STALKer: Terribly Accurate, Horribly Fast, Deep Learning Pipeline for Automatic Hemp Stalks Segmentation and Grasping Point Detection

Mateusz Piechocki[1], Dominik Pieczyński[1], Marek Kraft[1]

[1] Poznan University of Technology, Institute of Robotics and Machine Intelligence, Poznań, Poland

{mateusz.piechocki, dominik.pieczynski, marek.kraft}@put.poznan.pl

Abstract

The number of people working in the fields or during the harvest is decreasing year by year. That labor requires intense physical activity, often under unfavorable weather conditions such as high temperature, sunlight, dust, or rain. Agricultural work is repetitive in most cases, which makes it a suitable candidate to apply automation. One such example is the disposal of male fiber hemp specimens, which is the cause of the reduction of the seed quality. Currently, this process is performed manually by uprooting or pruning the plant. This method requires planters to oversee the entire crop and seek unwanted plant types. Our approach aims at automating and robotizing this procedure using sophisticated computer vision algorithms, edge computing, and autonomous mobile robots. The developed unmanned vehicle should patrol, detect and neutralize male fiber hemp representatives. The presented poster describes the part of work related to image processing, primarily focusing on the stalk segmentation and grasping point detection tasks. A semantic segmentation model is designed to achieve high throughput based on RGB images from a camera mounted on a robotic arm. The segmented stalks are further processed to select the one closest to the center of the sensor field of view. Then, the straight line is iteratively fitted to the chosen contour to obtain stalk pose - position and orientation in 3D space. The distance to the identified stem is read from a depth camera calibrated with a vision unit. In addition, to attain autonomous onboard operation, the semantic segmentation algorithm is optimized and quantized to fit the resource-constrained, power-limited hardware available in the mobile platform and maximize system performance. Finally, the proposed pipeline's correctness, accuracy, and robustness have been proven by a wide range of simulator experiments, laboratory trials, and field tests completed using the target robotic arm with a self-designed gripper.

Project Purpose

The project goal is a development of an intelligent agricultural robot to increase the quality of fiber hemp seeds. The STALKer pipeline is the last stage of the proposed application, aimed to increase the certainty of detection of a male cannabis individual by segmenting the plant stalk and, in the loop, providing a grasping point pose for the gripper with poison.

Main Objectives

1. Accurate identification of position and orientation of the grasping point on the stalk.
2. Energy-efficient, low-latency method with near real-time processing on edge device.
3. High throughput of the pipeline providing vision corrections to the control loop.

