Wheat ear detection in RGB and thermal images using deep neural networks

Željana Šarić, Oskar Marko, Marko Panić, Sanja Brdar and Vladimir Crnojević

BioSense Institute, University of Novi Sad, Serbia

Abstract

The number of farmers who use smartphones is increasing and along with RGB, thermal cameras are becoming more and more available either as smartphone gadgets or as integrated parts of the smartphone. Using them, farmers could have early information on the wheat yield in a few milliseconds. Currently, yield prediction by counting ears and extrapolating the values for the whole field requires ears to be counted manually on the field, which is prone to subjective evaluation, takes a lot of time and requires large human resources. In the case of larger fields samples must be taken from more than one location, which additionally slows down the process and has influence on the accuracy of yield prediction.

Here we present the first results of field phenotyping experiments regarding ear segmentation in wheat based on RGB and thermal imaging. The segmentation algorithm is based on a convolutional neural network with 3-layers, 40 feature maps per layer and 7x7 convolution filters. We selected Parameterized Rectified Linear Unit as an activation function in hidden layers, while SoftMax was used for the output. We used binary masks with manually segmented ears as the ground-truth for training and assessment of the results. Network was trained with the learning rate of 0.1 and cross-entropy cost function. The study showed that this architecture was the optimal trade-off point between accuracy and time required for calculations. The size of filter was chosen in consideration to the size of visible ears in images. Too big dimensions of filters have a negative effect on preservation of important details, while too small dimensions do not provide information that is distinctive enough. To quantify the performance of image segmentation methods and validate the algorithm we used Sorenson-Dice index (F1 score) that essentially measures similarity between two objects. In our case it was defined as the size of the overlap between blobs in manually and automatically segmented images, divided by the total size of the objects.

Proposed deep learning architecture for ears segmentation achieved F1 score of 85.51% in thermal and 63.37 % in RGB images. Thermal imaging was more appropriate for ear detection due to substantial temperature difference between ears and other parts of the plant that is the result of processes happening inside the ear and its big thermal capacitance compared to the stem and leaves. This makes ears more visible and thus much easier to detect in thermal than in RGB bands. In the places where ears were well exposed and did not overlap each other, the algorithm showed high accuracy. Segmentation errors were made in the events of ear occlusion or if the size of ears were relatively small and need to be addressed in the next research stages. Ear detection is just the first step towards the wheat yield prediction. Future experiments will tackle the challenge of correlating the density and the number of detected ears in the images with the actual yield on experimental locations. Development of such comprehensive system would have a huge impact on the field monitoring and improvement of the decision making in wheat production.