Pattern Recognition (LITRP), Universidad Católica del Maule, Talca 3466706, Chile; jnaranjo@ucm.cl

² Department Agricultural Science, Universidad Católica del Maule, Curicó 3480112, Chile; cfredes@ucm.cl

³ Department of Economy and Administration, Universidad Católica del Maule, Talca 3466706, Chile; avalen@ucm.cl

* Correspondence: mmora@ucm.cl

† These authors contributed equally to this work.

Abstract: Raspberries are fruit of great importance for human beings. Their products are segmented by quality. However, estimating raspberry quality is a manual process carried out at the reception of the fruit processing plant, and is thus exposed to factors that could distort the measurement. The agriculture industry has increased the use of deep learning (DL) in computer vision systems. Non-destructive and computer vision equipment and methods are proposed to solve the problem of estimating the quality of raspberries in a tray. To solve the issue of estimating the quality of raspberries in a picking tray, prototype equipment is developed to determine the quality of raspberry trays using computer vision techniques and convolutional neural networks from images captured in the visible RGB spectrum. The Faster R-CNN object-detection algorithm is used, and different pretrained CNN networks are evaluated as a backbone to develop the software for the developed equipment. To avoid imbalance in the dataset, an individual object-detection model is trained and optimized for each detection class. Finally, both hardware and software are effectively integrated. A conceptual test is performed in a real industrial scenario, thus achieving an automatic evaluation of the quality of the raspberry tray, in this way eliminating the intervention of the human expert and eliminating errors involved in visual analysis. Excellent results were obtained in the conceptual test performed, reaching in some cases precision of 100%, reducing the evaluation time per raspberry tray image to 30 s on average, allowing the evaluation of a larger and representative sample of the raspberry batch arriving at the processing plant.

Keywords: raspberries; deep learning; disease and defect detection



Citation: Naranjo-Torres, J.; Mora, M.; Fredes, C.; Valenzuela, A. Disease and Defect Detection System for Raspberries Based on Convolutional Neural Networks. *Appl. Sci.* **2021**, *11*, 11868. <https://doi.org/10.3390/app112411868>

Academic Editor: Keun Ho Ryu

Received: 2 November 2021

Accepted: 9 December 2021

Published: 14 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Raspberries (*Rubus idaeus* L.), as with other berries, are fruit of great importance for human beings since they are a fundamental part of healthy nutrition and powerful allies for human health; they are also good sources of nutrients and antioxidant phytochemicals that contribute to the prevention of chronic diseases [1–3]. Even the wastes produced from raspberry processing in the industry are rich in bioactive compounds, whose enhanced downstream production processes have the potential to be used in the food and pharmaceutical industries, generating economic and environmental benefits [4,5]. According to the Food and Agriculture Organization of the United Nations (FAO) [6], the world production of raspberries is more than 800,000 tons, of which almost 98% is split between Europe and the Americas, as shown in Figure 1.

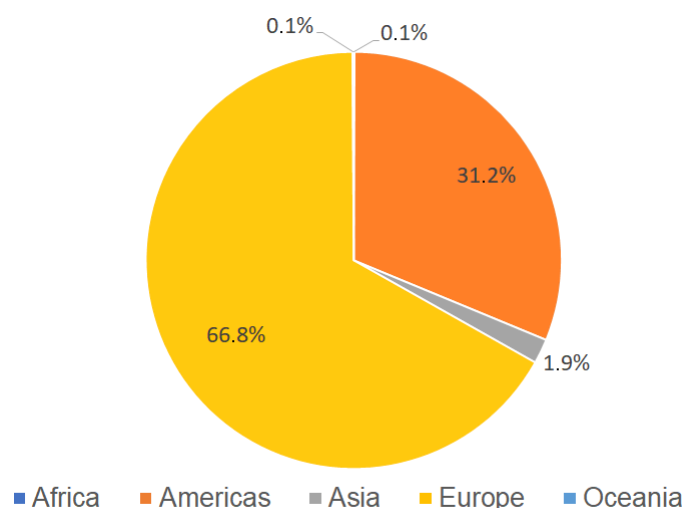


Figure 1. Production share of raspberries by region [6].

Berry production is of great importance in different regions of Chile. According to the Fruit Bulletin of June 2021 published by the Chilean Office of Agricultural Studies and Policies, during the period January–May 2021, a total of 1.74 million tons of fruit exports was registered [7]. The raspberry is the second most important berry in Chile, which currently ranks tenth in terms of raspberry production worldwide [6].

The market for raspberry products is segmented by quality standards. Fruit with the highest quality is destined for export as fresh fruit, medium quality is destined for frozen fruit, and low quality is used in the production of juice and jams. This quality assessment defines the price of raspberries for both producers and exporters.

Currently, in the agro-industries, estimating raspberry quality is a manual process, carried out by specialized personnel at the reception area of the fruit processing plant, as shown in Figure 2, this inspection takes between 5 and 10 min for each harvesting tray. Raspberries arrive at the fruit processing plant in harvesting trays. The procedure is as follows: (i) a random sample is taken from the batch of raspberries (harvesting trays) that arrive at the fruit processing plant; then (ii) specialized personnel at the reception area carry out an inspection, visually separating and counting the raspberries that present defects or diseases; and (iii) the quality of the lot is defined according to parameters pre-set by the food industry. Some critical defects or diseases to consider are fungus-rust, albinism, overripeness, and peduncle attached to the fruit.

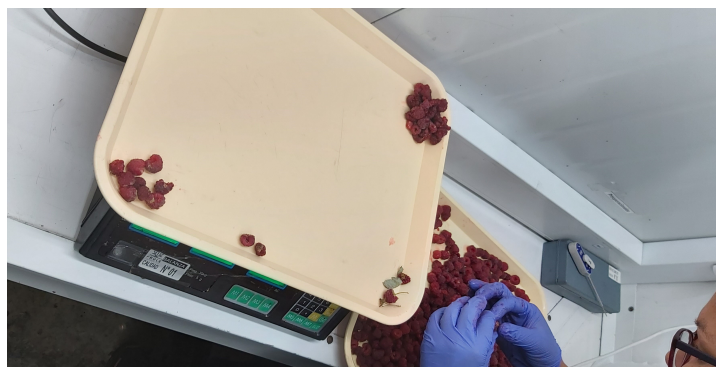


Figure 2. Manual quality measurement of raspberries in a harvesting tray.

The visual inspection process that is carried out has several drawbacks, such as being subjective because it depends on the criteria of the expert present at the time, prone to errors and not representative, since it usually considers a very limited sample of raspberry trays per lot due to the slow measurement process. In addition, errors in the estimation

of quality imply economic losses for both the producer and the exporter. Therefore, it is proposed to resolve this problem, which consists of estimating or measuring the quality of raspberries in harvesting trays, directly and automatically without the intervention of a human expert.

The development and application of non-destructive methods to streamline and automate the quality criteria of fruit for human consumption is of great interest [8]. Therefore, non-destructive testing is considered one of the fastest and most efficient tests that do not affect the appearance and quality of fruit by analyzing them [9]. In this sense, Computer Vision Systems (CVS) are very useful for non-destructive testing. CVSs are being widely used in agribusiness for quality control in the fruit packing process. CVSs are generally made up of a controlled lighting environment, a sensor or camera to acquire images, and a computer containing the algorithms for image processing.

Moreover, in recent years the agricultural industry has been progressively increasing the practical use of Artificial Intelligence (AI) algorithms, such as Machine Learning (ML) and Deep Learning (DL), in planting, harvesting and processing of fruit and vegetables, in particular using computer vision systems (CVS), processing images of fruit or vegetables to classify different types of fruit and vegetables, detection for harvesting, quality control and accurate detection of diseases or defects in fruit [10].

Considering the above exposure, a prototype CVS equipment is proposed to solve the problem of estimating the quality of raspberries in the harvesting tray at the entrance of the fruit processing plant, contributing to a more objective estimation of the quality of raspberries in the harvesting tray.

1. Develop and construct a controlled environment that allows imaging of the raspberry trays.
2. Capture color images (RGB) in the controlled lighting environment.
3. Develop software to process the images with object-detection algorithms based on convolutional neural networks (CNN) and determine the quality of the raspberry trays.
4. Perform classification according to defects or diseases present in raspberries and determine the quality of raspberry trays.

