Article

Development of an Optimal Algorithm for Detecting Damaged and Diseased Potato Tubers Moving along a Conveyor Belt Using Computer Vision Systems

Sergey Alekseevich Korchagin ^{1,*} , Sergey Timurovich Gataullin ¹, Aleksey Viktorovich Osipov ¹, Mikhail Viktorovich Smirnov ¹, Stanislav Vadimovich Suvorov ² , Denis Vladimirovich Serdechnyi ³ and Konstantin Vladimirovich Bublikov ⁴

¹ Department of Data Analysis and Machine Learning, Federal State Budgetary Institution of Higher Education, Financial University under the Government of the Russian Federation, 105187 Moscow, Russia; STGataullin@fa.ru (S.T.G.); AVOsipov@fa.ru (A.V.O.); MVSmirnov@fa.ru (M.V.S.)

² Federal State Autonomous Educational Institution of Higher Education, Moscow Polytechnic University, 107023 Moscow, Russia; ssw1168@mail.ru

³ Federal State Budgetary Educational Institution of Higher Education, State University of Management, 109542 Moscow, Russia; dv_serdechnyj@guu.ru

⁴ Institute of Electrical Engineering Slovak Academy of Sciences, Researcher, Dúbravská cesta 3484/9, Karlova Ves, 841 04 Bratislava, Slovakia; elecubl@savba.sk

* Correspondence: SAKorchagin@fa.ru; Tel.: +7-929-77-33-555

Citation: Korchagin, S.A.; Gataullin, S.T.; Osipov, A.V.; Smirnov, M.V.; Suvorov, S.V.; Serdechnyi, D.V.; Bublikov, K.V. Development of an Optimal Algorithm for Detecting Damaged and Diseased Potato Tubers Moving along a Conveyor Belt Using Computer Vision Systems. *Agronomy* **2021**, *11*, 1980. <https://doi.org/10.3390/agronomy11101980>

Academic Editor: Peng Fu

Received: 1 September 2021

Accepted: 28 September 2021

Published: 30 September 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/>).

Abstract: The article discusses the problem of detecting sick or mechanically damaged potatoes using machine learning methods. We proposed an algorithm and developed a system for the rapid detection of damaged tubers. The system can be installed on a conveyor belt in a vegetable store, and it consists of a laptop computer and an action camera, synchronized with a flashlight system. The algorithm consists of two phases. The first phase uses the Viola-Jones algorithm, applied to the filtered action camera image, so it aims to detect separate potato tubers on the conveyor belt. The second phase is the application of a method that we choose based on video capturing conditions. To isolate potatoes infected with certain types of diseases (dry rot, for example), we use the Scale Invariant Feature Transform (SIFT)—Support Vector Machine (SVM) method. In case of inconsistent or weak lighting, the histogram of oriented gradients (HOG)—Bag-of-Visual-Words (BOVW)—neural network (BPNN) method is used. Otherwise, Otsu's threshold binarization—a convolutional neural network (CNN) method is used. The first phase's result depends on the conveyor's speed, the density of tubers on the conveyor, and the accuracy of the video system. With the optimal setting, the result reaches 97%. The second phase's outcome depends on the method and varies from 80% to 97%. When evaluating the performance of the system, it was found that it allows to detect and classify up to 100 tubers in one second, which significantly exceeds the performance of most similar systems.

Keywords: neural networks; defects detection; crop; potato disease; potato classification; fast detection; machine learning

1. Introduction

Rapid population growth requires an increase in the efficiency and productivity of the agricultural sector. To increase productivity in agriculture, farmers use pesticides widely, which negatively affect the human body. Recently, chemical pesticides are gradually being replaced by pesticides based on bacteria, fungi, and viruses [1]. However, this process is hampered, both at the level of legislation of some states and the level of farms [2]. In this situation, the automatic selection of vegetables, root crops, and fruits at the time of laying them for storage and during the preparation of seed material using computer vision can reduce the volume of pesticides applied and storage costs.

Potatoes are among the most disease-prone crops. Potato infection occurs through mechanical damage, insects, infected tubers, and soil. Infections are carried by wind and rain, transmitted with infected garden tools, by contact with infected potato tops or tubers. Identification and removal of damaged or diseased tubers will allow, firstly, to protect healthy tubers from infection, and secondly, to use discarded tubers in various types of industries where this kind of raw material is applicable.

The detection and classification of potato diseases can be tracked both during the growing phase and during the harvest and storage phase. During the growing phase, we have identified two categories of scientific work. The first category examines the separate leaves [3], and their images reveal the disease. The second category concerns images of entire areas of plant fields [4], reveals foci of potato disease and the degree of its damage. In the first case, a laptop with an installed and trained program and a web camera is sufficient, and in the second case, the use of a drone is still required.

The peculiarities of these two types of problems are that no real-time algorithms are required for image analysis. The solutions used in these problems are not applicable for fast processes, where the data processing must be carried out almost instantly.

At the stage of harvesting and storage, the tubers themselves are examined directly. The examination of tubers can also be divided into two categories. These are laboratory studies that are carried out with selected tubers or studies in natural or close to natural conditions. The first, for example, is the use of MRI to determine internal defects in potato tubers [5]. Determination of damaged and diseased tubers in natural production conditions is very different from studies of piece and dirt-free material. Therefore, most of the methods applied in the laboratory are not suitable in the field. Deep neural networks, which work great for single instances, are practically useless in the conditions faced by vegetable storage workers. These are insufficient and uneven illumination, the presence of adhered dirt, the overlap of tubers, etc. Only a few methods can be used in these conditions. These include, for example, the SVM method, which in combination with other methods, shows good results when classifying potato tubers [6,7].

There are also unique expert systems, which, based on many input characters using trained neural networks, make it possible to determine the potato variety or disease that infects the tuber [8].

Different versions of conveyor belts are used when passing vegetables, fruits, and root crops in vegetable stores, depending on the type of activity, like a bookmark for storage, bulkhead, or unloading from a vegetable store. It is at this stage that it is advisable to organize the selection of healthy agricultural food products for the injured and sick [9–17].

After examining the works devoted to sorting agricultural products using computer vision systems, we found that most of the research is related to artificial restrictions imposed on the system. For example, J.F. Bautista et al. [9] developed a mechanism for sorting tomatoes by quality. Each fruit that comes into the field of view of a video camera stops the conveyor, and the computer vision system determines the class of the fruit. Following the class, the sorting mechanism directs it to the appropriate container. The processing time for one fruit is 9 s.

J.J. Jijesh et al. [10] proposed a system for sorting agricultural products without stopping the conveyor while processing images of fruits. This system matches the speed of the conveyor to the rotation of the sorting bar. Despite the constant speed of the conveyor, the sorting system has a low throughput.

