



A Pixel-wise Segmentation Model to Identify Bur Chervil (*Anthriscus caucalis* M. Bieb.) Within Images from a Cereal Cropping Field

Hadi Karimi¹ · Hossein Navid² · Karl-Heinz Dammer³

Received: 22 May 2022 / Accepted: 5 October 2022 / Published online: 7 December 2022
© The Author(s) 2022

Abstract

Because of insufficient effectiveness after herbicide application in autumn, bur chervil (*Anthriscus caucalis* M. Bieb.) is often present in cereal fields in spring. A second reason for spreading is the warm winter in Europe due to climate change. This weed continues to germinate from autumn to spring. To prevent further spreading, a site-specific control in spring is reasonable. Color imagery would offer cheap and complete monitoring of entire fields. In this study, an end-to-end fully convolutional network approach is presented to detect bur chervil within color images. The dataset consisted of images taken at three sampling dates in spring 2018 in winter wheat and at one date in 2019 in winter rye from the same field. Pixels representing bur chervil were manually annotated in all images. After a random image augmentation was done, a Unet-based convolutional neural network model was trained using 560 (80%) of the sub-images from 2018 (training images). The power of the trained model at the three different sampling dates in 2018 was evaluated at 141 (20%) of the manually annotated sub-images from 2018 and all (100%) sub-images from 2019 (test images). Comparing the estimated and the manually annotated weed plants in the test images the Intersection over Union (Jaccard index) showed mean values in the range of 0.9628 to 0.9909 for the three sampling dates in 2018, and a value of 0.9292 for the one date in 2019. The Dice coefficients yielded mean values in the range of 0.9801 to 0.9954 for 2018 and a value of 0.9605 in 2019.

Keywords Bur chervil · Deep learning · Image analysis · Precision farming · Rye · Wheat

Hadi Karimi
h_karimi@areeo.ac.ir, karimi.areeo@gmail.com

✉ Hossein Navid
navid@tabrizu.ac.ir

✉ Karl-Heinz Dammer
KDammer@atb-potsdam.de

¹ Agricultural Engineering Research Department, Kerman Agricultural and Resource Research and Education Center, Areeo, Kerman, Iran

² Department of Biosystems Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

³ Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Max-Eyth-Allee 100, 14469 Potsdam, Germany

Ein pixel-basiertes Segmentierungsmodell zur Identifizierung von Hunds-Kerbel (*Anthriscus caucalis* M. Bieb.) in Farbbildern eines Getreidefeldes

Zusammenfassung

Nach einer Herbizidbehandlung im Herbst zur Kontrolle eines breiten Spektrums von Unkrautarten hat sich in den vergangenen Jahren aufgrund deren geringen Wirkung Hunds-Kerbel (*Anthriscus caucalis* M. Bieb.) zum Problemunkraut im Frühjahr in Getreidefeldern entwickelt. Durch die Klimaerwärmung kann diese Unkrautart den ganzen Winter über keimen, was zu deren Ausbreitung zusätzlich beiträgt. Eine teilflächenbezogene Herbizidapplikation im Frühjahr wäre aus ökologischen und ökonomischen Gründen sinnvoll. Der Einsatz von Farbbildkameras an landwirtschaftlichen Maschinen oder unbemannten Fluggeräten bietet eine preiswerte Möglichkeit zum lückenlosen Monitoring kompletter Felder. Im vorliegenden Beitrag wurde ein Unet-basiertes künstliches neuronales Netz verwendet, um in Farbbildern das Unkraut zu identifizieren. Im Frühjahr 2018 erfolgte in einem Winterweizenfeld an 3 Terminen eine Farbbildaufnahme an 38 Beprobungspunkten. Im Folgejahr wurden die Bilder an einem Termin an 36 Punkten im gleichen Feld mit Winterroggen generiert. Eine manuelle Markierung (Annotation) von Hunds-Kerbel erfolgte in allen Bildern unter zur Hilfenahme einer Software. Nachdem die Originalbilder in kleinere Bilder geteilt wurden, geschah das Trainieren des künstlichen neuronalen Netz Modells an 560 (80 %) der Teilbilder von 2018 (Trainingsbilder). Die Klassifizierungsgüte des trainierten Modells wurde für 2018 anhand 141 (20 %) der Teilbilder und für 2019 anhand aller (100 %) Teilbilder (Testbilder) durchgeführt. Der Vergleich der durch das Modell geschätzten mit den annotierten Hunds-Kerbelpflanzen ergaben zu den drei Aufnahmezeitpunkten 2018 Intersection over Union (Jaccard index) Werte von 0.9628 bis 0.9909 sowie einen Wert von 0.9292 für 2019. Der Dice Koeffizient ergab Werte von 0.9801 bis 0.9954 für 2018 und einen Wert von 0.9605 für 2019.

Schlüsselwörter Bildverarbeitung · Hunds-Kerbel · Maschinelles Lernen · Präziser Pflanzenschutz · Roggen · Weizen

Introduction

Herbicide spraying in cereals under Central European conditions is usually done in autumn. To control a broad spectrum of weeds, in autumn farmers in Germany used to spray herbicides containing biocidal chemical compounds like chlortoluron, metsulfuron, diflufenican, pendimethalin, and tribenuron methyl. Most of these substances show weak effectiveness against bur chervil (*Anthriscus caucalis* M. Bieb.) from the family *Apiaceae*. The germination of this weed usually occurs from autumn to spring. Due to climate change, temperatures in Central Europe have been unusually high in recent years. With favorable temperatures, germination runs throughout the winter. The occurrence of this weed species in cereal crops has increased in recent years. Even as neophyte in Oceania the plant is abundant (Rawnsley 2005).

After common herbicide spraying in autumn, in spring often *A. caucalis* is the only weed species that is present in cereal fields in typical patches. Precise, selective spraying of only those patches would make economic and ecological sense. In spring, farmers often use metsulfuron and thifensulfuron-based herbicide products that are effective against *A. caucalis*. If winter rape or inter-tillage crops are sown after a dry summer, problems with the germination of those crops can occur due to the persistence of the active agents of herbicide products in the soil. Therefore, in addition to economic and ecological aspects, precise spot spraying would reduce this risk to a smaller area of the field.

Prior to a precise herbicide spraying, the target must be identified. In recent years, efforts have been made to use camera-sensors to identify weeds. For example, if they are arranged close to each other at the spray boom an online spot spraying of herbicides (weed detecting and spraying in one operation) can be done (Anonymous 2022).

