Blackleg Detection in Potato Plants using Convolutional Neural Networks ⋆

Manya Afonso * **Pieter M. Blok** * **Gerrit Polder** *
Jan M. van der Wolf * **Jan Kamp** *

* *Wageningen University & Research, Wageningen, The Netherlands (e-mail: {manya.afonso, pieter.blok, gerrit.polder, jan.vanderwolf, jan.kamp}@ wur.nl).*

Abstract: Potato blackleg is a tuber-borne bacterial disease caused by species within the genera Dickeya and Pectobacterium that can cause decay of plant tissue and wilting through the action of cell wall degrading enzymes released by the pathogen. In case of serious infections, tubers may rot before emergence. Management is largely based on the use of pathogen-free seed potato tubers. For this, fields are visually monitored both for certification and also to take out diseased plants to avoid spread to neighboring plants. Imaging potentially offers a quick and non-destructive way to inspect the health of potato plants in a field. Early detection of blackleg diseased plants with modern vision techniques can significantly reduce costs. In this paper, we studied the use of deep learning for detecting blackleg diseased potato plants. Two deep convolutional neural networks were trained on RGB images with healthy and diseased plants. One of these networks (ResNet18) was experimentally found to produce a precision of 95 % and recall of 91 % for the disease class. These results show that convolutional neural networks can be used to detect blackleg diseased potato plants.

Keywords: Neural networks, Machine learning, Image processing, Detection algorithms, Agriculture

1. INTRODUCTION

Softrot and blackleg are bacterial diseases in seed potatoes, which can harm the potato tuber during the storage and in the field, respectively (Chang et al., 2017). When tubers infected with the causative agents, i.e. bacteria belonging to Dickeya and Pectobacterium species, are planted in the field, there is a risk that blackleg will develop and that the pathogen will be spread. At high densities, the pathogen produce cell wall degrading enzymes, which macerate plant tissues. This causes growth stagnation, plant wilting, stem rot or even plant destruction. Thus, this disease lowers both the quantity and quality of the potato yield, leading to an economic loss to the grower (Czajkowski et al., 2015). Disease management is primarily based on the use of pathogen-free, certified seed. Seed certification is based on field inspections for which qualified inspectors are required. An additional way to manage the disease is by roguing symptomic plants, in order to lower disease pressure in the field. To find diseased plants the grower needs to recognize diseased plants, preferably already in an early stage of disease development.

Methods for detecting the bacteria in the tubers use various techniques such as Polymerase Chain Reaction (Smid et al., 1995), bio-electric conductance (Fraaije et al., 1997), and bionic electronic noise that reflects change in volatile substances (Chang et al., 2017). These technologies can only be applied post-harvest. In this work, we investigated the use of image analysis for detecting blackleg in the field, since imaging with cameras is a relatively quick, non-invasive, and non-destructive way of measurement.

Recently, convolutional neural networks (CNNs) (LeCun et al., 2015) has emerged as the state-of-the-art technique for image analysis, from image classification (Krizhevsky et al., 2012) to pixel-wise or semantic segmentation (Long et al., 2015), and object detection (Ren et al., 2015; He et al., 2017). These methods are being used increasingly in agriculture for applications such as automatic harvesting, yield estimation, phenotyping, disease and pest detection (Kamilaris and Prenafeta-Boldú, 2018). The advantage of training CNNs over traditional machine vision methods is their ability to discern discriminative features that would be difficult to hand-craft with data with a lot of variability (LeCun et al., 2015).

Most recent works using CNNs for plant disease detection, work at the level of images of leaves. CNN-based classifiers have been recently applied for classifying images of individual plant leaves into healthy or diseased (Mohanty et al., 2016). These researchers used classifiers trained on the fully convolutional layers of the AlexNet (Krizhevsky et al., 2012) and LeNet (LeCun et al., 1998) architectures, and used the Plant Village Leaf Disease Classification [1] dataset. The leaves in this dataset were separated from the plant and were easy to segment from the background using

⋆ This project was funded by: Branche Organisation Arable Farming (BO-Akkerbouw), Dutch Farmers Organisation (LTO-Nederland), Kverneland Mechatronics, Agrico, HZPC, NAK and the Dutch Topsector AgriFood "Op naar precisielandbouw 2.0" (AF-14275).