For precise fixation of objects on the conveyor belt, J.E. Cruz et al. [11] proposed installing on the conveyor infrared sensors at regular intervals. The system captures piece objects well. It can be used to organize the counter of objects, but it is not applicable for the bulk processing of agricultural products.

We believe that:

- Computer vision systems working with piece objects will not be able to realize the throughput required for a vegetable store;

- It is necessary to develop a sorting mechanism that will allow removing damaged objects from a moving conveyor belt at the recommended speeds of its movement (Table 1), following their current coordinates established by the computer vision system.

Table 1. Recommended speeds of the conveyor for root crops [18].

Belt Speed v , m/s, with Belt Width B , mm								
B	300–500	650	800	1000	1200	1400	1600	2000
v	0.8	0.8	1	1	1	1	1	1

Each disease has its characteristic features and reveals itself at different times of growth and storage of potatoes [19] (Figure 1). We chose the most harmful of them—late blight as the subject of our research (Figure 1(1)). The main danger of the disease is the high rate of its development. The annual shortage of potato tubers due to late blight is about 10% of the gross harvest. In the case of late blight’s significant and early damage to the potato tops, the yield shortfall reaches 70% or more. Storing potatoes with a large number of infected tubers often leads to the rotting of the entire batch.

On the affected tubers, lead-gray or brown (depending on the variety and color of the skin), slightly depressed hard spots are formed, extending inward in the form of uneven brown smudges (“tongues”).

Late blight rot often turns into dry Fusarium rot (Figure 1(4)) by the middle or end of storage.

For a selection of tubers for planting, tubers damaged by rodents can be considered acceptable (Figure 2(4–6)). If potatoes are selected for sale, they must be discarded.

Figure 1. Potato Disease Identification (1. Late blight, 2. Skin spot, 3. Gangrene, 4. Dry rot, 5. Powdery scab, 6. Tobacco Necrosis Virus, 7. Common scab, 8. Silver scurf, 9. Potato Virus Y).

The authors faced mainly three problems:

- Tubers moving on the conveyor are captured by the camera several times, and each time they are perceived by the system as different. So, re-capturing the image slows down the system significantly.
- While moving on a conveyor, tubers may overlap each other and thereby hinder precise shape recognition.
- The problem of direct identification of tubers affected by disease or rodents.

When solving the identified problems, the authors studied articles on computer vision that solve similar tasks in various spheres of human activity to some extent.

The first task is to process one object once. We believe that the problem is similar to the problem of vehicle detection in a complex driving environment [20–23].

S. Aqel et al. [20] split this task into several subtasks. First, moving vehicles are detected using the background subtraction method. The researchers used morphological operators to reduce false areas and remove moving shadows, and finally, they performed classification using invariant Charlier moments. F. Liu et al. [21] studied the problem of real-time updating the background. They suggested using an indicator to avoid excessive background refresh. A. Soin and M. Chahande [22] found solutions to vehicle detection and recognition using deep neural networks (DNNs).

Y. Wei et al. noted that when vehicles move between rows, one of them may overlap the other. In their paper [23], they proposed using object tracking algorithms. This made it possible to significantly reduce the processing time of the video signal since the processed objects were not subjected to repeated analysis. The vehicle detection algorithm itself consisted of two stages: target extraction using the strong descriptive power of the HOG and segmentation of the region of interest (ROI region) based on Haar cascades.

The authors of [24] successfully implemented the problem of partial overlapping of objects. To accurately capture moving objects on weak mobile devices, they used the Boosted HAAR Cascade method. It significantly outperforms deep neural networks in image processing speed and allows the detection of partially closed objects.

We believe that:

- The background subtraction method is applicable in conveyor conditions;
- Getting rid of shadows is carried out by selecting the position of light sources;
- Elimination of partial overlapping of objects is carried out by metering the supply of these objects to the conveyor and increasing the conveyor speed.

Figure 2. Examples of potato tubers: (1–3) dry rotted tubers; (4–6) tubers affected by rodents; (7–9) healthy tubers.

To recognize damaged tubers, we must consider a set of factors related to the conditions in which tubers are selected. Most often, convolutional neural networks (CNNs) are used to solve such problems, which have recently significantly improved their performance [25–31]. However, as the authors of [32] have shown, convolutional neural networks working with high-resolution images are not intended to be implemented on devices with weak processors. It is necessary to use large kernels (for example, 7×7 or 9×9) or a large number of layers to obtain an acceptable field susceptibility with convolutional layers [33]. Both of these schemes lead to a very significant slowdown in the system. Therefore, most low-performance systems are limited to image sizes of less than 41×41 pixels to achieve

acceptable image processing time on low-performance devices. Moreover, the processing time for each such frame can reach several seconds. In the conditions of a continuously moving conveyor, this is unacceptable.

Under these conditions, algorithms with a favorable speed/resources and power ratio deserve special interest [34]. For this reason, several methods of equipment optimization have recently been proposed [35–38].

We believe that attention should be paid to methods that use descriptor structures. A descriptor is a method that identifies a certain area of an image based on a set of features. Having studied the material over the past three years, we concluded that it is advisable to use the following descriptors in our work: HOG [38,39] and HAAR features [40].

Saad Abouzahir et al. [38] used HOG to identify weeds in crop images. They showed that the method in its classical form does not work well with images of damaged and deformed leaves. The use of BOVW in combination with HOG has significantly improved the quality of the material supplied to the classifier. With the BPNN, the weed detection rate ranged from 93% to 98%, depending on the crop. This combined approach significantly exceeded the results of other descriptors considered in their work.

Various works have shown that image processing speed and object classification quality increase significantly if the images are subjected to preliminary processing. Image transformation is implemented using filters [41–43], thresholding operations [44–46], morphological operations [41,47], and artificial neural networks [48,49].

M. Yogeshwari et al. [44] used the Adaptive Otsu's thresholding algorithm to binarize images of crop leaves. This algorithm made the architecture much easier and accelerated the operation of the convolutional neural network used as a classifier.

L.P. Saxena [46] used Niblack's local binarization method and its modifications in real-time applications. These techniques work to distinguish objects in their background. The author noted that these methods are helpful in various applications, such as restoration of damaged documents or manuscripts, text search in video frames, license plate recognition, and reading product barcodes.

Artificial neural networks process images mainly to recognize the depicted objects. A training sample is collected, the neural network is trained, if the necessary objects are detected on the test image, they are replaced with the corresponding analogs. S. Kang et al. [48] used cascading modular U-networks (CMU-Nets) to binarize the document image. The researchers solved in their work the problem of the insufficient number of training samples. In the case of text documents, the method showed high accuracy and productivity higher than analogs [49]; however, despite this improvement, the productivity of the method does not allow it to be used in real-time applications.

