Autonomous Mapping of Grass-Clover Ratio Based on Unmanned Aerial Vehicles and Convolutional Neural Networks

Larsen D[1], Skovsen S[2], Steen A. K[1], Grooters K[1], Eriksen J[3], Green O[1], Jørgensen N. R[2]

[1]AGROINTELLI, Aarhus, Denmark

[2]Aarhus University, Department of Engineering, Aarhus, Denmark

[3]Aarhus University, Department of Agroecology, Aarhus, Denmark

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract. *This paper presents a method which can provide support in determining the grass-clover ratio, in grass-clover fields, based on images from an unmanned aerial vehicle. Automated estimation of the grass-clover ratio can serve as a tool for optimizing fertilization of grass-clover fields. A higher clover content gives a higher performance of the cows, when the harvested material is used for fodder, and thereby this has a direct impact on the dairy industry. An android application is implemented to make the drone fly fully autonomously and collect images at different locations within the field. In this android application it is possible to specify what location the drone should collect images from, which height, and upload the images to a server, which analyze the data based on a convolutional neural network. The convolutional neural network performs a semantic segmentation and thereby pixelwise classify the different classes: grass, clover, soil and weed. The classification of the pixels is used to determine the final grass-clover ratio. The results, presented in this paper, show that the CNN is able to segment the images into the different classes: grass, clover, soil and weed. It is possible to identify the different classes based on images captured at a height up to five meters. Thus, this paper shows a way to use UAVs to perform mapping of actual clover and grass ratio in dense grass-clover fields.*

Keywords. *Autonomous, CNN, Image processing, Unmanned Aerial Vehicle*

Introduction

Automated estimation of the grass-clover ratio can serve as a tool for optimizing fertilization of grass-clover fields. When the harvested material is used for fodder, cows achieve high performance when the clover content is high. Thereby this has a direct impact on the dairy industry.

Clover and grass are often grown in mixtures, as grass-clover leys with different species increase the yield stability (Eriksen et al. 2014; Frankow-Lindberg et al. 2009) and herbage quality (Phelan et al. 2015) compared to fertilized grass-only leys. This is due to niche complementarity (Nyfeler et al. 2009) and the greater protein content of the clover (Egaard 2009; Suter et al. 2015). The clover is able to fix nitrogen (N) from the air, but if there is enough available N in the soil, the grass will outperform the clover. It is thus possible to change the grass-clover ratio by adding more or less N in fertilizer. In order to implement targeted fertilization in practice, it is important to firstly estimate actual clover content, and secondly optimize fertilization based on clover content.

Previous research within automated clover-content estimation from images include pixel-wise classification of grass-clover images based on color indexes, edge detection, and morphological operations (Bonesmo et al. 2004). Himstedt et al. (2009) used digital image analysis on images of grass-legume mixtures from a pot experiment to determine the relative legume dry matter contribution. The image analysis was used to determine the legume coverage (red clover, white clover, or alfalfa) by applying a sequence of morphological erosions and dilations. In each image, the legumes were manually encircled to determine the actual legume coverage. An improvement of this work was presented in (Himstedt et al. 2012), where images were filtered in hue, saturation, and value (HSV) space and applying color segmentation to separate plants and soil after a sequence of morphological operations. McRoberts et al. (2016) used local binary patterns to estimate the grass fraction from color images converted to grayscale. The images were tiled into 64 × 64 pixel blocks, which were manually labeled as either legumes, grass, or unknown.

Current methods experience uncertainty in the image recognition process. Although the use of morphological operations has been shown to correlate with the clover content in the images, it lacks robustness with regards to parameters such as field conditions, scale invariance and estimation uncertainty. Due to varying clover sizes, vegetation heights, or camera resolutions, the parameters of the existing methods need adjustments, to avoid a drop in performance, as seen in the work of Himstedt et al. (2010), where the change from using green house pots to field conditions reduced performance.

In this paper, a system, which can provide support in determining the grass-clover ratio based on images from an unmanned aerial vehicle, is presented. The system is an extension of the work presented in (Skovsen et al. 2017), and thus, the underlying algorithmic approach is similar to the one presented in (Skovsen et al. 2017). A deep convolutional neural network, based on the fully convolutional network (FCN) architecture (Long et al. 2015), is utilized to directly classify relevant plant species visible in the images (Dyrmann et al. 2016; Mortensen et al. 2016). The convolutional neural network is designed to output a semantically segmented image, specifying the plant species of every pixel in the image. Based on the detected composition of coverage of grasses, clovers, and weeds in the image, the clover content is estimated. The system consists of an android application, which is used to design a route plan for the UAV, and make sure that the images are captured at the desired locations and height. Captured images are processed after the image acquisition process, based on automated image upload from the android application.