This article is organized as follows: Section 2 corresponds to related work. It first shows some recent scientific papers using computer vision methods and DL algorithms applied to fruit studies. This is followed by a review of patents showing the different existing solutions for quality control of fruit at different stages of processing. Section 3 presents a description of the quality control prototype, an overview of the Faster R-CNN object detector algorithm, and finally describes the dataset development process. Section 4 presents the results of the training and evaluation process of the object detector models for each class of defect or disease to be detected. Then, with the trained models, the quantitative evaluation of the best detector models for each class is presented, together with the conceptual test. Section 5 presents the conclusions of the work carried out and the future work to be developed.

2. Related Works

2.1. Scientific Publications

There are numerous previous studies using CNN-based DL models and algorithms such as object detection, semantic segmentation, and instance segmentation applied to different areas of the agricultural process, harvesting, disease control, and quality control, among other applications [10,11]. Some of the recent scientific publications applying object-detection models in agribusiness are reviewed below.

In [12] the authors propose to produce a real-time pear counter system for mobile applications using RGB video and the variants of the YOLOv4 object detector model and the multiple object-tracking algorithm Deep SORT, obtaining an accuracy of 98% with a AP@0.50 with YOLOv4-CSP. They obtained better computational cost and speed with YOLOv4-tiny, concluding that the best model is the YOLOv4 in accuracy and speed. In [13], the authors properly evaluate various detector models based on the CNN, such as YOLO

v4, region-based fully convolutional network (R-FCN) and Faster R-CNN, in order to detect visible mechanical damage in sugar beet during harvesting in a harvester machine from RGB videos images. The better experimental results showed a recall, precision and F1-score of about 92, 94 and 93% respectively, and higher speed of around 29 frames per second.

In [14], the authors propose an improved object-detection algorithm based on YOLOv3 for early real-time detection of tomato diseases and pests from video images. They used a dilated convolution layer to replace the convolution layer in the backbone network. The results obtained were an F1 value of 94.77%, an AP value of 91.81%, a false detection rate of only 2.1%, and a detection time of 55 ms. In [15], in order to automatically recognize the graspable and ungraspable apples in an apple tree image, an apple target detection method was proposed for a picking robot using improved YOLOv5s and compared with original YOLOv5s, YOLOv3, YOLOv4, and EfficientDet-D0 model. The experimental results indicated that the graspable apples and the ungraspable apples could be identified effectively using the proposed improved network model, with an average recognition time of 0.015 s per image.

In [16], a synthetic image dataset of soybean leaf disease synthetic images is developed to first address the problem of an insufficient dataset. The study designs a fusion of multiple Faster R-CNN (MF3 R-CNN) functions. The results show average mean accuracy of 83.34% on the actual test dataset. In [17] the improved Faster R-CNN algorithm with the ResNet-50 network is proposed for the detection of maturity stages for coconuts. The detection performance achieved using the improved Faster R-CNN model was better than that single-shot detector (SSD), YOLO-V3 and Region-based Fully Convolutional Networks (R-FCN).

In [18], in order to detect tomatoes in environmental conditions, the authors use a modified YOLOv3 model called YOLO-Tomato to be used in a harvesting robot. Results show YOLO-Tomato-A at AP 98.3%, YOLO-Tomato-B at AP 99.3%, and YOLO-Tomato-C at AP 99.5% with detection times 48 ms, 44 ms and 52 ms, respectively. In [19], the authors apply YOLO v2 model (with different optimization algorithm and backbones), YOLOv3 model and Kalman filter, in order to automatically detect pear and apple in videos, managing to automatically count pears and apples in the videos, with an absolute error of 10%. A YOLO v2 network with a larger input image size and data augmentation method contributed to average precision of 0.97 in the pear and apple detection.

In [20], the authors apply instance segmentation on strawberry crops based on the use of a fully convolutional neural network, adding two new channels to the network output so that each strawberry pixel predicts the centroid of its strawberry comparing the results with the Mask R-CNN algorithm, in order to improve the development of automatic harvesting machines. The results show mAP and mean instance intersection over a union of 52.61 and 93.38, respectively, with a processing speed of 30 fps. In [21] the authors develop an automated image-processing methodology to detect, count and estimate the size of citrus fruit on individual trees, using images captured from an unmanned aerial vehicle (UAV) from a commercial citrus grove. To carry out this, a Faster R-CNN Deep-Learning model was used. The obtained result shows an average standard error of 6.59% between visual counting and fruit detection by the model.

In [22] the authors use Mask R-CNN, YOLOv2 and YOLOv3 for wine grape cluster detection. The dataset contains RGB images. The authors conclude that the Mask R-CNN network shows the best results as compared to the YOLO networks. In [23] a deep-learning approach for pixel-wise fruit detection and segmentation based on instance segmentation with the Mask R-CNN algorithm is presented using RGB and HSV images of the scene obtained from an orange grove under natural lighting conditions. The authors find that inclusion of HSV data improves the precision to 0.9753 from 0.8947, after use of RGB data, and obtained the overall F1 score of 0.89 using RGB + HSV.

In [24], the authors combine visual processing color-opponent theory in humans with one-stage deep-learning networks to detect ripe soft fruit (strawberries) in orchards. They

use the RetinaNet algorithm with the ResNet-18 network as feature extractor, achieving a F1 score of 0.793 in controlled conditions and in the real-world dataset F1 score of 0.744.

As is clear from the previously analyzed articles and the summary shown in Table 1, these recent studies applying object-detection methods to the study of fruit cover a very diverse area. Most of them use different versions of the YOLO algorithm and adaptations of it, because these applications need to detect objects quickly in videos, a task for which the YOLO algorithm has been shown to be very efficient. However, in certain cases, it is necessary to detect more accurately the objects in an image, in this case static RGB images, the other widely used algorithms are the Mask R-CNN that allows the segmentation of instances, and the Faster R-CNN algorithm.

Table 1. Summary of the reviewed studies.

Article	Data Type	Model
Parico et al. [12]	RGB video	YOLOv4, YOLOv4-CSP, YOLOv4-tiny.
Nasirahmadi et al. [13]	RGB video	YOLO v4, R-FCN and Faster R-CNN
Wang et al. [14]	RGB video	YOLOv3
Yan et al. [15]	RGB video	YOLOv5s, YOLOv5s, YOLOv3, YOLOv4 and EfficientDet-D0
Zhang et al. [16]	RGB image	Faster R-CNN
Parvathi et al. [17]	RGB image	Faster R-CNN, SSD, YOLOv3, R-FCN
Lawal et al. [18]	RGB images	YOLOv3 modified: YOLO-Tomato-A, YOLO-Tomato-B and YOLO-Tomato-C
Itakura et al. [19]	RGB video	YOLOv2
Perez-Borrero et al. [20]	RGB video	Mask R-CNN
Apolo Apolo et al. [21]	RGB images	Faster R-CNN
Santos et al. [22]	RGB images	Mask R-CNN, YOLOv2 and YOLOv3
Ganesh et al. [23]	RGB and HSV images	Mask R-CNN
Kirk et al. [24]	RGB video	RetinaNet (ResNet-18 feature extractor)

The proposal developed in this article is based on automating a manual process that takes up to 10 min to complete. As it is a process that does not need to be performed at high speed, it is proposed to use the Faster R-CNN object-detection algorithm to detect objects in real time, which has already been widely used [25–28].

2.2. Revision of Patents

In addition to the aforementioned items, there are several intellectual property solutions (patents) that attempt to improve the fruit quality estimation stage. In particular, the following patents can be mentioned:

WO2012039597 [29] relates to an oil palm fruit ripeness grading system using color detection by computer vision. US6610953B1 [30] discloses a method and apparatus incorporating NIR and visible cameras for simultaneously capturing images of apples passing through a packing line to be sorted according to their defects. CN101125333A [31] describes a method for sorting fruit according to the colors presented. CN102928357B [32] presents a technical solution of a non-destructive in-line detection system, from near-infrared and visible light, to establish fruit quality. The solution comprises a mechanical transport device, an illumination device and a detection device. CN205484102U [33] proposes a method using computer vision processing but focuses specifically on finding scars on the fruit and reducing these elements in the packing line.

Another solution is proposed by CN106568784A [34], which discloses a multispectral imaging system for in-line detection of surface defects of fruit and vegetables. On the other hand, CN101832941B [35] describes a quality assessment device based on multispectral images, which tries to detect a certain content of a component inside some fruit. The idea

is to perform an inspection that allows automatic sorting and packing. CN107833223A [36] describes an image-processing method, and specifically a method of segmenting hyperspectral images of fruit based on spectral information. Developed primarily to assist post-harvest processes and promote the development of post-harvest fruit treatments and address quality, standardization and industrialization.