The accuracy and efficiency of weed identification, determine the performance of the selective spraying technique (Xu et al. 2020). Some visible and non-visible characteristics can distinguish weeds from the main crop. Some of them are leaf shape and leaf margins, the arrangement of leaves on the stem, the presence or absence of hairs on leaves or other parts of the plant, flower structure, color and size. The leaf color is the main characteristic that can make the identification process easier or more difficult. Wu et al. (2011) examined a weed detection method based on position and edge of leaves. The algorithm could only determine image pixels belonging to weeds between the rows and was unable to discriminate the crop from the weeds. Xu et al. (2020) used an Absolute Feature Corner Point (AFCP) algorithm and developed Harris corner detection to extract the individual weed and crop corners as well as the crop rows in field images. The major challenge in developing a weed detection system is the large variability in the visual appearance that occurs in different fields. Thus, an effective classification system has to robustly handle substantial environmental changes, including varying weed pressures, various weed species, different growth stages of crop and weeds, and soil types (Lottes et al. 2018). Environmental

conditions like clouds and wind can drastically affect the characteristics of images (Ghosal et al. 2019).

Artificial Intelligence (AI) is fast becoming ubiquitous due to its robust applicability to challenges that neither humans nor conventional computing structures can efficiently solve. Agriculture is a dynamic field where situations cannot be generalized to propose a common solution. Apply AI techniques can grasp the intricate details of each situation and can provide suitable solutions for farming problems. (Bannerjee et al. 2018). Meanwhile, deep learning, coupled with improvements in computer technology, particularly in embedded graphical processing units (GPUs), has achieved remarkable results in various areas such as image classification and objection detection (Gu et al. 2018; Schmidhuber 2015). The use of convolutional neural networks (CNNs) can take advantage of the recent rapid increase in processing power and memory, allowing for the timely training of large sets of images. (Dos Santos Ferreira et al. 2017). Advances in deep learning and machine vision techniques have been improved and perfected over time mainly through CNNs. The CNNs have been leaders in training algorithms and can visualize and identify patterns in the image data with minimal human intervention (Pouyanfar et al. 2018). Dyrmann et al. (2017) developed a fully convolutional network (FCN) to identify weeds in images captured with an All-Terrain Vehicle (ATV) mounted camera of winter wheat fields. The results showed that the algorithm detected 46% of the weeds in a field. The algorithm faced problems detecting very small weeds like grasses, and weeds that were exposed to a large amount of overlapping. Dos Santos Ferreira et al. (2017) developed software based on CNN that performs weed detection in images of soybean fields, as well as differentiating between grasses and broadleaf weeds. A drone equipped with an RGB camera sensor created an image database of more than fifteen thousand images including soil, the soybean crop, broadleaf weeds and grasses. Using the CaffeNet architecture to train and develop the neural network model resulted in greater than 98% accuracy in the classification. Lottes et al. (2020) developed an algorithm that jointly learns to recognize the stems and the pixel-wise semantic segmentation, taking into account image sequences of local field strips. Their experiments showed the system worked well even in previously unseen sugar beet fields under varying environmental conditions. There are classic and modern network architectures in CNN. Modern architectures such as ResNet, DenseNet and U-Net provide innovative ways for constructing convolutional layers to deepen learning.

It is time-consuming and tedious to prepare a properly annotated image database for training a deep learning system. While many sample images may be needed to be carefully and precisely annotated to form a successful deep network system that can recognize objects in the same way

as a human. The higher the quality and quantity of annotations, the more likely it is that the trained deep learning models will perform well. To use the available annotated images more effectively, Ronneberger et al. (2015) introduced U-Net networks and a training procedure for biomedical image segmentation based on heavy data augmentation. They noted that such a network can be trained end to end from very few images and can effectively solve image segmentation problems. The U-Net architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. Karimi et al. (2021) successfully developed a U-Net-based model to determine two-class pixels of the emergence point of plants and background.

Due to heavy leaf occlusion and close color of *A. caucalis* and cereal plants, simple methods of image analysis have difficulties to discriminate them. In addition, there would be some issues in considering FCN object detection approaches. The trained bounding box predictor usually has trouble generating optimal bounding boxes that cover the entire plant. The problem could be more acute when the object detection model is confronted with high crop density and small weeds which is common in cereal fields. Therefore, in this study, a model for identifying image pixels belonging to bur chervil based on semantic segmentation is proposed. Due to the limitation in providing a large number of annotated images subject to excessive convergence of plants in cereal fields and U-Net's proven ability to resolve semantic segmentation issues, the U-Net architecture was chosen for the development of the pixel-wise weed detection model. The effectiveness and validity of the pixel-wise weed detection model was investigated at three different dates in 2018 and at one sampling date in 2019, while crop and weed conditions, background composition, lighting, etc. vary over time.

Material and Methods

Experimental Sites

The farmer's field is situated in the Eastern Germany South of Berlin (N 51.919175, E 13.151163). The winter wheat variety Nordkap™ was sown on October 1, 2017, with 300 seeds m⁻². The previous crop was winter rape. Before sowing, the soil was tilled with a field cultivator 15 cm in depth (without turning tillage system) after the application of 2.51 ha⁻¹ Roundup™ PowerFlex (480 g l⁻¹ glyphosat) on September 5, 2017. In autumn herbicides were applied twice: October 16, 2017: 1.01 ha⁻¹ Viper™ Compact (100 g l⁻¹ diflufenican, 15 g l⁻¹ penoxsulam, 3.57 g l⁻¹ florasulam), November 3, 2017: 26 g ha⁻¹ Trimmer™ SX (284 g kg⁻¹ tribenuron methyl).

On October 4, 2018 winter rye of the variety Daniello™ with 180 seeds m^{-2} was sown. Before sowing, the soil was tilled with a field cultivator 20 cm in depth after applying 1.71 ha^{-1} Plantaclean Label™ XL (360 g l^{-1} glyphosat) on September 2, 2018. A common herbicide spraying that is effective against a broad variety of weeds was performed on October 17, 2018, with 2.01 ha^{-1} Trinity™ (300 g l^{-1} pendimethalin, 250 g l^{-1} chlortoluron, 40 g l^{-1} diflufenican). All other measures such as nutrition, fungicide and growth regulator application were done in both years according to the rules of good agricultural practice.

Image Datasets

In 2018, 32 images were taken within 50 meters distance along two adjacent transects (machine tracks). Image acquisition continued at the same sampling points on March 27, April 12 and April 18. A different development of the weed and crop, as well as a low to high leaf overlap, was guaranteed. In 2019, 12 images each were taken along three transects (machine track) A Canon EOS 550D camera with a 33 mm lens was used to capture the color (RGB) images. The shutter speed was set on automatically. Each image in this dataset consists of 3-channel color (RGB) images with a resolution of 5184×3456 pixels. Fig. 1 shows examples of images during the three sampling times in 2018 and the one time in 2019.