The paper is organized as follows; the materials and methods section describe the procedure for data acquisition in the parcel trials. Methods for pixelwise segmentation and training of the convolutional neural networks for semantic segmentation are also presented in this section. This is followed by a results section, where both the pixelwise segmentation and its relation to actual grass-clover dry matter is presented. Furthermore, a workflow for utilizing the presented methods in an actual agricultural production, is presented. The paper ends with a discussion and conclusion of the presented work.

Materials and methods

In this section, the process of image acquisition in parcel trials is presented. This is followed by a short description of the methods used for automated pixel-wise segmentation of the captured images. For a detailed description, please refer to (Skovsen et al. 2017).

Data acquisition

The images were captured over two different days, with very different weather conditions. One day, the weather was cloudy, and the other day, the sun was shining from a bright sky. This is clearly visible in the examples shown in figure 1. The images originate from two different locations in Denmark. Images were captured both at Aarhus University, Foulum and at the DLF Seed & Science grass-clover breeding facility in Stevns. UAV images were captured at different heights, to measure the expected drop in performance when the UAV were capturing images at a coarser scale.

(a) (b)

Fig 1. Examples of the images captured by the UAV. (a) Image captured at Foulum, (b) Image captured at Stevns

In Foulum, the images where captured in a field trial, which consisted of 60 grass-clover plots with five different fertilization strategies (0, 50, 100, 200, and 300 kg total-N in catlle slurry/hectare) applied in a randomized design.

In Stevns, a total of 50 plots were selected by visual inspection to maximize the spread in clover fraction, total yield, time since establishment, stress levels, and phenotypes.

At both locations, patches within the plots were manually cut and clover, grass and weed dry matter were measured. The patched were marked with a quadratic visual indicator (visible in figure 1) to enable comparison between the image based, and the manual grass-clover dry matter estimation. As such, only the part of the image inside the visual indicator is analyzed in the results section of this paper.

Semantic segmentation

Semantic segmentation is understanding an image at pixel level, so it is possible to assign each pixel in the image to an object class. An example of this is shown in figure 2. Here, all the pixels in the original colour image has been assigned to an object class, defined by the colour coding in the rightmost image.

(a) (b)

Fig 2. Example of the pixel-wise classification through semantic segmentation. Clover pixels are labeled with orange, grass pixels are labeled with green and weed pixels are labeled with yellow (Skovsen et al. 2017).

As with image and object classification, convolutional neural networks (CNN) have had enormous success within the task of semantic segmentation. In 2014 Fully Convolutional Networks (FCN) by Long et al. (2015) popularized the use of CNN architectures for dense pixel-wise predictions without any fully connected layers. This allows segmentation maps to be generated for images of any size.

The architecture of the network used in this paper is the Fully Convolutional Network (FCN-8s) architecture by Long et al. (2015). This model is a modification of the VGG16 architecture (Simonyan and Zisserman 2014) which is made fully convolutional by replacing fully connected layers with convolutional layers. The network then produces spatial feature maps instead of single labels (as in image classification). The last convolutional layer and the two intermediate layers are followed by deconvolutional layers that upsamples the network output to the same size as the input image. The architecture can be seen in figure 3.

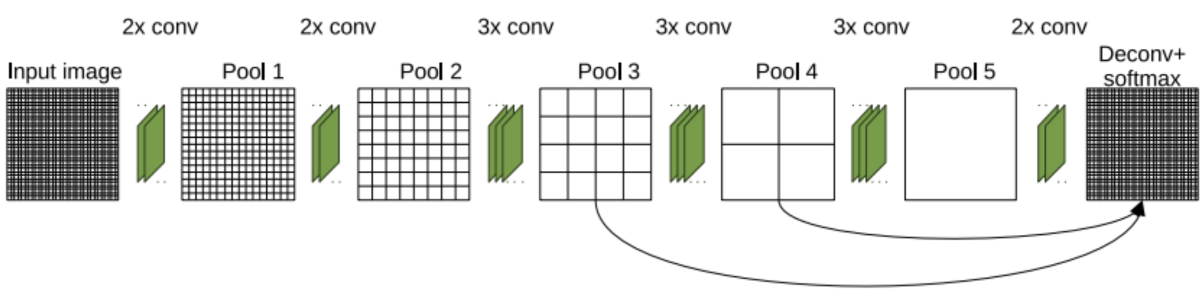


Fig 3. Overall overview of the utilized FCN model (Skovsen et al. 2017)