The simplest morphological operations are used primarily for noise reduction. However, they can be used to implement more complex image processing techniques. E. Imani et al. [47] managed with high precision to separate lesions of the eyeball and blood vessels. They used the Morphological Component Analysis (MCA) algorithm. This method allows separating structurally different objects, for each of which different transformations work effectively.

Der-Chang Tseng et al. [41] proposed to break down the image noise removal process into stages. The initially noisy image is analyzed using the enhanced MCA algorithm. After a series of transformations, the image is broken down into texture, structure, and image edges. In turn, noise removal is performed for each individual part. The image texture is removed using the BM3D algorithm, while part of the structure is removed using the ANLM method with an adaptive search box. The edges of the image are indicated by the K-SVD method. The first two are filters. The third (K-SVD) is a generalization of the k-means clustering method. The main idea of K-SVD is to split data into small chunks and alternate sparse coding and dictionary updates.

Using filters can be extremely useful to improve the results and speed of the method. D. Sharifrazi et al. [42] showed that using a Sobel filter improves the performance of a

convolutional neural network for detecting COVID-19 using X-ray images. They ranked this combination of methods as the best of the wide variety of options.

G. Ravivarma et al. [43] compared Sobel, Canny, Prewitt, Roberts, and fuzzy logic methods for edge detection. They noted that the Sobel filter has the smallest temporal and spatial complexity in comparison with the others. Testing was carried out on pest-affected leaves of crops.

We believe that we should use real-time algorithms for image preprocessing. Threshold binarization algorithms (Adaptive Otsu's, Niblack's) and the simplest morphological operations for noise suppression (erosion, closure, opening, etc.) are suitable. And we can improve the result with a filter (one of the most promising is the Sobel filter). It is not advisable to use artificial neural networks at this stage.

U. Ahmad et al. considered in their work [50] the decrease in the accuracy of image processing depending on the speed of the conveyor belt. The speed varied from 0.08 to 0.3 m/s. The researchers found a significant increase in the error at the lower and upper limits of the speed (on average, two times) in determining the color index, saturation, and intensity of images, which significantly influenced the method's result for the classification of mango fruits. In an actual vegetable store, this conveyor speed is unacceptable. When analyzing the design of the fetal image-recording camera [50] (Figure 3), we concluded that we could significantly improve its characteristics by using pulsed light with a frequency that coincides with the frequency of taking pictures.

Figure 3. Original and processed images of potatoes (1. The original image; 2. The image converted to grayscale; 3. Threshold binarization applied (Otsu's method); 4. The adaptive binarization applied (Niblack's method)).

For an apparent fixation of defects in vegetables, fruits, and root crops, in some studies, the researchers used video cameras operating in the infrared range [51–55].

P.V. Balabanov et al. [51] studied the emission spectra of healthy and diseased fruits. They used a Vis-NIR (Visible—Near Infrared) hyperspectral camera in the 400–1000 nm range. Studies have shown that all significant differences in the radiation of healthy and diseased fruits are in the visible part of the spectrum. A. Ibrahim et al. [54] examined the possibility of identifying internal damage to potatoes resulting from impacts during harvest (blackspot) by examining the absorption of a wavelength of 730 nm (near-infrared). They found that the tuber's damaged and undamaged inner parts have similar characteristics for most of the samples. They concluded that infrared radiation, located at the border with the visible part of the spectrum, is not suitable for solving such problems.

A. López-Maestresalas et al. [52] studied subsurface damage to potatoes (blackspot) using hyperspectral systems like Vis-NIR and SWIR (shortwave infrared) in the 1000–2500 nm range. They found out that on the three studied potato varieties, the SWIR system allows determining the presence of blackspot with an accuracy of higher than 93%, five hours

after harvesting; Vis-NIR also detects subsurface damage, but with less accuracy. It was noted that at this stage of research, hyperspectral systems of this kind are not suitable for operation in the industry and can only be used in laboratory conditions.

P.V. Balabanov et al. [56,57] suggested using thermal methods of indestructible and non-contact control with technical vision systems in the infrared spectral range of 8–14 microns when sorting agricultural products. The method is based on the fact that damaged, diseased and healthy plant tissues have different thermophysical characteristics. During the implementation of contactless measurements, the surface of the object under the study was heated using a laser [56] and an IR radiation source [57]. A FLIR A35 thermal imager was used to obtain information.

After analyzing the results of [51–57], we came to the following conclusions:

- It is enough to use video cameras operating in the visible part of the spectrum to identify diseased tubers;
- Some laboratory methods are available to detect subsurface damage, but these methods are not suitable for vegetable stores.

2. Materials and Methods

Grayscale transition and threshold binarization

For binarization, we considered two algorithms: Otsu's method and Niblack's method.

Otsu's method is a threshold binarization algorithm. The choice of the method is due to the following properties:

- ease of implementation;
- adaptation to various kinds of images by choosing the optimal threshold;
- fast lead time.

With this method, a threshold t is calculated to minimize the average segmentation error, i.e., the average error of deciding whether the image pixels belong to an object or a background. The image pixels' brightness values can be considered random, and their histogram can be taken as an estimate of the probability distribution density. If the probability distribution densities are known, then it is possible to determine the optimal (in the sense of the minimum error) threshold for image segmentation into two classes c_0 and c_1 (objects and background).

The histogram is plotted according to the values $p_i = n_i/N$. In this formula, N is the total number of image pixels, n_i is the number of pixels with a brightness level i ($0 \leq i \leq L$). The threshold t is an integer value from 0 to $L = \max$. With the help of the histogram, we can divide all the pixels into "useful" (object) and background ones. Relative frequencies W_0 and W_1 correspond to each type:

$$W_0(t) = \sum_{i=1}^t p_i \quad (1)$$

$$W_1(t) = \sum_{i=t+1}^L p_i = 1 - W_0(t) \quad (2)$$

Next, we calculate the average levels for each type of image using the formulas:

$$\mu_0(t) = \sum_{i=1}^t \frac{ip_i}{W_0} \quad (3)$$

$$\mu_1(t) = \sum_{i=t+1}^L \frac{ip_i}{W_1} \quad (4)$$

Next, we find a threshold that reduces the variance of pixels of a particular type, determined by the following formula:

$$\delta_W^2(t) = W_1(t)\delta_1^2(t) + W_2(t)\delta_2^2(t) \quad (5)$$

The next step is to determine the interclass variance using the formula below:

$$\sigma_{cl}^2(t) = W_0(t)W_1(t) * (\mu_1(t) - \mu_0(t))^2 \quad (6)$$

Then the maximum value is calculated to assess the quality of dividing the image into two parts, which corresponds to the desired threshold:

$$\eta(t) = \max \left[\frac{\sigma_{cl}^2(t)}{\sigma_W^2(t)} \right] \quad (7)$$

Figure 3(3) and Figure 4(3) show a fragment of the image binarization by Otsu's method.

