Towards Autonomous Plant Production using Fully Convolutional Neural Networks

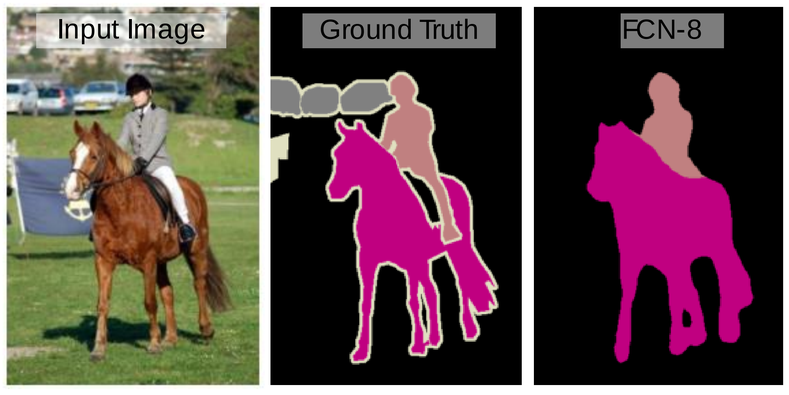
Peter Christiansen[a], René Sørensen[a], Søren Skovsen[a], Claes D. Jæger[b], Rasmus Nyholm Jørgensen[a], Henrik Karstoft[a] and Kim Arild Steen[b]

[a]Dep. of Engineering, Aarhus University, DK
[b]AgroIntelli, Aarhus, DK

1. Problem and aim:

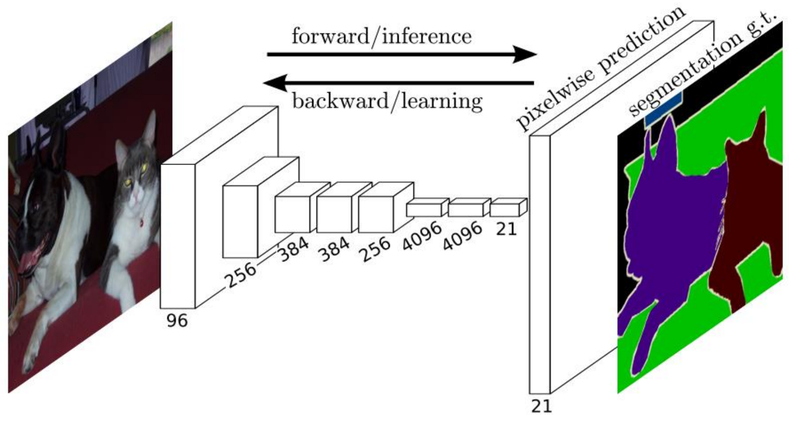
- **Aim**: Obstacle detection for autonomous agricultural vehicles using an RGB camera. Object detection using a fully convolutional neural network for per-pixel classification (semantic segmentation) using existing data.
- **Problem 1**: Object detection algorithms are not able to precisely delimit shelterbelt, ground, crop and water with a bounding box. Secondly, object detection algorithms are challenged in agriculture, where obstacles are likely to be heavily or partly occluded by the crop.
- **Problem 2**: Deep learning, semantic segmentation requires the comprehensive task of providing whole scene per-pixel labelling on a large data set.

Abstact: To avoid the task of creating per-pixel labelled data we investigate using a network trained on two already existing non-agricultural databases (ImageNet+Pascal-context). A pre-trained network performs pixel wise classification on 59 classes and remaps to 11 agricultural specific classes (e.g. field, shelterbelts, animal, and human) to investigate the potential of semantic segmentation in agriculture for two case studies: grass mowing and row crop operations.

