Article

Detecting Apples in the Wild: Potential for Harvest Quantity Estimation

Artur Janowski [1], Rafał Kaźmierczak [2,†], Cezary Kowalczyk [2,*] and Jakub Szulwic [3,†]

[1] Department of Geodesy, Faculty of Geoengineering, University of Warmia and Mazury in Olsztyn, 10-719 Olsztyn, Poland; artur.janowski@uwm.edu.pl

[2] Department of Spatial Analysis and Real Estate Market, Faculty of Geoengineering, University of Warmia and Mazury in Olsztyn, 10-719 Olsztyn, Poland; rafal.kazmierczak@uwm.edu.pl

[3] Faculty of Civil and Environmental Engineering, Gdansk University of Technology, 80-233 Gdansk, Poland; szulwic@pg.edu.pl

* Correspondence: cezary.kowalczyk@uwm.edu.pl; Tel.: +48-502-28-91-93

† These authors contributed to the work and should be regarded as the first author.

Abstract: Knowing the exact number of fruits and trees helps farmers to make better decisions in their orchard production management. The current practice of crop estimation practice often involves manual counting of fruits (before harvesting), which is an extremely time-consuming and costly process. Additionally, this is not practicable for large orchards. Thanks to the changes that have taken place in recent years in the field of image analysis methods and computational performance, it is possible to create solutions for automatic fruit counting based on registered digital images. The pilot study aims to confirm the state of knowledge in the use of three methods (You Only Look Once—YOLO, Viola–Jones—a method based on the synergy of morphological operations of digital imagesand Hough transformation) of image recognition for apple detecting and counting. The study compared the results of three image analysis methods that can be used for counting apple fruits. They were validated, and their results allowed the recommendation of a method based on the YOLO algorithm for the proposed solution. It was based on the use of mass accessible devices (smartphones equipped with a camera with the required accuracy of image acquisition and accurate Global Navigation Satellite System (GNSS) positioning) for orchard owners to count growing apples. In our pilot study, three methods of counting apples were tested to create an automatic system for estimating apple yields in orchards. The test orchard is located at the University of Warmia and Mazury in Olsztyn. The tests were carried out on four trees located in different parts of the orchard. For the tests used, the dataset contained 1102 apple images and 3800 background images without fruits.

Keywords: computing image analysis; deep learning; yield mapping in an orchard; fruit counting; computer vision

1. Introduction

The yield forecasting process can start in two stages. The first estimation may take place during the flowering of trees, which is particularly important for the estimation of future harvest [1,2]. The second stage, which was analyzed in the article, is counting the fruit on the tree [3,4]. Naturally, the future income is correlated with the number, size and quality of apples [5–7]. The fruit supply chain is long and complex, and numerous stakeholders are involved, including farm input suppliers, orchardists, collectors, packing stations, transporters/shipping companies, retailers/food service providers and the government and authorities, among others [8]. Several steps are included moving from upstream (production) to downstream (trade, storage, processing). In practice, the question of harvest size is revealed at several production stages, which is necessary for the preparation

of the harvest itself and further affords the fruit commercial campaign. At this moment, the producer has to estimate the harvest size to contract the receipt of the fruit [7]. According to data from the Central Statistical Office of Poland, in 2017, the area of agricultural land was 16,414,831 hectares, including 361,965 hectares of orchards (2.2%); a total of 156,995 farms had orchards. Taking into account the existence of statistical systems for forecasting fruit yields in individual countries and at the European level, counting fruit in orchards and sending it to the central data exchange center would be an extremely good complement to the data. Storage and packing stations and processing companies that sign contracts require a forecast about the quantity of received fruit. For distribution planning, it is a requirement to determine the number of transported products and their recipients early enough. The optimization of fruit distribution will allow a change in the communication model of fruit producers and consumers (distribution companies, refrigerators, supermarkets, etc.) [8,9]. The main problem is the current information about the yield forecast [10,11], and this can lead to the conduction of transactions without mutual knowledge, which leads to asymmetry in the decision-making process [12,13].

On a national scale, it is also important to forecast the quantities of apples produced on the market. Estimating yields based on previous harvests is not particularly accurate. We propose a solution called Fruit Calculation System (FCS). The first task is determining the number of apples. In the next stages of the study, the possibility of forecasting yields based on flowers and qualitative evaluation of apples (size, color, spots) should be examined. It should be added, at flowering time, that yield forecast is strongly impaired by the uncertainty of flower pollination, fruit set and further June drop.

Due to the possibilities for technical devices to be used by orchardists themselves, in this research, a system based on independent digital image material acquisition and transmission to a server was considered, where calculations will be carried out or the photo materials will be collected by a trained local representative. The end-user will have access to the final reports based on an application that communicates with the server that stores estimated results. The article tests three methods of counting apples for use in FCS.

Estimating the number of fruits before harvesting provides useful information for logistics. Although significant progress has been made in fruit detection, it is difficult to estimate the actual number of fruits on a tree. In practice, fruits often overlap in the image and are partially or completely hidden by leaves. Therefore, methods that detect fruits do not offer a general solution for estimating the exact number of fruits [14]. In the typical image classification process, the task is to specify the presence or absence of an object; however, the counting problem requires one to reason how many instances of an object are present in the scene. The counting problem arises in several real-world applications such as cell counting in microscopic images [15], wildlife counting in aerial images [16], fish counting [17] and crowd monitoring [18,19] in surveillance systems. Most modern research focuses on one of the components of the proposed system, i.e., fruit counting on the registered image. A non-destructive method was proposed to count the number of fruits on a coffee branch by using information from digital images of a single side of the branch and its growing fruits [20]. Recent years have seen significant advances in computer vision research based on deep learning. The algorithm efficiently counts fruits even if they are in the shade, occluded by foliage or branches, or if there is some degree of overlap amongst fruits [21–23], fruit diseases or damage [24–26].

