Potato Research
https://doi.org/10.1007/s11540-025-09873-x

Automated Detection and Localization of Potato Blackleg Using a Convolutional Neural Network and Activation Maps

G.W. Hulsman[1] · J.A. van der Heijden[2] · R. da Silva Torres[3]

Received: 22 December 2023 / Accepted: 28 April 2025

Abstract

The quality of a seed potato lot is determined by the number of blackleg-affected potato plants. To enhance the quality, affected plants are both detected and removed manually which is laborious and time-consuming. This paper investigates the potential of using a state-of-the-art machine method for automating the blackleg detection task, a problem overlooked in the literature. In particular, we assess for the first time the use of an efficient convolution neural network (EfficientnetV2) for this problem and elaborate on the explainability of obtained results based on activation maps computed using Grad-CAM. Conducted experiments considering a dataset comprising 4000 images depicting potato plants of six different varieties and affected by three different kinds of *Pectobacteriaceae*. The EfficientnetV2's use led to effective detection results (0.81 recall and 0.91 precision). The qualitative analysis of results using gradient activation maps (Grad-CAM) suggests that the trained model learned to classify blackleg on the characteristic symptoms of blackleg. Our results suggest the feasibility of the practical automation of potato blackleg detection. The investigated detection model can be embedded in a robot platform or mounted on a vehicle to assist farmers in scouting blackleg.

Keywords Activation maps · Efficientnetv2 · Grad-CAM · Potato blackleg · Seed potatoes · Soft rot

✉ G.W. Hulsman
andre.hulsman96@gmail.com

1 Field Crops, Department of Plant Sciences, Wageningen University and Research, Edelhertweg 1, Lelystad 8219PH, The Netherlands

2 Agricultural Biosystems Engineering, Department of Plant Sciences, Wageningen University and Research, Wageningen 6700 AA, The Netherlands

3 Artificial Intelligence Group, Department of Plant Sciences, Wageningen University and Research, Wageningen 6708PB, The Netherlands

Published online: 28 May 2025

Highlights

- We demonstrate the potential of the use of EfficientnetV2 for potato blackleg detection.
- We train EfficientnetV2 on a relatively more diverse and general disease dataset in comparison to similar studies in potato disease detection.
- We apply an image classifier for potato blackleg detection in densely populated fields.
- We investigate Grad-CAM activation maps for symptom-based explainability and localization of detection results.

Introduction

Declassification of seed potatoes as a result of the presence of blackleg[1] is one of the main reasons for economic loss in seed potato farming (Tsror et al. 2009; Kamp et al. 2020). The impact of seed potato declassification varies between countries, as disease certification tolerances differ on a national level (Toth et al. 2011). Seed potatoes are categorized into three categories: certified seed potatoes used to produce tubers for consumption or industry; basic seed potatoes to produce certified seed potato tubers; and pre-basic seed potatoes to produce basic seed potatoes. The categories are subdivided into classes with tolerances for variety mixture, virus symptoms, and blackleg (Haverkort and Delleman 2018). In the Netherlands, blackleg tolerances for pre-basic and basic seed potatoes are 0% and range from 0.1–0.3% for certified seed potatoes (NAK 2012). The Dutch seed crop inspection organization called NAK[2] is responsible for classifying seed potato fields. Common procedures include inspecting seed potato fields twice to spot symptoms of diseases. Later, a post-harvest inspection of tubers and a lab-based analysis of tuber samples are conducted (Bus et al. 1996). To ensure positive assessment during a field inspection, Dutch seed potato farmers inspect their fields at least twice a year to spot blackleg and other diseases. Suspect potatoes are removed from the field in enclosed bags to prevent them from being a source of contamination for other potato plants (Bus et al. 1996).

Blackleg is a result of infection with *Pectobacteriaceae* of genera *Dickeya* and *Pectobacterium* and are latently present in the tuber (Adeolu et al. 2016; Czajkowski et al. 2011). The predominance of different species *Dickeya* and *Pectobacterium* is different per country and also varies over time (van der Wolf et al. 2017). Since 2004, *Pectobacterium brasiliense* has become the most predominant infection source for blackleg in seed potatoes in western Europe reporting 70% of disease incidence in Switzerland in 2014 (de Werra et al. 2015). In Europe, the other main species causing blackleg are *Pectobacterium wasabiae*, *P. atrosepticum*, *Dickeya dianthicola*, and *D. solani* (Toth et al. 2011; van der Wolf et al. 2017; Waleron et al. 2013; Nunes Leite et al. 2014; Van Gijsegem et al. 2021). The most characteristic symptom of potato

[1] Older synonym: Erwinia.

[2] Nederlandse Algemene Keuringsdienst voor zaaizaad en pootgoed van landbouwgewassen—Dutch general inspection service for seeds and seed potatoes in agricultural crops.

blackleg is a slimy, wet, black rot lesion (Fig. 1a) spreading from the rotting mother tuber up the stems, especially under wet conditions (Czajkowski et al. 2011). These symptoms can also occur secondarily as a result of for example late blight, freezing, and flooding (Bus et al. 1996). However, when conditions are dry, symptoms tend to be stunting (Fig. 1b), yellowing (Fig. 1c), wilting (Fig. 1d), and desiccation of stems and leaves (Fig. 1e and f) (Czajkowski et al. 2011; Pérombelon and Kelman 1980). Also, in some cases, potato plants do not show symptoms during cultivation but still hold the bacteria. *Pectobacteriaceae* spread over tubers via specific soil conditions, presence in soil, precipitation or irrigation, storage conditions, or contaminated farming equipment (Chang et al. 2017; van Doorn and van der Wolf 2005; Pérombelon and Kelman 1980). Consequently, it is difficult to eliminate bacterial infection completely in seed potato farming.