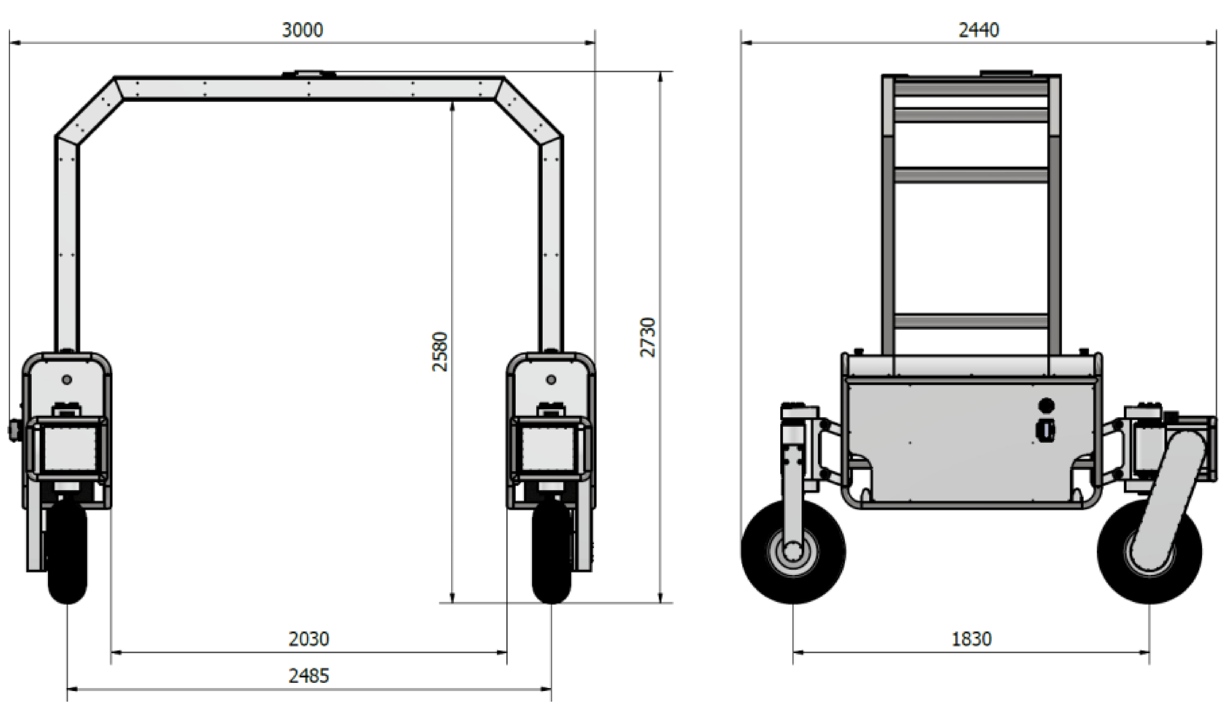

Figure 1: Jarowit mobile robot

STALKer Pipeline

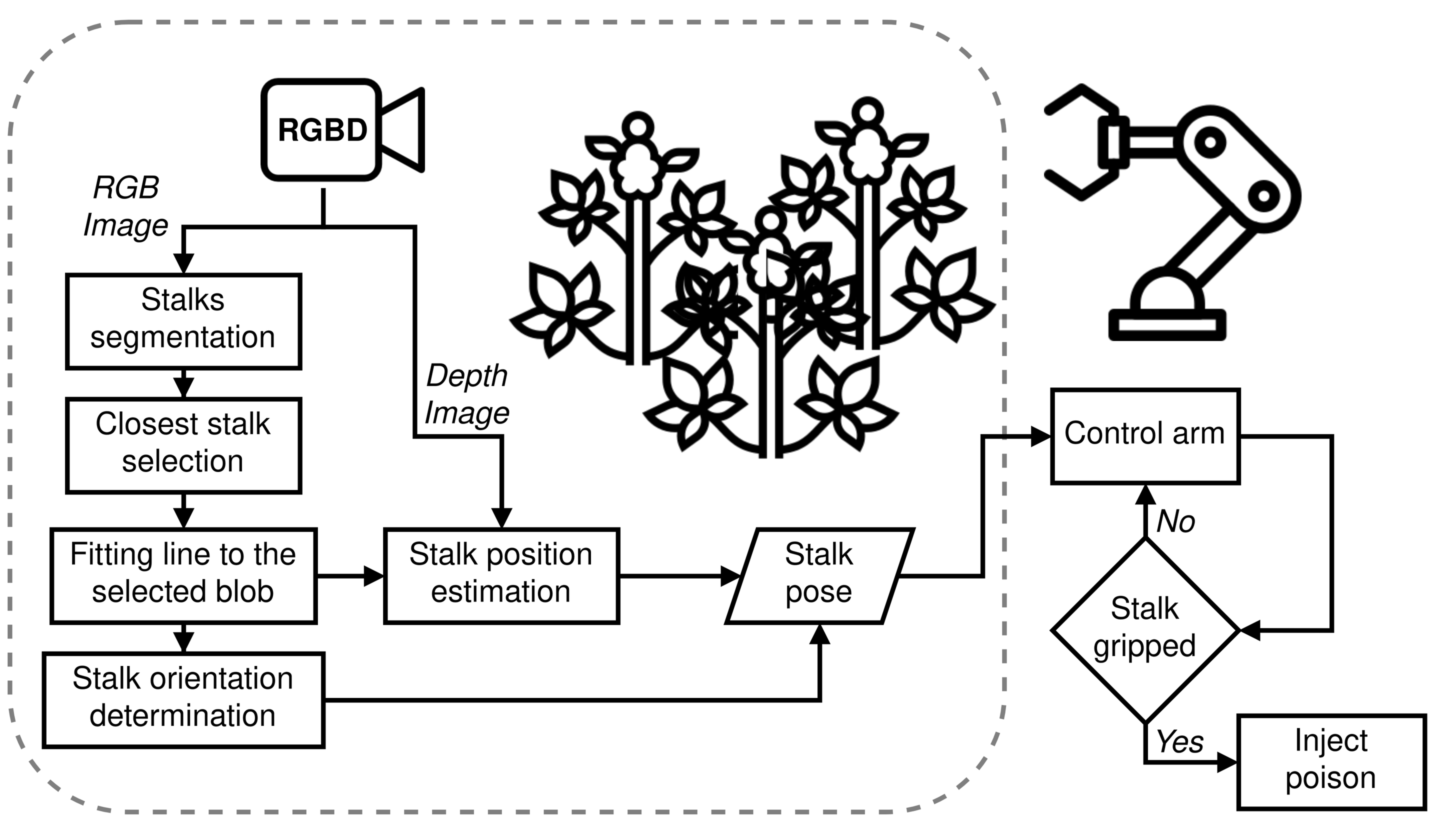

Figure 2: STALKer pipeline

Stalks Segmentation

The hemp stalk segmentation model is based on the UNet architecture [1] with ResNet 18 [2] encoder and Tversky [3] loss function. The dataset for training purposes contains 250 samples and was collected both in the greenhouse and in the field. The achieved neural network metrics are presented in Table 1 whereas Figure 3 depicts model output representing segmented stalks.

