

## Article

# Sugar Beet Damage Detection during Harvesting Using Different Convolutional Neural Network Models

Abozar Nasirahmadi \*, Ulrike Wilczek and Oliver Hensel 

Department of Agricultural and Biosystems Engineering, University of Kassel, 37213 Witzenhausen, Germany; wilczek@uni-kassel.de (U.W.); agrartechnik@uni-kassel.de (O.H.)

\* Correspondence: abozar.nasirahmadi@uni-kassel.de

**Abstract:** Mechanical damages of sugar beet during harvesting affects the quality of the final products and sugar yield. The mechanical damage of sugar beet is assessed randomly by operators of harvesters and can depend on the subjective opinion and experience of the operator due to the complexity of the harvester machines. Thus, the main aim of this study was to determine whether a digital two-dimensional imaging system coupled with convolutional neural network (CNN) techniques could be utilized to detect visible mechanical damage in sugar beet during harvesting in a harvester machine. In this research, various detector models based on the CNN, including You Only Look Once (YOLO) v4, region-based fully convolutional network (R-FCN) and faster regions with convolutional neural network features (Faster R-CNN) were developed. Sugar beet image data during harvesting from a harvester in different farming conditions were used for training and validation of the proposed models. The experimental results showed that the YOLO v4 CSPDarknet53 method was able to detect damage in sugar beet with better performance (recall, precision and F1-score of about 92, 94 and 93%, respectively) and higher speed (around 29 frames per second) compared to the other developed CNNs. By means of a CNN-based vision system, it was possible to automatically detect sugar beet damage within the sugar beet harvester machine.

**Keywords:** convolutional neural network; damage; deep learning; harvester; sugar beet



check for updates

**Citation:** Nasirahmadi, A.; Wilczek, U.; Hensel, O. Sugar Beet Damage Detection during Harvesting Using Different Convolutional Neural Network Models. *Agriculture* **2021**, *11*, 1111. <https://doi.org/10.3390/agriculture11111111>

Academic Editor:  
Rafael-Ruben Sola-Guirado

Received: 28 October 2021  
Accepted: 7 November 2021  
Published: 9 November 2021

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



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

## 1. Introduction

Sugar beets (*Beta vulgaris* L.) are cultivated and used for sugar production globally. However, it is one of the main industrial crops in Europe and in particular Germany. Sugar beet are usually harvested mechanically with multi-row self-propelled harvesters. In this context, the leaves along with the petioles and crown are removed from the beet, and the root is lifted from the soil in the harvesting process [1]. To remove soil from the roots, they are bounced over chain or rollers in the cleaning area [1] and then transported to a tank in the harvester. After the harvesting, sugar beets are stored in large windrows on the field and covered in the event of frost. The quality of sugar beet is strongly affected in terms of root injuries, bruises, and breaks during harvesting [2]. This damage may lead to loss of mass due to pieces that are lost, adverse effects on the product and sugar qualities as well as storage capability [3]. If the roots are damaged, the respiration of sugars increases due to energy needed for wound healing; as a consequence, the sugar yield decreases to a greater extent than in the case with undamaged roots [3]. Mechanical damage is considered as one of the main challenges affecting on the quality of processing and sugar yield, but as a part of the production process that is difficult to avoid. Sugar beet harvesters have improved significantly in recent years; however, the mechanical damage still occurs during harvesting in the harvester machines. According to [4], the mechanical damage of sugar beets can occur in all processes and components of the harvester and can be categorized into: (i) damage and breaking of the beet due to harvesting and improper trimming in the topper unit, (ii) damage in cleaning and transporting sections, (iii) those from roots falling during transport and unloading stages in the tank or on the field. Damage may be

increased by mechanical stress during cleaning [5] and transport to the tank of the harvester. Furthermore, the intensity and speed of cleaning during harvesting may increase root damage in sugar beet [5]. Due to the short harvesting period and the weather conditions at the time of harvesting, the harvesters are often not optimized for each farming condition. Furthermore, the damage is assessed randomly by the machine operators. In this context, the direct observation method by humans can be very subjective, labor-intensive and time consuming [6] and in some cases impossible. Since the mechanical damage of sugar beet plays an important role in overall product quality, the question arises whether an automatic monitoring system could be developed to inspect damages during harvesting. Therefore, in order to make the sugar beet damage inspection automatic, in this study, state-of-the-art machine vision and deep learning-based algorithms were developed to monitor sugar beet in a harvester.

Machine vision and image processing techniques are used as alternatives and cheap solutions for human direct observation for a wide variety of applications in agriculture, e.g., for fruit and vegetable classification, variety detection, sorting and grading [6]. However, performance of machine vision models may be influenced by different lighting conditions, a high level of noise, different shapes of target objects as well as the quality of captured images. Therefore, in this study to tackle these problems different machine learning approaches (in particular deep learning) along with a visual system were developed for external damage detection of sugar beet in digital two-dimensional (2D) images during the harvesting.

There are several research studies in recent years where machine vision techniques, particularly hyperspectral imaging systems, along with machine learning models, have been applied for damage detection of agricultural products. For instance, near-infrared spectral images and artificial neural networks were used to detect and classify mechanical damage in mushrooms [7]. In another project, mechanical damage in blueberries was characterized and classified using hyperspectral images and logistic regression, multilayer perceptron-back propagation techniques [8]. Another research group has used hyperspectral imaging technique to detect micro damage types of litchi fruit [9]. In their study, partial least squares discriminant analysis was used to predict quality features of the litchi fruit with different damages. Ref. [10] proposed a model to classify five mechanical damages in sugar beet seeds based on a multispectral imaging data. They showed that machine vision techniques have the capability to assess the quality of sugar beet seeds with high accuracy.

Recently, some more state-of-art machine learning techniques using 2D imaging have been utilized in an agricultural context to detect damage in agricultural products. A convolutional neural network (CNN) was developed using mobile phone image data to detect and diagnose jackfruit fruit damage [11] and the proposed method could detect damages with high accuracy of around 98%. For detection of mechanically damaged potatoes in 2D images, different machine learning algorithms were developed and tested by [12]. In their research, the Viola-Jones algorithm was applied to find potato tubers on a conveyor belt, then a support vector machine (SVM) model was developed to detect damage. Their developed models allowed users to classify and detect mechanical damage of up to 100 tubers in one second. In a similar study, the deep learning-based detection of potato damage and defects was investigated by [13]. In that study, transfer learning through various deep-CNNs, i.e., single shot multibox detector (SSD) inception v2, faster regions with convolutional neural network (Faster R-CNN) ResNet101, and region-based fully convolutional network (R-FCN) ResNet101 were developed. Their results showed that the R-FCN ResNet101 had the best overall performance in detection speed and accuracy [13]. Likewise, a CNN and a computer vision method to classify defects and damages of green plums was adopted [14]. The developed CNN model was based on the VGG network architecture combined with a stochastic weight averaging optimizer and trained weights on ImageNet.