A promising alternative to the "manual" detection of potato blackleg relies on the use of automated methods based on computer vision technologies. In fact, computer vision has been successfully employed in several relevant applications in the agricultural sector, including automatic harvesting, quality testing, monitoring farmland information, crop growth monitoring, and disease detection or prevention (Tian et al. 2020). A recent example of automating disease detection is found in Tulip Breaking virus detection where conducted research evolved from theoretical studies (Polder et al. 2010, 2014, 2019b) to prototypes (van Vilsteren et al. 2023). According to a large literature review by Abade et al. (2021), only two studies (Pardede et al. 2018; Khamparia et al. 2020) focused on disease detection in potatoes using RGB images. Both

Fig. 1 Six examples of potato blackleg and its symptoms. Figure 1a shows a black leg or stem (highlighted in red) of a potato plant, Fig. 1b shows stunting, Fig. 1c shows yellowing, Fig. 1d shows wilting, and Fig. 1e and f show the desiccation of the stem and leaves of the potato plant. Figures refer to cropped images from our dataset

studies focused on training a convolutional neural network (CNN) on the PlantVillage dataset that contains among others healthy and infected potato plant leaves. Some other studies (Polder et al. 2019a; Griffel et al. 2023, 2018) focused on potato virus Y detection using hyperspectral or multispectral data. However, virus symptoms can be very subtle in potato plants and only trained experts can distinguish them from non-infected plants in dark weather conditions, which is the most favorable lighting condition for observing the virus (Bus et al. 1996). Only Afonso et al. (2019) investigated the potato blackleg detection problem. Their approach relied on using a ResNet-based image classifier, which led to promising results for potato blackleg detection. However, their research used a relatively small dataset (532 images). Also, their dataset comprised images associated with a single potato variety and was obtained in one season. In addition, the images displayed a single potato plant showing clear symptoms of blackleg. In real potato fields, however, potato plants are densely planted, and the symptoms of blackleg are not always clearly visible. Additionally, the appearance of blackleg depends on weather conditions and potato variety as mentioned by Haverkort and Delleman (2018) and Bus et al. (1996).

When it comes to localization of disease in crops, Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al. 2017) has shown promising results in disease detection in wheat (Ennadifi et al. 2020; Genaev et al. 2021) and detecting nutrient deficiency in sugar beets and orange trees (Espejo-Garcia et al. 2022). To the best of our knowledge, Grad-CAM has not been investigated in the context of the blackleg detection problem.

In this article, we advance the practical automation of potato blackleg detection. Unlike previous studies, our focus is on collecting data under realistic field conditions and developing effective models for blackleg detection. In actual potato fields, plants are densely populated, so our captured data consists of images featuring multiple potato plants. We utilized a state-of-the-art image classifier, EfficientNetV2 (Tan and Le 2021), to determine whether an image contains a blackleg-infected potato plant. Furthermore, we demonstrate that a positive classification for blackleg is linked to the visual symptoms using Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al. 2017). Potentially, the investigated solution could be embedded in a robot platform or mounted on a vehicle to assist a seed potato farmer in scouting blackleg in an automated way.

Materials and Methods

This section introduces the materials and methods used in this study. To classify the collected images as either blackleg or healthy, we utilized a state-of-the-art convolutional neural network (CNN), specifically EfficientNetV2L (Tan and Le 2021). Additionally, we employed Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al. 2017) to localize blackleg within the images. In this section, we describe the employed research pipeline according to the flowchart shown in Fig. 2.

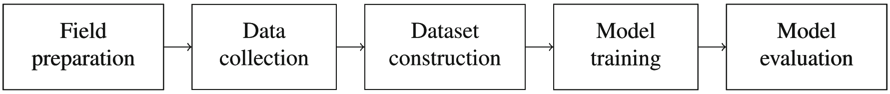


Fig. 2 Flowchart with the methodological steps, starting with the field preparation and data collection in the field, followed by the dataset construction, model training, and the Grad-CAM-based analysis

Field Preparation

During the 2021 and 2022 potato growing seasons, data was collected from healthy and infected potato plants in a field plot of Wageningen University & Research Field Crops in Lelystad, The Netherlands. The field consisted of six potato varieties (Agria, Esmee, Fontane, Kuras, Kuroda, and Riviera), with 1250 tubers each. To study the effect of bacterial inoculation on potato plants, the tubers were divided into four groups and inoculated with three different *Pectobacteriaceae* (*D. solani* 250/1250, *P. atrosepticum* 250/1250, and *P. brasiliense* 250/1250) or just water (500/1250) to avoid the effect of the inoculation process. Potatoes inoculated with *Pectobacteriaceae* were grown in separate plots from the water-inoculated potatoes. The tubers were planted at a distance of 0.49 m.

In both 2021 and 2022, we counted the number of plants showing symptoms as shown in Tables 1 and 2. Visually, we noticed that the water-inoculated plants grew larger in comparison to *Pectobacteriaceae*-inoculated plants whether they showed symptoms or not. In 2021, all varieties were inoculated with CFU[3] 10E6 ml^{-1} of a specific bacterium inoculated. This amount of inoculum is similar to that employed in other studies (van der Wolf et al. 2017; de Haan et al. 2008). In 2022, it was decided to decrease the inoculation concentration with a factor of 10 to CFU 10E5 ml^{-1} for the varieties Agria, Esmee, Fontane, and Riviera. It was expected that with a lower amount of inoculation, the potato plants would show less severe symptoms. Both Kuras and Kuroda were still treated with CFU 10E6 ml^{-1} since Table 1 suggests that these varieties were less sensitive for showing symptoms. In 2021, for variety Agria, the tubers were naturally contaminated with *Pectobacteriaceae* before lab inoculation which is also shown in Table 1 (column water). Consequently, in the field, it was observed that not all tubers developed into potato plants since the tubers were too sick to do so.

Data Collection and Dataset Construction

Data collection was performed using a modified selection cart (Fig. 3a) equipped with a camera system (3× IDS GV-5280FA-C-HW with a RICOH FL-CC0814-5 M 8 mm lens), four led lights (7400 lumens each) to eliminate shadows, and an rtk-GPS (Emlid reach RS2+) (Fig. 3b). During a 6-week period, a qualified NAK inspector was hired to inspect the potato field to mark potato plants with blackleg infection with an orange stick. During data collection, we stopped at an infected plant, removed the demarcation

[3] Colony forming unit: An estimate of the number of microbial cells (among others bacteria) in a sample.