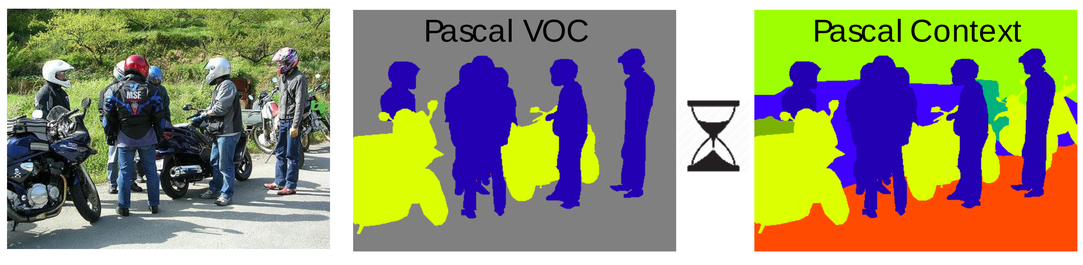


2. Material and method:

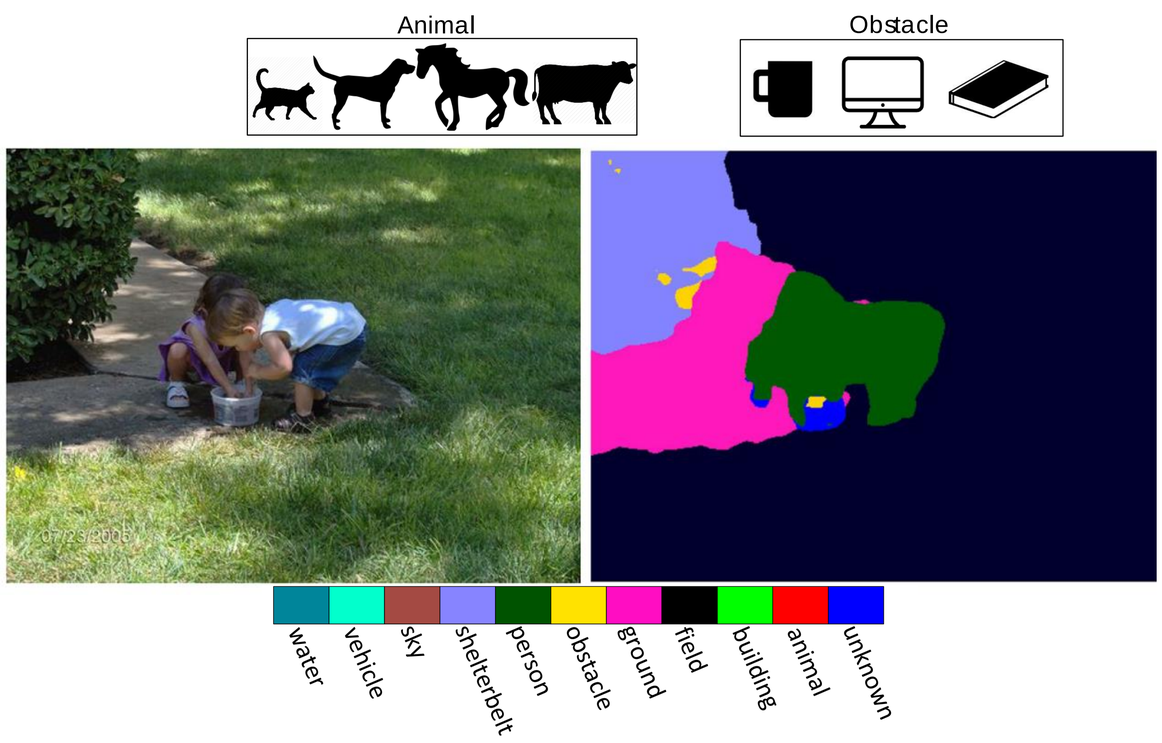
- Using a fully convolutional neural network for semantic segmentation [1].



