Assessment of Ground Conditions in Grassland on a Mower with Artificial Intelligence

Christoph Manss[1], Viktor Martel[2], Roman Weisgerber[2]

Abstract: Process-monitoring for autonomous mowers in agriculture is crucial to establish an online quality assessment. Here, neural networks (NNs) are employed to classify ground conditions, distinguishing between dry, mowed, unplanted, and grass. The data comprises RGB images that are captured by a camera mounted on a mower. These images are then used to train various NNs, with EfficientNet_V2_s emerging as the most accurate network and with ResNet18 to be the most efficient network in terms of training duration and accuracy. The study also reveals for this use-case that employing transfer learning enhances the overall network performance. The developed NNs is intended for deployment on mowers, enabling them to adjust their mowing blades, conserve energy, and enhance the quality of mowed grass. Beyond mowing, the NN can be applied in process control and the identification of other plant species or weeds in the agricultural field, contributing to biodiversity assessments and more sustainable farming practices.

Keywords: Classification, process control, autonomous systems

1 Introduction

Autonomously driving mowers are becoming more relevant in agriculture as they lower the demand of humans that are elsewise required for driving the mowers. Thus, the mowers must detect their surrounding and plan a driving pattern depending on the farmland that must be mowed. Consequently, the focus currently lies on path planning, object detection, and obstacle avoidance [Da20, Ko18]. Yet, the quality of the grass before and after mowing is currently not considered for automation. However, depending on the grass quality (dry, wet, or many unwanted plants) the arrangement of the mowing device can be changed, or the mowing blades can be stopped to save energy during the process. This is particularly useful, if the grassland consists of several dry spots such that mowing is not reasonable as well as when the mower drives over already mowed grass.

[1] DFKI Niedersachsen, Marine Perception, Marie-Curie-Str. 1, 26129 Oldenburg, christoph.manss@dfki.de, https://orcid.org/0000-0003-4851-2622
[2] Maschinenfabrik Bernard Krone GmbH & Co. KG, Heinrich-Krone-Straße 10, 48480 Spelle, Viktor.Martel@krone.de, Roman.Weisgerber@krone.de

Das DFKI Niedersachsen (DFKI NI) wird gefördert im Niedersächsischen Vorab durch das Niedersächsische Ministerium für Wissenschaft und Kultur und die VolkswagenStiftung (FKZ: ZN3683). Die Arbeit entstand im Projekt Agri-Gaia, das durch das Bundesministerium für Wirtschaft und Klimaschutz gefördert wird (FKZ: 01MK21004A).

Additionally, this enables the mower to inject quality assessment into its workflow for automatized mowing.

Currently, remote sensing is utilized to assess the quality of the grassland [DRD21, Mo21]. For example, in [Mo21], the authors reviewed the current research on aboveground biomass (AGB) of grasslands estimation through remote sensing and machine learning. The authors concluded that remote sensing hits a plateau regarding the accuracy of AGB estimation, and that the proximity of the sensor system to the ground is most influential for the accuracy. On the detection of mowing events with remote sensing, the authors in [DRD21] also identified similar limitations. In another study [Zu20], the authors detect the grass length at the mower to control the motor blades in the mower by using a random forest algorithm that considers seven different features of a built-in 9-axis inertia measurement unit. Yet, for bigger mowing devices this approach can be too sensitive to noise.

Therefore, in this paper, RGB images that have been collected with a camera on top of the cabin of a mower are used to detect ground conditions of the grassland in front of the mower. The considered conditions are dry, mowed, unplanted, and grass. The data was collected during different seasons and over two years to ensure that both, relatively dry and wet weather conditions appear in the data. For the classification of these ground conditions neural networks (NNs) are trained and evaluated. We used networks readily available in Pytorch and optimized the augmentations with the help of SMAC [Li22]. To understand the classification results, methods of explainable AI are used to further understand the classification results.

2 The Dataset

The dataset[3] was detected with a camera that was mounted on the top of a mower. The height of the camera was about 4m above ground, and the camera was facing forward with an angle of 45°. In , the view of the camera is displayed. As the mower uses three blade rows, the recorded image is split into three equally large sub-images, which address one of the mowing blades; the left most image corresponds to the left mowing blade, etc. As the same neural network will be used for all sub-images, a dataset is created that comprises only of the sub-images. The classes are addressed by sorting the images into folders with the name of the corresponding class (dry, mowed, unplanted, and grass). At the end, 10.213 sub-images for grass, 670 sub-images for dry, 4.195 sub-images for mowed, and 198 sub-images for unplanted have been collected. Figure 2 displays examples of the classes. The classes dry and unplanted are underrepresented in the dataset, because no mowing has been done on very dry fields and likewise no mowing has been done on

[3] The dataset has been published on Zenodo.org under the title „forefield_grassland“. DOI: 10.5281/zenodo.10371371

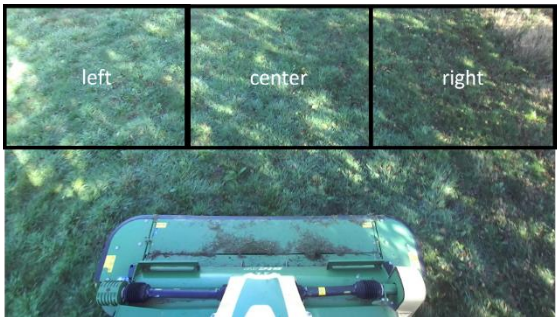

Figure 2: View of the RGB camera from the top of the mower. The black boxes indicate the regions within the image, which are cropped and passed to the neural network.