Model training

CNNs require large amounts of labeled training data in order to adjust the millions of parameters in the model. For semantic segmentation the labeled training data consist of images where each pixel in the image has been labeled as one of the output classes. This is a very time-consuming process, especially with respect to the requirements for the number of labeled images.

As described in (Skovsen et al. 2017), another method of getting labeled images is by generating new, artificial, images from already labeled data. The trained network is based on high quality images, captured in a static setup, with a high-quality DSLR camera (Nikon D810A). The simulated data from these images does not represent the images which are to be captured by a commercially available UAV, such as the DJI Phantom 4. Therefore, new artificial images have been simulated, based on images captured by the UAV, and used in the model training.

Fig 4. Examples of the image crops being used to generate simulated data with known pixel labels.

The artificial images are generated is by cropping examples of the object classes, which are to be detected in the UAV images. In the presented work, the classes are soil, weed, clover and grass. A total of 40 different weeds, clover and grass has been cropped and used to generate images according to the method presented in (Skovsen et al. 2017). Some examples of the different classes used for creating the training images can be seen in figure 4. All of the image crops are based on images from Foulum, and does not include any clover, weed or grass from the Stevns dataset. An example of a simulated image image is shown in figure 5.

Fig 5. An example of a simulated image

Results

This section presents the achieved results. First, a visual illustration of the obtained semantic segmentation is presented. This is followed by an evaluation of the relationship between the pixel-wise estimated grass-clover ratio and the actual dry matter in within the patch. Lastly, the workflow of utilizing these algorithms in real-life scenarios, in the fields, is presented.

Segmentation

In figure 6 an example of a segmentation at 2 meter flying height is shown. In the visual inspection it is seen that most of the grass (green), weed (yellow) and clover (orange) have been labeled correctly. A small part of the image has not been labeled, as the model was uncertain to which class to assign the pixels. However, this is a small part of the image, and often in areas with very dark shadows.

Fig 6. Pixel-wise classification in an image captured by the UAV

As described in the data acquisition section, UAV images were captured at different height to test the performance in various settings. A higher capture height would allow the UAV to cover larger parts of the field with one image, thus improving efficiency with respect to battery usage per hectare. In figure 7 it is seen that the model is able to estimate the presence of clover and grass at all measured heights. It is seen that in low capture heights, more weeds are present in the image. This is because the rotors from the UAV have blown some of the lighter grass to the side, allowing for weeds to be visible.