Taking into account the rapid technological development related primarily to the miniaturization of measuring devices and the increase in computing power in mobile devices, it is possible to undertake the task of creating an apple-counting system based on a smartphone or an image obtained from a drone camera. To realize this hypothesis, preliminary studies were carried out in the natural environment. To verify the hypothesis, in our pilot study, was tested three methods of counting apples to create an automatic system for estimating apple yields in orchards.

2. Methods and Methodology

Detection of apple fruit by reference to color [27,28], shape [29], visibility [30] and size requires the use of appropriate computer vision techniques [31–33]. The selection of appropriate techniques depends on the goal to be achieved through the digital acquisition of an apple image [9,34,35]. The goal may be to assess the number of fruits or estimate their condition or their size [31,32]. Therefore, it is not easy to separate the obtained image of an apple on sub-surfaces with pixels unequivocally (uniformly) connected with fruit and other pixels (so-called background). Variable observation and environmental conditions were indicated as the main reason. Unfortunately, none of the classic methods offer direct high (satisfying) efficiency. The goal can be defined as two main tasks:

- fruit counting [36–41],
- information about chosen fruits, such as color and quality rate, resulting from the counting.

Despite the impressive results achieved by these approaches, all of them need strong supervision information during the training phase. Based on literature research, the following groups of methods can be distinguished [27–33,38–41]:

- Convolutional Neural Networks (CNN)—This method is more accurate than the latest one based on Gaussian Mixture Models (GMMs). The multi-class classification approach used in this method provides an accuracy of 80% to 94% without the need for any pre-or post-processing steps. The deep learning network reacts to different fruit colors and lighting conditions. To check the suitability of the method for yield estimation, tests were conducted. The described method allows the achievement of approximately 96% accuracy concerning the actual number of apples [14].
- Deep Simulated Learning (DSL)—Automatic number of fruits alone estimation based on robotic agriculture provides a real solution in this area. The network is fully trained on synthetic data and tested on real data. To capture functions on multiple scales, a modified version of the Inception-ResNet architecture was used. The algorithm counts effectively even if the fruits are hidden in the shade, obscured by leaves or branches, or if the fruits overlap to some extent. Experimental results show 91% average test accuracy in real images and 93% in synthetic images. The proposed methodology works effectively even if the variant has a lighting deviation [23].
- Mixed method—This method combines deep segmentation, frame-to-fruit tracking and 3D location to accurately count the visible fruits in the image sequence. Segmentation is performed on the monocular camera image stream, both in natural light and under controlled night-time lighting. The first step is to train a fully revolutionary network (FCN) and to divide the video frame images into fruit and fruit pixels. Then, frame-by-frame fruit is tracked using a Hungarian algorithm, in which an objective result is determined based on the improved Kalman filter, i.e., Kanade–Lucas–Tomasi (KLT) [42].

Detection using the general descriptor YOLOv3-608 COCO TRAINVAL, although effective, can be improved by creating a customized set of weights and classes based on a specific spectrum of possible detection images. The Training Dataset contained 1102 apple images and 3800 background images without fruit. Each picture, named after the source, was pre-processed manually. Non-apple elements have been removed. The size of the image was then changed to the box of the fruit in the image. Thanks to this, the parameters for the proper scaling of the source image and its background were known. This allowed for more flexible preparation of images for machine learning, which was performed by overlaying the source images on any background—here, in the form of pictures of leaves, branches, etc. As a result of such overlapping combined with the changing of the scales of the vertical and horizontal axes, rotation, adding noise and blur, 16,530 images were created.

Despite the high performance of object detection using YOLO, it has been decided to use it as a parallel solution—dividing the main image into smaller sections processed by separate Central Processing Units or Graphics Processing Units.

The actual number of apples on the tree was determined manually. This approach has made it possible to establish a clear reference level. Each result obtained by the tested methods was visually verified in the image. As part of the verification, it was checked whether the counted objects are apples (which groups of pixels on the tested objects qualified as apples).

Three methods of counting objects in photos were tested in the research.

2.1. The Use of Image Filtration and Hough Transform—Solution A

In this solution, several steps were taken to move from a simple picture of the fruit to counting its shapes (Figure 1).
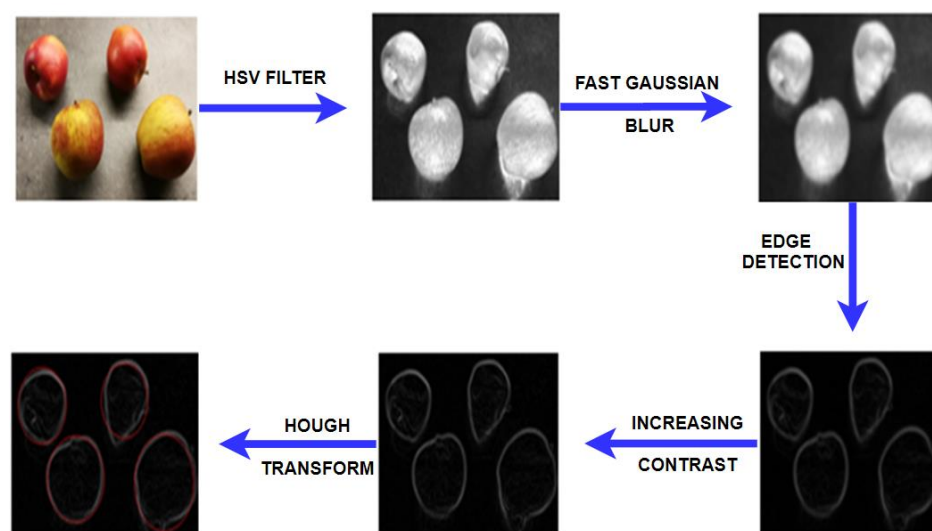

Figure 1. Steps are required to detect a number of fruit shapes from a digital image. Source: own study.