Figure 4. Inverted image and its processing (1. Inverted image; 2. Inverted image converted to grayscale; 3. Threshold binarization applied (Otsu's method); 4. Adaptive binarization applied (Niblack's method)).

In addition, we considered the possibility of using adaptive binarization (Niblack's algorithm), where the global binarization threshold for the entire image is not sought, but local information is used.

The idea behind this method is to vary the brightness threshold B of binarization from dot to dot based on the local value of the standard deviation. The brightness threshold at the (x, y) dot is calculated as follows:

$$B(x, y) = \mu(x, y) + k * s(x, y), \quad (8)$$

where $\mu(x, y)$ is the mean and $s(x, y)$ is the standard deviation of the sample for some neighborhoods of the dot. The size of the neighborhood should be as small as possible but sufficient to preserve local image details. At the same time, the size should be large enough to reduce the effect of noise on the result. The value of k determines which part of the object's border to take as the object itself. A value of $k = -0.2$ specifies a fairly good separation of objects represented in black and $k = +0.2$ when objects are in white.

Figure 3(4) and Figure 4(4) show a fragment of the image binarization by Niblack's method.

The authors applied the considered two binarization methods to both usual and inverted images of potatoes (Figures 3 and 4)

Figure 3 shows that tubers in this representation are difficult to distinguish as separate objects, but if we knew in advance the location of each tuber in the image, damage to the tubers themselves is very noticeable. Moreover, it is easy to calculate the damaged area.

Based on this, the authors proposed determining the tubers' location in the image, not in the usual color but the inverted one with subsequent binarization.

Thus, by comparing the two representations of tubers in Figures 3 and 4, in one case, it is quite easy to identify the locations of tubers, and in the other case, to identify tubers with damage.

Using filters to update boundaries

The Sobel operator is used to define the boundaries of objects. This operator is based on the convolution of the image with integer filters. The operator uses the G_x and G_y kernels, with which the image is convolved to calculate the horizontal and vertical derivatives:

$$G_x(t) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad G_y(t) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (9)$$

This operator is used to approximate the gradient of the pixel intensity function. To detect the presence of gradient discontinuity, we can calculate the gradient change at the dot (i, j) . This can be done by finding the following value:

$$|G^{ij}| = \sqrt{(G_x^{ij})^2 + (G_y^{ij})^2} \quad (10)$$

The following expression determines the direction of the gradient Q :

$$Q^{ij} = \arctan\left(\frac{G_y}{G_x}\right) \quad (11)$$

Figure 5 shows the result of applying the Sobel filter to the image shown in Figure 2. When implementing the method, the kernels (9) were used.

Figure 5. Processing the image of potato tubers in Figure 2 using the Sobel filter: (1–3) dry rotted tubers; (4–6) tubers affected by rodents; (7–9) healthy tubers.

The use of the descriptor

The HOG method assumes that the type of distribution of image intensity gradients makes it possible to accurately determine the presence and shape of objects present on it.

The image is split into cells. The histograms h_i of the directional gradients of the interior dots are calculated in the cells. They are combined into one histogram ($h = f(h_1, \dots, h_k)$), after which it is normalized to brightness. We can obtain the normalization factor in several ways, but they show approximately the same results. We will use the following equation:

$$h_L = \frac{h}{\sqrt{\|h\|_2^2 + \epsilon^2}}, \quad (12)$$

where $\|h\|_2$ is the used norm, ε is some small constant.

When calculating the gradients, the image is convolved with the kernels $[-1, 0, 1]$ and $[-1, 0, 1]^T$, resulting in two D_x and D_y matrices of derivatives along the x and y axes, respectively. These matrices are used to calculate the angles and values (moduli) of the gradients at each dot in the image.

Figure 6 shows the result of applying the HOG method to the image shown in Figure 2. The gradient value only is shown for clarity (the brighter the pixel, the larger the gradient).

Figure 6. Processing the image of potato tubers in Figure 2 by the HOG method: (1–3) dry rotted tubers; (4–6) tubers affected by rodents; (7–9) healthy tubers.

The SIFT descriptor is used to extract feature points from the image, which are later used in classifiers. The key point in finding them is building a pyramid of Gaussians and the difference of Gaussians. Gaussian—image blurred with a Gaussian filter:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \quad (13)$$

where $L(x, y, \sigma)$ is the value of the Gaussian at the dot with coordinates (x, y) and blur radius σ ; $G(x, y, \sigma)$ —Gaussian kernel; $I(x, y)$ —the value of the original image; $*$ —convolution operation.

The difference of Gaussians is an image obtained with pixel-by-pixel subtraction of the Gaussian of the original image from a Gaussian with a different blur radius ($k\sigma$):

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (14)$$

A pyramid of Gaussians and Gaussian differences is built. When moving from one level of the pyramid to another, the dimensions of the images are halved.

After building the pyramids, key points are determined, which are the local extrema of the differences between the Gaussians. False key points are discarded, and for the remaining ones, their orientation is calculated. We determine the gradient's value m and direction θ from the formula:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (15)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (16)$$

The SIFT method operates the descriptor as a vector. The method takes a 4×4 square area centered at the particular dot and rotates it according to the singular point's direction. Each element of the area indicates the value of the gradient in eight directions.

Viola-Jones method

This algorithm uses Haar features [40] to classify objects in the image. These features are similar to convolution kernels and are rectangular regions composed of several adjacent parts (Figure 7). The Viola-Jones method uses the AdaBoost algorithm to construct a cascading classifier. When forming each new level, the AdaBoost algorithm selects the most informative features. The formation of the classifier ends when a predetermined target quality of the classifier is reached.

Figure 7. Haar features.

Figure 8 shows the result of applying the Viola-Jones method.

Figure 8. Identification of potato tubers by the Viola-Jones method.

For the correct operation of the method, the article's authors investigated several options for image preprocessing for subsequent training and a combination of trained classifiers.

3. Results

For video shooting, we took a 4k action camera with a frequency of 25 fps. For reliable fixation of the image, we used pulsed light of the same frequency. A rectangular pulse generator controlled frequency and duty cycle. Several white-LED strips were fixed on a rectangular frame of 800×500 mm, having a total luminous flux of 4500 LM in regular operation. At this luminous flux value, the duty cycle was selected manually and varied from 15% to 25% to obtain a clear image. Pulsed light made using a conventional action camera possible instead of an expensive industrial camera [51]. The number of frames per second was adjusted so that the camera would shoot each tuber two times. This value is four frames per second with an aspect ratio of 16:9, a conveyor width of 800 mm, and a conveyor speed of 1 m/s.