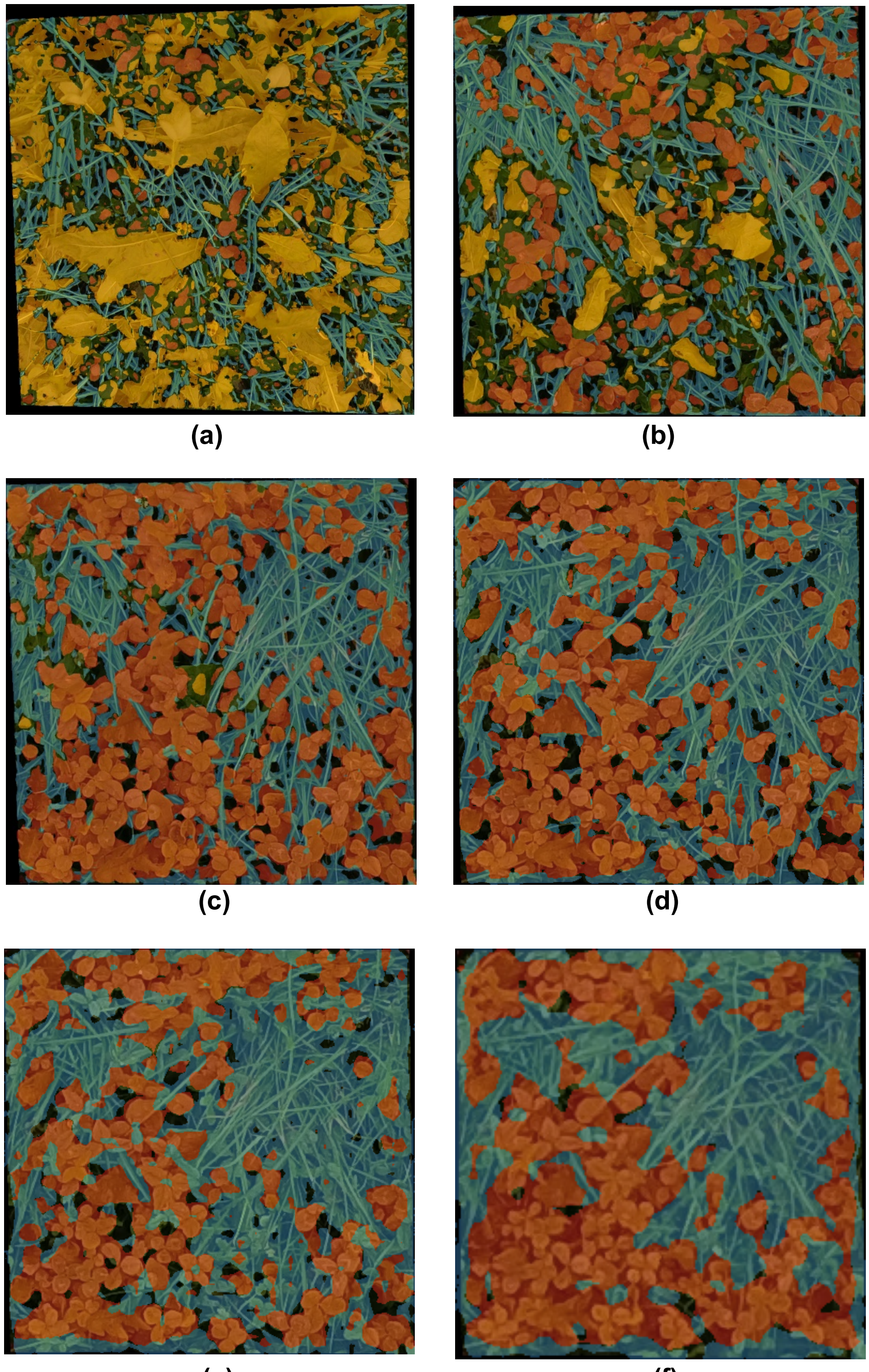
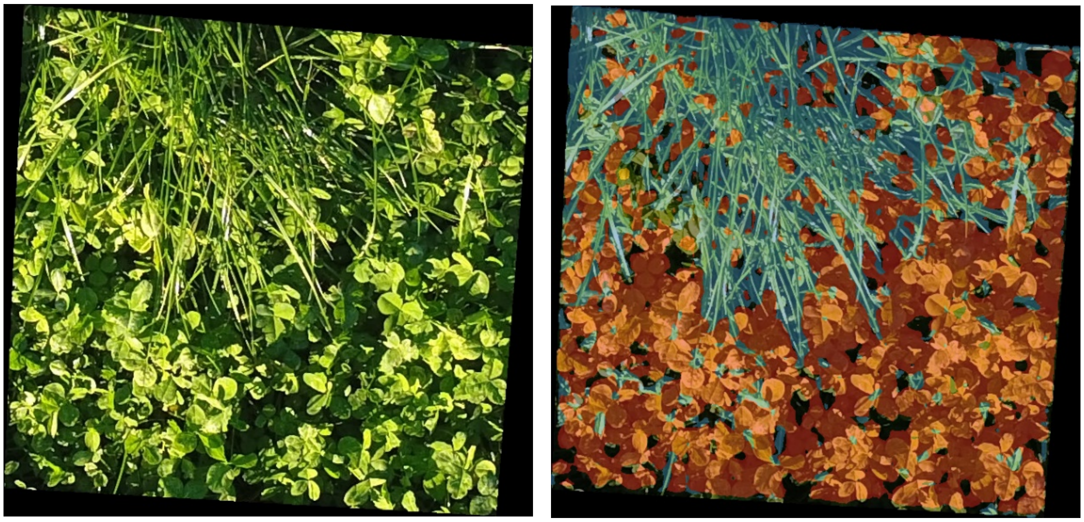