[1] https://www.crowdai.org/challenges/1

Fig. 1. Blackleg diseased plants were geometrically stored with a RTK-GNSS rover.

image processing. Image classification using the LeNet architecture was also used in Wallelign et al. (2018) to detect three possible soybean leaf diseases from the Plant Village dataset. A recent work (Fuentes et al., 2017) used FasterRCNN (Ren et al., 2015) with the VGG (Simonyan and Zisserman, 2014) and ResNet (He et al., 2016) convolutional architectures to detect leaves affected by diseases such as mold and powdery mildew. In analogous research, a CNN was used to detect the Y-virus in potato plants using hyperspectral imaging (Polder et al., 2019).

In this work, we aim to classify healthy or blackleg diseased potato plants at individual plant level, rather than detecting at the level of leaves. Our scope is therefore a binary image classification problem, with an image of a plant being classified as either healthy or blackleg diseased. We train a binary classifier on the fully convolutional layer of the ResNet architecture (He et al., 2016).

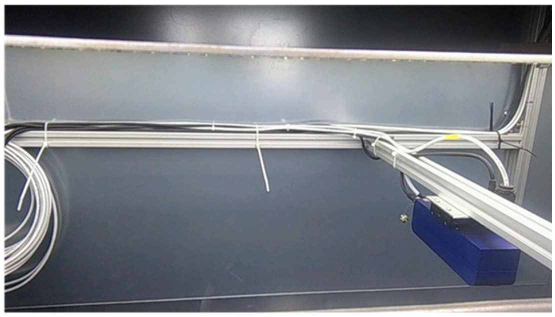
In section 2, we describe our experimental trial, imaging hardware, dataset, and details of our CNN setup. We present the results of our CNN classification in section 3, and conclude with some recommendations to future work in section 4.

2. MATERIALS AND METHODS

2.1 Experimental field

The images were acquired in 2018 on an experimental field that was located in Tollebeek (The Netherlands). The experimental trial involved two rows that contained 1052 potato plants that were grown 0.50 m apart inside the row. To enhance enough examples of both classes, the potato tubers were inoculated with suspensions of Pectobacterium carotovorum subsp. brasiliense before planting (using 107 cells/ml). A crop expert had to visually inspect all 1052 plants on a weekly basis, as some of the inoculated tubers did not develop to a blackleg diseased plant, but remained healthy. Plants that showed disease symptoms were geometrically stored with a RTK-GNSS rover (Hiper Pro, Topcon). A VRS signal (06-GPS, Sliedrecht, Netherlands) was used to guarantee a 0.02 m accuracy on the position estimate. The crop expert obtained the position of a diseased plant by placing the rover at the center point of the plant (figure 1).

(a)



(b)

Fig. 2. Imaging setup: (a) agricultural vehicle with imaging system, (b) camera.

2.2 Image acquisition

To acquire the color images, we used an industrial RGB camera (IDS UI-5240FA-C-HQ). The camera was mounted inside an enclosed box in front of the tractor (figure 2). The camera acquired top-view images of the potato plants in one row in the field. Inside the box, LED strips (OSRAM VFP2400S-G3-865-03) were installed that provided uniform illumination conditions. Black rubber strips formed an apron around the region being imaged, to prevent ambient light from entering.

During image acquisition, the real-world positions of the images were obtained by a RTK-GNSS receiver (Viper 4, Raven) that was installed on the cabin of the tractor. The tractor was stopped in case the RTK-fix signal of the GNSS receiver was lost. In this way the exact world location of all images could be determined with a precision of 0.02 m. The RGB images and the GNSS coordinates were automatically stored on an embedded PC (Nexcom NISE3500) for offline processing.