In terms of real-time detection of damage in agricultural products, different approaches were developed using machine vision and CNN models. For instance, in order to

detect broken corn at a conveyor belt of a corn harvester You Only Look Once (YOLO) v3, YOLO v3-tiny and SSD models were developed in a research study [15]. A digital camera was mounted on a fixed bracket above the conveyor belt to capture image of corn before the peeling process. Amongst the developed networks, YOLO v3 had the highest accuracy (90.24%) compared to the other models for broken and non-broken corn detection.

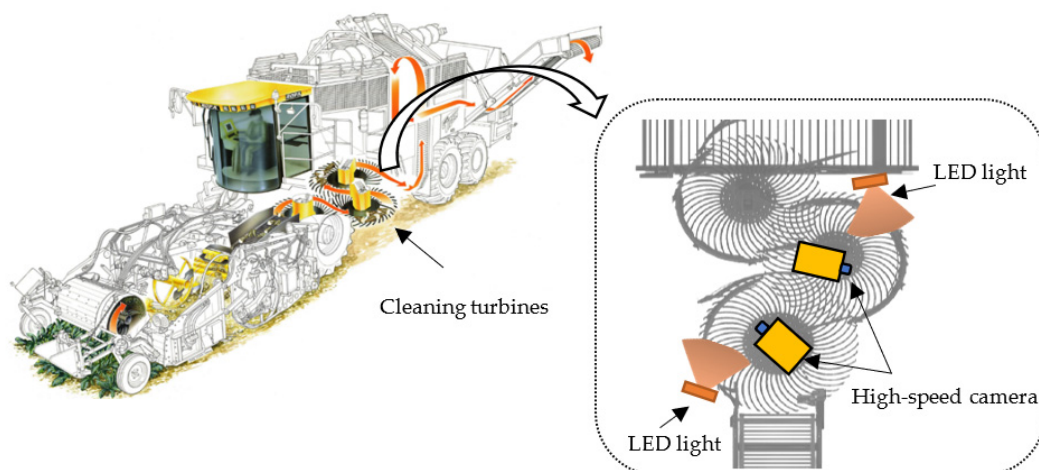
More recently, [16] developed a CNN along with image processing models to identify and classify cracked chili fruits in a sorting machine. A digital camera with LED lights was attached to the sorting machine to capture images from the top part of the chili fruit (the position where the calyx connects to the fruit). It is reported that an accuracy of 97 and 95% for static and working conditions of the sorting machine (respectively) was obtained.

However, no studies have yet been reported for real-time detection of sugar beet mechanical damage during harvesting. Addressing the challenge of real-time monitoring of damages in harvesters is a key parameter in introducing advanced models into the design of automatic and optimized control system to improve the quality of the sugar beet during harvesting [17]. Hence, the aim of this study was to develop various CNN vision models, i.e., YOLO v4, Faster R-CNN and R-FCN for detection of sugar beet damage during harvesting using digital cameras installed in a sugar beet harvester.

## 2. Materials and Methods

### 2.1. Imaging and Data Recording

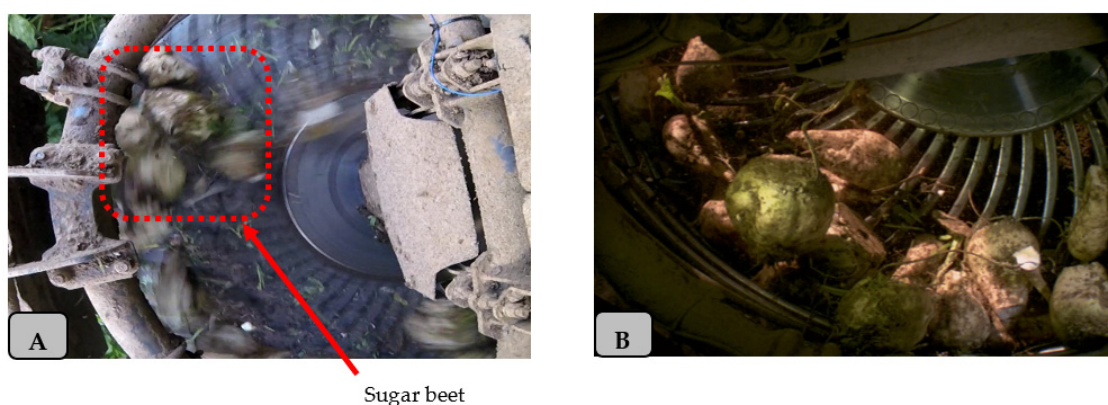
The data recordings for this study were conducted during harvesting days at different commercial sugar beet (variety BTS 440) farms in Lower Saxony (Friedland), in Germany, with a six-row sugar beet harvester (Euro-Tiger 6, 2017, ROPA Fahrzeug und Maschinenbau GmbH, Figure 1). The harvesting took place between October and November in 2018 and 2019 with daily mean temperature of 5.6–19 °C. The travel speed of the harvester, as well as the sieve star cleaning unit speed, were based on standard/usual operating (harvesting) procedure during the trials to give the possibility of mapping real harvesting conditions. In this study, due to the importance and impact of cleaning speed and intensity on the mechanical damage of sugar beet during harvesting [5], two cameras were attached to two cleaning turbines (Figure 1).



**Figure 1.** Schematic representation of the sugar beet harvester [https://www.ropa-maschinenbau.de/site/assets/files/5615/ropa\\_euro-tiger\\_v8-4\\_d.pdf](https://www.ropa-maschinenbau.de/site/assets/files/5615/ropa_euro-tiger_v8-4_d.pdf) (accessed on 7 November 2021), attachment of cameras and LED lights during field trials.

Due to high rotating speed of cleaning turbines, the feasibility of using a normal camera (GigE uEye UI-5240 SE, Imaging Development Systems GmbH, Obersulm, Germany) with maximum frames per second (fps) of 60 and a high-speed camera (Chronos 1.4 from Kron Technologies, Burnaby, BC, Canada) with maximum fps of 40413 was considered during a pre-test recording phase. Figure 2 presents images of sugar beet taken with a normal camera compared with a high-speed camera. The images with a normal camera

(e.g., Figure 2A) had the issue of motion blur since the speed of cleaning turbines was faster than the camera's fps. However, the high-speed camera recorded videos with clear sugar beet along with visible detailed damage in the image (e.g., Figure 2B). Therefore, further video images were recorded using the high-speed camera with remote controllable record timing during the trials. Resolution and fps of the high-speed camera are adjustable; the highest image resolution (1280 × 1024 pixels) could result in the lowest fps (1069); however, the lowest image resolution (320 × 96 pixels) could allow recording with the highest fps (40413) of the camera. In this study, based on the pre-test recording phase a resolution of 1024 × 768 pixels and fps of 1770 were selected. Due to the high fps and image resolution, each video clip was recorded for 16 s to avoid processor's memory error during data processing.