Figure 1: Example images of all classes in the dataset from left to right: dry, mowed, unplanted, and grass.

unplanted fields. Yet, these categories can be observed and are therefore part of the dataset. So far, the classes are superficial and are currently only intended to test the system and its performance, but the number of classes can also be extended in the future.

3 Development of the Artificial Intelligence

To make the implementation of NNs easily reproduceable and available, we focused on using NNs that are already implemented in Torchvision (Version 0.14.1). The tested networks are ResNet18, ResNet34, ResNet50 [He16], MobileNet_V2 [Sa19], MobileNet_V3_Small and MobileNet_V3_Large [Ho19], and EfficientNet_V2_s [TL21]. All networks are trained from scratch and with the provided weights in Torchvision, which are based on ImageNet [Ru15].
The input images are resized to squared images with a pixel size of 256 pixel and are also casted to float values with a range from 0.0 to 1.0. To optimize the augmentation of the data during the training of the networks, the optimizer SMACV3 [Li22] was employed with the search space: brightness, contrast, hue, and saturation. As a result, the brightness and the hue are uniformly varied within a range of 0.2, and the contrast and saturation are uniformly varied within a range of 0.1.
All networks are trained on a Nvidia V100-DGX with 32GB RAM. Each network was trained for 50 epochs, with the Adam optimizer, a batch size of 64, and a learning rate of 0.0001. For training, the data was split into 80% for training and 20% for validation.

4 Results

The results are presented in Table 1, where all entries are sorted according to the F1-Score in descending order. The F1-Score can be seen as the harmonic mean of the precision and recall scores. The F1-Score can also account for the class imbalance of the dataset. From Table 1 it can be concluded that using transfer learning increases the performance of the network on a different classification task. Here, the gain is about 1-2% in precision and recall.

In terms of the performance with respect to the F1-Score, the EfficientNet_V2_s is the best neural network. However, it is not the best regarding the duration it took to train the network. The duration for training is connected to the computational complexity of a neural network, and a low computational complexity often contradicts a good accuracy, as a smaller neural network is less able to learn complex features in the data. Thus, in the last column, we divided the F1-Score by the duration for training to get a quantification for performance vs complexity. Then, a high number indicates an efficient network. Accordingly, a good choice of a fast network with relatively low computational complexity would be the ResNet18 or the MobileNet_v3_Small with pretrained weights.

Model	Pretrained	Duration [h:mm]	Precision	Recall	F1-Score	F1-Score/duration
EfficientNet_V2_s	yes	2:20	0,9913	0,9912	0,9912	10,2
MobileNet_V3_large	yes	00:49	0,9908	0,9908	0,9908	29,1
ResNet34	yes	00:50	0,9908	0,9908	0,9908	28,5
MobileNet_v2	yes	00:47	0,9886	0,9885	0,9885	30,3
MobileNet_V3_small	yes	00:40	0,9882	0,9882	0,9882	35,6
ResNet18	yes	00:31	0,9879	0,9879	0,9879	45,9
ResNet50	yes	01:03	0,9874	0,9869	0,9870	22,6
ResNet18	no	00:31	0,9785	0,9761	0,9768	45,4
MobileNet_V3_large	no	00:48	0,9759	0,9755	0,9754	29,3
MobileNet_V3_small	no	00:40	0,9739	0,9735	0,9736	35,1
MobileNet_V2	no	00:46	0,9708	0,9699	0,9701	30,4
EfficientNet_V2_s	no	02:20	0,9696	0,9696	0,9695	10,0
ResNet34	no	00:50	0,9682	0,9673	0,9668	27,8
ResNet50	no	01:30	0,9625	0,9614	0,9608	15,4

Table 1: Training results of all trained network. The training results are sorted according to the F1-Score in descending order.

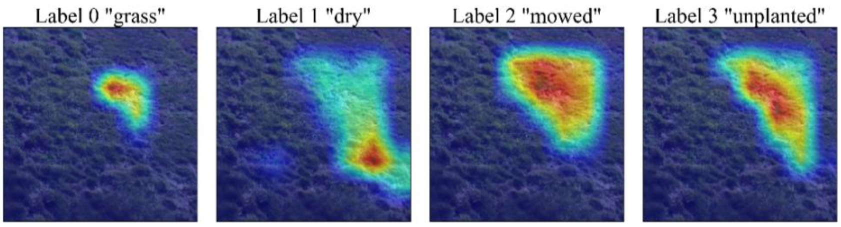

Figure 3: Saliency map for all classes created with GradCam++ for a misclassified example. Here, the true class of this example is *dry*, but the network classified it as *mowed*, with prediction accuracies of 0.39% and 0.996%, respectively. Thus, the saliency map yields stronger features for *mowed* instead for the other classes.