Table 1 Number of blackleg situations where visual symptoms were observed in the field at the last data collection moment in **2021**

Variety	*D. solani*	*P. atrosepticum*	*P. brasiliense*	water
Agria	55/250	89/250	125/250	57/500
Esmee	95/250	116/250	156/250	0/500
Fontane	93/250	100/250	155/250	0/500
Kuras	5/250	7/250	11/250	0/500
Kuroda	25/250	30/250	18/250	0/500
Riviera	16/250	29/250	83/250	0/500

Each potato variety *Pectobacteriaceae* combination contained 250 infected tubers. For variety Agria, we noticed that we had used a contaminated lot of potatoes and that therefore not all tubers developed into a potato plant

stick, and captured an image for annotation. In 2021, due to a wet season, we had to skip parts of the data collection since it was not always possible to drive in the field.

To construct our dataset, 200 images (if possible) for each variety each year spread were randomly selected for overall data collection moments. The final distribution of the dataset is shown in Table 3. As shown, in 2021, fewer images were selected for varieties Kuras and Kuroda due to limited data as a result of a lower disease incidence. The dataset contains a total of 4528 images of which 50% contain blackleg potato plants and 50% contain healthy potato plants. The column water in Tables 1 and 2 shows that there were some potato plants showing symptoms of blackleg even when they were inoculated with water. To select images of healthy potato plants, we excluded all photos taken within a 1.5-m radius of any infected potato plants marked by an inspector. This ensured that the images contained only healthy potato plants.

Model Training and Evaluation

We have implemented Efficientnetv2L using the TensorFlow model hub in TensorFlow 2.8.0. We built the model using a dropout (Srivastava et al. 2014) of 0.2 after the average pooling layer to prevent overfitting and a Dense layer of size 2 with softmax

Table 2 Number of blackleg situations where visual symptoms were observed in the field at the last data collection moment in **2022**

Variety	*D. solani*	*P. atrosepticum*	*P. brasiliense*	water
Agria	14/250	54/250	82/250	1/500
Esmee	57/250	38/250	100/250	0/500
Fontane	36/250	83/250	159/250	0/500
Kuras	6/250	32/250	50/250	1/500
Kuroda	15/250	37/250	36/250	1/500
Riviera	12/250	58/250	166/250	2/500

Each potato variety *Pectobacteriaceae* combination contained 250 infected tubers

(a) (b)

Fig. 3 The selection cart (Fig. 3a) equipped with sensors (Fig. 3b) for data collection: A, emlid reach RTK gps antenna; B, xsense IMU (not used); C, IDS camera lens; D, matronics led lights; and E, touchscreen PC

activation to observe the final output. Additionally, we applied transfer learning using the ImageNet dataset (Deng et al. 2009). The model was trained for 150 epochs with a batch size of 4 using an Adam optimizer with a learning rate of 10^{-6}. We trained the model using one NVIDIA v100 GPU. The training process was repeated five times, and we reported the average results of the trained model on the test set.

Our dataset (Table 3) was split into a training (70%), validation (15%), and test (15%) set. Both the training and validation sets underwent the following augmentation techniques: random rotation (−40 to 40 degrees), random translation (with factors of 0 and 0.2 for width and height, respectively), random zoom (with a width and height factor of 0.2), and random horizontal flipping.

Model Evaluation

The model performance was assessed using the precision (Eq. 1) and recall (Eq. 2), where TP stands for true positive (blackleg classified as blackleg), FP stands for false

Table 3 Content of the dataset with blackleg and healthy potato plants

Year	Class	Agria	Esmee	Fontane	Kuras	Kuroda	Riviera	Total
2021	Blackleg	200	200	200	76	188	200	
	Healthy	200	200	200	76	188	200	
	Total	400	400	400	152	376	400	2128
2022	Blackleg	200	200	200	200	200	200	
	Healthy	200	200	200	200	200	200	
	Total	400	400	400	400	400	400	2400

Whenever possible, 200 blackleg and 200 healthy potato plants per year were selected per variety positive (healthy classified as blackleg), and FN, for false negative (blackleg classified as healthy). The recall is used to evaluate the percentage of correctly classified images containing a plant showing blackleg symptoms, while precision is used to evaluate images containing healthy plants misclassified as having blackleg symptoms.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

We utilize Grad-CAMs to visualize the regions that a CNN focuses on when classifying an image containing multiple potato plants. Gradient-weighted Class Activation Mapping (Grad-CAM) is a tool used to understand how deep learning models produce their predictions for a particular class (Selvaraju et al. 2017). When the gradient is low, or there is no focus, the color will be a very light blue, when there is more focus the color moves from yellow to red which indicates the most important pixels. The color map is then overlaid on the original input image, showcasing the regions where the algorithm had the most focus.

Results

The results are organized as follows: first, the network effectiveness performance on the overall test set and across the varieties is presented in the "Network Effectiveness Performance" section, second, the detection results are evaluated using the gradient activation maps in the "Evaluation of Classification Using Grad-CAM" section.

Network Effectiveness Performance

Table 4 shows the average network performance for every single variety. Recall and precision were calculated for blackleg for each separate variety in the test set. On average, the use of the EfficientNetV2 led to a recall of 0.81 and a precision of 0.91. However, we observed a difference in network performance for different varieties ranging from a recall of 0.70 for Kuroda to a recall of 0.89 for Riviera. From a farmer's perspective, the recall will affect the quality of a seed potato lot whereas the precision will only affect the yield.