Image Annotation

In order to provide datasets for training, evaluation and validation processes with suitable truth masks in machine learn-

ing, all images from 2018 and 2019 were annotated with the GIMP—GNU Image Manipulation Program (GIMP Development Team 2019; Solomon 2009). Using a separate layer in the GIMP GUI, a segmentation mask was manually created for each image to precisely mark weeds from other objects such as background soil and cereal plants (Fig. 2). To implement the training process, two classes of the weed and the non-weed must be discrete in each mask frame. Using the OpenCV Python library (Mordvintsev and Abid 2014), the pixels in the mask that included weeds were binarized to white and given a value of 1, while the other pixels were set to black and given a value of 0 (Fig. 2). To ensure that each mask matches the corresponding training image when the CNN U-Net model is run, the individual mask and the corresponding image have been stored separately as a NumPy array (Van Der Walt et al. 2011). A NumPy array data set was then created by appending the relevant image data set for the mask and the training image in a For Loop.

Architecture Used

The U-Net architecture was chosen for the development of the pixel-wise weed detection model. The idea of identifying weed areas with the U-Net architecture can be explained as follows. The U-Net architecture was built with the three paths of contraction, bottleneck and expansion (Ronneberger et al. 2015). The contraction and expansion paths were built with six blocks. In the contraction path, each block included two layers of convolution, followed by a layer of down-sampling. After processing every pixel in each block, a new feature map of the result is stored in the

Fig. 1 Images examples: winter wheat on March 27, April 12 (a, b), and April 18, 2018 (c) as well in winter rye on March 29, 2019 (d)

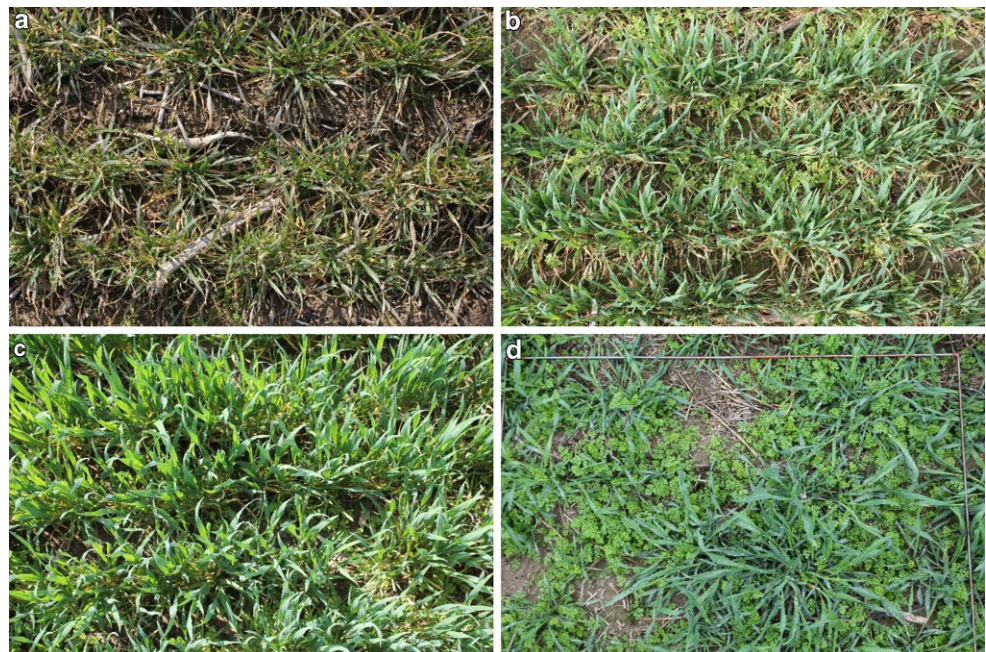
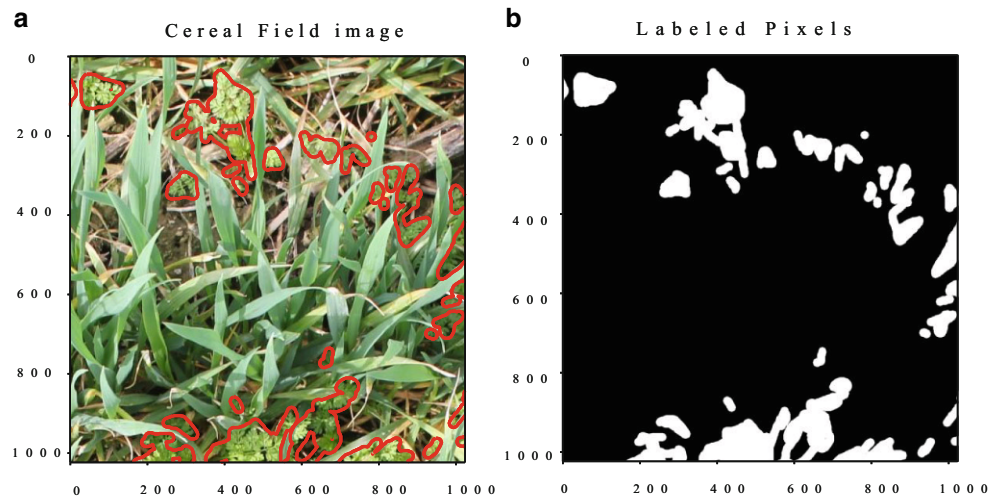


Fig. 2 Manually annotated pixels (red encircled: *A. caucalis*) in a sub-image (a) and transferred to white in the binary mask (b)



same order as the input image. The down-sampling layer was applied to abbreviate the dimension of the feature map so that only the most important parts of the feature map were retained. The reduced features map was then used as input for the next contraction block. The spatial dimensions of the feature maps were reduced by one half and the number of feature maps was duplicated several times by the down-sampling layer (Guan et al. 2019; Weng et al. 2019; Lin and Guo 2020a). The bottleneck, constructed with two layers of convolution, mediated the contraction and expansion sections. The spatial dimension of the data at the bottleneck was 32×32 and 2048 feature maps (Lin and Guo 2020b; Ronneberger et al. 2015). In the expansion path, the block comprised two convolution layers, followed by an up-sampling layer. After each up-sampling layer, the number of feature maps was divided in half and the spatial dimensions of the feature maps was doubled to keep the symmetry of the entire architecture. Once all extension blocks have been run, the final output feature map could provide the final distinction between weed-infested areas and the background. The proposed pixel-wise image segmentation architecture consisted of a total of 34,637,346 parameters, of which 34,623,090 were trainable and 14,256 non-trainable parameters. The non-trainable parameters are the number of weights that are not updated during training with backpropagation.