As shown above, due to the large number of computational operations required for convolutional networks, it is not feasible to use them for fast object recognition in high-resolution images.

We divided the diagnostic procedure into separate phases, allowing us to speed up identifying individual tubers in the video stream and direct analysis.

First phase: identifying individual tubers in the image. Because the system must process 4–5 images with a resolution of 3840×2160 pixels per second, the Viola-Jones

method [58–63] is the most suitable method for detecting tubers. But even taking into account the high performance of this algorithm, the processing of such an image format will take an unacceptably long time. Therefore, it is necessary to reduce the resolution by order of magnitude along each coordinate axis (384×216 pixels). For a medium tuber 60×50 mm, the photo now has 30×25 pixels. Thus, we choose a scanning window of 30×25 pixels.

We trained the Haar cascade and implemented the Viola-Jones recognition algorithm using Python's standard OpenCV Traincascade application. We used 850 images of tubers (positive images) and 1000 images of the working area of the conveyor without control objects (negative images) as training data.

However, in practical use, the classifier showed relatively low results (Figure 8). Up to half of the tubers are not detected by this method. In addition, the method classified objects that are not tubers as tubers.

To improve the classification results, the authors applied image preprocessing using the Sobel filter. As a result, the probability of detecting tubers in the case of their location in one layer reached 97%.

When recognizing tubers, they are marked, their images are recorded and sent for further processing. Considering that each tuber appears on the images at least twice, further work is performed with only one of them (not the cropped edge of the video frame). This is an optimization element that allows you to speed up the algorithm significantly.

Based on the results of phase 1, we receive the selected images of tubers for subsequent diagnostics.

Second phase:

Option 1: SIFT-SVM

Selecting tubers depends on the defined tasks. The use of the SIFT descriptor followed by classification using the SVM method identifies damage localized in small areas—over 95% (Figure 9(1)), but much worse identifies damage in large areas—52% (Figure 9(2)).

Figure 9. Identification of potato tubers by the Viola-Jones method: (1) damage localized in small areas, (2) damage in large areas.

Option 2: HOG-CNN and HOG-BOVW-BPNN

HOG is another promising candidate for the role of a reliable descriptor. According to the results of experiments, HOG with CNN, when identifying damaged tubers, shows results up to 75%, and in the case of diseased tubers, 67%. It should also be noted that the classic HOG method is relatively slow. The proposed BOVW-based HOG method works much faster and better.

For more accurate results, the image is grayscaled to minimize noise and brightness effects. Then the histogram is flattened. The size of training images is taken at a minimum of 384×384 and can be increased horizontally and vertically by a multiple of 64. The image is divided into small grids (128×128 with 50% overlap). Standard HOG is applied. The result is transferred to the stage of additional processing. Additional processing aims to reduce the further calculations significantly. The potato tuber has no predominant direction and is practically symmetrical concerning rotation. A simple transformation of matrices can reduce the possible options several times. This reduces the number of effective clusters and the number of letters in the visual dictionary (Figure 10).

Figure 10. Steps to create Visual-Words.

For a binary classification: damaged or not damaged tuber, we trained a BPNN.

Option 3: Otsu's Threshold Binarization—CNN

For binary classification, we used another method. To achieve greater accuracy, in addition to the usual images of the potato tubers, we also used their inverted copies. Otsu's threshold binarization was applied to both the images and their inverted copies. These images were transferred to two classifiers, for which we used convolutional networks with the dimension of convolutional kernels 3×3 (Figure 11). Reducing the dimension of convolutional kernels made it possible to significantly speed up the work of classifiers. The results of the work of both classifiers complement each other.

Figure 11. Generalized scheme of the algorithm.

Comparison and selection of options for implementing the second phase

Table 2 shows the implementation methods and the results of the second phase of the algorithm.

Table 2. Implementation methods and results of the second phase of the algorithm.

Preprocessing Methods and Descriptors	Used Classifiers	Result
SIFT	SVM	52–95% depending on damage types
HOG	CNN	67–75%
HOG-BOVW	BPNN	80–95%
Otsu threshold binarization	CNN	85–97%

It should be noted that the illumination and overlapping of the tubers had a strong influence. The displacement of the lamp at a considerable distance from the video camera, the mismatch between the flash firing frequency and the fps of the video camera leads to a significant deterioration in the results of the algorithms. Algorithms based on HOG turned out to be less sensitive to such mismatches of systems. It should also be noted that the HOG-BOVW algorithm is significantly faster than CNN.

Tubers overlap can be partially eliminated by increasing the conveyor speed. In this case, the recognition accuracy is significantly increased. However, you should take into account the recommended speeds of the conveyors for the transportation of root crops

(Table 1). In addition, an increase in speed leads to a deterioration in the clarity of the resulting images.

To speed up the operation of convolutional networks, the authors propose to reduce the dimension of convolutional kernels to 3×3 but apply convolutional networks not to the original images (Figure 2) but to the images subjected to differentiation using the Sobel filter (Figure 5).

4. Discussion

At the modern stage of the development of computer vision technologies, the question of object classification makes no longer a problem. Convolutional neural networks, decision trees, etc., have been performing it much more accurately than a human would have done. However, we drew attention to several limitations of these methods associated with their application in real conditions of vegetable storage. So, for example, based on the conclusions of modern researchers, convolutional networks that are most promising for solving computer vision problems are not able to process the stream of images from a video camera of rapidly moving objects of small size and a significant number [32]. The algorithm, launched on inexpensive computer hardware, simply cannot keep up with the conveyor belt. And obsolete, even if properly processed, data will be of little use if damaged objects have already left the conveyor belt.

We proposed using the Viola-Jones algorithm at the first stage of processing the image from a video camera, which, unlike convolutional neural networks, works in a real-time mode [58–63]. This method was created for recognizing human faces and did not give good results when used to detect potato tubers; however, by selecting preprocessing filters, we achieved a probability of 97%, which corresponds to the results of a convolutional neural network (from 91 to 95% in works on convolutional networks for the last three years) [25–31].

At the second stage, we work with an image in which the sizes and coordinates of the tubers have already been determined. The classification task is reduced to determining whether the tuber is suitable for further use or if it should be recognized as damaged or diseased. Dividing the image into small fragments simplifies the task greatly. When processing small images, a convolutional neural network installed on a personal computer is capable of processing up to a hundred images per second. This corresponds to a complete classification of tubers that slightly fill the conveyor and move at low speed. The result of correct image classification by the convolutional network was up to 97%. On average, 50% more tubers are processed by the combination of HOG-BOVW-BPNN methods, but its classification result is about 2% less. These results significantly exceed the processing speed of computer vision systems installed on conveyors [9–17]. We plan to continue improving the second stage of the image processing in the direction of increasing the number of classified tubers per second and bringing it to several hundred, which will allow installing the computer vision system on conveyors with almost any load and moving at the maximum permissible speed. For this, we propose to create systems of parallel computation and split the objects selected at the first stage into several parallel processed threads.