The solutions described above are related to the dynamic stage of the fruit, i.e., they present control systems or elements associated with the stage of fruit movement in some transport lines. However, there are no solutions associated with static fruit control, which corresponds to the stage after harvesting and before the fruit packing line, specifically in the quality control of fruit arranged in boxes or harvesting trays.

3. Materials and Methods

3.1. Description of the Prototype Quality Control Equipment

The prototype of the raspberry tray quality estimation equipment is a cubicle or cabinet (patent process Application Number: 1570-2021). The dimensions of the prototype are $1400 \times 800 \times 600$ millimeters as shown in the isometric view of Figure 3, the plinth measures 10 mm.

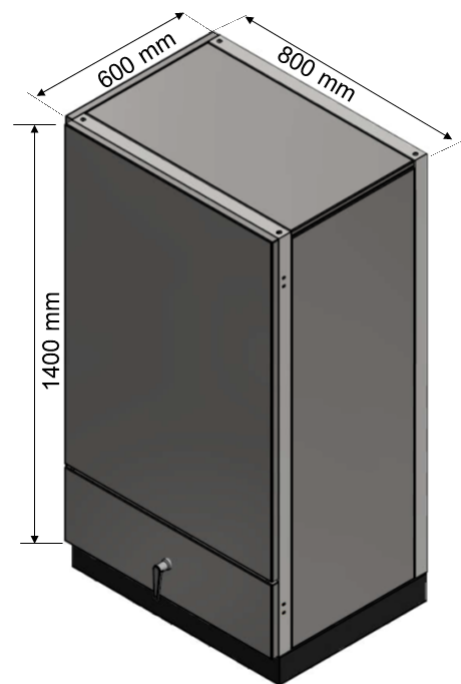


Figure 3. Isometric view of the raspberry tray quality estimation equipment.

The cubicle consists of a front door, a side door, and at the bottom of the cubicle a drawer where a tray with fruit is mounted. Figure 4a shows a front view of the quality control equipment, with the front door closed, Figure 4b shows the equipment with the door open, and the interior of the equipment can be observed.



(a) Front view of closed equipment. (b) Interior view of the equipment.

Figure 4. Front view of the raspberry tray quality estimation equipment.

The drawer that is observed in the lower front part of the equipment, when the equipment is installed, is extracted and the harvesting tray containing the fruit is placed to be introduced to the equipment with the objective of capturing the image. Figure 5a shows a top view of the equipment with the drawer open with a harvesting tray placed, Figure 5b shows a perspective front view of the equipment with the drawer open.



(a) Drawer top view.

(b) Drawer perspective view.

Figure 5. Top and front view of the lower drawer of the raspberry tray quality estimation equipment.

An RGB camera is mounted on the inside of the unit at the top, arranged on a camera support structure as illustrated in the image in Figure 6a. The RGB camera is positioned so that it points towards the fruit on the tray placed at the bottom of the equipment. The camera mounted on the device is an RGB camera in the visible spectrum Basler model acA4024-8gc GigE with a SONY IMX-304 CMOS color sensor of a maximum resolution of 4096×3000 pixels with a pixel size of 3.45 mm.

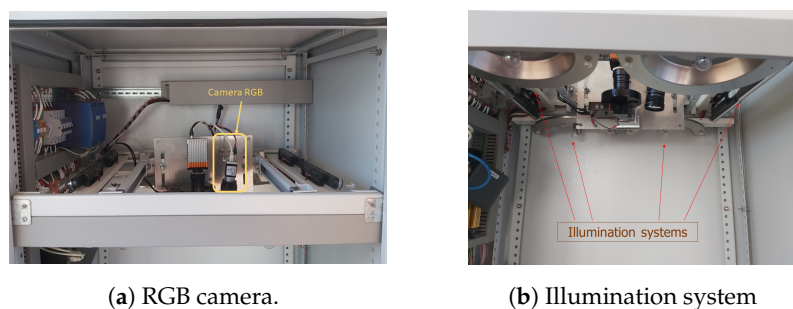


Figure 6. RGB camera and the illumination system of the raspberry tray quality estimation equipment.

The cubicle consists of a controlled illumination system arranged inside as shown in the image of Figure 6b. This lighting system consists of LED lights, which emit the illumination in a diffuse way to avoid light hits (white glows) and shadows (blacks) on the fruit, as well as halogen lights. The equipment also includes a connection to the mains for powering all internal systems; an electrical and communications systems assembly board placed on the side door of the equipment; and a modular RJ45 socket for connection to the computer that controls image capture and transfers to the computer for processing.

3.2. Object Detection—Faster R-CNN Algorithm

Object detection is the process of locating and classifying objects in an image. One approach to DL is the regions with convolutional neural networks (R-CNN) algorithm. This algorithm combines proposed rectangular regions with features extracted from convolutional neural networks. R-CNN is a two-stage detection algorithm. The first stage identifies a subset of regions in an image that may contain an object. The second stage classifies the object in each region.

The R-CNN family of object detectors uses region proposals to detect objects within images. The number of proposed regions dictates the time it takes to detect objects in an image. The most evolved and widely used version of the family is Faster R-CNN [37]. It consists of two modules. The first module is a fully convolutional deep network proposing regions called region proposal network (region proposal network—RPN) to generate region proposals directly in the network. The second module is the Fast R-CNN detector [38] that uses the proposed regions. The generation of region proposals in the network is faster and better adapted to data. The whole system is a single, unified network for object detection (Figure 7). The RPN module tells the Fast R-CNN module where to look.

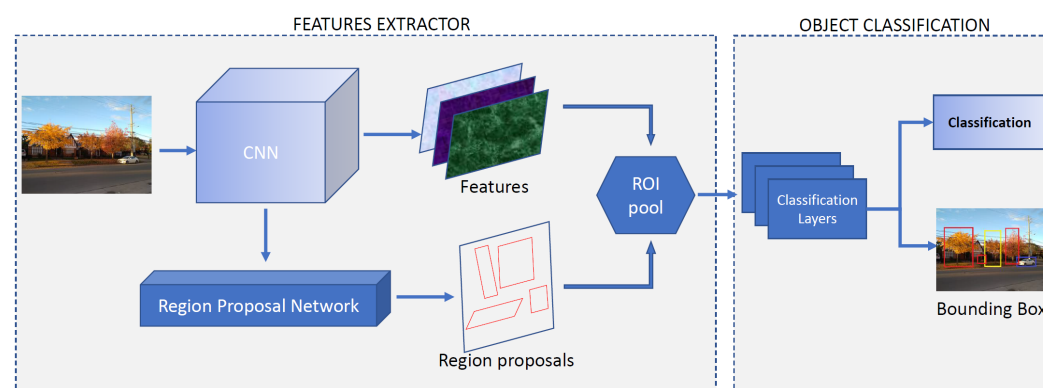


Figure 7. Structure of the Faster R-CNN object detector.

A fundamental part of the Faster R-CNN algorithm is the convolutional network to be used as the feature extraction network; it is typically a pretrained CNN. In this paper, we evaluate some widely known pretrained CNN models, which are the following:

- Inception-v3 [39]: The network is 48 layers deep, the architecture is built progressively, step by step, and has a combination of symmetric and asymmetric compiler blocks; it

is the third edition of Google's Inception Convolutional Neural Network, originally presented during the ImageNet Recognition Challenge.

- ResNet (ResNet-101 and ResNet-50) [40]: these two networks are from the ResNet family; both use residual blocks of direct access connections, reducing the number of parameters to be trained, with a good compromise between performance, structural complexity and training time.
- SqueezeNet [41]: the network is 18 layers deep, is a smaller network that was designed as a more compact replacement for AlexNet. It has almost 50x fewer parameters than AlexNet, yet it performs 3x faster.
- VGG-16 and VGG-19 [42] version, Developed by the Visual Geometry Group (VGG) of the University of Oxford, is an AlexNet enhancement replacing large kernel-sized filters with multiple 3×3 kernel size filters one after another, increasing network depth and thus being able to learn more complex features.

The software for the evaluation of raspberry trays is developed with the objective of classifying 4 *classes* of defects or diseases that occur in raspberries; all these classes can be seen in Figure 8. Similarly, an object-detection model is trained and optimized for each class. This facilitates the optimization of the parameters according to the class to be detected.