Hardware and Libraries Used

The semantic image segmentation model, which classified weed pixels from the background pixel in the image, was performed with Keras. Keras is an open-source software library running on top of the TensorFlow package (Abadi et al. 2016), provides a Python interface for solving machine learning problems. It comes with predefined layers, meaningful hyperparameters and a clear and simple API

(Application Programming Interface) that can be used to define, adapt and evaluate standard deep learning models (Géron 2019). By choosing Keras and using models from the open-source community, a supportable solution was developed. It allows for minimal startup time and a greater focus on the architecture of the network rather than the implementing neurons.

As deep learning models spend a lot of time for training, even powerful central processing units (CPUs) were not efficient enough to handle so many computations at once. Compute Unified Device Architecture (CUDA) as a software platform coupled with GPU hardware was used to accelerate computations with the parallel processing power of Nvidia GPUs. The training was carried out under Win 10 as the operating system on a desktop computer with an Intel Coffee Lake Core i7-8700K CPU, 24 GB DDR4 RAM and a GeForce GTX 1080 Ti GPU with 11 GB integrated DDR5 memory.

Loss Function

In deep learning, the optimization task is followed by an established neural network. Based on the data and features of a task, a loss function had to be created to outline the learning objective. Pixel class imbalance is the most common issue in image segmentation practices. The background often goes far beyond the proportion of the target class to be segmented. Therefore, it is important to choose a specified loss function taking into account the optimization goal of the class imbalance segmentation task (Tseng et al. 2021). A large segment of an image taken from a cereal field is expected to belong to crops and soils, resulting in significant differences between the pixels associated with the background and the pixels associated with the weed-infested areas. Thus, in the pixel-by-pixel classification, there is a class

imbalance between the pixels of the weed-infested areas and the pixels of the background.

The Dice coefficient, which appears suitable for image segmentation tasks, was chosen as the loss function. The Dice coefficient is a measure of the overlap between the model prediction results and a reality mask created through manual annotations (Zou et al. 2004). This measure ranges from 0 to 1, with a Dice coefficient of 1 indicating two regions that perfectly overlap and a Dice coefficient of 0 indicating no match. The image segmentation model is designed to make the predictions as close as possible to manual annotation. This means that the training process is essentially based on achieving a greater Dice coefficient. The Dice coefficient could be calculated as follows:

$$\text{DSC} = \frac{2 \times |X_{\text{pred}} \cap Y_{\text{label}}|}{|X_{\text{pred}}| + |Y_{\text{label}}|} \quad (1)$$

Where X_{pred} stands for the area predicted by the model and Y_{label} for the area manually labelled. In terms of the confusion matrix, it can be written using the definition of true positive (TP), false positive (FP), and false negative (FN):

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FN} + \text{FP}} \quad (2)$$

To formulate a loss function that can be minimized during the training process, simply the inverse of the Dice coefficient was applied. The Dice loss can be formulated as $1 - \text{Dice}$. This loss function is known as soft Dice loss because it uses the predicted probabilities directly rather than thresholding them and converting them to a binary mask.

Image Data Preprocessing and Augmentation

In the image pre-processing step, all original field images were randomly cropped and downscaled to 1024×1024 pixels (Fig. 2). The purpose was to reduce the computational load, reduce the risk of overfitting and adjust the image dimensions according to the chosen network architecture. This size limitation is also due to the memory limitation of the Nvidia GPU used for training. The success of developing deep learning models depends heavily on the amount of data available for training (Kebir et al. 2021). Due to the limited original image data of the training set and difficulty in the manual annotation of images, expanding the number of images could enhance the performance of CNNs models (Chatfield et al. 2014). Image data extension techniques have been used to solve the limitations that exist in a short number of training data. Data augmentation is a regularization process that relies on applying transformations of images. (Garcia-Garcia et al. 2017). The image

augmentation aims to apply some transformations to generate more training data (sub-images) from original images. This benefits the model to generalize better to unseen images and avoid over-fitting. (Stirenko et al. 2018; Espejo-Garcia et al. 2020). In this study, the Pixel-wise Segmentation model was trained on the heavy augmented images. In the training process, several transformations were carried out randomly on the original training images and also on the corresponding labeled masks. Tensor imagery batches with real-time data augmentation were produced employing the Keras library. The data were looped in batches. Taking into account the uncertainties such as the angle and position of crops and weeds in images, the image magnification was set to randomly rotate (45°), horizontal shift (0.2), vertical shift (0.2), shear transformation (0.2), zoom (0.2), horizontal and vertical flip, normalize the images (1/255), with the “Wrap” fill mode.

Training and Evaluation

After the image augmentation process, the U-Net algorithm was used to create a two-channel pixel classification model loading augmented training sub-images of the year 2018 as input and the corresponding manually identified binary mask of infested weed regions as output. The network’s Dice loss was determined by comparing the predictions with the reality from the manually labeled mask. Next, the network propagated the prediction error back and renewed the network parameters.

The model was trained using the RMSprop optimizer (Kurbiel and Khaleghian 2017) with a variable learning rate of 0.0001 (Learning rate policy: conditional decay). Eighty epochs were performed in the training process. The number of epochs was determined based on the training image size, training required time, and the overall performance of the model. Due to memory restrictions, the sub-images were uploaded to the network in two batches. Over 701 sub-images and associated annotated sub-image masks were provided for the development and validation of the segmentation model. For developing the model, after random cropping, 80% of the sub-images of the year 2018 in the provided data set were defined as training and the remaining 20% were used as validation and test. Such that, in a data set of 701 sub-images in total, 560 were used to train the model and the other 141 were used for validation and evaluation. The output of the developed network is a two-channel binary mask of 1024×1024 pixels, representing two classes of the weed-infested image area and the background area.

To assess the strength of the pixel-wise model in identifying weed-infested areas in cereal fields, metrics must be established to show whether weed pixels are recognized or not. Further overlap between predictions of the model

and truth (labeled pixels) indicates the greater capacity of the model in the pixel-by-pixel classification of field images. For this purpose, the Dice and the Intersection over Union (IoU), also known as the Jaccard index, were selected to evaluate the segmentation model. These metrics could quantify the percentage of agreement between the target mask and the predictions.

$$\text{JAC} = \text{IoU} = \frac{|X_{\text{pred}} \cap Y_{\text{label}}|}{|X_{\text{pred}} \cup Y_{\text{label}}|} \quad (3)$$

Where X_{pred} represents the predictions and Y_{label} represents the target mask. The intersection ($X_{\text{pred}} \cap Y_{\text{label}}$) consists of the pixels identified in both the prediction mask and the labeled mask, while the union ($X_{\text{pred}} \cup Y_{\text{label}}$) consists of all pixels identified in either the output or truth mask. The Jaccard index is closely related to the Dice coefficient and can be rewritten in terms of the confusion matrix with the definition of true positive (TP), false positive (FP) and false negative (FN):

$$\text{JAC} = \text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}} \quad (4)$$

In order to investigate the possibility of the influence of the growth stage of crops and weeds on the effectiveness of weed detection, the image data set was collected on three different dates from the same points in the field in 2018.