The used methods allow us to find the percentage of damaged tubers and the number of tubers. Taking into account the area of each tuber in the image, we can determine the mass of potatoes passed along the conveyor with an error of up to 10%. This is consistent with the materials of the article by A. Kalantar et al., who determined the weight of agricultural products from its image [64].

We believe that the proposed algorithm can be used in a picking robot installed on the conveyor belt of a vegetable store [65,66].

5. Conclusions

For a reliable video recording of objects moving on the conveyor, it is necessary to use pulsed light with a pulse frequency equal to the fps of a video camera. Selecting the duty

cycle and luminous flux can significantly improve the image characteristics even using a cheap action camera.

It is enough to use a video camera of the visible range of light to detect external signs of potato tubers disease. All developments with the identification of internal damage involve the use of expensive video cameras in the infrared spectrum; however, the methods proposed in the articles do not allow them to be used in a vegetable store.

When separating potato tubers from a conveyor image, most computer vision methods are not suitable. The Viola-Jones method turned out to be the most acceptable for this task, which works much faster than similar methods used to detect objects. Provided that the image is preprocessed, the quality of the selection of tubers is comparable.

When objects have already been selected, and you need to classify them, you can use various methods, each of which has advantages and disadvantages. The choice of the method may depend on the conditions of its use and the classified objects. For example, SIFT-SVM is great for detecting dry rot, and HOG-BOVW-BPNN is better for unbalanced or low light conditions.

Author Contributions: Conceptualization, D.V.S.; Data curation, S.A.K.; Formal analysis, S.V.S.; Project administration, S.T.G.; Resources, M.V.S.; Software, A.V.O.; Validation, K.V.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors are grateful to the Dean of the Faculty of Information Technologies and Big Data Analysis, Vladimir Soloviev, for a helpful discussion of this work and for comments on this study, which allowed us to improve the quality of the material. We also thank the Agronomy reviewers for quality peer review and related comments.

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

References

1. Arthurs, S.; Dara, S.K. Dara Microbial biopesticides for invertebrate pests and their markets in the United States. *J. Invertebr. Pathol.* **2019**, *165*, 13–21. [[CrossRef](#)] [[PubMed](#)]
2. Keswani, C.; Dilnashin, H.; Birla, H.; Pratap, S. Singh Regulatory barriers to Agricultural Research commercialization: A case study of biopesticides in India. *Rhizosphere* **2019**, *11*, 100155. [[CrossRef](#)]
3. Argüeso, D.; Picon, A.; Irusta, U.; Medela, A.; San-Emeterio, M.G.; Bereciartua, A.; Alvarez-Gila, A. Few-Shot Learning approach for plant disease classification using images taken in the field. *Comput. Electron. Agric.* **2020**, *175*, 105542. [[CrossRef](#)]
4. Gao, J.; Westergaard, J.C.; Sundmark, E.H.R.; Bagge, M.; Liljeroth, E.; Alexandersson, E. Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning. *Knowl. Based Syst.* **2021**, *214*, 106723. [[CrossRef](#)]
5. Hajjar, G.; Quellec, S.; Pépin, J.; Challos, S.; Joly, G.; Deleu, C.; Leport, L.; Musse, M. MRI investigation of internal defects in potato tubers with particular attention to rust spots induced by water stress. *Postharvest Biol. Technol.* **2021**, *180*, 111600. [[CrossRef](#)]
6. Payman, M.; Navid, R.; Mohsen, A. Computer vision-based potato defect detection using neural networks and support vector machines. *Int. J. Robot. Autom.* **2013**, *28*, 1–9.
7. Wang, C.; Li, X.; Wu, Z.; Zhou, Z.; Feng, Y. Machine vision detecting potato mechanical damage based on a manifold learning algorithm. *Nongye Gongcheng Xuebao/Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 245–252.
8. Przybyl, K.; Boniecki, P.; Koszela, K.; Gierz, L.; Lukomski, M. Computer vision and artificial neural network techniques for classification of damage in potatoes during the storage process. *Czech J. Food Sci.* **2019**, *37*, 135–140. [[CrossRef](#)]
9. Bautista, J.F.; Oceña, C.D.; Cabrerós, M.J.; Alagao, S.P.L. Automated Sorter and Grading of Tomatoes using Image Analysis and Deep Learning Techniques. In Proceedings of the 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Manila, Philippines, 3–7 December 2020; pp. 1–6.
10. Jijesh, J.J.; Shankar, S.; Ranjitha; Revathi, D.C.; Shivaranjini, M.; Sirisha, R. Development of Machine Learning based Fruit Detection and Grading system. In Proceedings of the 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), Bangalore, India, 12–13 November 2020; pp. 403–407.
11. De la Cruz, J.E.C.; Ramirez, O.J.V. Convolutional neural networks for the Hass avocado classification using LabVIEW in an agro-industrial plant. In Proceedings of the 2020 IEEE XXVII International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Lima, Peru, 3–5 September 2020; pp. 1–4.