2.3 Dataset

The imaging setup was used to acquire images of the two rows in our experimental trial. This was done on a weekly basis for six weeks in total. The ground truth of the acquired images was determined by matching the RTK-GNSS coordinates of the rover and the images. From the images, a subset of 532 images was selected across the six different dates. The following criteria were used to select the images:

(1) The potato plant must be isolated from other neighboring plants without overlap or contact.
(2) The selected images should be realistic representations of each class (healthy or blackleg diseased). As such, we avoided trivial classification, for instance healthy detections with large sized plants and blackleg symptoms with smaller sized plants.

Afterwards, the selected images were resized to 224 x 224 pixels using aspect ratio retention to prevent image distortion. The resolution of 224 x 224 pixels corresponds to the preprocessing transformation in the CNN software. The subset of images was randomly split into a training and testing set, with a train-test ratio of 80:20, leading to a training set of 426 images (218 healthy, 208 blackleg diseased) and a test set of 106 images (60 healthy, 46 blackleg diseased). The test set was used for independent benchmarking of the CNN classifier.

2.4 Software and Setup

The PyTorch framework was used to code the CNN classifier in Python, on a workstation with one NVIDIA GeForce GTX 1080 Ti 11GB GPU, 12 core Intel Xenon E5-1650 processor and 64GB DDR4 RAM running Linux Mint 18.2 and CUDA 9.0.

We trained two Residual Network (ResNet) architectures (He et al., 2016), one with 18 layers (ResNet18) and one with 50 layers (ResNet50). We chose for a residual architecture, because this network alleviates the vanishing gradient problem. The fully connected (FC) layer was redefined by a classifier that linearly combined the output of the FC layer into a vector with size, over which a rectified linear unit (RELU) activation was applied. The last network layer consisted of a two class linear classifier using logarithmic soft max activation to enable our binary classification (healthy versus blackleg).

For network weight initialization, we applied transfer-learning using a model that was pretrained on the ImageNet dataset (Russakovsky et al., 2015). For weight optimization, we used the Adam optimizer with a learning rate of 1e-4 for ResNet18 and 5e-5 for ResNet50. Both networks were trained for 100 epochs, using a mini-batch size of 12. To prevent model over-fitting, we applied two forms of explicit regularization. The first was random neuron disconnection (drop-out) and the second was network weight decaying (L2-regularization). For ResNet18, we used a drop-out with 1 % probability (0.01) and for ResNet50 a drop-out with 20 % probability (0.2). The L2-regularization parameter was 0.05 for both networks. Each network was trained for 5 times on random splits of the training set.

2.5 Evaluation procedure

The evaluation was performed on the test set. For each test image, the decision healthy or blackleg diseased was decided through majority voting over the 5 trained models of each network. We evaluated the classifier using confusion matrices. A confusion matrix shows the breakdown of true healthy and true blackleg diseased images, and how many of them were correctly classified. From the confusion matrices we calculated the recall and the precision. The recall is number of true positives divided by the sum of true positives and false negatives. The precision is number of true positives divided by the sum of true positives and false positives. The overall accuracy was calculated as the number of correct classifications divided by the sum of correct and incorrect classifications.

3. RESULTS

The binary classification results for each Resnet architecture are presented in the confusion matrices in tables 1 and 2. The recall values for each class is presented in parentheses in the right most column and the precisions are shown in the parentheses in the bottom row. The accuracy is indicated in parenthesis in the right-bottom corner of the matrices.

Table 1. Confusion matrix for potato plant classification using ResNet18

	Predicted class		
Ground truth	Healthy	Blackleg	Total
Healthy	58	2	60 (97)
Blackleg	4	42	46 (91)
Total	62 (94)	44 (95)	106 (94)

Table 2. Confusion matrix for potato plant classification using ResNet50

	Predicted class		
Ground truth	Healthy	Blackleg	Total
Healthy	50	10	60 (83)
Blackleg	9	37	46 (80)
Total	59 (85)	47 (79)	106 (82)

From the confusion matrix of ResNet18 in table 1, we see that 94% of the images were classified correctly. For the class healthy, the precision was 94% and the recall was 97%. For the class blackleg diseased, these metrics were 95% and 91%. Figure 3 shows some examples of correct classification using ResNet18, and figures 4 and 5 respectively show the four blackleg diseased plants being misclassified as healthy, and the two healthy plants misclassified as diseased. It can be seen from these visual examples that ResNet18 classification can robustly classify healthy or diseased plants, and that the few misclassified instances were very hard examples that visually offer no clue to their true class.