Evaluation of Classification Using Grad-CAM

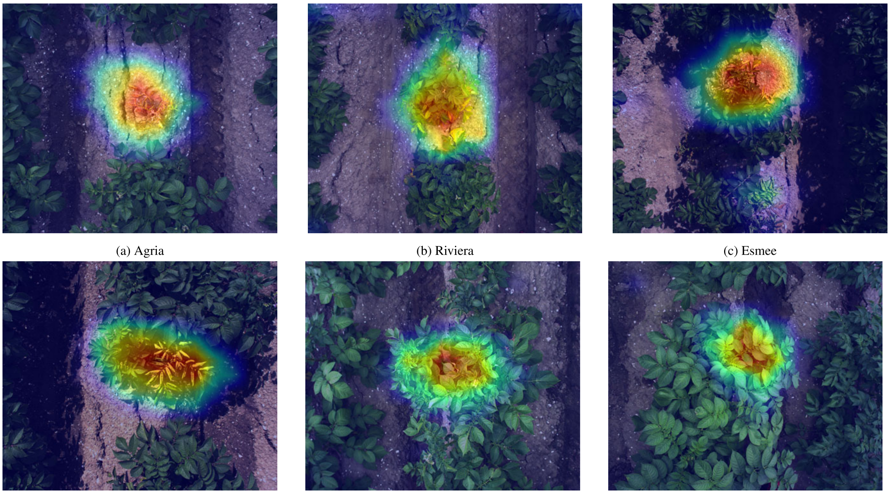
This section elaborates on the detection results observed in Table 4 using Grad-CAM of the last convolutional layer to explain the true classifications ("Correctly Classified Blackleg and Healthy Potato Plants" section) and false classifications ("Misclassification Cases" section).

Table 4 Network effectiveness performance after training for every single variety and all varieties together

	Agria	Esmee	Fontane	Kuras	Kuroda	Riviera	Total
TP	50	40	47	40	40	58	275
FN	7	10	13	11	18	7	65
FP	4	5	4	9	4	3	28
TN	58	55	56	44	45	46	305
Recall	0.88	0.80	0.78	0.78	0.70	0.89	0.81
Precision	0.93	0.89	0.92	0.81	0.92	0.95	0.91

Correctly Classified Blackleg and Healthy Potato Plants

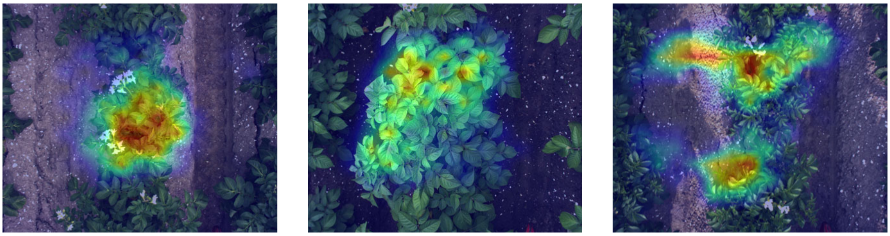
Grad-CAM was applied on the test set to observe how the network identifies blackleg. Figure 4 shows Grad-CAM for six images that contained potato blackleg and were classified correctly. Figure 4a (desiccation of stem and leaves) and b (yellowing/stunting) show that the network focused on the whole plant as the whole plant is showing symptoms of blackleg. Figures 4c (wilting/desiccation of stem and leaves) and 4d (wilting) show that the network focused on the symptomatic part of the potato plant. In Fig. 4e and f, it is hard to observe from the image whether this plant was infected with blackleg or not, the part where the Grad-CAM focused on does not seem to be associated with a symptom of blackleg. Figure 5 shows truly classified healthy potato plants on which we applied Grad-CAM for class healthy. The network focuses on either the plant (Fig. 5a and b) or the surroundings of the potato plant (Fig. 5c).



(a) Agria (b) Riviera (c) Esmee

(d) Kuroda (e) Fontane (f) Kuras

Fig. 4 Grad-CAM applied on images containing blackleg. For the shown potato plants, the network has focused on the whole plant (Fig. 4a and b), on symptomatic parts only (Fig. 4c and d), or specific parts of the plant that do not seem to show symptoms (Fig. 4e and f)



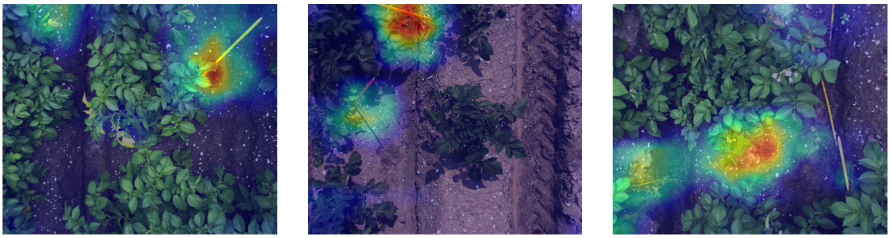
(a) Agria (b) Esmee (c) Kuras

Fig. 5 Grad-CAM applied on images that only contain healthy potato plants. Figure 5a, b, and c show the gradient activation map of class healthy

In the "Data Collection and Dataset Construction" section, it was noted that sticks were used to identify potato plants exhibiting symptoms of blackleg. To collect an image of a potato plant affected by blackleg, the demarcation stick was simply removed, and then an image was taken. However, in eight instances, the stick was inadvertently left in place, or the plant of interest was surrounded by other plants that were also marked with sticks. As shown in Fig. 6a and b, the network focused on the stick instead of the symptomatic part of the potato plant. This suggests that the network has learned to associate the presence of an upright orange stick with potato plants that are diseased with blackleg. On the other hand, Fig. 6c illustrates a case where the network correctly identified the plant despite sticks lying beside it.

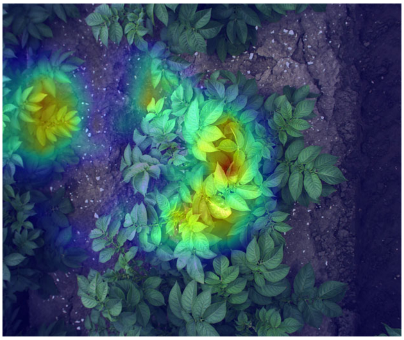
Misclassification Cases

This section elaborates on misclassification cases. In order to reach the highest possible quality standard for seed potatoes, it is important to eliminate all infected plants. Figure 7 shows four examples of potato blackleg being classified as healthy as a result of different issues. First, Fig. 7a shows an example (Kuras) where no blackleg symptoms are visible in the image. Secondly, the plant is completely dead and is barely visible in the image (Fontane, Fig. 7b), and finally, some wilting leaves are visible in the bottom part of the potato plant (Kuroda and Riviera, Fig. 7c and d).