Fig 7. Visual interpretation of results at different acquisition heights. (a) 1 meter (b) 1.5 meter (c) 2 meter (d) 2.5 meter (e) 3 meter and (f) 5 meter

Figure 7 are based on data captured in Foulum, where the weather was cloudy and therefore ideal for capturing images using the UAV. In Stevns, where the other dataset was collected, the sun was shining from a bright sky. This resulted in images with very bright patches and areas with hard shadows. Furthermore, the model was not trained on simulated data from Stevns images. In figure 8, and example from the segmentation in the Stevns dataset is shown. Here, it is seen that the model correctly estimates a large part of the clover correctly, and only as small part of the clover underneath the grass have been falsely labeled as grass.



(a) (b)

Fig 8. Pixel-wise classification in UAV image from Stevns

Using the method in real-life

In the visual interpretation of the segmentation presented above, some pixels are not classified correctly. Further development could and should improve this performance. However, with respect to the agricultural production two parameters is of interest; namely the actual dry matter within the field and the knowledge of where to apply more or less fertilizer.

In all plots, a small part of the plot has been manually cut, and the dry matter of grass, weed and clover has been measured. The small part, indicated by a visual indicator, has likewise been used as input to the CNN. In figure 9, the relationship between the manually measured dry matter clover fraction and the automatically estimated pixel-based clover fraction, is shown. The pixel-based clover fraction are results from capture heights of 2-4 meters, which provided the most consistent result throughout the parcel tests. From the graphs it is seen that there is a relation between the dry matter and the estimated clover content, however especially in the Foulum dataset, this linear relation is not very strong. In the Stevns dataset, the measured data is more distributed between low and high clover content, and this relation is also captured by the automated image processing.

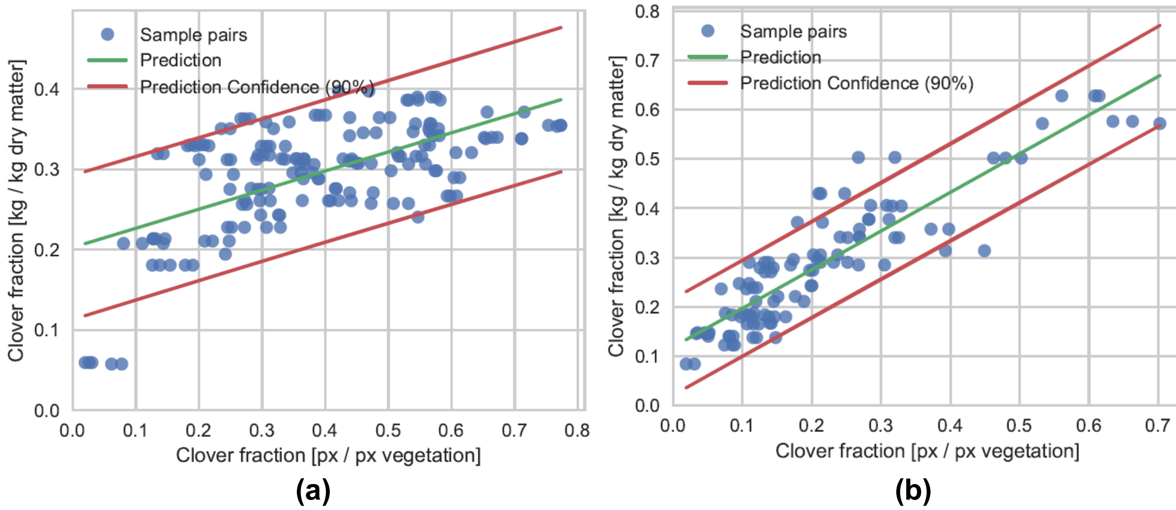


Fig 9. Relationship between measured dry matter clover fraction and estimates pixel-based clover fraction in Foulum (a) and Stevns (b)

The data that is currently available shows that is it currently not possible to estimate the ratio of clover dry matter, based on UAV based analysis. However, it is possible to detect the ratio between visible clover, grass and weed, which can be utilized to manage the relative fertilization strategy within a given field. Based on this, an android application has been developed in order to automate the process of generating grass-clover ratio maps of agricultural fields instead of field patches.