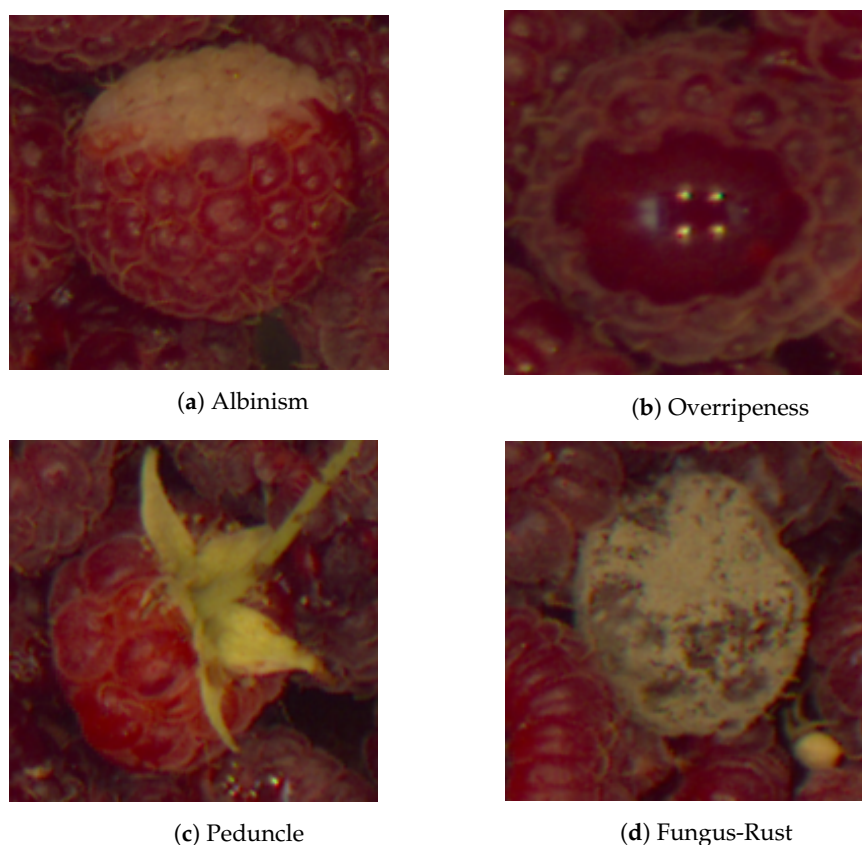


Figure 8. Sample of the different classes of defects and diseases of raspberries.

3.3. Datasets

The process of capturing the images of the trays with raspberries was carried out during the raspberry harvest period in Maule Region, Chile. In the reception area of the fruit processing plant, the prototype of the equipment was strategically placed. Figure 9 shows a diagram of the prototype and an example of the image capture setup and an example of the captured images. Images of 286 raspberry trays were taken with a resolution of 3948×2748 pixels. Eight of these images were selected for use in the conceptual test to be performed after proper training of the object detector models.

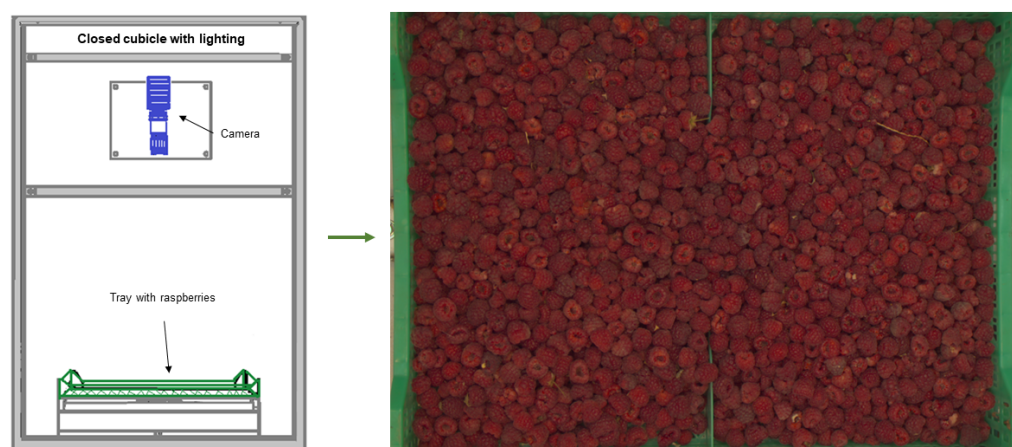


Figure 9. Diagram of closed cubicle with the lighting system and example of tray with raspberries photographed.

Labeled—Classes

Four classes of critical defects or diseases present in raspberries are properly considered for the labeling. These are albinism, overripeness, peduncle and fungus-rust. A sample of each of these defects or diseases present in the raspberries can be seen in Figure 8.

4. Results

The process of training, optimization and testing of the Faster R-CNN model for the different pretrained CNN networks by performing transfer learning by retraining the whole network were carried out with a computer server of the following characteristics: 2x Intel Xeon Gold 6140 CPU @ 2.30 GHz, 36 total physical cores, 24.75 MB L3 Cache 157 Memory 126 GB, Operating System Debian GNU/Linux 10 (buster) Kernel 4.19.0-10-amd64×86-64. We used the MatLab Deep-Learning Toolbox [43], which provides a framework for designing and implementing deep neural networks with pretrained models.

The database is constructed from the images of the raspberry trays captured by the equipment. To carry out the training of the object detector model, these images were preprocessed before labeling the corresponding classes. Each image is divided into four equal sections and resizing each of these sections to 680×480 pixels. This facilitates memory management at the time of training while maintaining good image quality. From this division, a database with 1100 images is built.

The database is divided into an 80% training database and a 20% database. This division is performed using a random permutation of integers from $1, \dots, n$ with no repeated elements, using the same sets for each of the object detector model training processes. The number of labeled objects for each class in the image set is shown in Table 2.

Table 2. Number of labeled objects for each class.

Class	N° Training Object	N° Test Object
Albinism	1978	577
Overripeness	452	104
Peduncle	348	73
Fungus-Rust	215	50

The number of objects for each class shown in Table 2 corresponds to those found in the training and test sets, taking into account that there are not raspberries of all classes in each image. As can be observed in Table 2, the number of objects contained in the images of the training set is different for each class. In order to solve this imbalance problem, an individual object-detection model is trained and optimized for each of the classes.

4.1. Object-Detection Models Training

A Faster R-CNN network is created given a pretrained feature extraction network, specifically one for each of the pretrained CNNs defined in Section 3.2. Network training was performed using the Stochastic Gradient Descent (SGD) optimizer with momentum. Parameter optimization is performed during training. The number of regions to sample, which controls the number of image regions per image, varies between 128 and 256 regions. Positive training samples used are those that overlap with the ground truth frames by 0.5 to 1.0, as measured by the intersection over union (IoU) metric of the bounding box. Negative training samples are those that overlap between 0 and 0.4. Furthermore, a sweep of anchor box pyramid scale factor values used to successively upscale anchor box sizes is performed. The values are from 1 through 2. High values produce faster results and small values produce greater accuracy.

The evaluation of the results of the object detector models trained for each class individually is performed as follows:

- The Average Precision (AP) calculated from the area under the Precision–Recall curve is determined.
- A value called Detection Rate (DR) is estimated; this value is given by:

$$DR = \frac{\text{number of Objects Detected}}{\text{Objects Ground Truth}} \quad (1)$$

where the number of detected objects is formed by the sum of True Positive (TP) and False Positive (FP). Thus, if $DR < 1$ indicates an underestimation of the number of detected objects, if $DR > 1$ it would be overestimating the number of real detected objects. Finally, $DR = 1$ or approximately one would be ideal in this case.

Table 3 shows a summary of the AP results obtained from the Precision–Recall curves after the optimization of the parameters in the training of the object detector models for each network used as the basis of the Faster R-CNN algorithm. The DR value obtained for each class and network used is also shown.

Table 3. AP and DR values for each class according to the CNN network used in the object detector training.

CNN	Albinism		Fungus-Rust		Peduncle		Overripeness	
	AP	DR	AP	DR	AP	DR	AP	DR
Inception V3	0.22	0.84	0.03	0.24	0.38	1.11	0.05	0.42
SqueezeNet	0.52	1.51	0.33	2.08	0.37	1.32	0.19	2.10
ResNet50	0.48	0.93	0.17	0.52	0.47	1.12	0.18	0.88
ResNet101	0.46	1.07	0.29	0.70	0.40	1.03	0.22	0.99
VGG16	0.55	1.10	0.42	0.90	0.46	1	0.27	1.33
VGG19	0.55	1.07	0.35	1.02	0.45	0.90	0.25	1.15

In the Albinism class, it can be seen from Figure 10 and the curves that the lowest-performing network is Inception V3, as indicated by the values in Table 3 with an AP = 0.22 and a DR = 0.84 indicating a significant underestimation of the possible objects in the images. SqueezeNet, with an AP = 0.52 is the second best in this class, but has a DR = 1.51, indicating that it detects approximately 51% more objects than are labeled in the test images, i.e., it produces a high number of FPs, making it unfeasible for use. ResNet50 and ResNet101 present very similar results, with an AP of 0.48 and 0.46 respectively, and a DR in both cases with an absolute deviation of 7%, being globally the second-best networks as backbone of the Faster R-CNN for this class of raspberry. The VGG nets for this class are the best choice. As can be seen in Figure 10 and Table 3, they present the best AP values and their DR are among the most adequate.