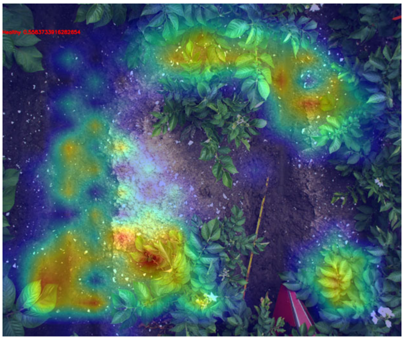


(a) Riviera (b) Agria (c) Agria

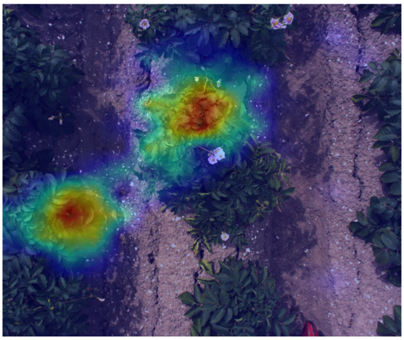
Fig. 6 Blackleg plants with orange demarcation stick visible in the image. Figure 6a and b show that the network is triggered by the stick, while Fig. 6c shows an example of a situation in which the network was not triggered by symptoms of blackleg



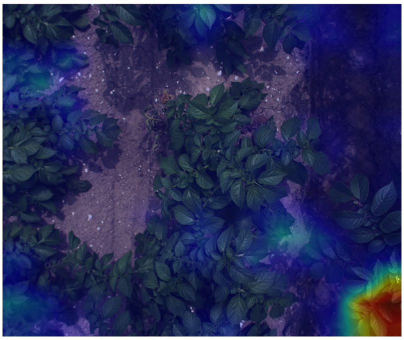
(a) Kuras



(b) Fontane



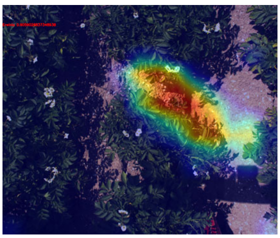
(c) Kuroda



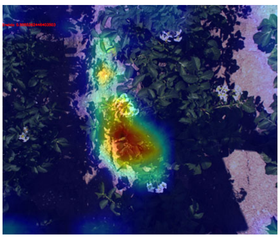
(d) Riviera

Fig. 7 Plant with potato blackleg classified as healthy with Grad-CAM applied for class healthy. Figure 7a shows an example of blackleg which is not visible, Fig. 7b shows an example of a completely dead potato plant, and Fig. 7c and d show an example of some wilting leaves at the bottom part of the plant

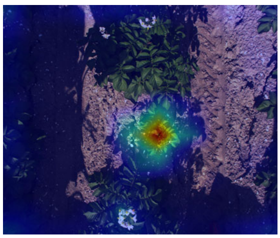
Figure 8 shows three examples of images that contain only healthy plants. Those images, however, were classified as a blackleg potato plant. For a seed potato farmer, these misclassifciations will affect the harvest yield. Figure 8a (Kuras) shows an example of space in between two potato plants, while Fig. 8b (Agria) presents a plant that shows light wilting symptoms. Figure 8c (Esmee) shows a healthy but underdeveloped potato plant. Furthermore, it was observed that 5 out of 24 misclassified healthy potato



(a) kuras



(b) Agria



(c) Esmee

Fig. 8 Three examples of healthy potato plants being classified as blackleg. Grad-CAM for class healthy was applied to the image plants were collected in the field in 2021 whereas the other 19 were images collected in the field in 2022.

Discussion

In this section, we discuss the development of blackleg in our field ("Blackleg Development in Our Field" section) and then the network effectiveness performance ("Detection Model Performance and Generalization" section). Finally, we examine the practical applicability of this research ("Practical Implementation and Future Improvements" section).

Blackleg Development in Our Field

Kwasiborski et al. (2013) pointed out that potato blackleg symptoms relate to weather circumstances. Our data collection took place in two different seasons with different weather circumstances. In 2021, we had an extremely wet planting season, which made us start our experiment 4 weeks later. During the data collection, we also had to cope with wet conditions, which hampered data collection as expected. In 2022, we had a dry season and had to apply irrigation twice. The distribution of misclassified healthy plants mentioned in the "Misclassification Cases" section could therefore be related to drought stress as symptoms can occur for both blackleg and drought stress (Haverkort and Delleman 2018; Bus et al. 1996). Different studies (Bus et al. 1996; Del et al. 2011; Bain et al. 1990) state that some varieties are less likely to show symptoms than others. We observed the same in the field during data collection. Furthermore, we noticed that the time of symptom development was different across the different species. According to van der Wolf and De Boer (2007) and Toth et al. (2011), different factors affect symptom development. For example, Toth et al. (2011) suggest that lower inoculum densities are required for *D. solani* than for *P. atrosepticum* for blackleg symptom development. Some other studies identified a relation between the severity of blackleg and soil moisture content (Pérombelon et al. 1989), temperature (Aleck and Harrison 1978), level of inoculum (Aleck and Harrison 1978), nutrient availability (Graham and Harper 1966), and cultivar susceptibility (Bain et al. 1990; Pérombelon and Hyman 1987).

Detection Model Performance and Generalization

Our trained model achieved a recall of 0.81 and a precision of 0.92 for the class labeled as blackleg. The best performance for specific varieties was observed with Riviera, which reached a recall of 0.89 and a precision of 0.95. According to Griffel et al. (2023), the effectiveness of a neural network tends to decrease when trained on multiple varieties instead of a single variety. Although our results show slightly lower performance compared to the study by Afonso et al. (2019), which reported a recall of 0.91 and a precision of 0.95, our dataset is more challenging, as it involves multiple plants within a single image. Finally, our network performance aligns with findings from studies on potato Y virus detection, where Polder et al. (2019a) reported recall values for different datasets ranging from 0.75 to 0.92 and precision values from 0.30 to 0.92.

In our analysis of misclassification cases ("Misclassification Cases" section), we identified that the trained network struggled to detect blackleg in dead potato plants and in plants where the symptoms were not clearly visible in the images. As noted in previous studies (Bus et al. 1996; Del et al. 2011; Bain et al. 1990), certain potato varieties are less likely to exhibit symptoms of blackleg than others. This may explain the lower recall observed for the Kuras variety, as shown in Table 4. In this case, the symptoms were difficult to discern in the captured images, even though they were visible to the human eye in the field. Different from our camera system, a field inspector can examine the plant from multiple angles or physically inspect it before making a decision. Furthermore, the misclassified healthy plants exhibited signs of drought stress or underdevelopment, which could also indicate the presence of blackleg, as discussed in the "Blackleg Development in Our Field" section.