For ResNet50, it can be seen from the confusion matrix in table 2 that the number of misclassifications is higher and the values of the accuracy, precision and recall per class are lower than ResNet18. It must be noted that ResNet50 has many more layers and weights than ResNet18 and therefore requires further optimization of deep learning hyperparameters, which will be addressed in future work. It can be seen from table 3 that the average image analysis time is fast enough for practical applications. The classifier can process 133 frames per second when using ResNet50, while 217 frames per second can be analyzed with ResNet18. These results show that ResNet18 is the preferred classifier for blackleg detection in the field.

4. CONCLUSIONS AND FUTURE WORK

The results from this research show that a CNN, and specifically ResNet18, can work as a robust detector for

Table 3. Computation times for training and testing

Network	ResNet18	ResNet50
Training (minutes) (100 epochs, 426 images)	3	9
Testing per image (ms)	4.6	7.5

blackleg diseased potatoes in the field. In this work, we have used images of individual potato plants that were isolated from neighboring plants. We acknowledge that this situation is only practice during the first weeks after plant emergence. In future research, we will extend our CNN with an object detector, such as FasterRCNN (Ren et al., 2015) or Mask-RCNN (He et al., 2017), to be able to deal with overlapping plants. This classifier is expected to improve the practicability in the field.

In a follow-up research, we also plan to investigate CNN classification using additional image channels other than RGB, such as the hyperspectral channels and depth (RGB-D). Increasing the dataset size and using data augmentation can also be expected to improve the detection performance. Yet another direction for future research would be another transfer learning strategy. Due to the lack of a publicly available dataset with images of potato plants, we used initial models pretrained on the ImageNet dataset (Russakovsky et al., 2015), but an improvement in detection may be possible by transfer learning on a dataset with images of plants similar to potato. We also aim to understand what features a CNN is actually learning (Toda et al., 2019).

ACKNOWLEDGEMENTS

We thank Toon Tielen of Wageningen University and Research (WUR) for building the image acquisition device, Gert-Jan Swinkels (WUR) and Angelo Mencarelli (WUR) for coding the acquisition software, Andre de Ridder (WUR) and Arie de Groot (NAK) for their help in the field trials and Kees Boons (NAK) for providing the experimental field and visual scoring of the potato plants.

REFERENCES

Chang, Z., Lv, J., Qi, H., Ma, Y., Chen, D., Xie, J., and Sun, Y. (2017). Bacterial infection potato tuber soft rot disease detection based on electronic nose. *Open Life Sciences*, 12(1), 379–385.

Czajkowski, R., Pérombelon, M., Jafra, S., Lojkowska, E., Potrykus, M., Van Der Wolf, J., and Sledz, W. (2015). Detection, identification and differentiation of pectobacterium and dickeya species causing potato blackleg and tuber soft rot: a review. *Annals of Applied Biology*, 166(1), 18–38.

Fraaije, B., Appels, M., De Boer, S., Van Vuurde, J., and Van den Bulk, R. (1997). Detection of soft rot erwinia spp. on seed potatoes: conductimetry in comparison with dilution plating, pcr and serological assays. *European journal of plant pathology*, 103(2), 183–193.

Fuentes, A., Yoon, S., Kim, S.C., and Park, D.S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9). doi:10.3390/s17092022. URL `https://www.mdpi.com/1424-8220/17/9/2022`.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2980–2988. IEEE.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Kamilaris, A. and Prenafeta-Boldú, F.X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147, 70–90.

Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., et al. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.

Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431–3440.

Mohanty, S.P., Hughes, D.P., and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7, 1419.

Polder, G., Blok, P.M., de Villiers, H., van der Wolf, J.M., and Kamp, J. (2019). Potato virus y detection in seed potatoes using deep learning on hyperspectral images. *Frontiers in plant science*, 10, 209.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, 91–99.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., and Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3), 211–252. doi:10.1007/s11263-015-0816-y.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Smid, E., Jansen, A., and Gorris, L. (1995). Detection of erwinia carotovora subsp. atroseptica and erwinia chrysanthemi in potato tubers using polymerase chain reaction. *Plant pathology*, 44(6), 1058–1069.

Toda, Y., Okura, F., et al. (2019). How convolutional neural networks diagnose plant disease. *Plant Phenomics*, 2019, 9237136.

Wallelign, S., Polceanu, M., and Buche, C. (2018). Soybean plant disease identification using convolutional neural network. In *The Thirty-First International Flairs Conference*.

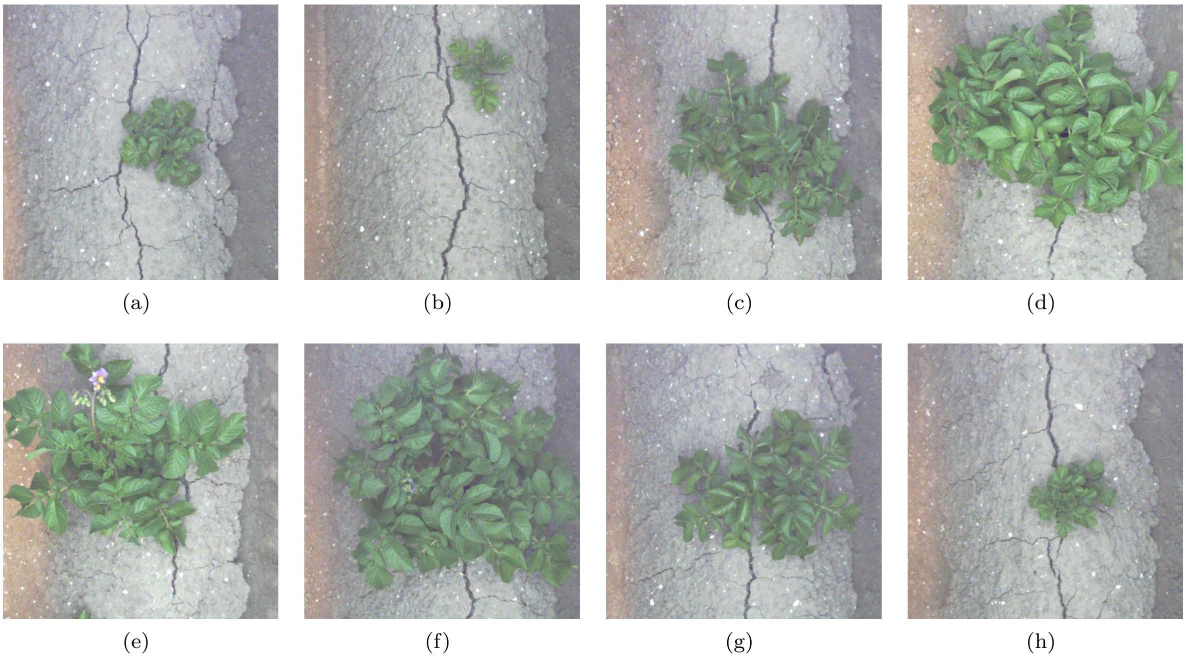


Fig. 3. Examples of correct classification using ResNet18: (a)-(d) blackleg diseased; (e)-(h) healthy.

(a)

(b)

(c)

(d)

Fig. 4. Blackleg diseased plants misclassified as healthy, by ResNet18.

(a)

(b)

Fig. 5. Healthy plants misclassified as blackleg diseased, by ResNet18.