In this case, the color image (stored in RGB—Red, Green, Blue components) is transformed to the HSV representation model (H—hue, S—saturation, V—value) [43–45]. The use of the HSV model makes it easier to indicate where the fruit pixels are by using the HSV value (after blurring the images with a Gauss filter; Figure 2). Work began in the autumn and these were the first attempts to acquire and process images.

Figure 2. HSV (hue, saturation, value) filtration mode for mockups of an apple fruit image efficiency ratio of 41% with many selection errors (Tree no. 1). Source: own study.

Appropriately selected filter edge parameters narrow the search area even more. It is possible to fit in circles (an approximation of apple shapes) by using the Hough transform method, for example.

Previous research has made comparisons of edge detection and Hough transformation techniques for the extraction of geologic features [46] or Msplit estimation [47,48].

2.2. Viola–Jones Object Detection — Solution B

Another approach involves using an object detection framework and finding objects by using a dataset of positive image objects (Figure 3) for training it. This process requires training the classifier on thousands of images and searching these images for target objects.

Figure 3. Positive image samples for database training. Source: own study.

The Viola–Jones algorithm was used because it has several advantages, such as a sophisticated feature selection and an invariant detector that determines scales. This results in scaled functions instead of scaling the image itself [49].

The use of the Viola–Jones algorithm [50] is based on the description of features rather than the pixels of the image directly. The analysis of the features proposed by Viola and Jones is performed in random rectangles, as in Figure 4.
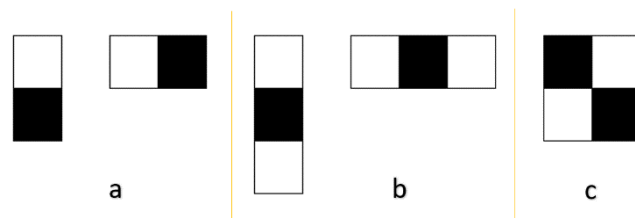

Figure 4. Example rectangle features (based on the original article in [49]) are shown relative to the enclosing detection window. The sum of the pixels that lie within the white rectangles is subtracted from the sum of the pixels in the grey rectangles. Rectangle features can contain two sub-rectangles (**a**), three rectangles (**b**), or four rectangles (**c**), and their size can be changed cascadingly.

Each feature result is a single value, which is calculated by subtracting the sum of the values of the pixels under white rectangles from the sum of the pixels under black rectangles.

Thanks to using such a generalization, it is possible to cascadingly increase the size of black and white rectangles, thus allowing for studying and comparing images with different scales.

Unfortunately, despite the promising initial assumptions, it turned out that the algorithm (the Viola–Jones algorithm) is not suitable for generalizing the classification of objects (creating classes)—it is used primarily to detect specific objects, which, in the case of apples, turned out to be an erroneous assumption. Additionally, even when detecting specific objects (not classes), it has a problem with torsion tilt and different lighting conditions. Fruit count tests were also performed for the selected apple tree (Tree no. 1). An efficiency of 55% was achieved. The result is presented in Figure 5.

Figure 5. Effective detection of apple objects using the Viola–Jones algorithm (Tree no. 1). Source: own study.

2.3. YOLO: Real-Time Object Detection—Solution C

The use of the modern real-time object detection system YOLO (You Only Look Once) is the third solution assessed. YOLO uses a single ConvNet (or CNN, convolution neural network) for classification and localizing by using bounding boxes. The advantage of this solution, as the authors indicated, is the reconstruction of object detection to a single regression problem, directly from image pixels to coordinates defining rectangular envelopes, and the probability of the occurrence of appropriate classes of objects [51].

The YOLO algorithm can be described in a few steps. The input image is divided into an SxS grid (Figure 6). Each cell in this grid is designed to predict the existence of only one object in it.
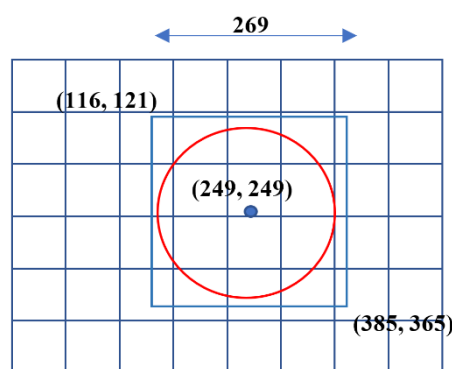

Figure 6. An example of the division of an image into a grid in the YOLO algorithm. Source: own study.

The blue line is the bounding box (bbox), which must be described by 5 components related to the selected cell of the grid, and these coordinates must be normalized, i.e., defined within the range of 0–1. The following parameters describe each field:

- x blue box = (385 − 116)/2 = 135 but normalized and related to the corresponding grid cell, here: (135 − 100)/100 = 0.35
- y blue box = (365 − 121)/2 = 122 but normalized and related to the relevant grid cell, here: (122 − 100)/100 = 0.22
- w width blue box = (385 − 116)/500 = 0.54 (normalization in relation to the width of the whole image)
- h height blue box = (365 − 121)/500 = 0.49 (normalization relative to the height of the whole image)
- c confidence, which is the Intersection over Union (IoU) between the predicted box and the ground truth (c = area of overlap/area of union).

It is assumed that only one type of class is assigned to one cell. The output vector is in the form of a tensor SxSx(C + B×5), where B stands for the number of blue boxes. The rest looks like a normal CNN, with convolutional and max-pooling layers. All details can be found in the source document [52].