Ruigrok et al. (2023) pointed out that there is a gap between test-set performance and field performance caused by a generalization error. Currently, model architecture development and generalization mainly focus on generalizing on benchmark datasets, such as, for example, ImageNet (Deng et al. 2009) and common objects in context (COCO) (Lin et al. 2014). However, different studies (Ruigrok et al. 2023; Blok et al. 2021) suggest the need for further research towards ensuring detection models are more generalizable in different agriculture contexts (e.g., different plant varieties or different soil types). For example, Ruigrok et al. (2023) investigated the generalization of volunteer potato detection in sugar beets. In their study, they needed as few as 25 images to avoid the generalization error in some fields. However, to reduce the generalization error by 90%, they needed on average 300 images. Concerning the factors related to the symptomatic appearance of blackleg in combination with a low infection density on the field, multiple hectares of seed potatoes could be needed to avoid the generalization error. With respect to our dataset, we learned that the network focused on symptoms of blackleg in the images. However, as shown in the "Misclassification Cases" section, the model also learned from specific visual properties. For example, in some cases, activation maps refer to upstanding orange sticks as the main visual cue to classify an image as containing blackleg (Fig. 6).

Practical Implementation and Future Improvements

This paper shows promising results for moving to a practical implementation toward the automatic detection and localization of blackleg in seed potatoes. Other studies (Polder et al. 2019b; van Vilsteren et al. 2023) in tulip disease detection have shown that it is feasible to move to a practical application of using a computer vision model in real field crop circumstances to detect disease. The practical use of an EfficientNetV2 model as proposed in this study or in the platforms mentioned in van Vilsteren et al. (2023) requires a vehicle or robot to move in the field. As a consequence in seed potatoes, diseases can be spread more easily, as mentioned by Bus et al. (1996). Alternatively, the use of drones for disease detection in potatoes, as proposed in two recent studies (Siebring et al. 2019; Sugiura et al. 2016), could be further explored for blackleg detection and localization.

Future field implementation using the approach investigated in our research should also consider the differences between on-field performance and on-dataset performance as pointed out by Ruigrok et al. (2020). In our study, the field used differed from typical farmers' fields. According to Haverkort and Delleman (2018), a general potato planting distance is in the range of 13 cm and 33 cm depending on tuber size, which indicates that a farmer's field is more densely populated with potato plants. Furthermore, blackleg potato plants in a typical farmer field are likely to have less severe symptoms, since the natural tuber inoculation regularly contains a lower CFU ml^{-1} than we used during laboratory inoculation (Aleck and Harrison 1978). We expect these conditions will make disease detection even more difficult for both in-field scouting and image-based approaches. The urgency of more investigations to address this issue is acknowledged by Zhang et al. (2020) where they reviewed different deep-learning approaches in dense agricultural scenes. We expect that training a model with images collected under more complex field conditions will be the next challenge to tackle, as disease symptoms will be less visible in the images.

Conclusions and Recommendations

In this paper, we demonstrated that a convolutional neural network (CNN) model can effectively detect potato blackleg in open-field conditions. In addition, we highlighted the potential of using gradient activation maps to analyze the network's decision-making process and to localize blackleg disease within images. In our analysis of misclassified blackleg images, we discovered that the network struggled to detect the disease when the symptoms were not clearly visible or when the potato plants were dead. However, in the case of correctly classified potato blackleg images, it became evident that the network was indeed responsive to the characteristic symptoms of blackleg, allowing it to localize the disease effectively.

This study is a step toward developing a practical solution for automated blackleg detection. However, building on recent research regarding the application of CNNs in open-field situations, we recommend that future studies focus on improving network generalization to overcome the variations caused by different field and weather conditions.

Acknowledgements This research would not be possible without the help of Andre de Ridder (WUR) and Peter Steenbergen (WUR), who modified the hardware of the selection cart for the purpose of data collection. We thank Wim-Peter Dirks (WUR) for his help with the software developed for open-field data acquisition. We also thank Jeroen de Bruijn (WUR), Andries van der Meer (WUR), and Thijs Ruigrok (WUR) for their help with data collection during the growing season in 2021 and 2022 and Kees Boons (NAK) for marking potatoes showing symptoms. With respect to the field used in this study, we thank Jan van der Wolf (WUR) for his support with the lab inoculation of the tubers and Johannes Schra (WUR) for planting and maintaining the field. At last, we thank Jan Kamp (WUR) for giving us the opportunity to write this paper within the project AGROS disease detection in seed potatoes.

Data Availability Data is not publically available, if interested in data contact, data owner (WUR Field crops, Lelystad).

Declarations

Competing Interests The authors declare no competing interests.

References

Abade A, Ferreira PA, de Barros Vidal F (2021) Plant diseases recognition on images using convolutional neural networks: a systematic review. Comput Electron Agric 185:106–125. https://doi.org/10.1016/j.compag.2021.106125

Adeolu M, Alnajar S, Naushad S, S Gupta R (2016) Genome-based phylogeny and taxonomy of the '*enterobacteriales*': proposal for *enterobacterales ord. nov*. divided into the families *enterobacteriaceae*, *erwiniaceae fam. nov., pectobacteriaceae fam. nov., yersiniaceae fam. nov., hafniaceae fam. nov., morganellaceae fam. nov., and budviciaceae fam. nov* Int J Syst Evol Microbiol 66:5575–5599. https://doi.org/10.1099/ijsem.0.001485

Afonso M, Blok PM, Polder G, van der Wolf JM, Kamp JALM (2019) Blackleg detection in potato plants using convolutional neural networks. IFAC-PapersOnLine 52:6–11. https://doi.org/10.1016/j.ifacol.2019.12.481

Aleck JR, Harrison MD (1978) The influence of inoculum density and environment on the development of potato blackleg. Am Potato J 55:479–494

