Automating Stone Detection in Agriculture with UAVs and AI based Object Detection

Maik Focke[1], Arnold Lange[1], Norman Rauthe[2], Tobias Hillmann[2], Mark Vehse[1]

[1]University of Applied Science Stralsund, School of Mechanical Engineering
Zur Schwedenschanze 15, 18435 Stralsund, Germany
[2]Neubrandenburg University of Applied Sciences, Faculty of Landscape Sciences and Geomatics
Brodaer Straße 2, 17033 Neubrandenburg, Germany
maik.focke@hochschule stralsund.de

Keywords: Object Detection, AI, UAV, Stone Detection

1 Introduction

Agricultural technology is evolving through the adoption of integrated data systems that streamline operations and foster innovation. This study details the collaborative development of a comprehensive data storage and processing system. This described system integrates data from diverse sources such as satellites, agricultural machinery, business software, and Unmanned Aerial Vehicles (UAVs), bolstering data accessibility and analytical capabilities. An example use case implemented is Stone Detection. This process addresses the problem of stones typically larger than 15 cm that surface during plowing and pose a threat to harvesting equipment. Conventionally, these stones are manually detected and removed to prevent equipment damage. This work aims to create the basis for future automation of the process. UAVs are utilized to map farming areas and collect images.

2 Image Acquisition

The process initiates by converting field boundaries into DJI (SZ DJI Technology Co., Ltd.) mission files for precise UAV mapping. Integrated services implement the DJI Cloud API (Application Programming Interfaces) to deploy these missions directly to pilots into their DJI Pilot app. Pilots execute these missions using drones such as the DJI Mavic 3M. After flight, the system automatically uploads and assigns the captured media to its respective field. The images were taken from a relative altitude of approximately 30m above ground. With an image sensor resolution of 5280×3956 pixels, an area of 15cm×15cm spans approximately 18×18px. Since down-sampling the images would result in significant information loss, they are sliced for further processing.

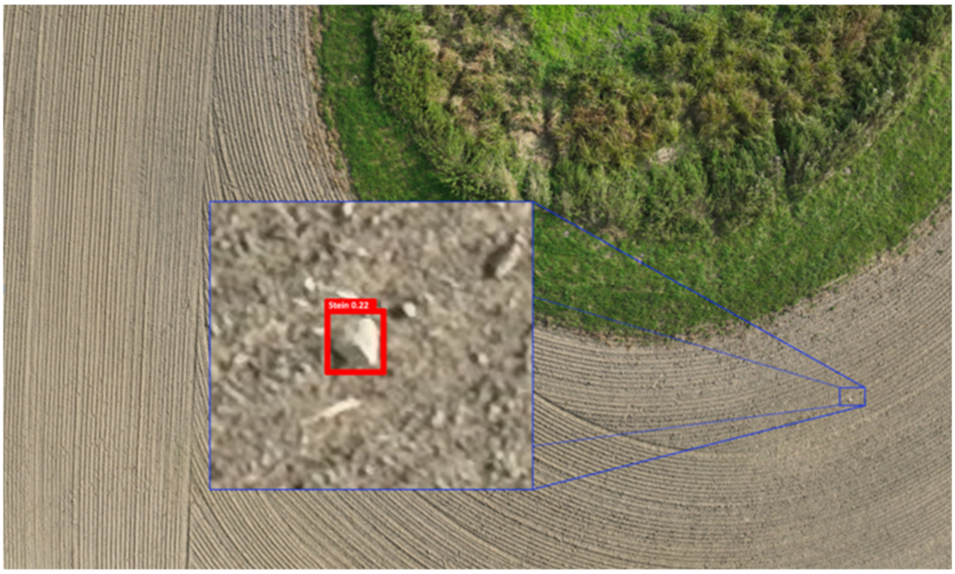

Figure 1: Visualization of a detected Stone in an aerial image taken by a DJI Mavic 3M

3 Object Detection

To detect stones a dataset was created and annotated manually. This dataset was used to fine tune a Conditional-DETR (End-to-End Object Detection with Transformers) [1] object detection model pre-trained on the COCO dataset (Common Objects in Context) [2]. The resulting model has shown good performance on the stone detection validation dataset and even disclosed flaws in it, as a manual review of seemingly false positives has revealed.

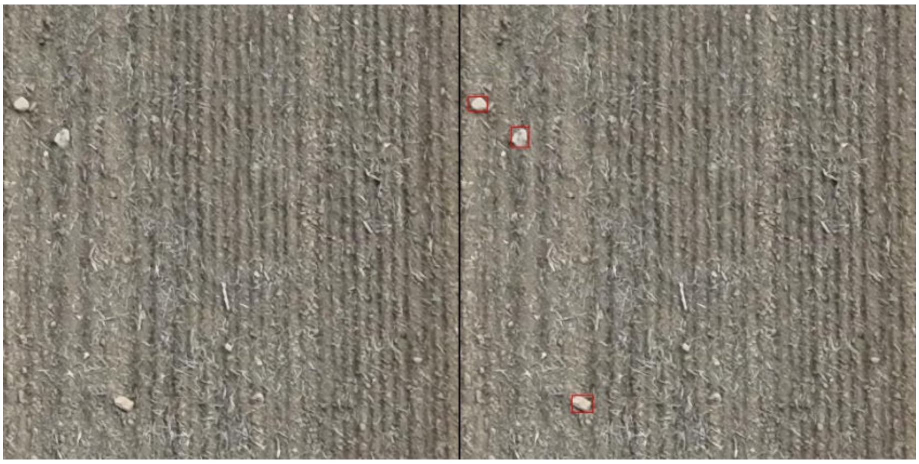

Figure 2: Raw image slice (left) and bounding boxes of detected stones (right)

4 Coordinate Transformation

Image coordinates of detected stones are transformed into GPS (Global Positioning System) coordinates. The image contains vendor specific metadata such as the UAV's position and camera angle (NED) to calculate the camera extrinsic as well as calibrated parameters to calculate the intrinsic. For simplification the ground model is a horizontal plane that contains the UAV's starting point and allows to use the relative altitude metadate.

5 Process Integration

This infrastructure not only supports current manual operations but also serves as a foundational step for forthcoming projects, aimed at developing robotic solutions for autonomous stone collection. Affordable and readily available UAVs are deployed to efficiently gather the necessary data, which will guide the design of a demonstrator that outlines the essential steps for automating stone collection.

6 Conclusions

The necessary steps from flight planning to stone detection and localization can be performed. The detection results are promising, though the dataset requires revision. Since the validation dataset was annotated by a human expert based on visual features in the images, there are uncertainties due to assumptions made from their experience and the low resolution of objects. These uncertainties introduce an error margin in the measured accuracy, which cannot be precisely estimated. To address this issue, a method needs to be developed for more timely and accurate annotation and verification of the images.

7 Acknowledgement

BigData4AF is a part of the WIR!-Alliance "ArtIFARM – Artificial Intelligence in Farming". The authors gratefully acknowledge the Bundesministerium für Bildung und Forschung (FKZ: 03WIR4804) for financial support.

8 References

[1] Meng D., Chen X., Fan Z., Zeng G., Li H., Yuan Y., Sun L., Wang J.: Conditional DETR for Fast Training Convergence. arXiv:2108.06152v3 [cs.CV], https://arxiv.org/abs/2108.06152, 29 Sep 2023

[2] Lin, TY. et al..Microsoft COCO: Common Objects in Context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, 2014