The main goal of our solution was to use real-time detection with image acquisition by a mobile device in practical implementation. Moreover, it was important to choose a neural network dedicated to the performance of mobile devices. Each image was divided into 4 or 16 parts depending on the resolution of the image, and each analyzed fragment was analyzed separately. This increased the digital detection of objects and reduced the memory load of the algorithm. This gives an insight into the future possibilities of analyzing images acquired in the form of video recordings as parallel, multithread computing.

3. Results

The use of YOLO allowed us to obtain better results than with other classifiers. At the same time, the working time was much shorter with YOLO. To evaluate the work and results, a set of YOLOv3-608 scales trained at the COCO was used http://cocodataset.org/ accessed on 10 January 2021.

With confidence threshold = 10% and assuming a search only for 47 classes (apples in the weighing file) and excluding overlapping of objects more than 30%, 66 objects were found for the above image, which represents 67% of detectable objects in such an image (Figure 7). This illustrates the multi-threaded detection of apple objects based on the numerically corrected image with:

- change of the clarity in the arbitrary range from −18 to 18 levels,
- noise removal using the non-local means algorithm,
- Gamma correction from 1.1 to 1.6,
- Increase in the number of detectable items from 27% to 80% (Figure 8).

Figure 7. Example of YOLO (You Only Look Once) operation on the selected tree (**a**—original image, **b**—counted apples). Source: own study.

Figure 8. Results of apple object detection based on the numerically corrected image (Tree no. 1). Source: own study.

Apple fruit detection, regarding their specific color (problem 1), shape (problem 2) visibility and size (problem 3), requires the use of appropriate computer vision techniques. The research carried out has led to the following conclusions:

1. The impact of the first problem can be significantly reduced by:

- Anti-noise filters: non-local means filters are suggested but it is possible to experiment with local ones,
- Edge detecting algorithms, for example, operators based on the first derivative (Prewitt, Sobel, Canny, Scharr or Roberts) or second derivative (Marr–Hildreth algorithms [10]),
- Use of a thermal imaging camera. The literature related to the subject of study includes an effective attempt to use a thermal infrared camera; however, its cost is extremely high, which limits the scale of the task, and the achievable resolution is still not satisfactory.

2. The solution to the second problem is to adopt the circular shape of a standard apple and use Hough transform or Msplit estimation to complement the incomplete shape of the circle.

3. The different distances of apple fruits from the camera during the acquisition of a digital image results in different sizes (numbers of pixels) of their digital representation. While this is not important when assessing the number of fruits, it is of high importance when it comes to interpreting fruit size or belonging to the examined apple tree. Hence, it is important to know the size of the expected single apple in the picture. The goal can be reached by using one of two methods separately or by compiling them. Using a fixed focal length camera, the known location of the camera and the apple tree allows for the approximation of the size of the fruit, and its assessment in terms of dimensions. The stability of the focal length and the positions of the camera and the tree guarantee a differential analysis of the development of inflorescence and, later, fruit; however, the parameters of the camera should be selected individually. A synthetic comparison of the three methods is presented in Table 1.

Table 1. Comparison of the systems used for object identification.

	Method A	Method B	Method C
	image filtration and Hough transformers	Viola–Jones object detection	YOLO: Real-Time Object Detection
Color	The color of this assumption is important. Initial image filtration is based on the ranges of individual color components.	Does not matter	Does not matter
Shape	Only shapes were similar to circles.	A well-prepared descriptor works properly on different shapes but you should put them yourself in the training set.	Using a well-trained network or having trained it with new images, there is no need to place fragments of the image of the fruit if its detection is desired. The algorithm does it.
Size	Relaxation of the radius causes considerable elongation of the object search operation.	Does not matter	Does not matter
Processing time	Very long	Medium (not in real-time)	In real-time

Table 2 presents the results of fruit counting efficiency using three methods on four test trees.

Table 2. Results comparison of the systems used for object identification.

Tree No.	The Real Number of Apples	Hough Transform		Viola–Jones		YOLO	
		Detected Apples	Effectiveness [%]	Detected Apples	Effectiveness [%]	Detected Apples	Effectiveness [%]
1	220	90	41	121	55	176	80
2	82	25	30	40	49	73	89
3	52	15	29	29	55	43	83
4	30	14	46	19	63	26	84
Average			37		55		84

The reference number of apples on the tree was determined manually. The results indicate the use of YOLO as an effective solution for counting the number of fruits on the objects presented in the article. Limitations in detecting more apples resulted from physical (partially overshadowing objects) and environmental conditions.

4. Discussion

Research work on issues related to fruit detection based on digital images has become extremely popular in recent years. This is primarily related to the development of innovative agricultural robots using modern image processing algorithms [52]. Concerning the effects of research work on various approaches of automatic apple counting based on images, the proposed approach has given satisfactory results. In terms of fruit detection, the obtained accuracy ranges between 80% and 96%. Naturally, such an accuracy range is related to the adopted method and the characteristics of the plants on which the fruits grow. Linker et al. in their approach reached the estimation accuracy of 85% [53]. They based their calculations on information about color and texture [1,54]. In the works of Wei et al. and Payne et al., among others, the results are also influenced by sunlight and color saturation [52,55]. Zhao et al. used a feature image fusion method to recognize mature tomatoes obtained, with 93% detection [56]. A similar level (92.4%) was reached by Qiang et al. [57]. Kelman et al. based their calculations on the shape of the detected objects, which resulted in 94.4% fruit detection in the pictures [58]. Similar results to those presented in the article were achieved by Kurtulmus et al. (84.6%) [59] and Yamamoto et al. (80% and 88%) [60]. The apple-counting method based on YOLO has limitations due to the operating algorithm. An erroneous definition of the detection bounding box causes a small error in interpretation for a large box to be insignificant, but for a small box to increase in insignificance. The biggest problem, regardless of the method, is that the fruit is covered by leaves and two fruits are in close proximity, therefore the system can interpret them as one object.