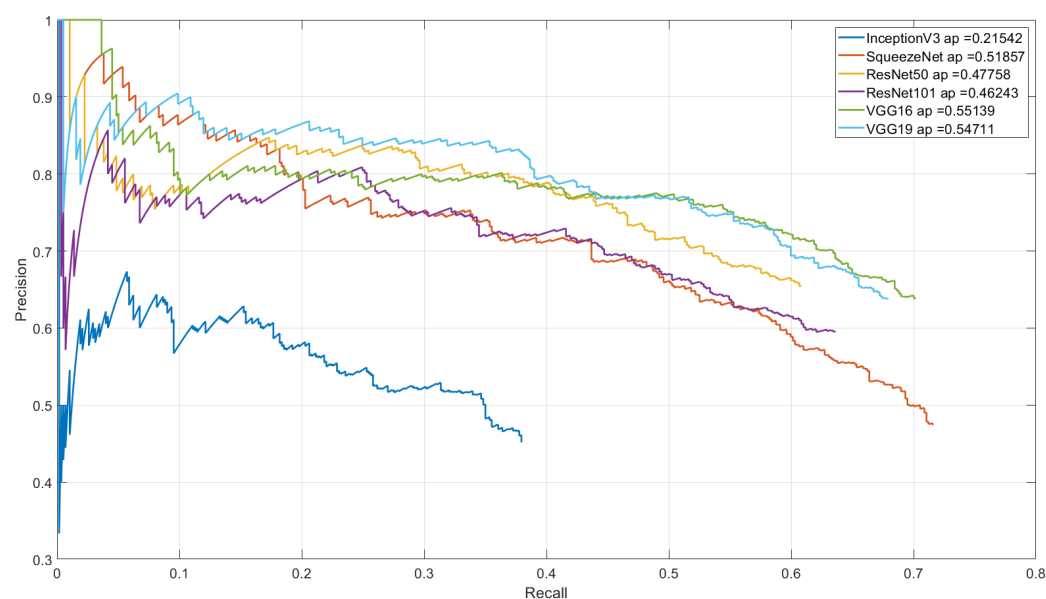


Figure 10. Precision–Recall curve for the albinism class.

In the Fungus-Rust class, Figure 11 shows the results for raspberries presenting fungus-rust. For this type of defect/disease it is observed that almost all the nets lowered the yield with respect to the AP values. VGG16 is the network that presents the best result for the pair of AP and *DR* values, followed by VGG19. Therefore, for this case the most suitable to test as the basis of the object detector model is the VGG16 network.

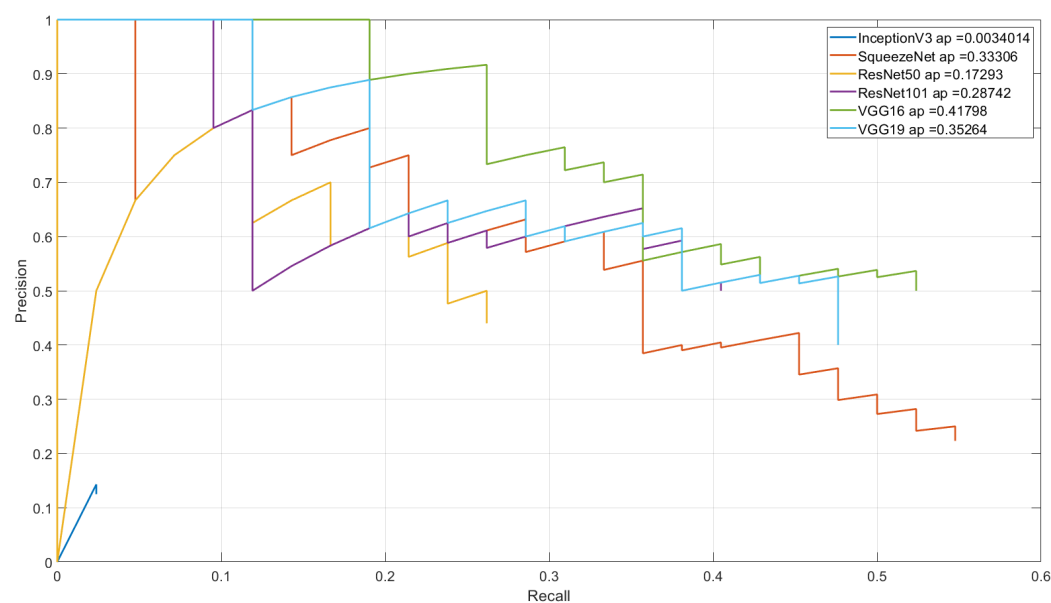


Figure 11. Precision–Recall curve for the fungus-rust class.

In the Peduncle class, the results as seen in Figure 12 and Table 3 are more similar to each other. In this case the Inception V3 and SqueezeNet networks show the lowest AP results, with the SqueezeNet network showing a significant overestimation of detected objects. The ResNet50, Vgg16 and Vgg19 networks present the best AP values, highlighting a value $DR = 1$ of VGG16 with the second-best AP = 1, making it the top choice.

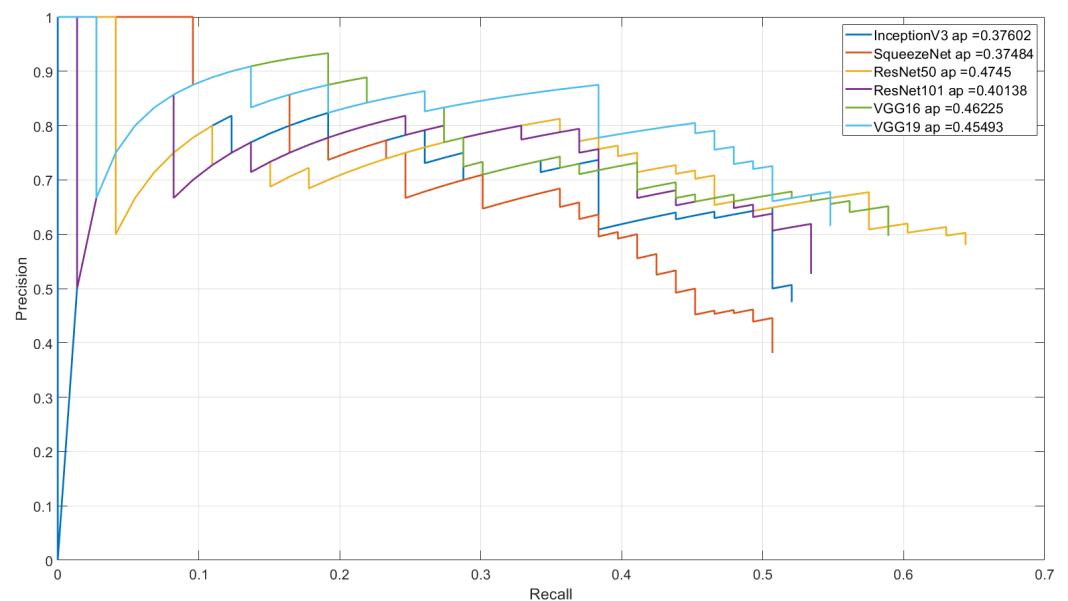


Figure 12. Precision–Recall curve for the peduncle class.

In the Overripeness class, in this case the results show how difficult it is for the algorithm to detect raspberries with overripeness. The Inception V3 network with a practically zero AP value and a very low object-detection value (as shown in Figure 13). The squeezeNet network has a DR of 2.10 indicating that it detected more than twice the number of objects that are actually labeled in the Test images with an AP of 0.19. The 3 best APs correspond to the VGG16, VGG19 and ResNet101 networks in that order of 0.27, 0.25 and 0.2 respectively; these networks are the possible candidates to use in the object-detection model.