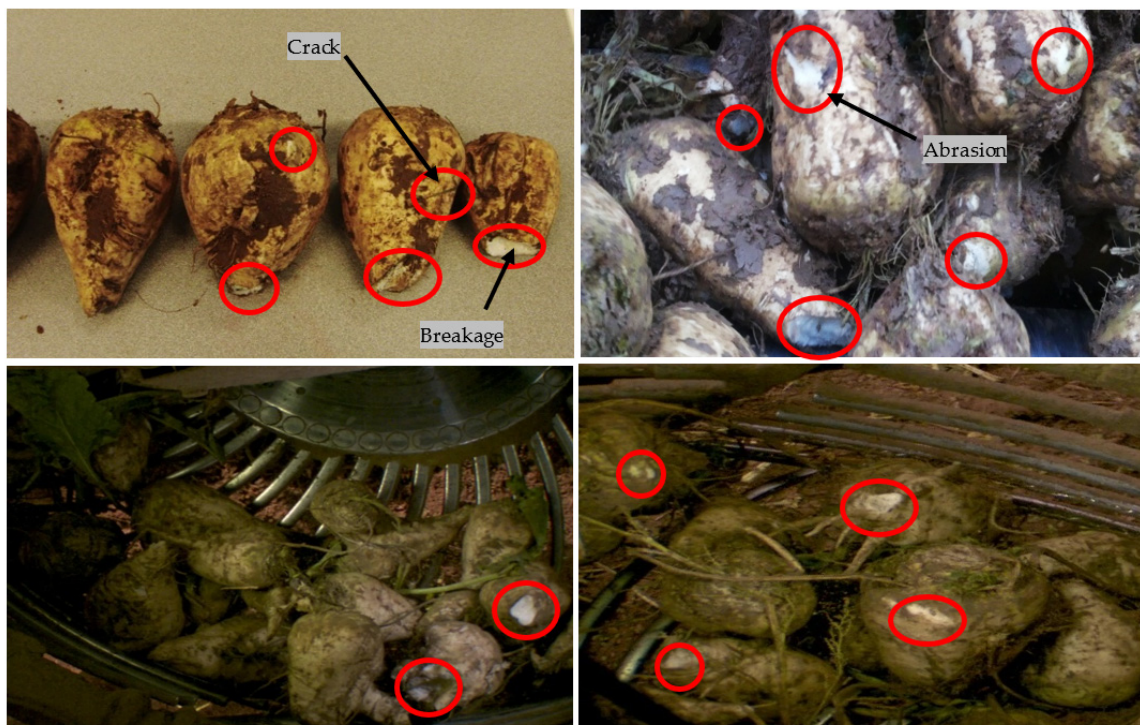
**Figure 2.** Example of sugar beet images in the cleaning unit using normal (A) and high speed (B) cameras in the sugar beet harvester.

In order to ensure constant lighting conditions during the trials, two LED lights with 48 watts and maximum luminous flux of 6500 (LM) were attached to the cleaning turbines (Figure 1). Furthermore, the outer part of the cleaning unit was covered by a black tarpaulin sheet to avoid direct sunlight effects on the image quality. To facilitate the development of a robust model, recordings were carried out in different harvesting conditions, allowing for the capture of sugar beet images with different damage, dirt, size and color.

## 2.2. CNN Methodologies

Due to the lack of available benchmark sugar beet damage data sets, cracks, breakage and surface abrasion were considered as damage in this work. Examples of damage types used in this research are shown in Figure 3. All important and visible mechanical damages of sugar beet in a harvester have been considered in the development of the detection models. To establish the damage data set, images from the recorded video was extracted for training and test phases. A set of 3425 images including sugar beet from various farms were taken. The training phase includes 80% of the total images (2740) and validation consists of 20% (685) of the images. Furthermore, 400 images (independent of the training and validation images), were randomly selected and used for evaluation (test) of the detection phase. All images were annotated using a graphical image annotation tool “LabelImg” [18] and saved as TXT files for the YOLO v4 model. However, to implement the labelled data sets in Faster R-CNN and R-FCN models, the TXT files were converted to PASCAL VOC XML annotation format by a self-developed code. The proposed detection methods were performed under the Python 3.6 environment and OpenCV 4.3. The training of the CNN models was performed by a Windows 10 with a NVIDIA GeForce RTX 2080 GPU with 8 GB of memory. Furthermore, to increase the effective size of the data set for training, data augmentation including geometrical transformation techniques, i.e., rotation and random horizontal flip methods were used. In this study, to find the best model for sugar beet damage detection in the harvester, various CNNs (i.e., YOLO v4 CSPDarknet53, Faster R-

CNN Inception v2, Faster R-CNN Neural Architecture Search (NAS) and R-FCN Residual Network (ResNet) 101 were developed and evaluated.



**Figure 3.** Examples of damage used for development of the CNN models.

#### 2.2.1. YOLO v4

The YOLO v4 network is a one-stage object detection model introduced by [19] and was used in this study to detect mechanical damage of sugar beet. It can also gain high accuracy and speed for object detection scenarios [20] using a single CNN to process and compute the classification result and find coordination of the object. The YOLO models divide input images into  $S \times S$  grids and compute the confidence score and predict the bounding box [21]. Furthermore, the probability of the presence of the object's center in each grid is used for detection of the object. The YOLO v4 network employs CSPDarknet53 as the main backbone network (which is a CNN with cross stage partial networks) to train the network and extract features from the input images [19]. In the YOLO v4 model, a path aggregation network (PAN) as a neck, aggregates feature from different backbone levels and ensures different important layers are fused [22]. Figure 4 illustrates the main stages of the YOLO v4 algorithm to detect sugar beet damage. In this study, the network input size was set to  $608 \times 608$  pixels and training of the network was conducted with a momentum of 0.94, learning rate of 0.001, decay of 0.0005 and iteration of 6000.

#### 2.2.2. Faster R-CNN

The Faster R-CNN method as proposed by [23] is a two-stage object detection algorithm and was used in this study to detect sugar beet damage. The Faster R-CNN models consist of a unique regional proposal network (RPN) in the first step using conventional layers to generate object proposals, and in the second stage it performs bounding box regression and generates feature maps to predict the classes [24]. The bounding box and classification adopt features of the target object candidates to generate region proposals and resize by pooling layers to have constant width and length [25]. Figure 5 shows the architecture and different steps of the Faster R-CNN utilized for sugar beet damage detection. In this study, two feature extraction techniques (i.e., Inception V2 and NAS) were adopted with the aim of finding the best Faster R-CNN sugar beet damage detection model.