The process of forecasting the number of apples in the future harvest can be divided into two basic stages. The first one is related to monitoring the condition of trees and counting the number of flowers on the trees [1,2], and counting the ripening apples. From the orchardists' point of view, a special role is played by the possibility to determine the size of the harvest [61], hence a large number of emerging scientific studies in this area [3–8]. In this study, several approaches to fruit (apple) number evaluation were analyzed in a practical way, which allowed for the compilation of the results presented below.

Fruit images, including apples, are characterized by a high degree of texture irregularities. The lack of surface uniformity results from the differences in fruit exposition and is a natural consequence of the fruit location within the tree crown, occultation by branches, leaves and others. Although the optimistic assumption of apple shape observation from any position and camera angle indicates the approximation of the circular shape, the overlapping of fruit images and the mentioned covering of fruits with other elements recorded in the images and with the shadows cast by them can also cause an unpredictable change in the shape of a single fruit in an image [2]. A single fruit can also be interpreted as two apples or more when the image of an apple is divided by a view of a branch.

The whole process of fruit counting, when it comes to one tree, is based on taking a series of pictures with the center of the projection shifted to a small longitudinal parallax. This allows the obtainment of a smudged image of a single tree. In this way, a full picture of the tree crown was obtained. A similar solution can be applied to the proposed schemes of the material image acquisition from a drone or mobile device for the whole orchard.

From the technical side of the image processing system, it is necessary to collect an appropriate number of apple images, on the basis of which the system can start its calculations. Häni et al. adopted 1000 high-resolution images acquired in orchards, together with human annotations of the fruit on trees. The fruits were marked with polygonal masks for each object, which helped to precisely detect, locate and segment the object [61]. For their research, Gao et al. authors acquired 800 images, which after processing gave a total input of 12,800 images [6]. An analogous number of images (800 images) was used by Fu et al. in their research using low-cost Kinect V2 sensors [7]. In this research, a similar number of input photos were taken as taken by other researchers. Our input base was 1102 images. In the field, three photos were taken for each tree on one side.

After choosing the method of counting the apples in digital images, it is necessary to propose the structure of the system for taking images in the orchard. The key assumption of the proposed solution was to minimize the costs of its creation and use of the system. Hence, it assumes the use of generally available mobile devices as a component of digital image acquisition—georeferenced images (determined on the basis of Global Navigation Satellite Systems (GNSS) technology). Such a solution offers the possibility of mass use in horticulture.

5. Conclusions

The main objective of this study was to verify the optimal method for identifying and counting apples on trees from photographs taken in the orchard. Based on the tests performed, it can be concluded that the best results are obtained using the YOLO method.

The reduced number of trees accepted for the test allowed manual counting of the number of fruits on each tree. With a larger test sample, without the ability to count and determine a reliable number of reference fruits, the tests would have low reliability. Therefore, for validation of individual object recognition methods, in the authors' opinion, the presented sample is sufficient. The adopted approach provided an unambiguous reference number of counted fruits. It allowed to unequivocally determine the level of counting accuracy. The obtained accuracy of individual methods was confirmed by literature review and achievements of other researchers. After carrying out the pilot experiments according to the assumptions presented above, the decision was made to implement the task using smartphones equipped with a camera with the required image acquisition accuracy and accurate positioning by GNSS (Figure 9).
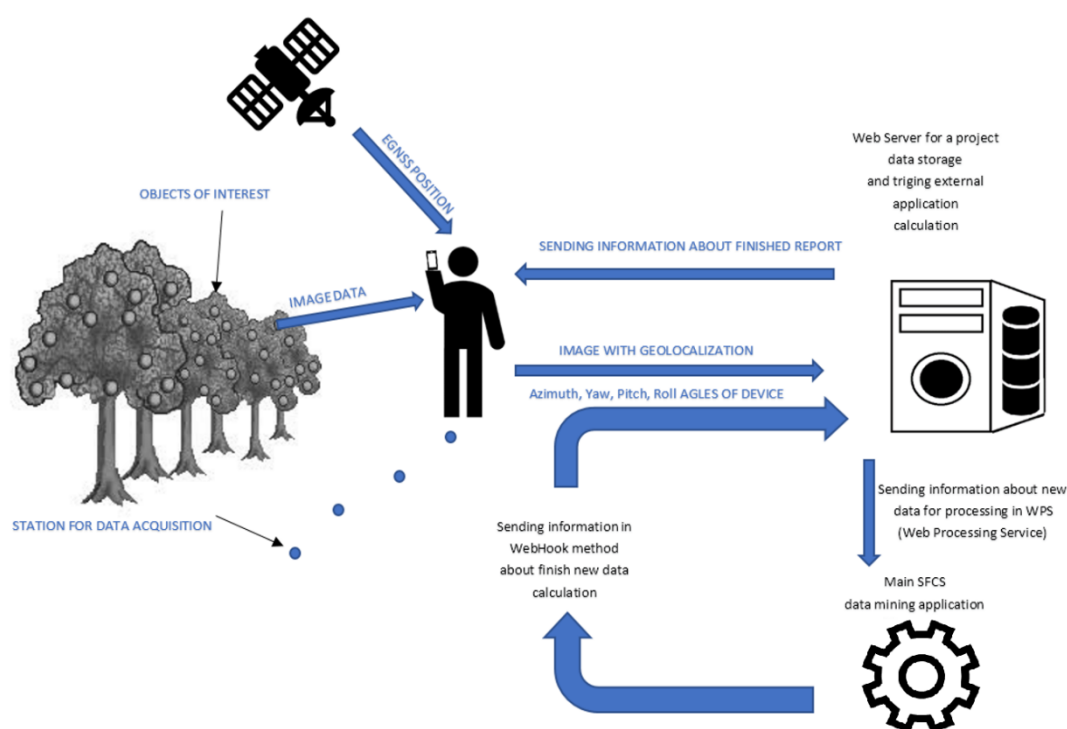
Figure 9. The global scheme of the system functioning. Source: own study.