A typical machine vision-based weed detection system faces various issues, such as color deviation due to shading and glare, daylight changes, background segmentation, overlapping of neighboring plants, etc. The overlapping of plants is even more severe in cereal fields as seeds are sown with random patterns and generally close row spacing. Therefore, a generalized system with the ability to detect weeds in different soil and plant conditions, background composition, lighting, etc. is very desirable. In this case, in addition to evaluating the effectiveness of weed detection at three different times in 2018, while crop and weed conditions change over time, the generalizability of the developed pixel-wise detection model was examined using 36 original images captured in the next year on March 29, 2019. In the image preprocessing step, the number increased to 306 sub-images after random cropping and downscaling to a size of 1024×1024 pixels. All new images that the network has not previously seen were fed in and corresponding predictions were made.

Results and Discussion

Training and Validation

The training process involved 80 runs of the entire input data from 2018, namely 79 epochs. As soon as learning was caught and stopped during training, the network benefited from decay in the learning rate by a factor of 0.0. Therefore, if no progress was perceived within 3 epochs, the learning rate decreased (Fig. 3).

The curve of Dice loss and Dice coefficient over the epochs are shown in Figs. 4 and 5. The best time of the learning process, when the model updated its weights for the last time, was signed on the diagrams. As can be observed, no progress has been perceived over a number of specific epochs and the updating of the model weight has been terminated. Such that Dice coefficient and Dice loss value did not respectively increase from 0.9797 and decrease from 0.0568 beyond 57th epochs. However, the best model was determined in the epoch of 72 with Dice loss validation of 0.048.

Evaluation of the Segmentation Model (Test Images 2018)

The output of the trained U-Net model after the activation function is a probable two-channel space with values varying from 0 to 1. This probability space reflects each pixel's possibility of belonging to the target class or the associated background. Fig. 6 shows a comparison between the manually labelled and the predicted weed-infested regions by the

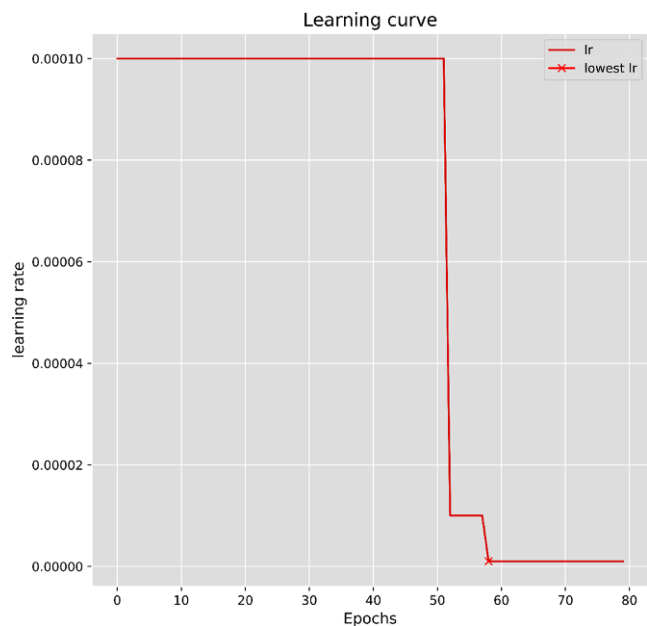


Fig. 3 Decrease in the learning rate by a factor of 0.0 if no progress was perceived within 3 epochs

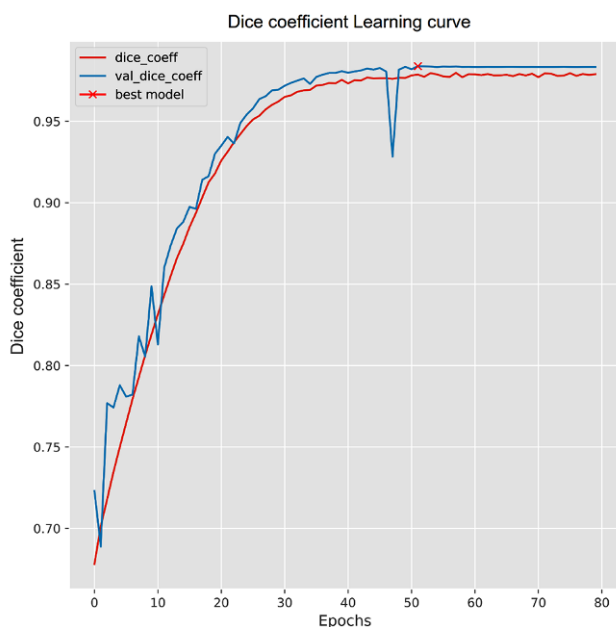


Fig. 4 The curve of the Dice coefficient over the epochs

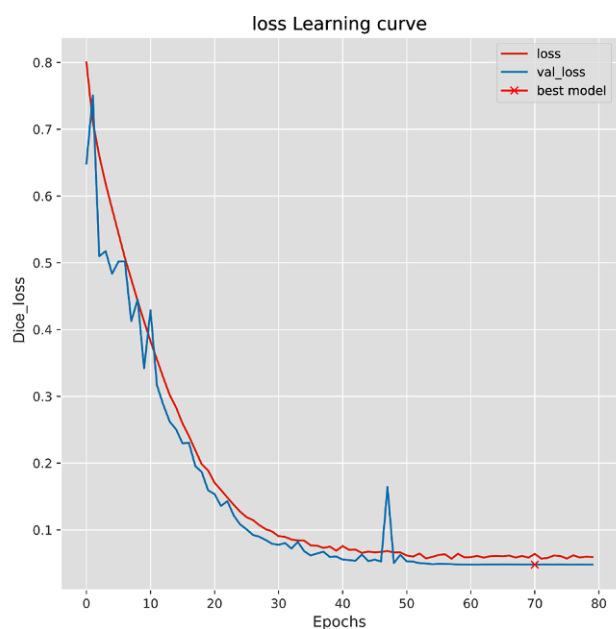


Fig. 5 The curve of the Dice loss over the epochs

model of three examples (with and without *A. caucalis*) of test sub-images of 2018 in winter wheat. Binary predicted points (pixels) in Fig. 6 right are predicted weed-infested points after applying a threshold of 0.25 to the probable predicted space. The closed red lines in Fig. 6 represent actual weed-infested regions that were previously manually labeled.

The mean IoU index and the Dice coefficient of the test data set were calculated by averaging the IoU and Dice coefficient of each sub-image and class. Fig. 7 shows the

results of the mean IoU index and the mean Dice coefficient at the three different dates from the same sampling points in the field in 2018. At the corresponding sampling dates, these metrics indicate a high percentage of agreement between the annotated *A. caucalis* image pixels and the model's predicted pixels.