- Use a model trained on the 59 most frequent classes of Pascal Context

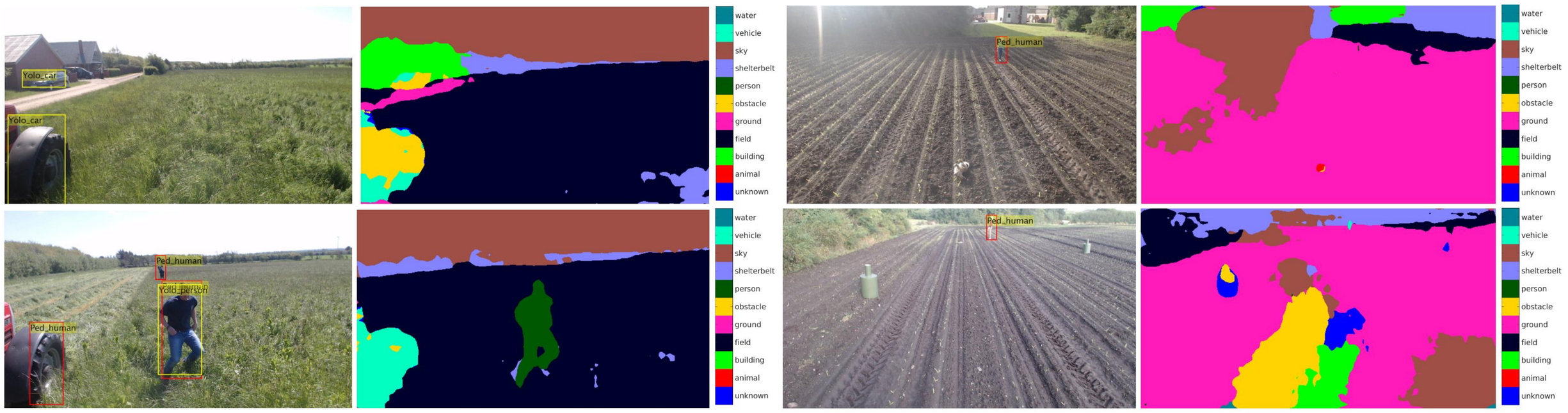


- Remapping of the 59 classes to 11 agricultural super-categories; animal, building, field, ground, obstacle, person, shelterbelt, sky, vehicle, water, and unknown.



3. Results:

- Evaluation is performed using the ground truth annotations for 10 test images taken from a grass and row crop use case.
- The total pixel classification accuracy is 82.81% .
- Evaluating the the grass and row crop individually shows - with a classification accuracy of respectively 95.25% and 70.54% - a significant spread between the two use cases.
- A qualitative comparison to bounding box object detection algorithms also shows the potential of semantic segmentation.



4. Discussion and conclusion:

- This preliminary study uses a simple remapping to show the application of deep learning semantic segmentation for autonomous vehicles in agriculture. For an algorithm trained on a completely different data set, a classification accuracy of 95.25%, show very convincing perceptive capabilities for a grass mowing use case. The row crop use case is less reliable with a classification accuracy of 70.54%. However, the inferior performance in row crops can be explained by the data from PASCAL-Context that do not contain a row crops class. However, we have showed that deep learning semantic segmentation trained on PASCAL-Context is able to generalize to a grass mowing use case, thus allowing us to avoid the comprehensive task of making per-pixel labelling. The preliminary study encourages us to train a new network only on agriculture specific classes from the PASCAL-Context data or alternatively remap all 407 classes to a few agricultural specific classes prior to training. Finally, whole scene annotations of agricultural images would provide even better results. Semantic segmentation is able to detect animal, human and vehicle obstacles - as an object detector. However, the benefit of semantic segmentation is both its ability to classify elements that are not precisely delimited with a bounding box and that it provides much denser information of the environment. This information can be used to detect non-traversable areas such as shelterbelts, water, buildings and even unknown obstacles as the barrel. However, a classification of traversable areas such as road, ground or field is also favorable to autonomous farming vehicles when performing navigation and path-planning. An autonomous vehicle in agriculture should - as in the automotive industry - rely on multiple algorithms and sensor technologies to get more reliable perception of especially visually hidden or camouflaged obstacles and obstacles at far ranges. However, the low cost of a camera and the power of deep learning perceptive algorithms makes it consumer affordable for unsupervised autonomous vehicles in agriculture.

5. References:

[1] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. 8 Mar, 2015. “Fully Convolutional Networks for Semantic Segmentation”
[2] P. Christiansen, R. Sørensen, S. Skovsen, C. D. Jæger, R. N. Jørgensen, H. Karstoft and K. A. Steen. 28 Jun 2016. “Towards Autonomous Plant Production using Fully Convolutional Neural Networks”, CIGR