Initially, a solution can be proposed based on the measurements with the mobile device, because of its advantage over the classical methods, used mainly by mass users.

Regarding the considerations related to fruit counting, YOLO was chosen for its:

- efficiency,
- possibility of implementation on mobile devices,
- effectiveness,
- ability to increase the effectiveness by constantly supplementing the YOLO training patterns requiring time for specific apple cultivars.

The main component obtaining the data is an orchardist or a person indicated by him/her. The measurement is made according to the assumptions that were initially set for the given orchard (depending on the way the trees are planted, density, number of rows, etc.).

The mobile application made available to the orchardist allows the user to take images with initial control.

The proposed solution preliminarily assesses the images in terms of chromatics, a histogram and its alignment and width, which makes it possible to reject completely incorrect photos (at the stage of acquiring them).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Database of images used to this article: https://doi.org/10.34808/g4t6-cm21 (Data set no. 1—multicolour) [62] and https://doi.org/10.34808/gx4e-bv72 (Data set no. 2—greyscale) [63] – accessed on 17 December 2020.

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

References

1. Wu, D.; Lv, S.; Jiang, M.; Song, H. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Comput. Electron. Agric.* **2020**, *178*, 105742, doi:10.1016/j.compag.2020.105742.
2. Wang, X.A.; Tang, J.; Whitty, M. Side-view apple flower mapping using edge-based fully convolutional networks for variable-rate chemical thinning. *Comput. Electron. Agric.* **2020**, *178*, 105673, doi:10.1016/j.compag.2020.105673.
3. Vasconez, J.P.; Delpiano, J.; Vougioukas, S.; Cheein, F.A. Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation. *Comput. Electron. Agric.* **2020**, *173*, 105348, doi:10.1016/j.compag.2020.105348.
4. Samiei, S.; Rasti, P.; Richard, P.; Galopin, G.; Rousseau, D. Toward joint acquisition-annotation of images with egocentric devices for a lower-cost machine learning application to apple detection. *Sensors* **2020**, *20*, 4173, doi:10.3390/s20154173.
5. Gao, F.; Fu, L.; Zhang, X.; Majeed, Y.; Li, R.; Karkee, M.; Zhang, Q. Multi-class fruit-on-plant detection for an apple in SNAP system using Faster R-CNN. *Comput. Electron. Agric.* **2020**, *176*, 105634, doi:10.1016/j.compag.2020.105634.
6. Fu, L.; Majeed, Y.; Zhang, X.; Karkee, M.; Zhang, Q. Faster R–CNN–based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting. *Biosyst. Eng.* **2020**, *197*, 245–256, doi:10.1016/j.biosystemseng.2020.07.007.
7. Przyborski, M.; Szczechowski, B.; Szubiak, W.; Szulwic, J.; Widerski, T. Photogrammetric Development of the Threshold Water at the Dam on The Vistula River in Wloclawek from Unmanned Aerial Vehicles (UAV). Water Resources, Forest, Marine and Ocean Ecosystems, SGEM2015 2015, I, 493–500. http://dx.doi.org/10.5593/SGEM2015/B31/S12.063.
8. Zhang, X.; He, L.; Karkee, M.; Whiting, M.D.; Zhang, Q. Field evaluation of targeted shake-and-catch harvesting technologies for fresh market apple. *Trans. ASABE* **2020**, *63*, doi:10.13031/trans.13779.
9. Jidong, L.; De-An, Z.; Wei, J.; Shihong, D. Recognition of apple fruit in natural environment. **2016**, *127*, 1354–1362, doi:10.1016/j.ijleo.2015.10.177.
10. Maldonado, W., Jr ; Barbosa, J.C. Automatic green fruit counting in orange trees using digital images. *Comput. Electron. Agric.* **2016**, *127*, 572–581, doi:10.1016/j.compag.2016.07.023.
11. Liu, J.; Yuan, Y.; Zhou, Y.; Zhu, X.; Syed, T.N. Experiments and analysis of close-shot identification of on-branch citrus fruit with realsense. *Sensors* **2018**, *18*, 1510, doi:10.3390/s18051510.
12. Kowalczyk, C.; Nowak, M.; Źróbek, S. The concept of studying the impact of legal changes on the agricultural real estate market. *Land Use Policy* **2019**, *86*, 229–237, doi:10.1016/j.landusepol.2019.05.012.
13. Renigier-Bilozor, M.; Wisniewski, R.; Bilozor, A. Rating attributes toolkit for the residential property market. *Int. J. Strateg. Prop. Manag.* **2017**, *21*, 307–317, doi:10.3846/1648715X.2016.1270235.
14. Hani, N.; Pravakar, R.; Volkan, I. Apple counting using convolutional neural networks. *Int. Conf. Intell. Robot. Syst.* **2018**, doi:10.1109/IROS.2018.8594304.
15. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems, Proceedings of the Neural Information Processing Systems Conference, Montreal, QC, Canada, 7–12 December 2015*; Neural Information Processing Systems Foundation, Inc.: Ljubljana, Slovenia, 2015; pp. 91–99.
16. Laliberte, A.S.; Ripple, W.J. Automated wildlife counts from remotely sensed imagery. *Wildl. Soc. Bull.* **2003**, *31*, 362–371.
17. Del Río, J.; Aguzzi, J.; Costa, C.; Menesatti, P.; Sbragaglia, V.; Nogueras, M.; Sarda, F.; Manuèl, A. A new colorimetrically-calibrated automated video-imaging protocol for day-night fish counting at the OBSEA coastal cabled observatory. *Sensors* **2013**, *13*, 14740–14753, doi:10.3390/s131114740.
18. Ryan, D.; Denman, S.; Fookes, C.; Sridharan, S. Crowd counting using multiple local features. In Proceedings of the Digital Image Computing: Techniques and Applications, Melbourne, Australia, 1–3 December 2009; pp. 81–88.
19. Ma, Y.; Wu, X.; Yu, G.; Xu, Y.; Wang, Y. Pedestrian detection and tracking from low-resolution unmanned aerial vehicle thermal imagery. *Sensors* **2016**, *16*, 446, doi:10.3390/s16040446.
20. Ramos, F.A.; Prieto, E.C.; Montoya, C.E. Oliveros, Automatic fruit count on coffee branches using computer vision. *Comput. Electron. Agric.* **2017**, *137*, 9–22, doi:10.1016/j.compag.2017.03.010.
21. Liu, X.; Chen, S.W.; Aditya, S.; Sivakumar, N.; Dcunha, S.; Qu, C.; Taylor, C.J.; Das, J.; Kumar, V. Robust fruit counting: Combining deep learning, tracking, and structure from motion. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 1045–1052, doi:10.1109/IROS.2018.8594239.
22. Patel, H.N.; Jain, R.K.; Joshi, M.V. Fruit detection using improved multiple features based algorithm. *Int. J. Comput. Appl.* **2011**, *13*, 1–5.
23. Rahnemoonfar, M.; Sheppard, C. Deep count: Fruit counting based on deep simulated learning. *Sensors* **2017**, *17*, 905, doi:10.3390/s17040905.
24. Bagheri, N.; Mohamadi-Monavar, H.; Azizi, A.; Ghasemi, A. Detection of Fire Blight disease in pear trees by hyperspectral data. *Eur. J. Remote Sens.* **2018**, *51*, doi:10.1080/22797254.2017.1391054.
25. Martinelli, F.; Scalenghe, R.; Davino, S.; Panno, S.; Scuderi, G.; Ruisi, P.; Villa, P.; Stroppiana, D.; Boschetti, M.; Goulart, L.R.; et al. Advanced methods of plant disease detection. A review. *Agron. Sustain. Dev.* **2015**, *35*, 1–25, doi:10.1007/s13593-014-0246-1.
26. Yang, Y.; Zhang, Y.; Deng, Y.; Yi. X. Endozoochory by granivorous rodents in seed dispersal of green fruits. *Can. J. Zool.* **2018**, doi:10.1139/cjz-2018-0079.