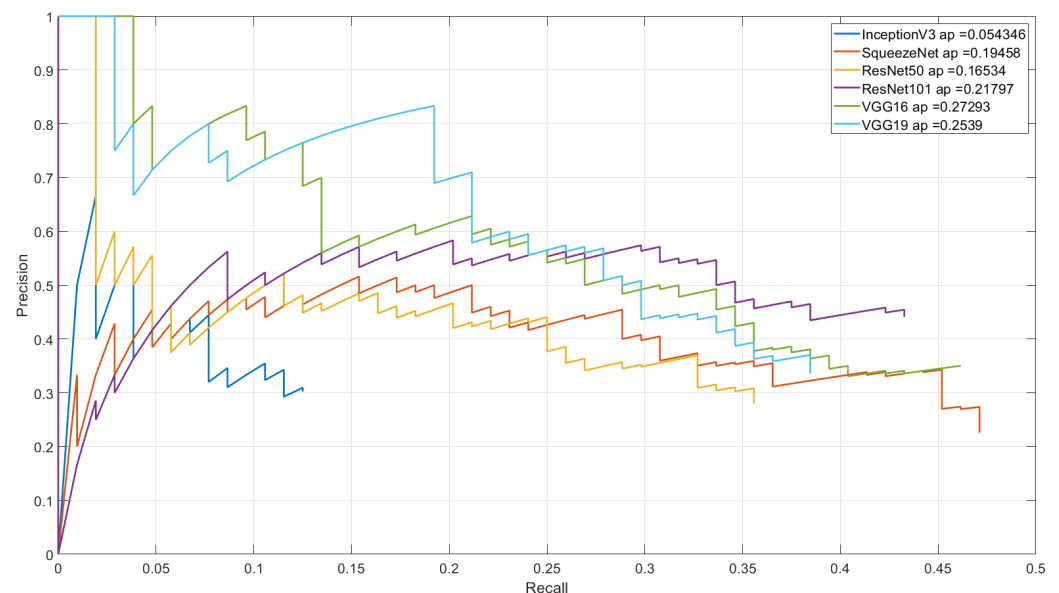


Figure 13. Precision–Recall curve for the overripeness class .

4.2. Quantitative Evaluation

The work process of the equipment once installed in the corresponding fruit reception area in the processing plant is shown in the flow diagram in Figure 14. As shown in the diagram, the process starts at the entrance of the fruit processing plant, and the shipment of raspberry trays is received, followed by the random selection of the sample of the entire shipment, i.e., the raspberry trays to be evaluated are selected. Next, once the sample

trays have been selected, a tray is introduced into the equipment. At this point the process is automatic, the image of the raspberry tray is captured, and the developed software performs the classification and displays the results.

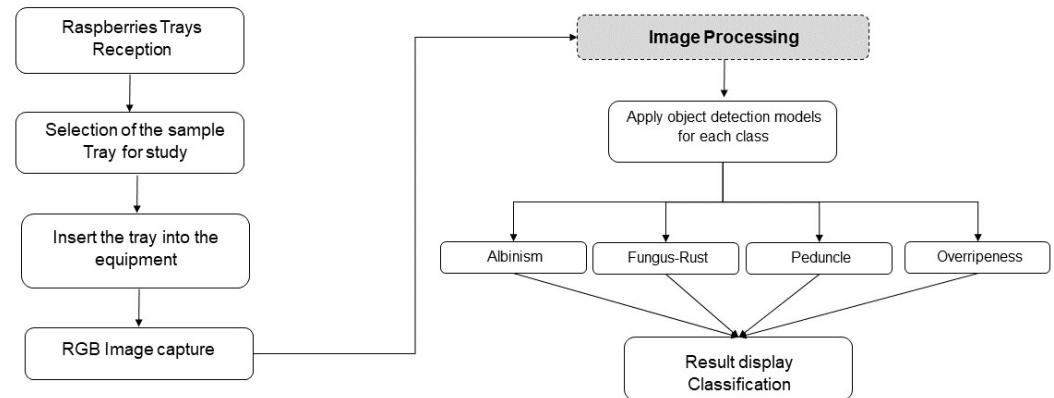


Figure 14. Flowchart of raspberry tray quality assessment team workflow.

4.3. Conceptual Test

A quantitative conceptual test is performed with eight images of raspberry trays. The software was developed with the individual trained object-detection models was implemented on a laptop connected to the equipment with the following characteristics: i7-9th Gen processor 128 GB SSD/1TB HDD, 16 GB RAM, NVIDIA GTX 1660Ti 6GB.

The process is started by selecting a sample raspberry tray, then it is placed inside the equipment and the RGB image capture system is activated. Next, the developed program that processes the image is launched.

1. A window with the image of the original unprocessed raspberry tray opens
2. Starts the image processing with each of the modules trained for each of the classes to be studied.
3. Once the image processing is done, a window opens for each of the classes defined with the image of the raspberry tray with a “box” around the raspberry of the detected class with a “label” with the name corresponding to the detected class.
4. Above the image in the window the class and the number of objects of that class detected by the model are indicated.

The results before process was performed over the image of tray-1 (Figure 15) are divided as follows: Figure 15a detects 105 raspberries of the class *Albinism*; those of the class *Fungus-Rust* are shown in Figure 15b, showing that 2 were detected; Figure 15c corresponds to the class *Peduncle* with 2 detected; and finally Figure 15d shows the class *Overripeness* with 9 detected. The entire raspberry tray image evaluation process takes an average of 30 s, which is a substantial improvement over the manual evaluation process that is the current industry standard, which takes approximately 5 to 10 min.

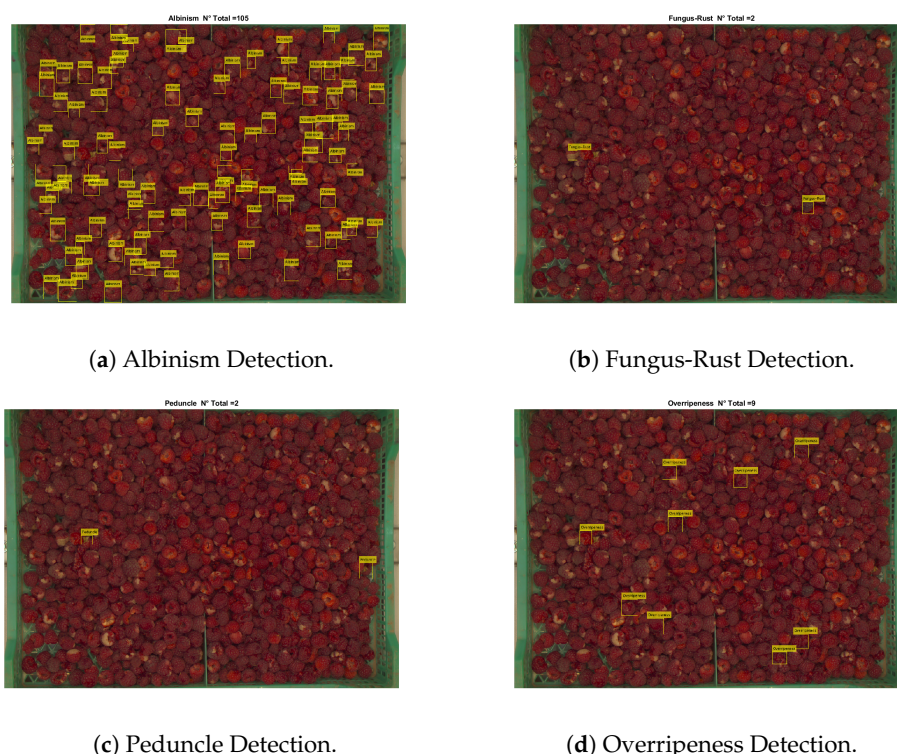


Figure 15. Images of the results of raspberry detection according to classes in the image of raspberries corresponding to tray-1.

Figure 16 shows the confusion matrices of the results for each of the images of raspberry trays used in this test.

Figure 16a shows the confusion matrix resulting from the evaluation of tray-1. The diagonal, highlighted, represents the number of hits or the raspberries correctly detected and labeled by the system. The system managed to correctly detect 105 raspberries with albinism out of the 110 visually counted, with a precision of 95.5%. There are no raspberries with fungus-rust. The peduncle class with two correctly detected obtains 100% precision. Finally, the overripeness class detected nine correctly and two not detected, achieving 82% detection precision.