12. Örnek, M.; Haciseferoğulları, H. Design of Real Time Image Processing Machine for Carrot Classification. *Yüziüncü Yıl Üniversitesi Tarım Bilimleri Derg.* **2020**, *30*, 355–366. [CrossRef]
13. Bazame, H.C.; Molin, J.P.; Althoff, D.; Martello, M. Detection, classification, and mapping of coffee fruits during harvest with computer vision. *Comput. Electron. Agric.* **2021**, *183*, 106066. [CrossRef]
14. Istiadi, A.; Sulistiyanti, S.R.; Fitriawan, H. Model Design of Tomato Sorting Machine Based on Artificial Neural Network Method Using Node MCU Version 1.0. *J. Phys.* **2019**, *1376*, 012026. [CrossRef]
15. Putra, K.T.; Hariadi, T.K.; Riyadi, S.; Chamim, A.N.N. Feature Extraction for Quality Modeling of Malang Oranges on an Automatic Fruit Sorting System. In Proceedings of the 2018 2nd International Conference on Imaging, Signal Processing and Communication (ICISPC), Kuala Lumpur, Malaysia, 20–22 July 2018; pp. 74–78.
16. Behera, S.K.; Jena, L.; Rath, A.K.; Sethy, P.K. Disease Classification and Grading of Orange Using Machine Learning and Fuzzy Logic. In Proceedings of the 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 3–5 April 2018; pp. 0678–0682.
17. Abbas, H.M.T.; Shakoar, U.; Khan, M.J.; Ahmed, M.; Khurshid, K. Automated Sorting and Grading of Agricultural Products based on Image Processing. In Proceedings of the 2019 8th International Conference on Information and Communication Technologies (ICICT), Karachi, Pakistan, 16–17 November 2019; pp. 78–81.
18. Manual Design Conveyor Transport Belt Conveyors (to SNIP 2.05.07-85), All-Union Design and Research Institute Industrial Transport (Promtransniiproekt) Gosstroya USSR, Moscow Stroyizdat 1988. Available online: https://xn--c1ahwb.xn--p1ai/uploadedFiles/files/Metodika_rascheta_lentochnykh_konveyerov_k_SNiP_2.05.07-85.pdf (accessed on 1 September 2021).
19. Khalil, I. Al-Mughrabi Biological Control of Fusarium Dry Rot and Other Potato Tuber Diseases Using *Pseudomonas fluorescens* and *Enterobacter Cloacae*. *Biol. Control* **2010**, *53*, 280–284. Available online: https://www.researchgate.net/publication/240445223_Biological_control_of_Fusarium_dry_rot_and_other_potato_tuber_diseases_using_Pseudomonas_fluorescens_and_Enterobacter_cloacae (accessed on 2 July 2021).
20. Aqel, S.; Hmimid, A.; Sabri, M.A.; Aarab, A. Road traffic: Vehicle detection and classification. In Proceedings of the Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, 17–19 April 2017; pp. 1–5.
21. Liu, F.; Zeng, Z.; Li, Z. A fast background update mechanism for vehicle detection in urban roads. In Proceedings of the 2017 9th Computer Science and Electronic Engineering (CEECE), Colchester, UK, 27–29 September 2017; pp. 60–64.
22. Soin, A.; Chahande, M. Moving vehicle detection using deep neural network. In Proceedings of the 2017 International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT), Dehradun, India, 17–18 November 2017; pp. 1–5.
23. Wei, Y.; Tian, Q.; Guo, J.; Huang, W.; Cao, J. Multi-vehicle detection algorithm through combining Harr and HOG features. *Math. Comput. Simul.* **2019**, *155*, 130–145. [CrossRef]
24. Raj, R.; Rajiv, P.; Kumar, P.; Khari, M.; Verdú, E.; Crespo, R.G.; Manogaran, G. Feature based video stabilization based on boosted HAAR Cascade and representative point matching algorithm. *Image Vis. Comput.* **2020**, *101*, 103957. [CrossRef]
25. Marino, S.; Beausery, P.; Smolarz, A. Weakly-supervised learning approach for potato defects segmentation. *Eng. Appl. Artif. Intell.* **2019**, *85*, 337–346. [CrossRef]
26. Afonso, M.; Blok, P.M.; Polder, G.; van der Wolf, J.M.; Kamp, J. Sydney, Australia.: Blackleg Detection in Potato Plants using Convolutional Neural Networks. *IFAC-Pap.* **2019**, *52*, 6–11.
27. Wu, A.; Zhu, J.; Ren, T. Detection of apple defect using laser-induced light backscattering imaging and convolutional neural network. *Comput. Electr. Eng.* **2020**, *81*, 106454. [CrossRef]
28. Kuznetsova, A.; Maleva, T.; Soloviev, V. Using YOLOv3 Algorithm with Pre- and Post-Processing for Apple Detection in Fruit-Harvesting Robot(Article). *Agronomy* **2020**, *10*, 1016. Available online: <https://www.mdpi.com/2073-4395/10/7/1016> (accessed on 2 July 2021). [CrossRef]
29. Korchagin, S.; Serdechny, D.; Kim, R.; Terin, D.; Bey, M. The use of machine learning methods in the diagnosis of diseases of crops. In *E3S Web of Conferences*; EDP Sciences: Les Ulis, France, 2020; Volume 176, p. 04011.
30. Marino, S.; Beausery, P.; Smolarz, A. Unsupervised adversarial deep domain adaptation method for potato defects classification. *Comput. Electron. Agric.* **2020**, *174*, 105501. [CrossRef]
31. Puno, J.C.V.; Billones, R.K.D.; Bandala, A.A.; Dadios, E.P.; Calilune, E.J.; Joaquin, A.C. Quality Assessment of Mangoes using Convolutional Neural Network. In Proceedings of the 2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), Bangkok, Thailand, 18–20 November 2019; pp. 491–495.
32. Yin, H.; Gong, Y.; Qiu, G. Fast and efficient implementation of image filtering using a side window convolutional neural network. *Signal Process.* **2020**, *176*, 107717. [CrossRef]
33. Shen, X.; Chen, Y.; Tao, X.; Jia, J. Convolutional neural pyramid for image processing. *arXiv* **2017**, arXiv:1704.02071v1.
34. Spagnoloa, F.; Perrib, S.; Corsonello, P. Design of a real-time face detection architecture for heterogeneous systems-on-chips. *Integration* **2020**, *74*, 1–10. [CrossRef]
35. Feng, X.; Jiang, Y.; Yang, X.; Du, M.; Li, X. Computer vision algorithms and hardware implementations: A survey. *Integration* **2019**, *69*, 309–320. [CrossRef]
36. Irgens, P.; Bader, C.; Lé, T.; Saxena, D.; Ababei, C. An efficient and cost effective FPGA based implementation of the Viola-Jones face detection algorithm. *HardwareX* **2017**, *1*, 68–75. [CrossRef]