27. Bhargava, A.; Bansal, A. Fruits and vegetables quality evaluation using computer vision: A Review. *J. King Saud Univ. Comput. Inf. Sci.* **2018**, doi:10.1016/j.jksuci.2018.06.002.

28. Al Ohali, Y. Computer vision based date fruit grading system: Design and implementation. *J. King Saud Univ. Comput. Inf. Sci.* **2011**, *23*, 29–36, doi:10.1016/j.jksuci.2010.03.003.

29. Blasco, J.; Aleixos, N.; Molto, E. Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm. *J. Food Eng.* **2007**, *81*, 535–543, doi:10.1016/j.jfoodeng.2006.12.007.

30. Lak, M.B.; Minaei, S.; Amiriparian, J.; Beheshti, B. Apple fruits recognition under natural luminance using machine vision. *Adv. J. Food Sci. Technol.* **2010**, *2*, 325–327.

31. Lingua, A.; Marenchino, D.; Nex, F. Performance analysis of the SIFT operator for automatic feature extraction and matching in photogrammetric applications. *Sensors* **2009**, *9*, 3745–3766, doi:10.3390/s90503745.

32. Luna, E.; San Miguel, J.C.; Ortego, D.; Martínez, J.M. Abandoned object detection in video-surveillance: Survey and comparison. *Sensors* **2018**, *18*, 4290, doi:10.3390/s18124290.

33. Zhou, R.; Zhong, D.; Han, J. Fingerprint identification using SIFT-based minutia descriptors and improved all descriptor-pair matching. *Sensors* **2013**, *13*, 3142–3156, doi:10.3390/s130303142.

34. Jiang, N.; Song, W.; Wang, H.; Guo, G.; Liu, Y. Differentiation between organic and non-organic apples using diffraction grating and image processing—A cost-effective approach. *Sensors* **2018**, *18*, 1667, doi:10.3390/s18061667.

35. Wang, Z.; Koirala, A.; Walsh, K.; Anderson N.; Verma, B. In field fruit sizing using a smart phone application. *Sensors* **2018**, *18*, 3331, doi:10.3390/s18103331.

36. Akin, C.; Kirci, M.; Gunes, E.O.; Cakir, Y. Detection of the pomegranate fruits on tree using image processing. In Proceedings of the 2012 First International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Shanghai, China, 2–4 August 2012; IEEE: New York, NY, USA; pp. 1–4, doi:10.1109/Agro-Geoinformatics.2012.6311724.

37. Sahu, D.; Potdar, R.M. Defect identification and maturity detection of mango fruits using image analysis. *Am. J. Artif. Intell.* **2017**, *1*, 5–14, doi:10.11648/j.ajai.20170101.12.

38. Blazek, M.; Janowski, A.; Kazmierczak, M.; Przyborski, M.; Szulwic, J. Web-cam as a means of information about emotional attempt of students in the process of distant learning. *ICERI Proc. ICERI2014* **2014**, 3787–3796, doi:10.13140/2.1.4699.2322.