Dataset	Dice Score	F1 Score	Precision	Recall
Training	0.6417	0.7350	0.6445	0.8552
Validation	0.6476	0.7302	0.7301	0.7312
Test	0.5734	0.6614	0.6514	0.6731

Table 1: Segmentation model results

Figure 3: Hemp stalks segmentation results

Grasping Point Detection

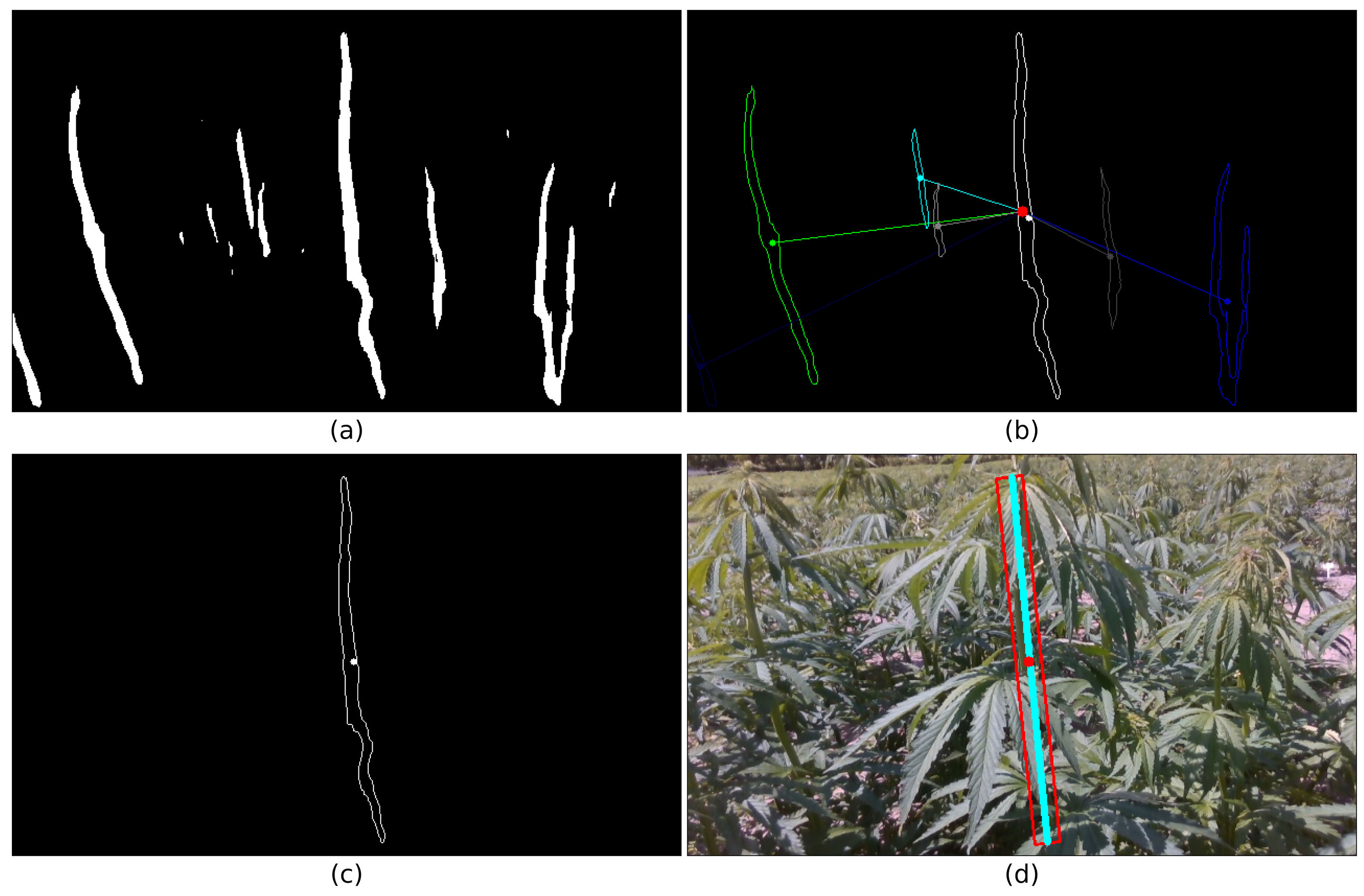

Figure 4: Grasping point detection process steps visualization: (a) neural network segmentation mask; (b) valid stalks contours; (c) closest stalk; (d) selected stalk blob with fitted line presented in RGB frame

Performance Results

The STALKer pipeline is designed to operate under field conditions onboard on the Jarowit mobile platform. Therefore the entire processing loop was benchmarked on NVIDIA Jetson AGX Orin and performance results, divided into individual stages, are presented in Table 2.

Pipeline function	Processing Time [ms]
Stalks Segmentation	28.68 ± 2.89
Closest stalk selection	0.99 ± 0.07
Grasping point calculation	4.26 ± 0.35
Whole pipeline	$\mathbf{33.93 \pm 3.31}$

Table 2: Benchmark results performed on NVIDIA Jetson AGX Orin

References

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[3] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3d fully convolutional deep networks," in *International workshop on machine learning in medical imaging*, pp. 379–387, Springer, 2017.

Acknowledgements

This work was supported by the National Center for Research and Development funding under Project POIR.01.01.01-00-2271/20.