Figure 16b corresponds to tray-2. This tray does not contain raspberries of the classes fungus-rust, peduncle and overripeness. Effectively the models did not detect raspberries of this class. The albinism class detected in 14 of 16 raspberries, achieving a precision of 87.5%. In the confusion matrix corresponding to tray-3 (Figure 16c) there are no raspberries of the classes fungus-rust, peduncle and overripeness. Of the albinism class it detected 9, which was the total of those present, giving a precision of 100%.

The confusion matrix in Figure 16d corresponding to tray four does not have raspberries of the Moho–Fungus class. It is observed that nine of the albinism class were detected and one was not detected, obtaining an accuracy of 90%. The raspberries with peduncle were all detected, 5 in total, achieving an accuracy of 100%. The overripe class detected 12 correctly, and 2 were not detected with an accuracy of 85.7%.

The confusion matrix corresponding to tray-5 is shown in Figure 16e. It can be seen that it classified or detected four raspberries as albinism correctly and 2 were not detected. The classes fungus-rust and peduncle show no elements in this tray of raspberries. Finally, the overripeness class managed to detect 11 and failed to detect four at this point. We observe for this image a decrease in the percentages of accuracy with respect to the previous images in the case of albinism and overripeness.

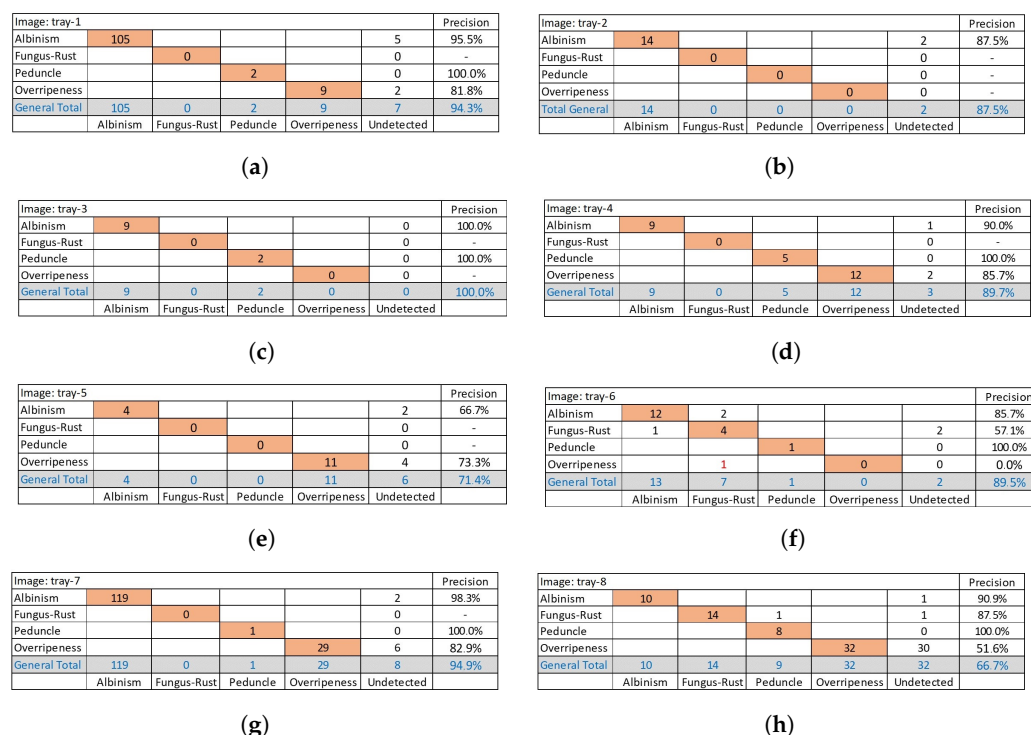


Figure 16. Confusion matrices of the results of the evaluation of the conceptual test images of raspberry trays.

The confusion matrices for tray-6, tray-7 and tray-8 are shown in Figure 16f–h, respectively. In these images it can be observed that for the albinism class the results present an accuracy always higher than 90%; in the case of tray-7, the result reaches 98% precision. The fungus-rust class in the three confusion matrices has good results, especially for tray-8, which reaches 87.5% of accuracy by detecting 14 correctly and only one undetected, and for tray-6, it is observed that it detects 4 correctly of the 7 present in the tray, classifying one with albinism; this occurs due to the similarities between the raspberries that present albinism and fungus-rust in some cases, failing to detect 2 raspberries with this class.

5. Conclusions and Future Work

It can be concluded that the main objective of the project was fully met. It was possible to develop functional equipment to determine the quality of raspberry trays using computer vision techniques and convolutional neural networks from images captured in the visible RGB spectrum. Effectively integrating both hardware and software and thus achieving an automatic onboard evaluation of raspberry tray quality.

The software developed based on object-detection algorithms, originally designed for autonomous driving, is able to adapt appropriately for the detection of raspberries with defects or diseases present in the harvesting trays with raspberries on entering the fruit processing plant, according to the parameters and standards of agribusiness, obtaining the quality of the fruit accurately, automatically and in real time.

For this, individual object detectors are trained for each class to be detected; in this case, we trained 4 detectors for 4 predefined classes. This process is done in this way, because each object detector can individually adjust its training parameters, optimizing the model training process. These class-based raspberry detection models generally achieve excellent results. Another important advantage of the modular per-class development of the software is that another model can be incorporated that has better results for a particular class, without affecting the models of the rest of the classes.

Regarding future work for the optimization of the complete system, it is necessary to work together with the industry, placing the device in the reception area of the fruit processing plant and jointly performing the evaluation with a human expert from the

company to evaluate the prototype quality for the raspberries in trays. In this way, we create a comprehensive database, having the image capture available together with evaluation made by an expert and jointly evaluate the results of both estimates, detailing similarities and errors to optimize the process, both in handling the equipment and in the accuracy of the quality assessment performed by the software.

Likewise, regarding future work with the objective of improving the detection of objects in the raspberry tray, the adaptation of depth cameras (RGB-D) is proposed. These would provide additional information for better image processing. Along with providing the image in the visible spectrum, it would provide the distance from the camera to the object, which would allow definition of a structure and achieve better segmentation or detection of edges of each object in the tray, and thus improve the results of the classification of raspberries contained in the tray.

Author Contributions: J.N.-T. Investigation, Software, Resources and Writing—original draft; M.M. Funding acquisition, Conceptualization, Methodology and Writing—review and editing; C.F. Writing—review and Funding acquisition; A.V. Writing—review, Funding acquisition and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Innovation Fund for Competitiveness—FIC, Government of Maule, Chile—Project Transfer Development Equipment Estimation Quality of Raspberry, code 40.001.110-0. (Esta investigación fue financiada por el Fondo de Innovación para la Competitividad—FIC, Gobierno de Maule, Chile—Proyecto de Transferencia de Desarrollo de Equipo de Estimación de Calidad de la Frambuesa, código 40.001.110-0.)

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study is available upon direct request to Marco Mora (mmora@ucm.cl) head of the Laboratory of Technological Research in Recognition of Patterns (www.litrp.cl, accessed on 2 June 2021).

Acknowledgments: The authors thank the Laboratory of Technological Research in Recognition of Patterns (www.litrp.cl, accessed on 2 June 2021) of the Universidad Catolica del Maule, Chile, for providing the computer servers where the experiments were carried out.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
FAO	Food and Agriculture Organization of the United Nations
CVS	computer vision systems

References

1. Rao, A.V.; Snyder, D.M. Raspberries and Human Health: A Review. *J. Agric. Food Chem.* **2010**, *58*, 3871–3883. [[CrossRef](#)] [[PubMed](#)]
2. Kalt, W.; Cassidy, A.; Howard, L.R.; Krikorian, R.; Stull, A.J.; Tremblay, F.; Zamora-Ros, R. Recent Research on the Health Benefits of Blueberries and Their Anthocyanins. *Adv. Nutr.* **2020**, *11*, 224–236. [[CrossRef](#)] [[PubMed](#)]
3. Seeram, N.P. Berries and Human Health: Research Highlights from the Fifth Biennial Berry Health Benefits Symposium. *J. Agric. Food Chem.* **2013**, *62*, 3839–3841. [[CrossRef](#)]
4. Ispiryan, A.; Viškelis, J. Valorisation of Raspberries By-Products for Food and Pharmaceutical Industries. 2019. Available online: <https://kosmospublishers.com/valorisation-of-raspberries-by-products-for-food-and-pharmaceutical-industries/> (accessed on 8 May 2021).
5. Beekwilder, J.; Hall, R.D.; De Vos, C. Identification and dietary relevance of antioxidants from raspberry. *Biofactors* **2005**, *23*, 197–205. [[CrossRef](#)] [[PubMed](#)]