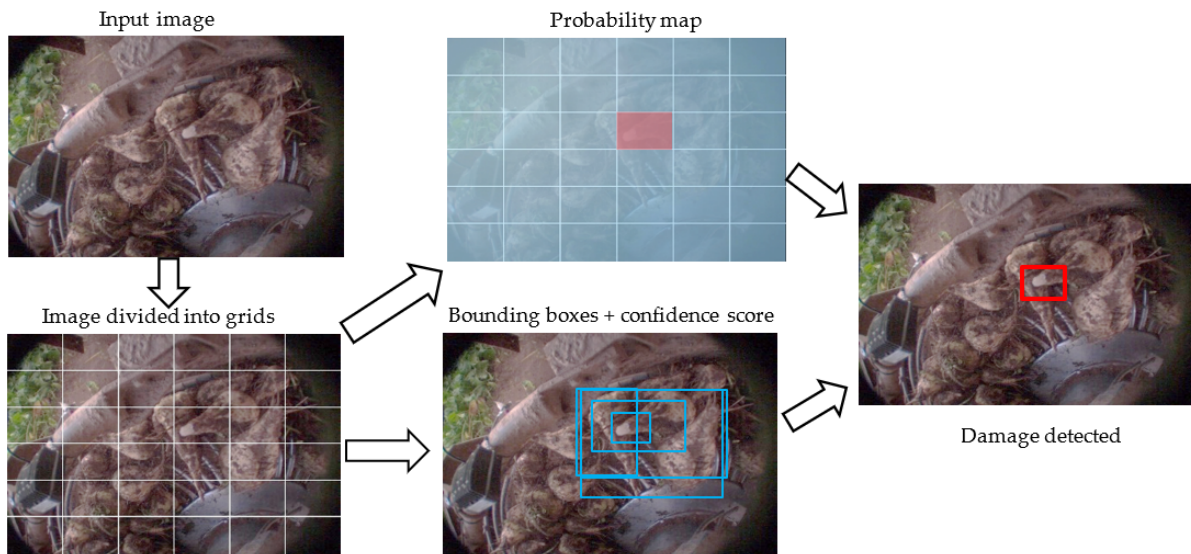


Figure 4. Schematic representation of YOLO v4 for sugar beet damage detection.

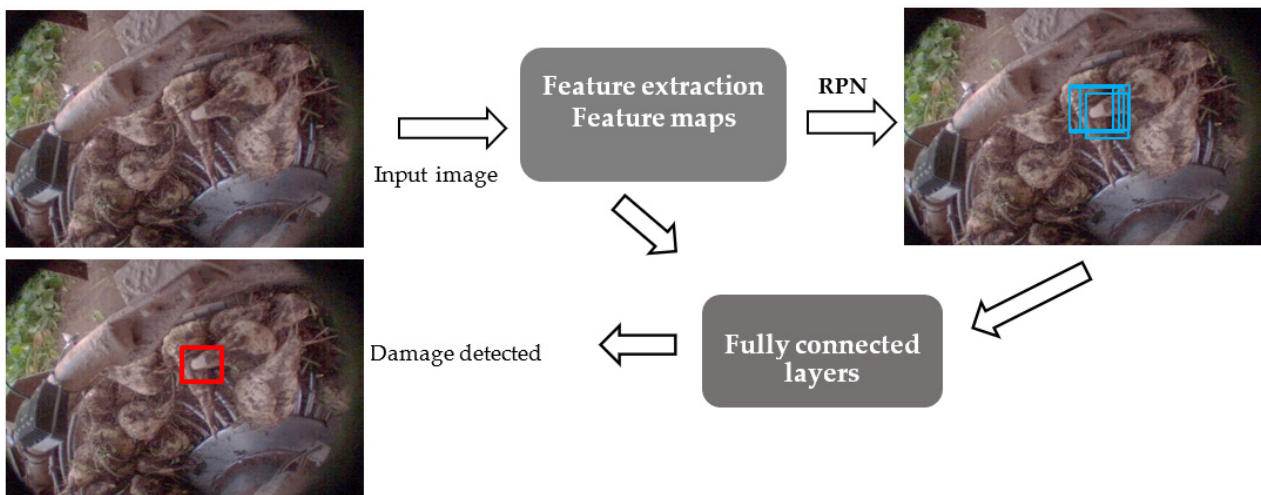


Figure 5. Schematic representation of Faster R-CNN for sugar beet damage detection.

### 2.2.3. R-FCN

R-FCN is a two-stage object detection network proposed by [26] and is an improvement of the Faster R-CNN model, used for the detection of damage in sugar beet. The schematic representation of the proposed R-FCN applied in this study is shown in Figure 6. The R-FCN which uses RPN to generate region proposals, is end-to-end approaches and shares previous feature maps with detection layers. The R-FCN network proposes position-sensitive region of interest pooling layers and position-sensitive score maps to address the issue of translation invariance [27]. In the R-FCN network, features are selected before prediction from the last layer of features and the computation is shared across the whole image by creating a deeper FCN. Therefore, it minimizes the computation memory [26]. In this work, R-FCN with a convolutional layer of ResNet101 was adopted for the detection of damage in sugar beet.

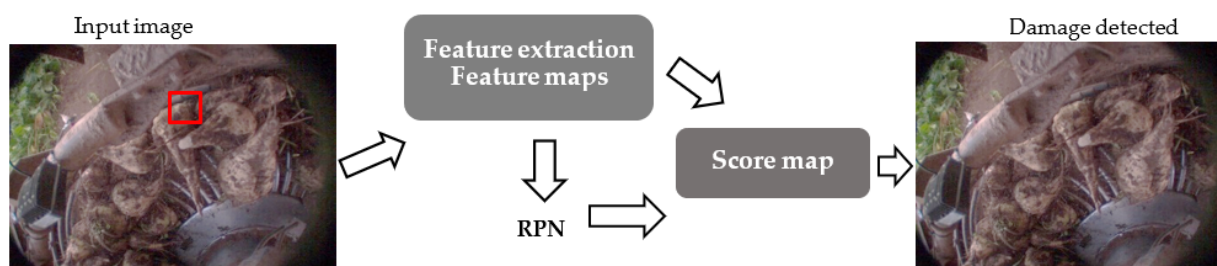


Figure 6. Schematic representation of R-FCN for sugar beet damage detection.

In this study, for both Faster R-CNN and R-FCN models an initial learning rate of 0.003 was selected. Training was conducted for both models with a momentum of 0.9 and an iteration number of 60,000 was chosen. In order to evaluate the performance of the developed models, the widely used evaluation metrics shown in Table 1 were computed. The intersection over union (IoU) was calculated to find performance of the bounding box position. In this context, the overlap between ground truth and the predicted bounding box was computed based on IoU definition [28]. In this regard, if the IoU is equal to or higher than the 0.5 threshold, the result is considered as a true positive (TP), and when the value is below 0.5, it is defined as a false positive (FP). However, the false negative (FN) for sugar beet damage detection means that the developed models present no damage in the image, but the image in fact contains damage [28].