Validation of the Segmentation Model (Test Images 2019)

The validity and generality of the developed pixel-wise segmentation model in 2018 were assessed on the sub-images from March 29, 2019. Taking into account the predicted weed pixels and the manually annotated weed pixels, the Dice similarity coefficient (DSC) and IOU index were calculated. The average values of the DSC- and IOU-index was 0.9605 and 0.9292, respectively. Fig. 8 shows a typical result and comparison of the manually labeled and predicted weed-infested image regions. Binary predicted points (pixels) in Fig. 8 are predicted weed-infested points after applying a threshold of 0.25 to the probable predicted space. The closed red lines in Fig. 8 represent previously manually labeled weed-infested regions. The results show that the model is acceptably valid for the classification of the images from 2019, which means that the detection system was generalized in a promising way.

Conclusions

A. caucalis is a undesirable weed species that grow in crops, competing for elements such as sunlight, water and nutrients, causing losses to crop yields. Site-specific herbicide application to cereal fields that avoids wasting resources and the destructive environmental effects is highly desirable. If a few individual plants appear, an initial control is needed to prevent further spreading. Color cameras can be attached to agricultural machinery, providing measures such as fertilization, spraying of crop protection products, etc., which can result in comprehensive monitoring of entire fields without gaps.

In this study, to determine *A. caucalis* within color images a fully convolutional and end-to-end trainable network model was developed. After creating a relevant annotated image database, the pixel-wise segmentation model was trained on heavy augmented images from 2018. To evaluate the trained segmentation model, the metrics of the Intersection over Union (IoU) index and the Dice similarity coefficient (DSC) were calculated for the test data set from 2018 and all images from 2019. At all sampling dates, these metrics indicated a high percentage of agreement, over 92% between the target, referred to as the labeled weed-infested areas and the model's predictions. The results suggest that

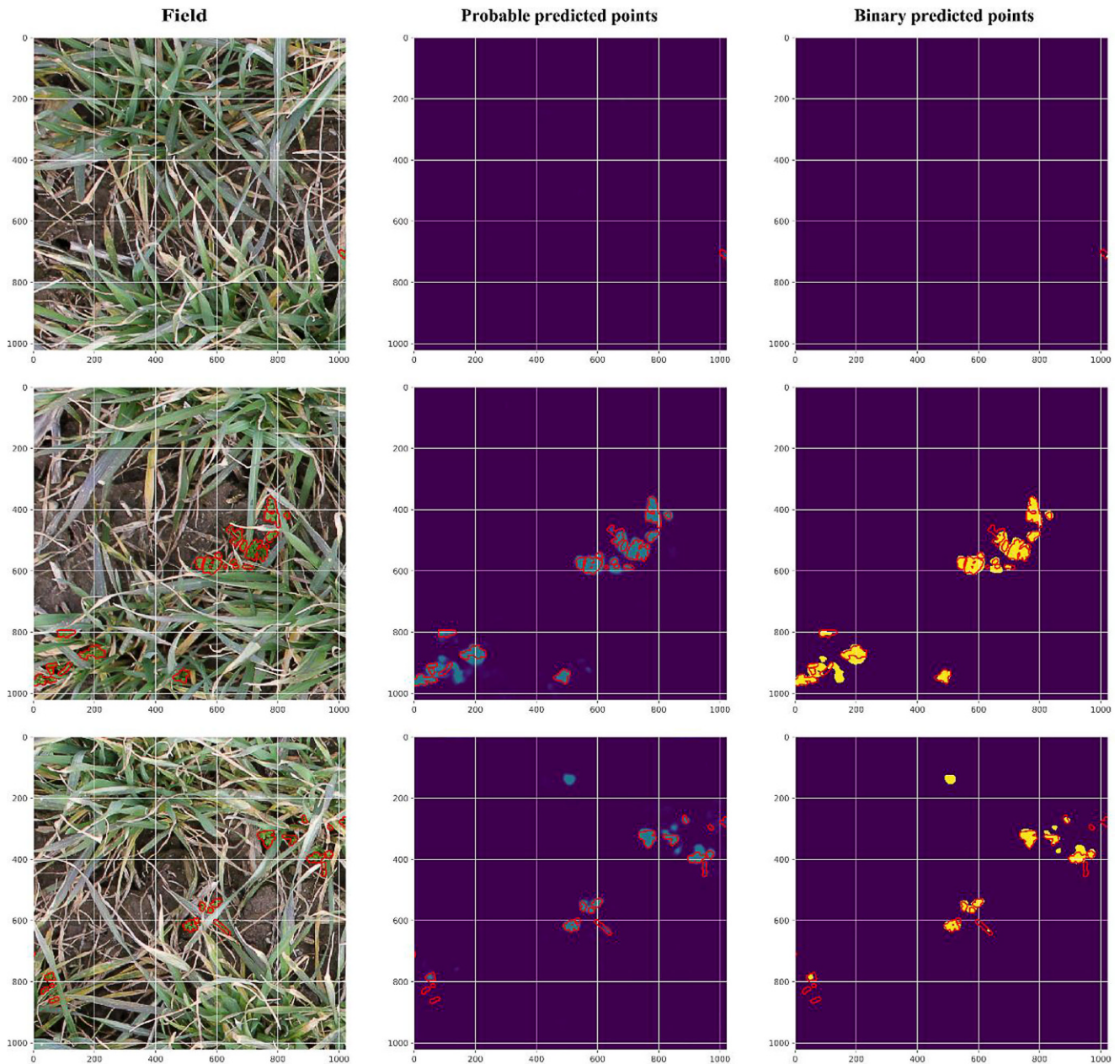


Fig. 6 Field sub-image (left row), probable (middle row), and binary spaces (right row) of predicted weed-infested regions in three examples (with and without *A. caucalis*) of test sub-images of 2018 in winter wheat

the developed pixel-wise segmentation model trained with the 2018 annotated images was favorably generalized and successfully identified image pixels belonging to *A. caucalis*, even within images from the following year with winter rye as a cereal crop. For the 2018 test sub-images, the early sampling date on March 27 showed slightly higher IoU and DSC Index than the two dates in April results (Fig. 7). One reason could be the increasing overlapping of cereal and weed crops with time.