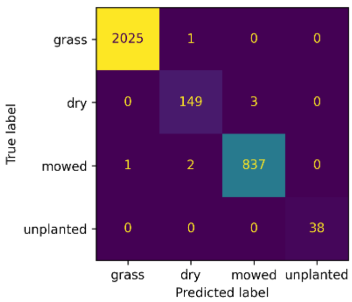

Figure 4: Confusion matrix of the validation set.

Explainability methods provide further insights on the detections of the NNs. Here, we used GradCam++ [Ch18][4], but other methodologies are also possible such as [He23]. GradCam++ analyses the second order gradients in the neural network to estimate a saliency map for the input image. The saliency map highlights regions on the input image that are most influential on the class. An example is shown in Figure 3, where the network detects *mowed* although the true label is *dry*. Through GradCam++ we can see that the classes are falsely classified in border region, i.e., an image where two classes are possible or where images have similar features. Here the dry area has also a low grass height and can therefore be seen as mowed. Another reason for false classification is that the classes *dry* and *unplanted* appear similar to the network. However, a misclassification happens rarely as can be seen in Figure 4. It would, however, be helpful to diversify the data more.

5 Conclusion and Outlook

This paper showed preliminary results on training a classification network to assess the quality of the grassland while mowing. With the gathered data, the network is able to achieve a high F1-Score and only few false classifications. We were able to achieve this by using an optimization framework on the augmentation and transfer learning. With the explainability method GradCam++, we identified that false classifications result from two possible classes in the image. Yet, the impact of such false classifications on the application has to be investigated.

In the future, this network will be deployed on the mower. Once the network can coarsely classify the grassland, the network should be improved for a more detailed assessment of

[4] https://github.com/vickyliin/gradcam_plus_plus-pytorch, accessed 14.12.2023

the grassland, e.g. through image segmentation [MGR23]. It is envisioned that the detection results can be used to adjust the mowing blades, reduce the energy consumption, and increase the quality of the mowed grass. Also, we would like to increase the size of the dataset and to increase the number of classes for a more detailed detection. Moreover, this network can be applied for automatized process control, and to detect other plants and weeds within the images like [Sm19] to determine the biodiversity on the field.

6 Bibliography

[Ch18] Chattopadhyay, A., et al.: Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 839–847 (2018).

[Da20] Daniyan, I., et al.: Development and Performance Evaluation of a Robot for Lawn Mowing. Procedia Manufacturing. 49, 42–48 (2020).

[DRD21] De Vroey, M., Radoux, J., Defourny, P.: Grassland Mowing Detection Using Sentinel-1 Time Series: Potential and Limitations. Remote Sensing. 13, 3, 348 (2021).

[He16] He, K., et al.: Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. (2016).

[He23] Hedström, A., et al.: Quantus: An Explainable AI Toolkit for Responsible Evaluation of Neural Network Explanations and Beyond. JMLR. 24, 34, 1–11 (2023).

[Ho19] Howard, A., et al.: Searching for mobilenetv3. Proceedings of the IEEE/CVF international conference on computer vision. (2019).

[Ko18] Korthals, T., et al.: Multi-Modal Detection and Mapping of Static and Dynamic Obstacles in Agriculture for Process Evaluation. Frontiers in Robotics and AI. 5, (2018).

[Li22] Lindauer, M., et al.: SMAC3: A Versatile Bayesian Optimization Package for Hyperparameter Optimization. Journal of Machine Learning Research. J. Mach. Learn. Res., 23, 1–9 (2022).

[MGR23] Manss, C., Guenther, I., Rahe, F.: Cropline and Weed Detection in Corn Using Weak Supervision. In: Bramer, M. and Stahl, F. (eds.) Artificial Intelligence XL. pp. 377–390 Springer Nature Switzerland, Cham (2023).

[Mo21] Morais, T.G., et al.: The use of machine learning methods to estimate aboveground biomass of grasslands: A review. Ecological Indicators. 130, 108081 (2021).

[Ru15] Russakovsky, O., et al.: ImageNet Large Scale Visual Recognition Challenge, http://arxiv.org/abs/1409.0575, (2015).

[Sa19] Sandler, M., et al.: Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE conference on computer vision and pattern recognition. (2018).

[TL21] Tan, M., Le, Q.: Efficientnetv2: Smaller models and faster training. International conference on machine learning. PMLR, (2021).

[Sm19] Smith, L.N., et al.: Weed classification in grasslands using convolutional neural networks. In: Applications of Machine Learning. pp. 334–344 SPIE (2019).

[Zu20] Zushida, K., et al.: Estimation of Lawn Grass Lengths based on Random Forest Algorithm for Robotic Lawn Mower. In: 2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE). pp. 1628–1633 (2020).