39. Nagrodzka-Godycka, K.; Szulwic, J.; Ziolkowski, P. State-of-the-art framework for high-speed camera and photogrammetric use in geometry evaluation of prestressed concrete failure process. In Proceedings of the 16th International Multidisciplinary Scientific Geoconference (SGEM 2016), Albena, Bulgaria, 30 June–6 July 2016.

40. Ziolkowski, P.; Niedostatkiewicz, M. Machine learning techniques in concrete mix design. *Materials* **2019**, *12*, 1256, doi:10.3390/ma12081256.

41. Bellocchio, E.; Ciarfuglia, T.A.; Costante, G.; Valigi, P. Weakly supervised fruit counting for yield estimation using spatial consistency. *IEEE Robot. Autom. Lett.* **2019**, *4*, 2348–2355, doi:10.1109/LRA.2019.2903260.

42. Xu, Y.; Yu, G.; Wang, Y.; Wu, X.; Ma, Y. A hybrid vehicle detection method based on Viola-Jones and HOG+ SVM from UAV images. *Sensors* **2016**, *16*, 1325, doi:10.3390/s1608132538.

43. Dorj, U.O.; Lee, M.; Yun, S.S. An yield estimation in citrus orchards via fruit detection and counting using image processing. *Comput. Electron. Agric.* **2017**, *140*, 103–112, doi:10.1016/j.compag.2017.05.019.

44. Al-Saddik, H.; Laybros, A.; Billiot, B.; Cointault, F. Using image texture and spectral reflectance analysis to detect yellowness and esca in grapevines at leaf-level. *Remote Sens.* **2018**, *10*, 618, doi:10.3390/rs10040618.

45. Yamamoto, K.; Guo, W.; Yoshioka, Y.; Ninomiya, S. On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors* **2014**, *14*, 12191–12206, doi:10.3390/s140712191.

46. Argialas, D. P., & Mavrantza, O. D. (2004). Comparison of edge detection and Hough transform techniques for the extraction of geologic features. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *34*(Part XXX).

47. Janowski, A. The circle object detection with the use of M split estimation. *E3S Web Conf.* **2018**, *26*, 00014, doi:10.1051/e3sconf/20182600014.

48. Janowski, A.; Bobkowska, K.; Szulwic, J. 3D modelling of cylindrical-shaped objects from lidar data—An assessment based on theoretical modelling and experimental data. *Metrol. Meas. Syst.* **2018**, *25*, 47–56, doi:10.24425/118156.

49. Aashish, K.; Vijayalakshmi, A. Comparison of Viola-Jones and Kanade-Lucas-Tomasi face detection algorithms. *Orient. J. Comput. Sci. Technol.* **2017**, doi:10.13005/ojcst/10.01.20.

50. Viola, P.; Jones, M.J. Robust real-time face detection. *Int. J. Comput. Vis.* **2004**, *57*, 137–154, doi:10.1023/B:VISI.0000013087.49260.fb.

51. Jarolmasjed, S.; Khot, L.R.; Sankaran, S. Hyperspectral imaging and spectrometry-derived spectral features for bitter pit detection in storage apples. *Sensors* **2018**, *18*, 1561, doi:10.3390/s18051561.

52. Redmon, J.; Divvala, S.; Grishik, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *arXiv* **2018**. Available online: https://pjreddie.com/media/files/papers/yolo.pdf ((accessed on 17 December 2020).

53. Wei, X.; Jia, K.; Lan, J.; Li, Y.; Zeng, Y.; Wang, C. Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot. *Optik* **2014**, *125*, 5684–5689, doi:10.1016/j.ijleo.2014.07.001.

54. Linker, R.; Cohen, O.; Naor, A. Determination of the number of green apples in RGB images recorded in orchards. *Comput. Electron. Agric.* **2012**, *81*, 45–57, doi:10.1016/j.compag.2011.11.007.

55. Kuznetsova, A.; Maleva, T.; Soloviev, V. Using YOLOv3 algorithm with pre-and post-processing for apple detection in fruit-harvesting robot. *Agronomy* **2020**, *10*, 1016.

56. Payne, A.; Walsh, K.; Subedi, P.; Jarvis, D. Estimating mango crop yield using image analysis using fruit at 'stone harden-ing'stage and night time imaging. *Comput. Electron. Agric.* **2014**, *100*, 160–167, doi:10.1016/j.compag.2013.11.011.

57. Zhao, Y.; Gong, L.; Huang, Y.; Liu, C. Robust tomato recognition for robotic harvesting using feature images fusion. *Sensors* **2016**, *16*, 173, doi:10.3390/s16020173.

58. Qiang, L.; Jianrong, C.; Bin, L.; Lie, D.; Yajing, Z. Identification of fruit and branch in natural scenes for citrus harvesting robot using machine vision and support vector machine. *Int. J. Agric. Biol. Eng.* **2014**, *7*, 115–121.

59. Kelman, E.E.; Linker, R. Vision-based localisation of mature apples in tree images using convexity. *Biosyst. Eng.* **2014**, *118*, 174–185, doi:10.1016/j.biosystemseng.2013.11.007.

60. Kurtulmus, F.; Lee, W.S.; Vardar, A. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precis. Agric.* **2014**, *15*, 57–79, doi:10.1007/s11119-013-9323-8.

61. Häni, N.; Roy, P.; Isler, V. A comparative study of fruit detection and counting methods for yield mapping in apple orchards. Journal of Field Robotics 2020; 37(2), 263-282. https://doi.org/10.1002/rob.21902.

62. Database of images used to this article - multicolour: https://doi.org/10.34808/g4t6-cm21 (accessed on 17 December 2020).

63. Database of images used to this article—greyscale: https://doi.org/10.34808/gx4e-bv72 (accessed on 17 December 2020).