Under cropping conditions like in this study, where only one weed species occurs, the proposed image analysis model offers the following summarized advantages:

- More effective use of available annotated images.
- High accuracy in the following year despite different crop conditions.
- Future suitability of the presented model to run on color cameras with an integrated computer (embedded system)

The spreading of bur chervil is increasing not only in Europe. As invasive species, this weed was reported for

Fig. 7 Mean Jaccard index (IoU) and the Dice coefficient (DSC) calculated for the test data set, collected at three different sampling dates in 2018 along the 2 transects in winter wheat

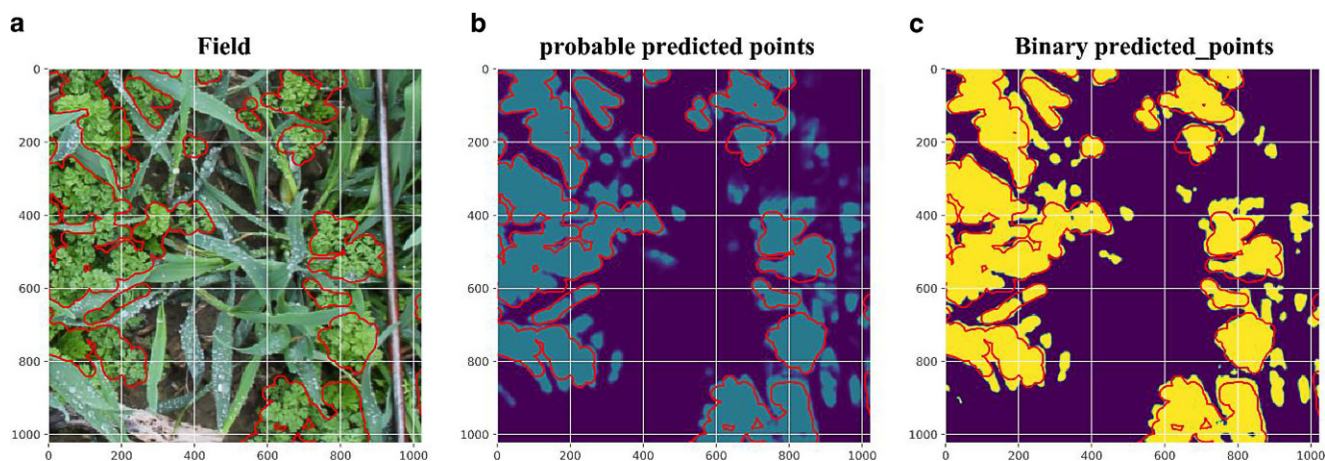
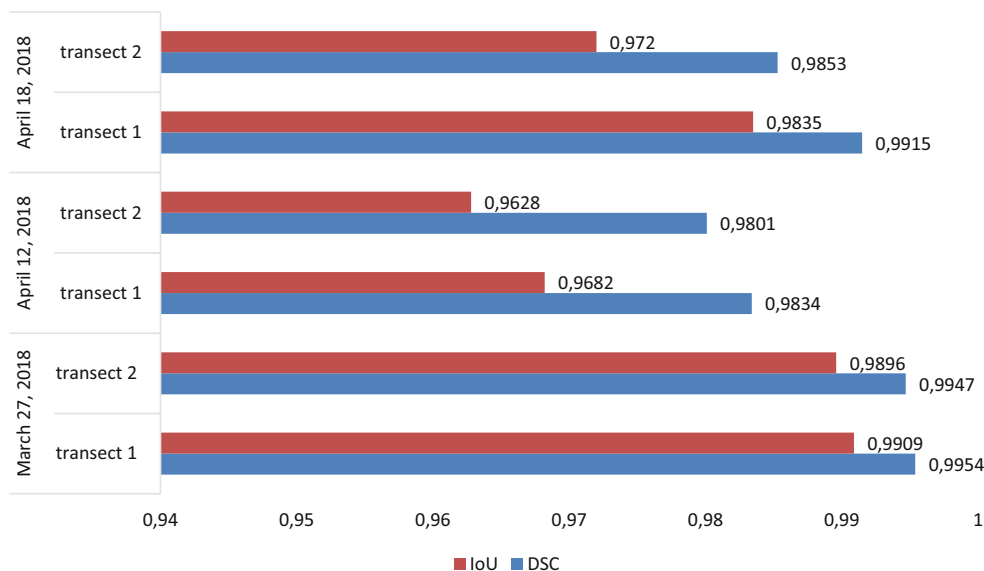


Fig. 8 Field sub-image (a), probable (b), and binary spaces (c) of predicted weed-infested regions in an example (with *A. caucalis*) of test images of 2019 in winter rye

example in Idaho in the western United States (Anonymous 2007). In recent years, the weed has even reached New Zealand (Anonymous 2018) and Tasmania (Rawnsley 2005). Therefore amongst others in future research, the following aspects should be bear in mind:

- Providing labeled image masks for the training of different areas (countries) and crops would help to generalize the *A. caucalis* detection model and to increase its performance in practical operation.
- A chemical control also in later growth stages would make sense to avoid flowering. Collecting image data for training in the later growth stages of the cereal could be reasonable.
- In recent years, unmanned aerial vehicles (UAVs) platforms that provide high-resolution, low-altitude images have shown great potential in weed detection and map-

ping (Torres-Sánchez et al. 2021). UAVs could cover large areas of land in a very short time and periodically provide historical and multispectral image data. In the future, it might be beneficial to study the ability of the developed methods to identify *A. caucalis* within UAV-based images.

Acknowledgements The work was funded by the German Federal Ministry of Education and Research within the joint Iran—German project “PrazisionsPS” (FKZ 01DK20084). We thank Dr. Christian Schulze for providing the field of his agricultural business.

Funding Open Access funding enabled and organized by Projekt DEAL.

Conflict of interest H. Karimi, H. Navid and K.-H. Dammer declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adapta-