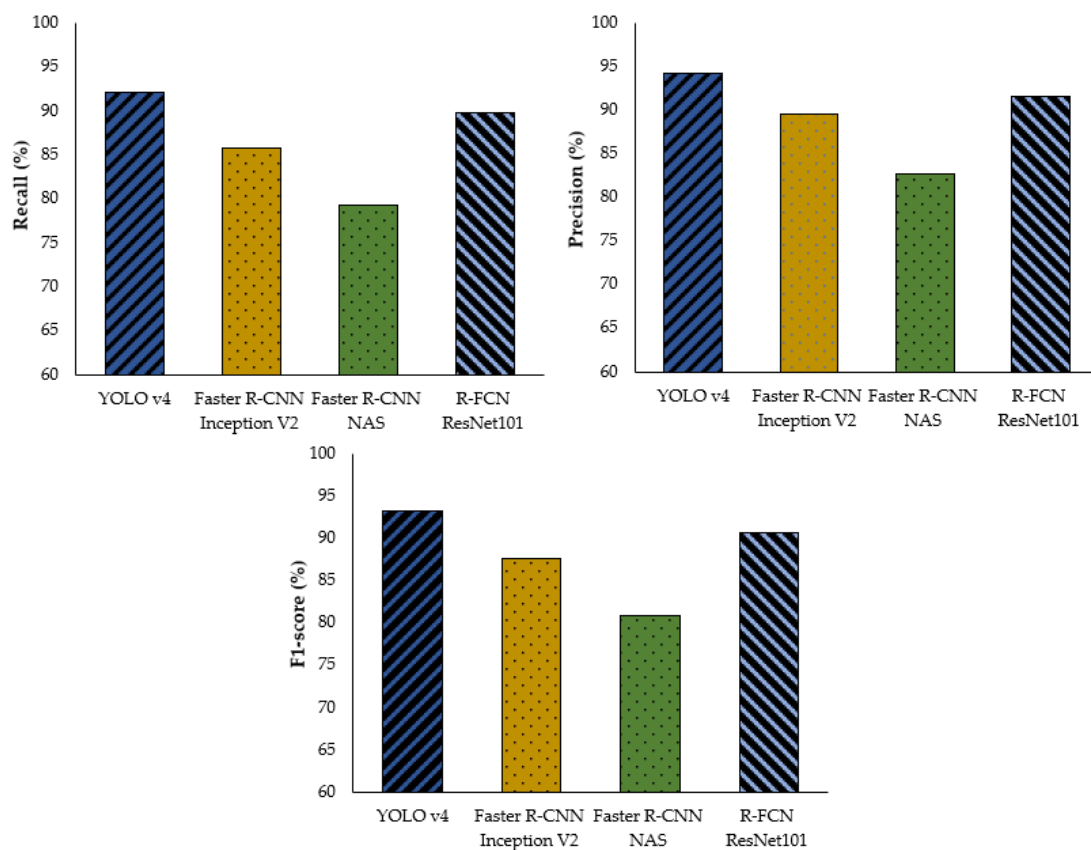
Table 1. Performance metrics for evaluation of the developed CNN models.

Scale	TP	IoU ≥ 0.5	FP	IoU < 0.5
Equation	$Recall = \frac{TP}{TP + FN}$	$Precision = \frac{TP}{TP + FP}$	$F1-score = 2 \times \frac{precision \times recall}{precision + recall}$	

### 3. Results and Discussion

One of the big challenges for sugar beet harvester is the quality of the product during the harvesting, which affects the sugar yield. The importance of automatic detection of damages during the harvesting of sugar beet has led to the development of CNN-based vision algorithms with the real-time ability to monitor a large number of sugar beet samples. However, due to the variability in shape of sugar beet, variety, farming and harvesting conditions, the development of robust detection techniques with the ability to detect damage in a real-time situation is examined in this study. In the test phase, new data (randomly selected 400 images with or without damages in each image and not used in training and validation phases) were used to evaluate the performance of the developed models.

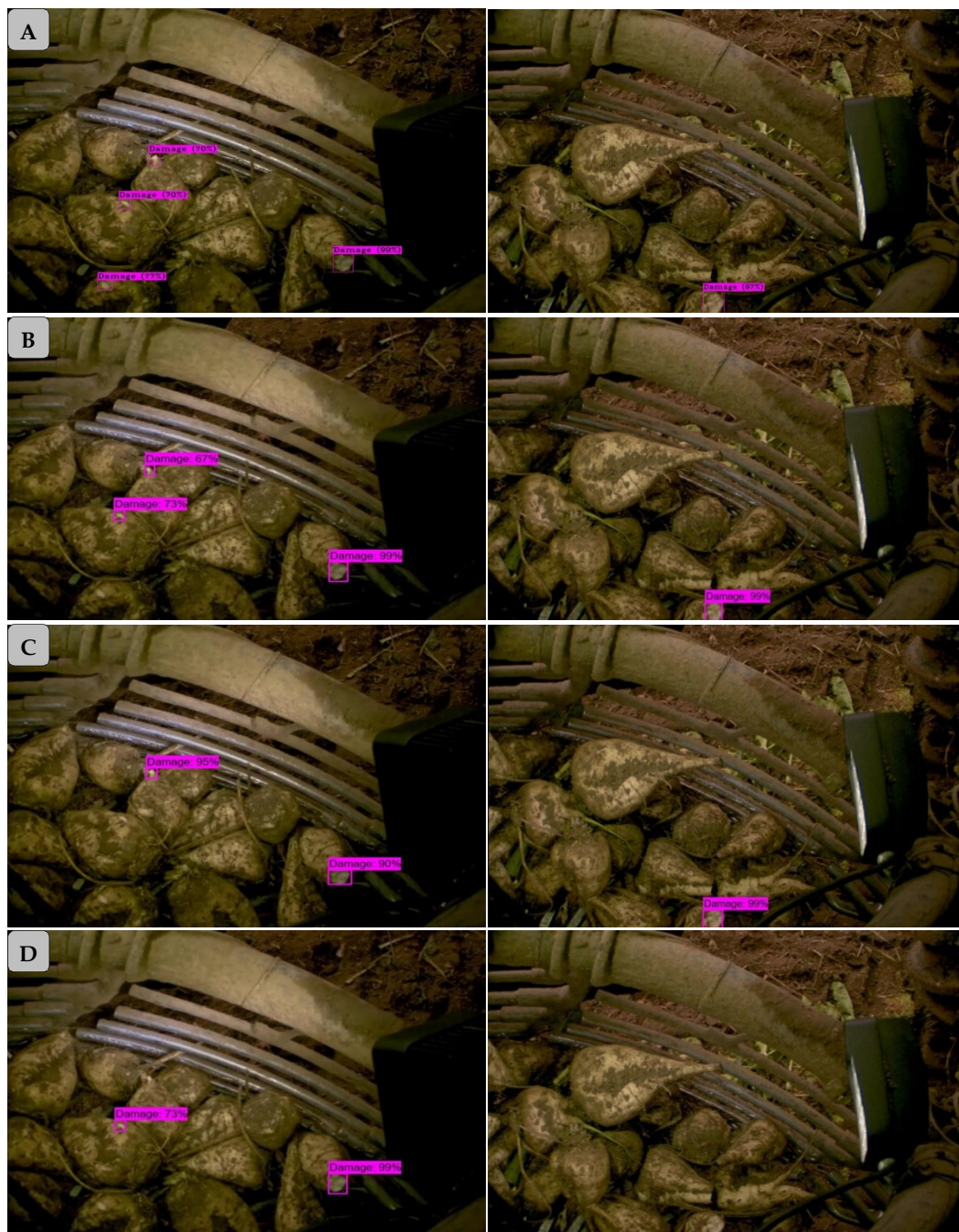
The performance obtained for each developed model in the test data set is presented in Figure 7. In order to evaluate the detection capability of the proposed models, the evaluation metrics (Table 1) were calculated. High values for the precision and recall show the acceptable performance of the YOLO v4, R-FCN ResNet101, Faster R-CNN Inception V2 models for the detection of visible mechanical damage in sugar beet. The F1-score, as a harmonic mean of precision and recall of a model, shows how robust the performance of a model is [29]. In this study, the YOLO v4 model shows better performance compared to the other developed networks. This finding is in line with previous studies conducted for detection of citrus in an orchard [30], apples in a farming complex environment [31], pests [32] and tree trunks in a forest [33]. However, the Faster R-CNN NAS model shows lower performance in this research. In this context, results of a reported study by [34] showed that the Faster R-CNN NAS model had higher precision than Faster R-CNN Inception v2 for kiwi fruit detection; however, Faster R-CNN Inception v2 showed better recall compared to the Faster R-CNN NAS model.