37. Chandana, P.; Ghantasala, G.P.; Jeny JR, V.; Sekaran, K.; Deepika, N.; Nam, Y.; Kadry, S. An effective identification of crop diseases using faster region based convolutional neural network and expert systems. *Int. J. Electr. Comput. Eng.* **2020**, *10*, 6531–6540. [\[CrossRef\]](#)
38. Abouzahir, S.; Sadik, M.; Sabir, E. Bag-of-visual-words-augmented Histogram of Oriented Gradients for efficient weed detection. *Biosyst. Eng.* **2021**, *202*, 179–194. [\[CrossRef\]](#)
39. Aslan, M.F.; Durdu, A.; Sabanci, K.; Mutluer, M.A. CNN and HOG based comparison study for complete occlusion handling in human tracking. *Measurement* **2020**, *158*, 107704. [\[CrossRef\]](#)
40. Soleimanipour, A.; Chegini, G.R. A vision-based hybrid approach for identification of Anthurium flower cultivars. *Comput. Electron. Agric.* **2020**, *174*, 05460. [\[CrossRef\]](#)
41. Tseng, D.-C.; Wei, R.-Y.; Lu, C.-T.; Wang, L.-L. Image restoration using hybrid features improvement on morphological component analysis. *J. Electron. Sci. Technol.* **2019**, *17*, 100014. [\[CrossRef\]](#)
42. Sharifrazi, D.; Alizadehsani, R.; Roshanzamir, M.; Joloudari, J.H.; Shoeibi, A.; Jafari, M.; Hussain, S.; Sani, Z.A.; Hasanzadeh, F.; Khozeimeh, F.; et al. Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of COVID-19 patients using X-ray images. *Biomed. Signal Process. Control* **2021**, *68*, 102622. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Ravivarma, G.; Gavaskar, K.; Malathi, D.; Asha, K.G.; Ashok, B.; Aarthi, S. Implementation of Sobel operator based image edge detection on FPGA. *Mater. Today: Proc.* **2021**, *45*, 2401–2407.
44. Yogeshwari, M.; Thailambal, G. Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks. *Mater. Today Proc.* **2021**. [\[CrossRef\]](#)
45. Khairnar, S.; Thepade, S.D.; Gite, S. Effect of image binarization thresholds on breast cancer identification in mammography images using OTSU, Niblack, Burnsien, Thepade's SBTC. *Intell. Syst. Appl.* **2021**, *10–11*, 200046.
46. Saxena, L.P. Niblack's binarization method and its modifications to real-time applications: A review. *Artif. Intell. Rev.* **2019**, *51*, 673–705. [\[CrossRef\]](#)
47. Imani, E.; Javidi, M.; Pourreza, H.-R. Improvement of retinal blood vessel detection using morphological component analysis. *Comput. Methods Programs Biomed.* **2015**, *118*, 263–279. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Kang, S.; KenjiIwana, B.; Uchida, S. Complex image processing with less data—Document image binarization by integrating multiple pre-trained U-Net modules. *Pattern Recognit.* **2021**, *109*, 107577. [\[CrossRef\]](#)
49. Pratikakis, I.; Zagori, K.; Kaddas, P.; Gatos, B. ICFHR 2018 competition on handwritten document image binarization (H-DIBCO 2018). In Proceedings of the 2018 International Conference on Frontiers in Handwriting Recognition, Niagara Falls, NY, USA, 5–8 August 2018; pp. 489–493.
50. Ahmad, U.; Adji, M.A.P. Accuracy in estimating visual quality parameters of mango fruits as moving object using image processing. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2019; Volume 542, p. 012008.
51. Balabanov, P.V.; Divin, A.G.; Belyaev, P.S.; Trapeznikov, E.V.; Egorov, A.S.; Zaharov, Y.A.; Yudaev, V.A.; Don-Uni, N.P. Technical vision system for quality control of objects of the ball-shaped form when sorting on the conveyor. *J. Phys.* **2020**, *1546*, 012001.
52. López-Maestresalas, A.; Keresztes, J.C.; Goodarzi, M.; Arazuri, S.; Jarén, C.; Saeys, W. Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control* **2016**, *70*, 229–241. [\[CrossRef\]](#)
53. Divin, A.; Egorov, A.; Balabanov, P.; Pozhidaev, Y.; Lyubimova, D. Robotic complex for agricultural products sorting. *Int. Multidiscip. Sci. GeoConference SGEM* **2018**, *18*, 557–564.
54. Ibrahim, A.; Grassi, M.; Lovati, F.; Parisi, B.; Spinelli, L.; Torricelli, A.; Rizzolo, A.; Vanoli, M. Non-destructive detection of potato tubers internal defects: Critical insight on the use of time-resolved spectroscopy. *Adv. Hortic. Sci.* **2020**, *34*, 43–51.
55. Lu, Y.; Lu, R. Detection of Surface and Subsurface Defects of Apples Using Structured Illumination Reflectance Imaging with Machine Learning Algorithms. *Trans. ASABE* **2018**, *61*, 1831–1842. [\[CrossRef\]](#)
56. Balabanov, P.V.; Divin, A.G.; Savencov, A.P.; Shishkina, G.V. Algorithms for Detecting Potato Defects Using Images in the Infrared Range of Spectrum. In Proceedings of the 2018 IEEE International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), St. Petersburg, Russia, 24–28 September 2018; pp. 417–419.
57. Balabanov, P.V.; Divin, A.G.; Egorov, A.S.; Yudaev, V.A. Mechatronic system for fruit and vegetables sorting. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2019; Volume 734, pp. 18–21.
58. Ekanayake, J. Bug Severity Prediction using Keywords in Imbalanced Learning Environment. *Int. J. Inf. Technol. Comput. Sci. (IJITCS)* **2021**, *13*, 53–60.
59. Alyushin, M.V.; Alyushin, V.M.; Kolobashkina, L.V. Optimization of the Data Representation Integrated Form in the Viola-Jones Algorithm for a Person's Face Search. *Procedia Comput. Sci.* **2018**, *123*, 18–23. [\[CrossRef\]](#)
60. Andriyanov, N.; Dementev, V.; Tashlinskiy, A.; Vasiliev, K. The Study of Improving the Accuracy of Convolutional Neural Networks in Face Recognition Tasks. In *Pattern Recognition; ICPR International Workshops and Challenges; Lecture Notes in Computer Science; ICPR: New Delhi, India, 2021; Volume 12665*, pp. 5–14.
61. Andriyanov, N. First Step towards Creating a Software Package for Detecting the Dangerous States during Driver Eye Monitoring. In *Pattern Recognition; ICPR International Workshops and Challenges; Lecture Notes in Computer Science; ICPR: New Delhi, India, 2021; Volume 12665*, pp. 314–322.
62. Cao, L.; Li, H.; Zhang, Y.; Zhang, L.; Xu, L. Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. *Inf. Fusion* **2020**, *53*, 196–208. [\[CrossRef\]](#)

-
63. Zhang, M.; He, Z.; Zhang, H.; Tan, T.; Sun, Z. Toward practical remote iris recognition: A boosting based framework. *Neurocomputing* **2019**, *330*, 238–252. [[CrossRef](#)]
 64. Kalantar, A.; Edan, Y.; Gur, A.; Klapp, I. A deep learning system for single and overall weight estimation of melons using unmanned aerial vehicle images. *Comput. Electron. Agriculture*. **2020**, *178*, 105748. [[CrossRef](#)]
 65. Kuznetsova, A.; Maleva, T.; Soloviev, V. Detecting Apples in Orchards Using YOLOv3. In *Computational Science and Its Applications; Lecture Notes in Computer Science; ICCSA: Cagliari, Italy, 2020; Volume 12249*, pp. 923–934.
 66. Kuznetsova, A.; Maleva, T.; Soloviev, V. YOLOv5 versus YOLOv3 for apple detection. *Studies in Systems. Decis. Control* **2021**, *338*, 349–358.