tion, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M (2016) Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:160304467. <https://doi.org/10.48550/arXiv.1603.04467>
- Anonymous (2007) Newly reported exotic species in Idaho. In: Research Progress Report of the Western Society of Weed Science. Hilton Portland & Executive—Tower Hotel March 13–15, 2007, ISSN-0090-8142, Portland Idaho, p. 144
- Anonymous (2018) Beaked parsley—have you seen this weed? In: From the ground up. Issue 95 I spring 2018 Foundation for Arable Research PO Box 23133 Hornby Christchurch 8441, New Zealand, pp 12–13
- Anonymous (2022) Amazone UX 5201 21 SmartSprayer. <https://amazone.de/de-de/agritechnica/neuheiten-details/amazone-ux-smartsprayer-997530>. Accessed 27 August 2022
- Bannerjee G, Sarkar U, Das S, Ghosh I (2018) Artificial intelligence in agriculture: a literature survey. *Int J Sci Res Comput Sci Appl Manag Stud* 7(3):1–6
- Chatfield K, Simonyan K, Vedaldi A, Zisserman A (2014) Return of the devil in the details: delving deep into convolutional nets. arXiv.1405.3531, 1–11, <https://doi.org/10.48550/arXiv.1405.3531>
- GIMP Development Team (2019) The GNU image manipulation program: the free and open source image editor (GIMP). <https://www.gimp.org>. Accessed 1 December 2022
- Dyrmann M, Jørgensen RN, Midtby HS (2017) RoboWeedSupport: Detection of weed locations in leaf occluded cereal crops using a fully convolutional neural network. *Adv Anim Biosci* 8(2):842–847. <https://doi.org/10.1017/S2040470017000206>
- Espejo-García B, Mylonas N, Athanasakos L, Fountas S, Vasilakoglou IJC et al (2020) Towards weeds identification assistance through transfer learning. *Computers and Electronics in Agriculture* 171:105306. <https://doi.org/10.1016/j.compag.2020.105306>
- García-García A, Orts-Escobedo S, Oprea S, Villena-Martínez V, García-Rodríguez J (2017) A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857, 1–23. <https://doi.org/10.48550/arXiv.1704.06857>
- Géron A (2019) Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media
- Ghosal S, Zheng B, Chapman SC, Potgieter AB, Jordan DR, Wang X, Singh AK, Singh A, Hirafuji M, Ninomiya S (2019) A weakly supervised deep learning framework for sorghum head detection and counting. *Plant Phenomics*. <https://doi.org/10.34133/2019/1525874>
- Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J (2018) Recent advances in convolutional neural networks. *Pattern Recognit* 77:354–377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- Guan S, Khan AA, Sikdar S, Chitnis PV (2019) Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal. *IEEE J Biomed Health Inform* 24(2):568–576. <https://doi.org/10.1109/JBHI.2019.2912935>
- Karimi H, Navid H, Seyedarabic H, Jørgensen NR (2021) Development of pixel-wise U-Net model to assess performance of cereal sowing. *Biosystems Engineering* 208:260–271, <https://doi.org/10.1016/j.biosystemseng.2021.06.006>
- Kebir A, Taibi M, Serradilla F (2021) Compressed VGG16 auto-encoder for road segmentation from aerial images with few data training. *Conference Proceedings ICCSA'2021*, 98 pages. <https://ceur-ws.org/Vol-2904/42.pdf>
- Kurbiel T, Khaleghian S (2017) Training of deep neural networks based on distance measures using RMSProp. preprint arXiv:170801911. <https://doi.org/10.48550/arXiv.1708.01911>
- Lin Z, Guo W (2020a) Sorghum panicle detection and counting using unmanned aerial system images and deep learning. *Front Plant Sci* 11:1346. <https://doi.org/10.3389/fpls.2020.534853>
- Lin Z, Guo W (2020b) Sorghum panicle detection and counting using unmanned aerial system images and deep learning. *Front Plant Sci* 11:534853. <https://doi.org/10.3389/fpls.2020.534853>
- Lottes P, Behley J, Chebrolu N, Milioto A (2018) Stachniss C Joint stem detection and crop-weed classification for plant-specific treatment in precision farming. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp 8233–8238 <https://doi.org/10.48550/arXiv.1806.03413>
- Lottes P, Behley J, Chebrolu N, Milioto A, Stachniss C (2020) Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming. *J Field Robotics* 37(1):20–34. <https://doi.org/10.1002/rob.21901>
- Mordvintsev A, Abid K (2014) Opencv-python tutorials documentation. Obtenido de. https://opencv24-python-tutorials.readthedocs.io/_downloads/en/stable/pdf/. Accessed 1 December 2022
- Pouyanfar S, Sadiq S, Yan Y, Tian H, Tao Y, Reyes MP, Shyu M-L, Chen S-C, Iyengar S (2018) A survey on deep learning: algorithms, techniques, and applications. *ACM Comput Surv* 51(5):1–36. <https://doi.org/10.1145/3234150>
- Rawnsley R (2005) A study of the biology and control of *Anthriscus caucalis* and *Torilis nodosa* in pyrethrum. University of Tasmania
- Ronneberger O, Fischer P, U-net BT (2015) Convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer, Berlin Heidelberg, pp 234–241 <https://doi.org/10.48550/arXiv.1505.04597>
- dos Santos Ferreira A, Freitas DM, da Silva GG, Pistori H, Folhes MT (2017) Weed detection in soybean crops using ConvNets. *Comput Electron Agric* 143:314–324. <https://doi.org/10.1016/j.compag.2017.10.027>
- Schmidhuber J (2015) Deep learning in neural networks: an overview. *Neural Netw* 61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Solomon RW (2009) Free and open source software for the manipulation of digital images. *Am J Roentgenol* 192(6):W330–W334
- Stirenko S, Kochura Y, Alienin O, Rokovyj O, Gordienko Y, Gang P, Zeng W (2018) Chest X-ray analysis of tuberculosis by deep learning with segmentation and augmentation. In: 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO). IEEE, pp 422–428 <https://doi.org/10.1109/ELNANO.2018.8477564>
- Torres-Sánchez J, Mesas-Carrascosa FJ, Jiménez-Brenes FM, de Castro AI, López-Granados FJA (2021) Early detection of broad-leaved and grass weeds in wide row crops using artificial neural networks and UAV imagery. *Agronomy* 11:749. <https://doi.org/10.3390/agronomy11040749>
- Tseng K-K, Zhang R, Chen C-M, Hassan MM (2021) DNetUnet: a semi-supervised CNN of medical image segmentation for su-

- per-computing AI service. *J Supercomput* 77(4):3594–3615. <https://doi.org/10.1007/s11227-020-03407-7>
- Van Der WS, Colbert SC, Varoquaux G (2011) The NumPy array: a structure for efficient numerical computation. *Comput Sci Eng* 13(2):22–30
- Weng Y, Zhou T, Li Y, Qiu X (2019) Nas-unet: Neural architecture search for medical image segmentation. *IEEE Access* 7:44247–44257. <https://doi.org/10.1109/ACCESS.2019.2908991>
- Wu X, Xu W, Song Y, Cai M (2011) A detection method of weed in wheat field on machine vision. *Procedia Eng* 15:1998–2003. <https://doi.org/10.1016/j.proeng.2011.08.373>
- Xu Y, He R, Gao Z, Li C, Zhai Y, Jiao Y (2020) Weed density detection method based on absolute feature corner points in field. *Agronomy* 10(1):113. <https://doi.org/10.3390/agronomy10010113>
- Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, Wells WM 3rd, Jolesz FA, Kikinis R (2004) Statistical validation of image segmentation quality based on a spatial overlap index. *Acad Radiol* 11(2):178–189. [https://doi.org/10.1016/s1076-6332\(03\)00671-8](https://doi.org/10.1016/s1076-6332(03)00671-8)
- Hossein Navid** works as a professor at the Department of Biosystems Engineering at the University of Tabriz. Hossein's research interests are the application of satellite imaging in agriculture and monitoring the performance of agricultural systems.