**Figure 7.** Precision, recall and F1-score for the developed CNN models.

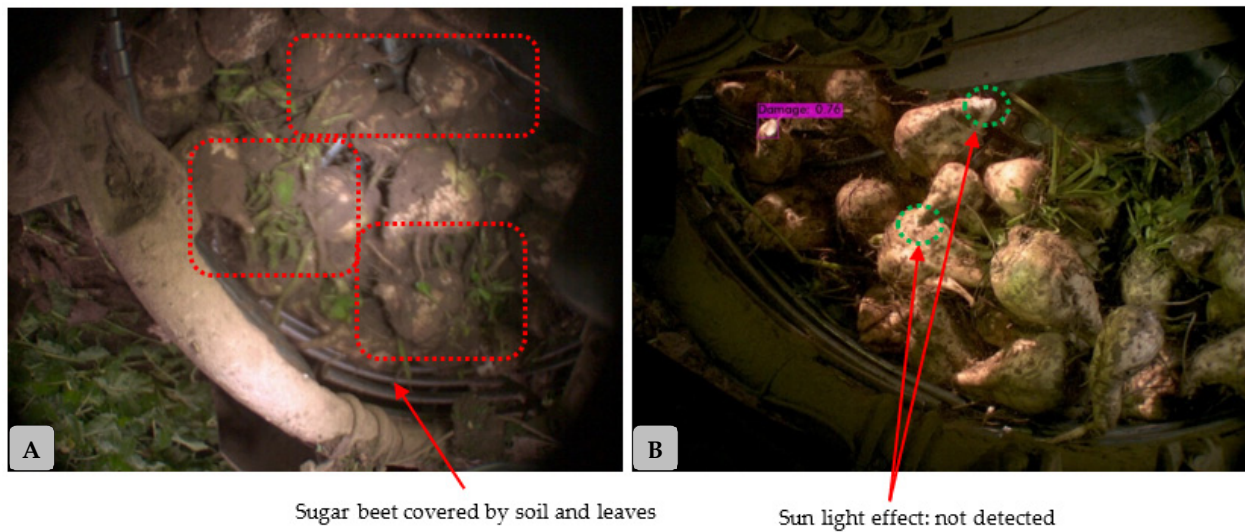
Examples of detected damage by YOLO v4, R-FCN ResNet101, Faster R-CNN Inception V2 and Faster R-CNN NAS models are illustrated in Figure 8. It can be seen from the result that, some of the developed models have the ability to detect mechanical damage of sugar beet in the test data set of the harvester machine. According to the Figures 7 and 8, the proposed YOLO v4 technique for the detection of sugar beet damages in the harvester machine under commercial farming conditions by digital cameras provided a high level of detection performance (e.g., 92, 94 and 93% recall, precision and F1-score, respectively). In a citrus detection study, a YOLO v4 model was able to detect the fruit with an accuracy of 96% using a Kinect v2 camera. In another study, pear fruit detection and counting using different YOLO v4 models (i.e., YOLO v4, YOLO v4-CSP, YOLO v4-tiny) resulted in achieved average precision of 98% [35]. Furthermore, [31] reported that YOLO v4 was able to detect apple fruit in a complex environment with recall and an average precision of around 93 and 88%, respectively, compared to Faster R-CNN with 90 and 83%, respectively. The use of R-FCN as a detector and ResNet101 as a feature extractor showed also high values of precision and recall of more than 98% in a study by [13] for potato surface defect detection. However, compared to our study, the images of the fruit were captured in a control situation. There were some misdetections of the test data set for damage detection in our study. It was observed during the visual assessment of the detection performance that, the major misdetections and causes of damage and lower detection performance were due to the effect of high moisture or wet soil covered sugar beets (Figure 9A), partial or incomplete removal of leaves in the topper during the harvesting. Furthermore, low resolution of the image, due to dry soil covering the camera lens during the harvesting, impacts on the performance of the developed models impacts, and the models are not able to extract enough features to perform accurate detection. Although the cleaning units were covered by a black tarpaulin sheet, the images were affected by sunlight during sunset and/or sunrise (Figure 9B), which has already been reported by [17] as an important parameter negatively affecting the quality of images in sugar beet harvesters.



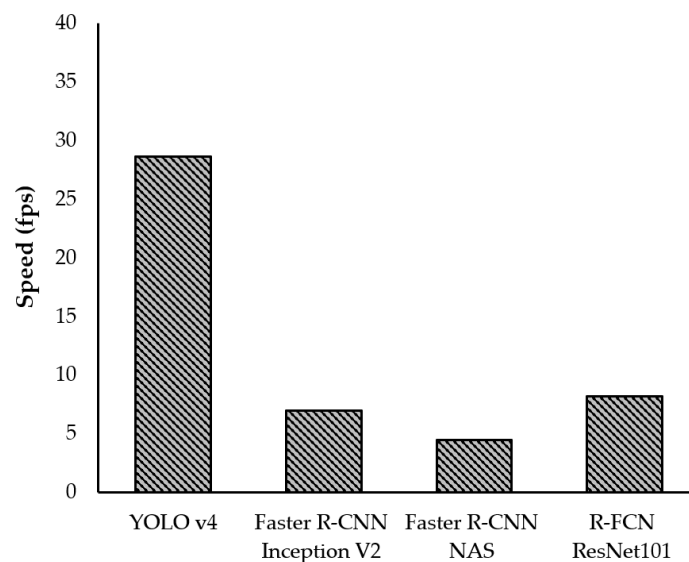


**Figure 8.** Sample of images of YOLO v4 (A), R-FCN ResNet101 (B), Faster R-CNN Inception V2 (C), and Faster R-CNN NAS (D) damage detection in a sugar beet harvester in different cleaning turbines and farms.