Bain RA, Pérombelon MCM, Tsror L, Nachmias A (1990) Blackleg development and tuber yield in relation to numbers of Erwinia carotovora subsp. atroseptica on seed potatoes. Plant Pathol 39:125–133

Blok PM, van Evert FK, Tielen APM, van Henten EJ, Kootstra G (2021) The effect of data augmentation and network simplification on the image-based detection of broccoli heads with mask R-CNN. J Field Robot 38:85–104. https://doi.org/10.1002/rob.21975 https://onlinelibrary.wiley.com/doi/pdf/10.1002/rob.21975

Bus CB, Loon CD van, Veerman A (1996) Teelt van pootaardappelen. In Ziekten en Plagen, pp 104–106. Lelystad: PAGV. http://edepot.wur.nl/252665

Chang Z, Lv J, Qi H, A YM, Chen D, Xie J, Sun Y (2017) Bacterial infection potato tuber soft rot disease detection based on electronic nose. Open Life Sci 12:379–385. https://doi.org/10.1515/biol-2017-0044

Czajkowski R, Pérombelon MCM, van Veen JA, van der Wolf JM (2011) Control of blackleg and tuber soft rot of potato caused by Pectobacterium and Dickeya species: a review. Plant Pathol 60:999–1013. https://doi.org/10.1111/j.1365-3059.2011.02470.x

Del M, Marquez-Villavicencio P, Charkowski A (2011) Soft rot disease severity is affected by potato physiology and pectobacterium taxa. Plant Dis 95:231–363. https://doi.org/10.1094/PDIS-07-10-0526

Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009). Imagenet: a large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, pp 248–255. Ieee. https://doi.org/10.1109/CVPR.2009.5206848

van Doorn J, van der Wolf JM (2005) Is Erwinia te beheersen? Een literatuurstudie over rotproblemen in diverse gewassen om met deze kennis Erwinia in bolgewassen beter te kunnen aanpakken. Technical Report PPO Bloembollen

Ennadifi E, Laraba S, Vincke D, Mercatoris B, Gosselin B (2020) Wheat diseases classification and localization using convolutional neural networks and gradcam visualization. In International conference on intelligent systems and computer vision (ISCV), pp 1–5. https://doi.org/10.1109/ISCV49265.2020.9204258

Espejo-Garcia B, Malounas I, Mylonas N, Kasimati A, Fountas S (2022) Using efficientnet and transfer learning for image-based diagnosis of nutrient deficiencies. Comput Electron Agric 196:106868. https://doi.org/10.1016/j.compag.2022.106868

Genaev MA, Skolotneva ES, Gultyaeva EI, Orlova EA, Bechtold NP, Afonnikov DA (2021) Image-based wheat fungi diseases identification by deep learning. Plants 10:1500. https://doi.org/10.3390/plants10081500

Graham DC, Harper PC (1966) Effect of inorganic fertilizers on the incidence of potato blackleg disease. Eur Potato J 9:141–145

Griffel LM, Delparte D, Edwards J (2018) Using support vector machines classification to differentiate spectral signatures of potato plants infected with potato virus y. Comput Electron Agric 153:318–324. https://doi.org/10.1016/j.compag.2018.08.027

Griffel LM, Delparte D, Whitworth J, Bodily P, Hartley D (2023) Evaluation of artificial neural network performance for classification of potato plants infected with potato virus y using spectral data on multiple varieties and genotypes. Smart Agric Technol 3:100101. https://doi.org/10.1016/j.atech.2022.100101

de Haan EG, Dekker-Nooren TCEM, van den Bovenkamp GW, Speksnijder AGCL, van der Zouwen PS, van der Wolf JM (2008) Pectobacterium carotovorum subsp. carotovorum can cause potato blackleg in temperate climates. Eur J Plant Pathol 122:561–569. https://doi.org/10.1007/s10658-008-9325-y

Haverkort AAJ, Delleman JJ (2018) Potato handbook : crop of the future. In Environment chapter 4. [Den Haag]: Potato World magazine

Kamp JALM, Blok PM, Polder G, van der Wolf JM, Jalink H (2020) Smart disease detection seed potatoes 2015-2018: detection of virus and bacterial diseases using vision and sensor technology. Technical Report WPR-793 Stichting Wageningen Research, Wageningen Plant Research, Business Unit Field Crops. https://doi.org/10.18174/494707

Khamparia A, Saini G, Gupta D, Khanna A, Tiwari S, de Albuquerque VHC (2020) Seasonal crops disease prediction and classification using deep convolutional encoder network. Circ Syst Signal Process 39:818–836. https://doi.org/10.1007/s00034-019-01041-0

Kwasiborski A, Mondy S, Beury A, Faure D (2013) Genome sequence of the Pectobacterium atrosepticum strain cfbp6276, causing blackleg and soft rot diseases on potato plants and tubers. Genome announcements 1. https://doi.org/10.1128/genomeA.00374-13

Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL (2014) Microsoft coco: Common objects in context. In Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds) Computer vision – ECCV 2014, pp 740–755. Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48

NAK (2012) De keuring van pootaardappelen. https://www.nak.nl/wp-content/uploads/archief/2012/Zakelijke%20informatie/Publicaties/Brochures/De%20keuring%20van%20pootaardappelen%202016.pdf. last Accessed 20 Feb 2023

Nunes Leite L, De Haan EG, Krijger M, Kastelein P, van der Zouwen PS, Van Den Bovenkamp GW, Tebaldi ND, Van der Wolf JM (2014) First report of potato blackleg caused by Pectobacterium carotovorum subsp. brasiliensis in the Netherlands. New Dis Reports 29:24–24. https://doi.org/10.5197/j.2044-0588.2014.029.024

Pardede HF, Suryawati E, Sustika R, Zilvan V (2018) Unsupervised convolutional autoencoder-based feature learning for automatic detection of plant diseases. In 2018 International conference on computer, control, informatics and its applications (IC3INA) pp 158–162. https://doi.org/10.1109/IC3INA.2018.8629518

Pérombelon MCM, Hyman LJ (1987) Epidemiology and etiology of blackleg. Rept Scott Crop Res Inst 1986:102–103