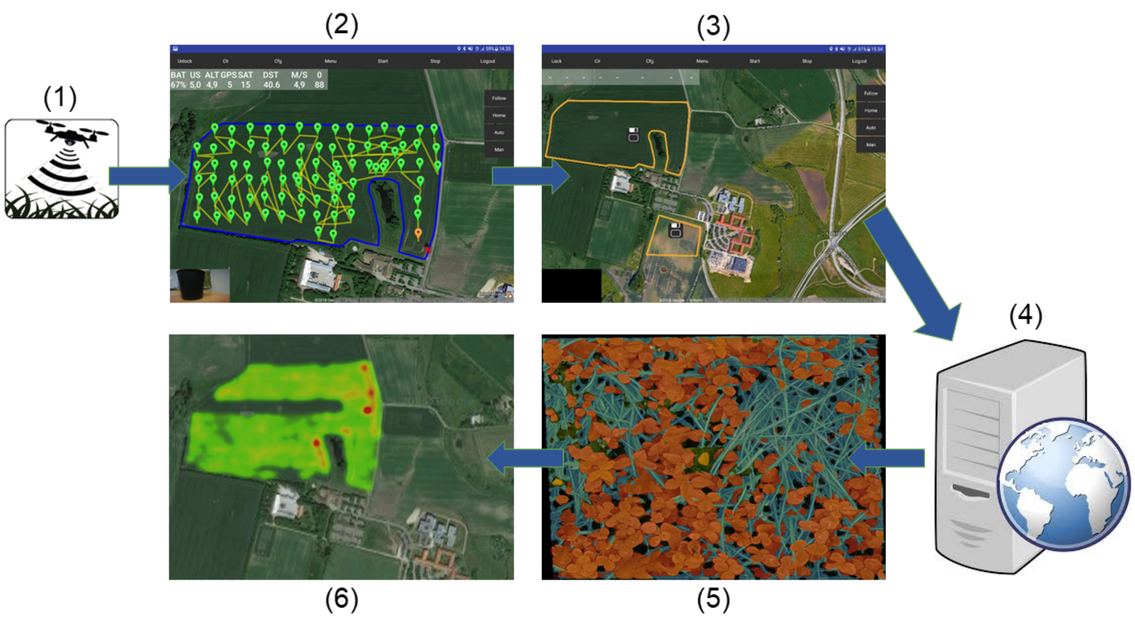


Fig 10. Dataflow of utilizing the presented model together with a commercially available UAV. The flight plan and context are controlled by an Android application and the analysis is done via cloud computing. The result is a grass-clover ratio map, which can be utilized for management of fertilization

The dataflow is depicted in figure 10. The android application (1) automatically generate a flight plan with waypoints for the location of image captures (2). The boundary of the fields is either pulled from an online database or manually generated. After the flight operation, where images are captured at a defined height at every waypoint, the field information is stored in the application (3). Based on this information, the images, with their GPS coordinates, are uploaded to a computing server in the cloud (4). Here all images are processed using the presented model (5), and all resulting grass-clover ratios are collected into a resulting grass-clover ratio map (6).

Discussion

The proposed system is capable of performing pixel-wise classification of grass-clover images into four classes: grass, weed, clover and soil. As the image acquisition height is increased, it is seen that more miss-classifications are introduced. In order in increase performance in CNN models, more training data is useful. A procedure for simulating training data have been presented in this paper, however the current method only utilizes clover, grass and weeds from one field at one type of weather condition. In the future, this could be increased.

As presented in the result section, it is currently not possible to predict the actual ratio of clover dry matter within the fields. It was seen that the relation between the actual dry matter and the, from image processing, estimated fraction, was different for the two locations. As seen in (Skovsen et al. 2017) are clear relation between the ratios have been shown. However, the results presented by Skovsen et al. (2017) arise from data collected at an earlier stage in the season, and they are therefore not directly comparable. More research needs to be done to clarify if and when the UAV can be used to estimate the clover dry matter ratio. Mortensen et al. (2017) presents a method for incorporating climate data to improve upon dry matter estimation. This could also be part of further development for utilizing UAVs for this task.

Based on the visual interpretation of the resulting segmentation, it is seen that the UAV images are suitable for estimating whether a part of the field has more or less clover than other parts of the field. Using this knowledge, it is possible to build a clover ratio map, which can be used for the subsequent fertilization strategies. The workflow for obtaining this has been presented and is currently based on commercially available UAVs coupled with on-line cloud computing.