Furthermore, in this study the testing time of the detection models was computed and is shown in Figure 10. The average speed of each proposed method was calculated based on the fps. YOLO v4 was the fastest (28.6 fps) due to the single-stage object detector architecture of the YOLO network and can achieve real-time performance [36]. However, the other two-stage object detectors, R-FCN ResNet101, Faster R-CNN Inception V2 and Faster R-CNN had fps of 8.2, 7 and 4.5, respectively.



**Figure 9.** Examples of images affecting detection performance; covered by soil (A) and sun light effect (B).



**Figure 10.** Average detection time of the proposed CNNs for sugar beet damage detection.

According to the performance of YOLO v4, it achieved a good balance between precision, recall, F1-score and speed which could be considered as the best model for sugar beet damage detection during harvesting. This finding is in line with reported studies that the YOLO networks could achieve higher speed and better overall performance e.g., [31,33,36]. Furthermore, our finding is in agreement with [37] who reported that YOLO models achieved higher speed and F1-score compared with SVM, Faster R-CNN for apple surface defect detection.

Comparing with the methodologies used for broken corn detection at a conveyor belt of a corn harvester based on different YOLO v3 models [15], the proposed YOLO v4 network in our study achieved a higher speed (fps) and performance in the detection phase. This is in line with previous findings that the YOLO v4 model is superior and faster than the YOLO v3 model [30]. In addition, ref. [14] used an improved VGG network to detect and classify green plum defects into rot, cracks, rain spots, scars, and intact skin. Compared to our study, their developed VGG model achieved lower recall, precision, and F1-score of 78, 93 and 85% for crack detection in green plums under controlled imaging conditions. However, in another study [13], their developed R-FCN ResNet101 model for

scratch detection in potato had a higher precision (98.1%), recall (99%) and F1-score (98.6%) compared with our proposed networks. This could be due to the different conditions used in two studies, e.g., imaging situations, the type of damage and samples, as well as the controlled conditions of trial in the study [13].

In this study, due to the use of different CNN models for visible mechanical damage of sugar beet during harvesting in a sugar beet harvester, a high enough performance for practical use was obtained. The YOLO v4 detection model described in this study could be a valuable tool to detect changes in the number of damaged sugar beet in real-time conditions during harvesting to improve quality and yield of end-product (sugar). A similar result was reported by [38] that supports the use of YOLO v4 as a reliable network for real-time object detection in agricultural context. The output of the developed model can be used as an input of a control system to automatically adjust setting (e.g., moving and cleaning speeds) of the machine to reduce losses and damage during harvesting. However, this new method needs to be adapted and evaluated with a wider range of harvesting conditions (e.g., various farming conditions, different harvesters, sugar beet varieties, environmental conditions) in future, which may require other imaging and lighting systems (or natural ambient light), a greater number of images, different feature extraction methodologies and CNN models.

#### 4. Conclusions

In sugar beet harvesters, it is essential to monitor and assess the amount number of damage during the harvesting process. However, this is a labor intensive and time-consuming procedure, and in most cases it is not possible for the operator to assess during the harvesting. Therefore, in this study, an alternative methodology based on machine learning and machine vision techniques was developed and tested. Two high-speed cameras were attached to the cleaning unit of a sugar beet harvester to record video data during the harvesting. The 2D images were extracted from the video data and used for training of different CNN models, i.e., YOLO v4, R-FCN and Faster R-CNN. These CNN models were trained using 2470 of the image data and validated by 685 of the 2D images. The trained model was then tested using 400 new images. The findings of the test phase showed a high level of performance (recall, precision and F1-score of about 92, 94 and 93%, respectively) and good processing speed (about 29 fps) for the YOLO v4 compared to the other CNN models. This developed CNN model performed well in the detection of visible mechanical damage of sugar beet in the cleaning unit of a sugar beet harvester, which has been a big challenge for human visual inspection of the machine during the harvesting process. The proposed YOLO v4 model has robustness and flexibility, which can be an important step towards development of an automatic, real-time, and computer-based control system for the sugar beet harvesters.

**Author Contributions:** Methodology, U.W. and A.N.; software, A.N.; validation, A.N. and U.W.; data curation, U.W. and A.N.; writing—original draft preparation, A.N.; writing—review and editing, A.N., U.W. and O.H.; visualization, A.N.; supervision, O.H.; project administration, O.H. and U.W.; funding acquisition, O.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Bundesanstalt für Landwirtschaft und Ernährung (BLE), the German Federal Office for Agriculture and Food, grant number “28-1-57.064-15”.