Pérombelon MCM, Kelman A (1980) Ecology of the soft rot erwinias. Annu Rev Phytopathol 18:361–387. https://doi.org/10.1146/annurev.py.18.090180.002045

Pérombelon MCM, Lumb VM, Zutra D, Hyman LJ, Burnett EM (1989) Factors affecting potato blackleg development. In: Tjamos EC, Beckman CH (eds) Vascular wilt diseases of plants: basic studies and control. Springer, pp 421–431

Polder G, Blok PM, de Villiers HAC, van der Wolf JM, Kamp JALM (2019) Potato virus y detection in seed potatoes using deep learning on hyperspectral images. Front Plant Sci 10. https://doi.org/10.3389/fpls.2019.00209

Polder G, Van der Heijden G, Van Doorn J, Clevers J, Van der Schoor R, Baltissen A (2010) Detection of the tulip breaking virus (TBV) in tulips using optical sensors. Precision Agric 11:397–412

Polder G, van der Heijden GW, van Doorn J, Baltissen TA (2014) Automatic detection of tulip breaking virus (TBV) in tulip fields using machine vision. Biosyst Eng 117:35–42

Polder G, van de Westeringh N, Kool J, Khan HA, Kootstra G, Nieuwenhuizen A (2019b) Automatic detection of tulip breaking virus (TBV) using a deep convolutional neural network**this work was partially funded by the Dutch product board for horticulture and the dutch ministry of economic affairs, agriculture and innovation. the research is part of the public private partnership bollenrevolutie 4.0 funded under the grant tki-tu-1806. IFAC-PapersOnLine 52:12–17. https://www.sciencedirect.com/science/article/pii/S2405896319323912 https://doi.org/10.1016/j.ifacol.2019.12.482. 6th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture AGRICONTROL 2019

Ruigrok T, van Henten EJ, Booij J, van Boheemen K, Kootstra G (2020) Application-specific evaluation of a weed-detection algorithm for plant-specific spraying. Sensors 20. https://doi.org/10.3390/s20247262

Ruigrok T, van Henten EJ, Kootstra G (2023) Improved generalization of a plant-detection model for precision weed control. Comput Electron Agric 204:107554. https://doi.org/10.1016/j.compag.2022.107554

Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: visual explanations from deep networks via gradient-based localization. In 2017 IEEE International conference on computer vision (ICCV), pp 618–626. https://doi.org/10.1109/ICCV.2017.74

Siebring J, Valente J, Domingues Franceschini MH, Kamp JALM, Kooistra L (2019) Object-based image analysis applied to low altitude aerial imagery for potato plant trait retrieval and pathogen detection. Sensors 19:5477. https://doi.org/10.3390/s19245477

Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15:1929–1958

Sugiura R, Tsuda S, Tamiya S, Itoh A, Nishiwaki K, Murakami N, Shibuya Y, Hirafuji M, Nuske S (2016) Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. Biosyst Eng 148:1–10. https://doi.org/10.1016/j.biosystemseng.2016.04.010

Tan M, Le QV (2021) Efficientnetv2: smaller models and faster training. In International conference on machine learning 2021. https://doi.org/10.48550/arXiv.2104.00298

Tian H, Wang T, Liu Y, Qiao X, Li Y (2020) Computer vision technology in agricultural automation –a review. Inf Process Agric 7:1–19. https://doi.org/10.1016/j.inpa.2019.09.006

Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Hélias V, Pirhonen M, Tsror L, Elphinstone JG (2011) Dickeya species: an emerging problem for potato production in Europe. Plant Pathol 60:385–399. https://doi.org/10.1111/j.1365-3059.2011.02427.x

Tsror L, Erlich O, Lebiush S, Hazanovsky M, Zig U, Slawiak M, Grabe G, van der Wolf JM, van de Haar JJ (2009) Assessment of recent outbreaks of Dickeya sp. (syn. Erwinia chrysanthemi) slow wilt in potato crops in Israel. Eur J Plant Pathol 123:311–320. https://doi.org/10.1007/s10658-008-9368-0

van der Wolf J, de Haan E, Kastelein P, Krijger M, de Haas B, Velvis H, Mendes O, Kooman-Gersmann M, van der Zouwen P (2017) Virulence of Pectobacterium carotovorum subsp. brasiliense on potato compared with that of other Pectobacterium and Dickeya species under climatic conditions prevailing in the Netherlands. Plant Pathol 66:571–583. https://doi.org/10.1111/ppa.12600

van der Wolf JM, De Boer SH (2007) Bacterial pathogens of potato. In Vreugdenhil D, Bradshaw J, Gebhardt C, Govers F, Mackerron DK, Taylor MA, Ross HA (eds) Potato Biology and Biotechnology chapter 27. Amsterdam, Elsevier Science B.V, pp 595–617 https://doi.org/10.1016/B978-044451018-1/50069-5

Van Gijsegem F, Toth IK, van der Wolf JM (2021) Soft rot Pectobacteriaceae: a brief overview. In Van Gijsegem F, van der Wolf JM, Toth IK (eds) Plant Diseases Caused by Dickeya and Pectobacterium Species chapter 1. Springer International Publishing, pp 1–11. https://doi.org/10.1007/978-3-030-61459-1_1

van Vilsteren S, Hageraats S, van Marrewijk B, Nieuwenhuizen A (2023) TomBollenrevolutie 4.0: eindrapportage werkpakket veld. Technical Report Stichting Wageningen Research, Wageningen Plant Research, Business unit

Ma Waleron, Waleron K, Lojkowska E (2013) Occurrence of Pectobacterium wasabiae in potato field samples. Eur J Plant Pathol 137:149–158. https://doi.org/10.1007/s10658-013-0227-2

de Werra P, Bussereau F, Keiser A, Ziegler D (2015) First report of potato blackleg caused by Pectobacterium carotovorum subsp. brasiliense in Switzerland. Plant Dis 99:551–551. https://doi.org/10.1094/PDIS-07-14-0742-PDN

Zhang Q, Liu Y, Gong C, Chen Y, Yu H (2020) Applications of deep learning for dense scenes analysis in agriculture: a review. Sensors 20:1520. https://doi.org/10.3390/s20051520

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.