6. The Food and Agriculture Organization of the United Nations (FAO). Food and Agriculture Data. Online. 2021. Available online: <http://www.fao.org/faostat/en> (accessed on 1 August 2021).
7. Lepe, J.P. *Boletín de la Fruta Junio 2021*; Technical Report; Oficina de Estudios y Políticas Agrarias, ODEPA: Santiago, Chile, 2021. Available online: <https://www.odepa.gob.cl/> (accessed on 1 August 2021).
8. Blackmore, S. New concepts in agricultural automation. In Proceedings of the R&D Conference “Precision in Arable Farming: Current Practice and Future Potential”, Grantham, UK, 28–29 October 2009; pp. 127–137.
9. Bhargava, A.; Bansal, A. Fruits and vegetables quality evaluation using computer vision: A review. *J. King Saud-Univ.-Comput. Inf. Sci.* **2021**, *33*, 243–257. [[CrossRef](#)]
10. Naranjo-Torres, J.; Mora, M.; Hernández-García, R.; Barrientos, R.J.; Fredes, C.; Valenzuela, A. A Review of Convolutional Neural Network Applied to Fruit Image Processing. *Appl. Sci.* **2020**, *10*, 3443. [[CrossRef](#)]
11. Maheswari, P.; Raja, P.; Apolo-Apolo, O.E.; Pérez-Ruiz, M. Intelligent Fruit Yield Estimation for Orchards Using Deep Learning Based Semantic Segmentation Techniques—A Review. *Front. Plant Sci.* **2021**, *12*, 684328. [[CrossRef](#)]
12. Parico, A.I.B.; Ahamed, T. Real Time Pear Fruit Detection and Counting Using YOLOv4 Models and Deep SORT. *Sensors* **2021**, *21*, 4803. [[CrossRef](#)]
13. Nasirahmadi, A.; Wilczek, U.; Hensel, O. Sugar Beet Damage Detection during Harvesting Using Different Convolutional Neural Network Models. *Agriculture* **2021**, *11*, 1111. [[CrossRef](#)]
14. Wang, X.; Liu, J.; Zhu, X. Early real-time detection algorithm of tomato diseases and pests in the natural environment. *Plant Methods* **2021**, *17*. [[CrossRef](#)] [[PubMed](#)]
15. Yan, B.; Fan, P.; Lei, X.; Liu, Z.; Yang, F. A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5. *Remote Sens.* **2021**, *13*, 1619. [[CrossRef](#)]
16. Zhang, K.; Wu, Q.; Chen, Y. Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN. *Comput. Electron. Agric.* **2021**, *183*, 106064. [[CrossRef](#)]
17. Parvathi, S.; Selvi, S.T. Detection of maturity stages of coconuts in complex background using Faster R-CNN model. *Biosyst. Eng.* **2021**, *202*, 119–132. [[CrossRef](#)]
18. Lawal, M.O. Tomato detection based on modified YOLOv3 framework. *Sci. Rep.* **2021**, *11*. [[CrossRef](#)] [[PubMed](#)]
19. Itakura, K.; Narita, Y.; Noaki, S.; Hosoi, F. Automatic pear and apple detection by videos using deep learning and a Kalman filter. *OSA Contin.* **2021**, *4*, 1688. [[CrossRef](#)]
20. Perez-Borrero, I.; Marin-Santos, D.; Vasallo-Vazquez, M.J.; Gegundez-Arias, M.E. A new deep-learning strawberry instance segmentation methodology based on a fully convolutional neural network. *Neural Comput. Appl.* **2021**, *33*, 15059–15071. [[CrossRef](#)]
21. Apolo-Apolo, O.; Martínez-Guanter, J.; Egea, G.; Raja, P.; Pérez-Ruiz, M. Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV. *Eur. J. Agron.* **2020**, *115*, 126030. [[CrossRef](#)]
22. Santos, T.T.; de Souza, L.L.; dos Santos, A.A.; Avila, S. Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association. *Comput. Electron. Agric.* **2020**, *170*, 105247. [[CrossRef](#)]
23. Ganesh, P.; Volle, K.; Burks, T.F.; Mehta, S.S. Deep Orange: Mask R-CNN based Orange Detection and Segmentation. *IFAC-PapersOnLine* **2019**, *52*, 70–75. [[CrossRef](#)]
24. Kirk, R.; Cielniak, G.; Mangan, M. L*a*b*Fruits: A Rapid and Robust Outdoor Fruit Detection System Combining Bio-Inspired Features with One-Stage Deep Learning Networks. *Sensors* **2020**, *20*, 275. [[CrossRef](#)] [[PubMed](#)]
25. Chen, X.; Lian, C.; Deng, H.H.; Kuang, T.; Lin, H.Y.; Xiao, D.; Gateno, J.; Shen, D.; Xia, J.J.; Yap, P.T. Fast and Accurate Craniomaxillofacial Landmark Detection via 3D Faster R-CNN. *IEEE Trans. Med Imaging* **2021**, *40*, 3867–3878. [[CrossRef](#)]
26. Chen, Y.; Wang, H.; Li, W.; Sakaridis, C.; Dai, D.; Gool, L.V. Scale-Aware Domain Adaptive Faster R-CNN. *Int. J. Comput. Vis.* **2021**, *129*, 2223–2243. [[CrossRef](#)]
27. Su, Y.; Li, D.; Chen, X. Lung Nodule Detection based on Faster R-CNN Framework. *Comput. Methods Programs Biomed.* **2021**, *200*, 105866. [[CrossRef](#)] [[PubMed](#)]
28. Nguyen, C.C.; Tran, G.S.; Nguyen, V.T.; Burie, J.C.; Nghiem, T.P. Pulmonary Nodule Detection Based on Faster R-CNN With Adaptive Anchor Box. *IEEE Access* **2021**, *9*, 154740–154751. [[CrossRef](#)]
29. Rashid, M.S.A.; Zaid, A.M.; Hamiruce, M.M.; Suhaidi, S.; Salem, M.A.M.; Din, A.M. Fruit Ripeness Grading System. WO Patent 2012/039597 A3, 21 June 2012.
30. Tao, Y.; Wen, Z. Item Defect Detection Apparatus and Method. U.S. Patent 6,610,953 B1, 26 August 2003.
31. Ying, Y.; Rao, X.; Jiang, H.; Wang, J. Fruit Classifying Method according to Surface Color. CN Patent 101125333 A, 20 February 2008.
32. Xie, L.; Ying, Y.; Wang, A.; Jie, D.; Rao, X. Rapid Nondestructive On-Line Detection SYSTEM for Fruit Quality Based on Near Infrared/Visible Light. CN Patent 102928357 B, 10 December 2014.
33. Fruit Quality Control Surveys SYSTEM Based on Computer Vision. CN Patent 205484102 U, 17 August 2016. Available online: <https://www.vipzhuanli.com/patent/201620105803.8/> (accessed on 8 May 2021).
34. Multispectral Imaging System and Implementation Method Which Are Used for Fruit and Vegetable Surface Defect On-Line Detection. CN Patent 106568784 A, 19 April 2017. Available online: <https://patents.google.com/patent/CN106568784A/en> (accessed on 8 May 2021).
35. Fruit Quality Evaluation Device Based on Multispectral Image. CN Patent 101832941 B, 13 March 2013. Available online: <https://patentimages.storage.googleapis.com/48/1d/00/6cfd3da9b6576/CN101832941B.pdf> (accessed on 8 May 2021).

36. Fruit Hyperspectral Image Segmentation Method Based on Spectral Information. CN Patent 107833223 A, 23 March 2018. Available online: <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=SCPD&dbname=SCPD2018&filename=CN107833223A&uniplatform=NZKPT&v=qSqr0C61Ual4VtslCrOUPHPyf5Exiw6N7qWo0YegS5fmU58OUJ2AoNfO1yudQDnc> (accessed on 8 May 2021).
37. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
38. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 142–158. [[CrossRef](#)] [[PubMed](#)]
39. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA, USA, 27–30 June 2016; pp. 2818–2826.
40. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA, USA, 27–30 June 2016; pp. 770–778.
41. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. *arXiv* **2016**, arXiv:cs.CV/1602.07360.
42. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations (ICLR2015), San Diego, CA, USA, 7–9 May 2015. Available online: <https://arxiv.org/abs/1409.1556> (accessed on 9 December 2021).
43. MathWorks. Deep Learning Toolbox™-Matlab. Available online: <https://www.mathworks.com/products/deep-learning.html> (accessed on 9 December 2021).