**Acknowledgments:** We thank the funding organizations and partners of the SmartBeet project.

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

## References

1. Fugate, K.K.; Ribeiro, W.S.; Lulai, E.C.; Deckard, E.L.; Finger, F.L. Cold temperature delays wound healing in postharvest sugarbeet roots. *Front. Plant Sci.* **2016**, *7*, 499. [CrossRef]
2. Bentini, M.; Caprara, C.; Rondelli, V.; Caliceti, M. The use of an electronic beet to evaluate sugar beet damage at various forward speeds of a mechanical harvester. *Trans. ASAE* **2002**, *45*, 547. [CrossRef]
3. Huijbregts, T.; Legrand, G.; Hoffmann, C.; Olsson, R.; Olsson, Å. Long-Term Storage of Sugar Beet in North-West Europe. Coordination Beet Research International. Report No. 1-2013. 2013. Available online: <https://www.nordicbeet.eu/wp-content/uploads/2016/04/COBRI-storage-report-2013-final-131004.pdf> (accessed on 7 November 2021).
4. Kołodziej, P.; Gołacki, K.; Boryga, M. Impact characteristics of sugar beet root during postharvest storage. *Int. Agrophysics* **2019**, *33*, 355–361. [CrossRef]
5. Hoffmann, C.M.; Schnepel, K. Susceptibility to root tip breakage increases storage losses of sugar beet genotypes. *Sugar Ind.* **2016**, *141*, 625–632. [CrossRef]
6. Nasirahmadi, A.; Ashtiani, S.H.M. Bag-of-Feature model for sweet and bitter almond classification. *Biosyst. Eng.* **2017**, *156*, 51–60. [CrossRef]
7. Rojas-Moraleda, R.; Valous, N.A.; Gowen, A.; Esquerre, C.; Härtel, S.; Salinas, L.; O'donnell, C. A frame-based ANN for classification of hyperspectral images: Assessment of mechanical damage in mushrooms. *Neural Comput. Appl.* **2017**, *28*, 969–981. [CrossRef]
8. Hu, M.H.; Zhao, Y.; Zhai, G.T. Active learning algorithm can establish classifier of blueberry damage with very small training dataset using hyperspectral transmittance data. *Chemom. Intell. Lab. Syst.* **2018**, *172*, 52–57. [CrossRef]
9. Xiong, J.; Lin, R.; Bu, R.; Liu, Z.; Yang, Z.; Yu, L. A micro-damage detection method of litchi fruit using hyperspectral imaging technology. *Sensors* **2018**, *18*, 700. [CrossRef] [PubMed]
10. Salimi, Z.; Boelt, B. Classification of processing damage in sugar beet (*Beta vulgaris*) seeds by multispectral image analysis. *Sensors* **2019**, *19*, 2360. [CrossRef]
11. Oraño, J.F.V.; Maravillas, E.A.; Aliac, C.J.G. Jackfruit Fruit Damage Classification using Convolutional Neural Network. In Proceedings of the 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines, 29 November–1 December 2019; pp. 1–6.
12. 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. [CrossRef]
13. Wang, C.; Xiao, Z. Potato Surface Defect Detection Based on Deep Transfer Learning. *Agriculture* **2021**, *11*, 863. [CrossRef]
14. Zhou, H.; Zhuang, Z.; Liu, Y.; Liu, Y.; Zhang, X. Defect classification of green plums based on deep learning. *Sensors* **2020**, *20*, 6993. [CrossRef] [PubMed]
15. Liu, Z.; Wang, S. Broken corn detection based on an adjusted YOLO with focal loss. *IEEE Access* **2019**, *7*, 68281–68289. [CrossRef]
16. Huynh, Q.K.; Nguyen, C.N.; Vo-Nguyen, H.P.; Tran-Nguyen, P.L.; Le, P.H.; Le, D.K.L.; Nguyen, V.C. Crack Identification on the Fresh Chilli (Capsicum) Fruit Destemmed System. *J. Sens.* **2021**, *2021*, 8838247. [CrossRef]
17. Schwich, S.; Schattenberg, J.; Frerichs, L. Development of a Machine Learning-based Assistance System for Computer-Aided Process Optimization within a Self-Propelled Sugar Beet Harvester. In Proceedings of the 2020 ASABE Annual International Virtual Meeting, 13–15 July 2020; p. 2000952. Available online: <https://elibrary.asabe.org/abstract.asp?aid=51512> (accessed on 7 November 2021).
18. Tzutalin. LabelImg. Git Code. 2015. Available online: <https://github.com/tzutalin/labelImg> (accessed on 7 November 2021).
19. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934. in preprint.
20. Du, S.; Zhang, P.; Zhang, B.; Xu, H. Weak and occluded vehicle detection in complex infrared environment based on improved YOLOv4. *IEEE Access* **2021**, *9*, 25671–25680. [CrossRef]
21. Mahurkar, R.R.; Gadge, N.G. Real-time COVID-19 Face Mask Detection with YOLOv4. In Proceedings of the 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 4–6 August 2021; pp. 1250–1255.
22. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.
23. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1137–1149. [CrossRef] [PubMed]
24. Nasirahmadi, A.; Sturm, B.; Edwards, S.; Jeppsson, K.-H.; Olsson, A.-C.; Müller, S.; Hensel, O. Deep Learning and Machine Vision Approaches for Posture Detection of Individual Pigs. *Sensors* **2019**, *19*, 3738. [CrossRef]
25. Dhiraj; Jain, D.K. An evaluation of deep learning based object detection strategies for threat object detection in baggage security imagery. *Pattern Recognit. Lett.* **2019**, *120*, 112–119. [CrossRef]
26. Dai, J.; Li, Y.; He, K.; Sun, J. R-fcn: Object detection via region-based fully convolutional networks. In *Advances in Neural Information Processing Systems*; International Barcelona Convention Center: Barcelona, Spain, 2016; pp. 379–387.

27. Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* **2017**, *17*, 2022. [[CrossRef](#)] [[PubMed](#)]
28. Chang, Y.-L.; Anagaw, A.; Chang, L.; Wang, Y.C.; Hsiao, C.-Y.; Lee, W.-H. Ship Detection Based on YOLOv2 for SAR Imagery. *Remote Sens.* **2019**, *11*, 786. [[CrossRef](#)]
29. Ye, X.; Duan, L.; Peng, Q. Spatiotemporal Prediction of Theft Risk with Deep Inception-Residual Networks. *Smart Cities* **2021**, *4*, 204–216. [[CrossRef](#)]
30. Chen, W.; Lu, S.; Liu, B.; Li, G.; Qian, T. Detecting Citrus in Orchard Environment by Using Improved YOLOv4. *Sci. Program.* **2020**. Available online: <https://www.hindawi.com/journals/sp/2020/8859237/> (accessed on 7 November 2021).
31. Ji, W.; Gao, X.; Xu, B.; Pan, Y.; Zhang, Z.; Zhao, D. Apple target recognition method in complex environment based on improved YOLOv4. *J. Food Process. Eng.* **2021**, *44*, e13866. [[CrossRef](#)]
32. Chen, J.-W.; Lin, W.-J.; Cheng, H.-J.; Hung, C.-L.; Lin, C.-Y.; Chen, S.-P. A Smartphone-Based Application for Scale Pest Detection Using Multiple-Object Detection Methods. *Electronics* **2021**, *10*, 372. [[CrossRef](#)]
33. da Silva, D.Q.; Dos Santos, F.N.; Sousa, A.J.; Filipe, V. Visible and Thermal Image-Based Trunk Detection with Deep Learning for Forestry Mobile Robotics. *J. Imaging* **2021**, *7*, 176. [[CrossRef](#)]
34. Lim, J.; Ahn, H.S.; Nejati, M.; Bell, J.; Williams, H.; MacDonald, B.A. Deep Neural Network Based Real-time Kiwi Fruit Flower Detection in an Orchard Environment. *arXiv* **2020**, arXiv:2006.04343. in preprint.
35. Parico, A.I.B.; Ahamed, T. Real Time Pear Fruit Detection and Counting Using YOLOv4 Models and Deep SORT. *Sensors* **2021**, *21*, 4803. [[CrossRef](#)] [[PubMed](#)]
36. Li, P.; Zhao, W. Image fire detection algorithms based on convolutional neural networks. *Case Stud. Therm. Eng.* **2020**, *19*, 100625. [[CrossRef](#)]
37. Xin, Y.; Ma, S.; Wei, Y.; Hu, J.; Ding, Z.; Wang, F. Detection of Apple Surface Defect Based on YOLOv3. In Proceedings of the 2021 ASABE Annual International Virtual Meeting, 12–16 June 2021; p. 2100611.
38. 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. Available online: <https://elibrary.asabe.org/abstract.asp?aid=52438> (accessed on 7 November 2021). [[CrossRef](#)]