Conclusion

This paper has demonstrated how a UAV can be utilized to capture images of grass-clover fields and estimate the ratio between clover and grass visible in the image. The presented model is able to detect clover in images from the UAV at acquisition heights up to 5 meters. A workflow for utilizing this has been presented. The workflow consists of an automated waypoint route plan for a commercially available UAV, which captures images at a predefined height. Based on these images, a grass-clover ratio map is generated. This map can subsequently be part of the fertilization strategy within a given field.

References

Bonesmo, H.; Kaspersen, K.; Bakken, A.K. Evaluating an image analysis system for mapping white clover pastures. Acta Agric. Scand. Sect. B—Soil Plant Sci. 2004, 54, 76–82.

Dyrmann, M.; Mortensen, A.K.; Midtiby, H.S.; Jørgensen, R.N. Pixel-wise classification of weeds and crops in images by using a fully convolutional neural network. In Proceedings of the International Conference on Agricultural Engineering, Aarhus, Denmark, 26–29 June 2016.

Egaard, K.S. Nitrogen fertilization of grass/clover swards under cutting or grazing by dairy cows. Acta Agric. Scand. Sect. B—Soil Plant Sci. 2009, 59, 139–150.

Eriksen, J.; Askegaard, M.; Søegaard, K. Complementary effects of red clover inclusion in ryegrass-white clover swards for grazing and cutting. Grass Forage Sci. 2014, 69, 241–250.

Frankow-Lindberg, B.E.; Halling, M.; Höglind, M.; Forkman, J. Yield and stability of yield of singleand multi-clover grass-clover swards in two contrasting temperate environments. Grass Forage Sci. 2009, 64, 236–245

Himstedt, M.; Fricke, T.; Wachendorf, M. Determining the contribution of legumes in legume-grass mixtures using digital image analysis. Crop Sci. 2009, 49, 1910–1916.

Himstedt, M.; Fricke, T.; Wachendorf, M. The relationship between coverage and dry matter contribution of forage legumes in binary legume-grass mixtures. Crop Sci. 2010, 50, 2186–2193

Himstedt, M.; Fricke, T.; Wachendorf, M. The benefit of color information in digital image analysis for the estimation of legume contribution in legume-grass mixtures. Crop Sci. 2012, 52, 943–950.

Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440

McRoberts, K.C.; Benson, B.M.; Mudrak, E.L.; Parsons, D.; Cherney, D.J. Application of local binary patterns in digital images to estimate botanical composition in mixed alfalfa-grass fields. Comput. Electron. Agric. 2016, 123, 95–103.

Mortensen, A.K.; Dyrmann, M.; Karstoft, H.; Jørgensen, R.N.; Gislum, R. Semantic segmentation of mixed crops using deep convolutional neural network. In Proceedings of the International Conference on Agricultural Engineering, Aarhus, Denmark, 26–29 June 2016.

Mortensen, A. K., Karstoft, H., Søegaard, K., Gislum, R., & Jørgensen, R. N. (2017). Preliminary Results of Clover and Grass Coverage and Total Dry Matter Estimation in Clover-Grass Crops Using Image Analysis. Journal of Imaging, 3(4), 59.

Nyfeler, D.; Huguenin-Elie, O.; Suter, M.; Frossard, E.; Connolly, J.; Lüscher, A. Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. J. Appl. Ecol. 2009, 46, 683–691.

Phelan, P.; Moloney, A.P.; McGeough, E.J.; Humphreys, J.; Bertilsson, J.; O'Riordan, E.G.; O'Kiely, P. Forage Legumes for Grazing and Conserving in Ruminant Production Systems. Crit. Rev. Plant Sci. 2015, 34, 281–326

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

Skovsen, S., Dyrmann, M., Mortensen, A. K., Steen, K. A., Green, O., Eriksen, J., ... & Karstoft, H. (2017). Estimation of the Botanical Composition of Clover-Grass Leys from RGB Images Using Data Simulation and Fully Convolutional Neural Networks. Sensors, 17(12), 2930.

Suter, M.; Connolly, J.; Finn, J.A.; Loges, R.; Kirwan, L.; Sebastià, M.T.; Lüscher, A. Nitrogen yield advantage from grass-legume mixtures is robust over a wide range of legume proportions and environmental conditions. Glob. Chang. Biol. 2015, 21, 2424–2438